Table of contents

Abstract

The majority of the world's children grow up learning two or more languages. The study of early bilingualism is central to current psycholinguistics, offering insights into issues such as transfer and interference in development. From an applied perspective, it poses a universal challenge to language assessment practices throughout childhood, as typically-developing bilingual children usually underperform relative to monolingual norms when assessed in one language only. We measured vocabulary with Communicative Development Inventories for 372 24-month-old toddlers learning British English and one Additional Language out of a diverse set of 13 (Bengali, Cantonese, Dutch, French, German, Greek, Hindi-Urdu, Italian, Mandarin, Polish, Portuguese, Spanish and Welsh). We furthered theoretical understanding of bilingual development by showing, for the first time, that linguistic distance between the child's two languages predicts vocabulary outcome, with phonological overlap related to expressive vocabulary, and word order typology and morphological complexity related to receptive vocabulary, in the Additional Language. Our study also has crucial clinical implications: we have developed the first bilingual norms for expressive and receptive vocabulary for 24-month-olds learning British English and an Additional Language. These norms were derived from factors identified as uniquely predicting CDI vocabulary measures: the relative amount of English versus the Additional Language in child-directed input and parental overheard speech, and infant gender. The resulting UKBTAT tool was able to accurately predict the English vocabulary of an additional group of 58 bilinguals learning an Additional Language outside our target range. This offers a pragmatic method for the assessment of children in the majority language when no tool exists in the Additional Language. Our findings also suggest that the effect of linguistic distance might extend

beyond bilinguals' acquisition of early vocabulary to encompass broader cognitive processes, and could constitute a key factor in the study of the debated bilingual advantage.

Vocabulary of 2-Year-Olds Learning English and an Additional Language: Norms and Effects of Linguistic Distance

Chapter 1 – Introduction

Parents eagerly await the moment their one-year-old infant produces their first word, and professionals working with young children track the appearance of the two-word combination stage at around age 2. To developmental psycholinguists, these two milestones in healthy language acquisition are supported by an impressive range of achievements, from the attuning of perceptual abilities (Werker, Yeung, & Yoshida, 2012), the development of word segmentation skills (Bergmann & Cristia, 2016) and the retrieval of word meaning (Stevens, Gleitman, Trueswell, & Yang, 2017), to the discovery of syntactic, morphological and conversational rules (Gleason & Ratner, 2017; Hoff, 2013). The complexity of the task that young children naturally solve places the study of language development at the heart of the debate about the nature of the human mind (Pinker, 1995; Tomasello, 2009). In addition, the co-occurrence - and interdependence - of language development and that of other domains such as motor coordination (Iverson, 2010), object perception (Jones & Smith, 2005), or social skills (Carpenter, Nagell, Tomasello, Butterworth, & Moore, 1998), indicates that the study of language learning is fundamental to the understanding of child development as a whole.

Even more impressive than an infant acquiring her maternal language is an infant learning two or more languages, an achievement of the majority of the world's children (Kohnert, 2010), including an increasing proportion in the United Kingdom (UK) (17.5% in primary schools; NALDIC, 2012). Although knowledge about language development is mostly built around the monolingual child model, the study of early bilingualism has become a central issue in psycholinguistics. This is because it allows us to address theoretical key questions such as specificity, differentiation, transfer and interference in development, but also because, from an applied perspective, it poses a universal challenge to language assessment practices throughout childhood. Indeed, language development in bilinguals is notoriously difficult to predict, due to a variety of situational factors related to the proportion and properties of dual-language exposure, making the use of monolingual norms largely inadequate for this population (Cattani et al., 2014). This is further complicated by the variety of language pairs being spoken in the world. For example, in the United States (US), whilst Spanish constitutes the most common other-than-English language spoken at home (62% speakers), the remaining 38% speakers share up to at least 350 languages, mostly in large metropolitan areas (United States Census 2020). Given the mounting evidence that children develop their two languages in separate, yet interfering, systems (Hoff, 2013), the degree of influence between two languages should theoretically depend upon the linguistic distance between the two languages.

The central aim of this paper is to investigate the role of linguistic distance in early language development, and through this, we address two key questions: first, we bring new theoretical knowledge by directly testing hypotheses regarding the existence of language-to-language influence in early bilingual development; second, we offer a pragmatic solution to the widespread problem of how to assess bilingual toddlers' language skills when no tools are available in the home language. Although the study was conducted in the UK, the theoretical conclusions we reached about the role of linguistic distance are not community-specific, and the principles behind the assessment tool we have designed can be exported to any other community.

We collected data about vocabulary development and contextual variables in a cohort of 372 bilingual two-year-old toddlers learning British English and one of 13 Additional Languages. We investigated how vocabulary development is modulated by linguistic distance between British English and the Additional Language, bringing new evidence about language-to-language influence in early acquisition, and the first demonstration to date of an effect of structural similarity between languages on early vocabulary development (Study 1). In addition, we generated a vocabulary model able to predict the lexicon size of 2-year-old bilinguals in English and one the 13 target Additional Languages, which fed into a new assessment tool, the UKBTAT (UK Bilingual Toddler Assessment Tool) (Study 2). Finally we showed that this model could reliably predict the English scores of a new cohort of 58 British-English 2-year-olds learning a new, non-target Additional Language (Study 3). In what follows, we review the literature pertaining to the theoretical (Study 1) and applied contributions of this work (Studies 2 and 3). In Chapter 2 we present the methods for the cohort data collection, followed by the analyses pertaining to Study 1 in Chapter 3, Studies 2 and 3 in Chapter 4, ending with Chapter 5 where we discuss how linguistic distance shapes language development and the limiting factors of this work.

Study 1: Understanding the Role of Linguistic Distance in Bilingual Development

Most bilingual studies are conducted with homogeneous samples, with Spanish-American English bilinguals constituting the largest cohort (e.g., Marchman, Fernald, & Hurtado, 2010; Place & Hoff, 2011), followed by Canadian English-French (e.g., Paradis, Crago, Genesee, & Rice, 2003), Barcelona Catalan-Spanish (e.g., Bosch & Sebastián-Gallés, 2001) and Welsh-British English (e.g,. Gathercole & Thomas, 2009). Although it is often reasonable to generalise results obtained with one language pairing to different language pairs, is it safe to assume that bilingual children from different backgrounds are confronted with the same linguistic problems to solve, or that they are able to solve them within the same learning span?

There is growing evidence that, according to the language they acquire, monolinguals' learning paths can differ, including for early lexical prosodic processing (Adam & Bat-El, 2009), word segmentation (across dialects: Floccia et al., 2016; Nazzi, Mersad, Sundara, Iakimova, & Polka, 2014), and phonological processing (Nazzi, Floccia, Moquet, & Butler, 2009; Mani & Plunkett, 2007; Delle Luche, Floccia, Granjon, & Nazzi, 2016; Bouchon, Floccia, Fux, Adda-Decker, & Nazzi, 2015), culminating in differences in vocabulary growth (Bleses et al., 2008; Thordardottir, 2005). Bilingualism is likely to exacerbate these language-specific differences, adding not only a new language but also the complexity of interactions between languages.

Why should we expect interactions between the two language systems in a bilingual child? In adult bilinguals, evidence for automatic cross-language activation in visual word recognition (Lauro & Schwartz, 2017), and in spoken word recognition (Mercier, Pivneva, & Titone, 2014; Spivey & Marian, 1999) strongly point to the interdependence of the two language systems (see French & Jacquet, 2004). A common position nowadays, following the original proposal of a primary, undifferentiated language system (Volterra & Taeschner, 1978), is that early bilinguals develop two independent language systems (Genesee, 1989; Genesee, Nicoladis, & Paradis, 1995). However, accumulating evidence suggest early influences from one language system onto the other, mirroring the structure of the adult system. Interaction between the two language systems is perhaps most obvious in production, where intra-sentential code-mixing, i.e., including elements of each language at the sentence level, is frequently observed in toddlers at the phonological, lexical and morphosyntactic level (e.g., Gildersleeve-Neumann, Kester, Davis, & Peña, 2008; Paradis & Genesee, 1996). More convincing however is non-selective lexical access in comprehension, that is, the fact that speech presented in one language activates word recognition in the two languages, as demonstrated in adults (e.g., Dijkstra, 2005), and observed at least as young as from three years of age (e.g., Poulin-Dubois, Bialystok, Blaye, Polonia, & Yott, 2013; Von Holzen & Mani, 2012; see the review by DeAnda, Poulin-Dubois, Zesiger, & Friend, 2016). However, even if there were an initial separation of lexicons, bilinguals would still demonstrate language-specific differences from the parallel learning of two language systems. For example, French infants rely more on consonants than vowels for lexical processing from the age of 8 months (Nishibayashi & Nazzi, 2016), whereas British English learners process consonants and vowels equally (Mani & Plunkett, 2007; Floccia, Nazzi, Delle Luche, Poltrock, & Goslin, 2014), at least until the age of 30 months (Nazzi et al., 2009), and Danish children rely more on vowels than consonants at 20 months of age (Højen & Nazzi, 2016). How do such differences translate to the case of bilingual learners? Will language-specific routes for vowel-consonant processing be delayed until the onset of separate language processing, or will one linguistic strategy be adopted, at an efficiency cost to the other language?

In sum, each language pairing will necessarily produce a different linguistic learning problem for bilingual infants to solve, which is likely to translate to variable delays and/or adapted pathways (see Polka, Orena, Sundara, & Worrall, 2017, for word segmentation outcomes differing in bilingual and monolingual 8-month-olds). In this project we conducted the first systematic evaluation of the impact of differences between languages, as measured through metrics of linguistic distance, on vocabulary acquisition in both British English and the Additional Language.

Measuring linguistic distance. As adults, second language learning seems easier if the language is intuitively similar to our own (e.g., English/German vs English/Cantonese), which is supported by studies in second language learning for both adults and school-aged children (e.g., Lado, 1957; Lindgren & Muñoz, 2012; Van der Slik, 2010). For example Lindgren and Muñoz (2012) showed that a cognate-based measure of language distance is the most important predictive factor for formal second language learning in schools, above differences in the exposure of the languages at home. These results support the idea that in second-language learning, the knowledge and structure of L1, i.e. the maternal language, provide scaffolding for the acquisition of L2, i.e. the Additional Language (see also the literature on cross-linguistic transfers in second language reading acquisition, e.g., Genesee, Geva, Dressler, & Kamil, 2006).

However, in early simultaneous bilinguals the effects of language distance are more complex as the languages are acquired in parallel from infancy. While similarities between the two languages may reinforce phonological, lexical and syntactic acquisition across the two languages, it would also reduce the perceptual separation between languages. Knowledge of the interaction between reinforcement and separation is crucial to our understanding of bilingual acquisition, but the complexity of multi-dimensional representations of language means that unitary measures of the seemingly intuitive notion of 'linguistic distance' are difficult to evaluate. This complexity is reflected in the many distance metrics which have been proposed, including cognate distance (e.g., Dyen, Kruskal, & Black, 1992), genetic linguistic distance (Harding & Sokal, 1988; Ruhlen, 1987), phonetic distance (Nerbonne et al., 1996), distance in terms of linguistic rhythm (Ling, Grabe, & Nolan, 2000; Ramus, Nespor & Mehler, 1999) and second language learnability (Chiswick & Miller, 2005).

Of all linguistic distance measures, cognate distance is probably the metric that has the widest currency, at least at the lexical level. Traditionally, this refers to the proportion of translation equivalents sharing common historical origins, such as *lait* in French and *leche* in Spanish (*milk*, sharing the Latin root *lac*). In an influential cognate database (e.g., Dyen et al., 1992, adapted by McMahon & McMahon, 2005), the index of linguistic cognate distance is obtained from the compilation of 200 frequent culture-neutral words in 84 Indo-European languages and dialects, and for each language pairing. However, the definition of cognates in Dyen et al.'s database makes it difficult to generalise to languages without a clear common history. Approaches based on automatic methods have been proposed to refine the definition of cognates, for example by using intra-language similarity (Ellison & Kirby, 2006) or cross-language orthographic similarity measures (Serva & Petroni, 2008). While some of these metrics may be suitable for the adult speech environment, child-directed speech differs in lexical, prosodic and pragmatic content from adult-directed speech (e.g., Thiessen, Hill, & Saffran, 2005). Importantly, infants do not share adults' meta-linguistic and orthographic knowledge.

Given the young age of our participants it was necessary to base our distance metrics on a set of child-familiar basal words and to consider phonetic, phonological and metrical similarities rather than historical origins or orthographic properties. To this end we developed a measure of linguistic distance which focussed upon corpora related to toddlers' speech environment. We used the Oxford Communicative Developmental Inventory of Oxford CDI (Hamilton et al., 2000), which supplied us with a list of words that should be known to British English children of our target age group. Here a short introduction to the CDIs is necessary, as we will use these tools not only to develop a metric of phonological overlap, but also to collect vocabulary data from toddlers.

The first Communicative Developmental Inventories, which are parental reports of their children's vocabulary on a checklist of familiar words, were developed for American English children (Fenson et al, 1994), with norms published in two separate forms for different age ranges (8-18 months for the Words and Gestures form, and 16-30 months for the Words and Sentences form). An updated norming sample of 2,550 US children was produced later for the two forms (MacArthur-Bates CDI, Fenson et al., 2007). Crucially, the normed CDI parent reports have been adapted (not translated) in a multitude of languages, with the purpose of mirroring the structure of the reference language as much as possible. The availability of CDIs in many languages has created new opportunities for cross-linguistic studies of language development (see e.g., CLEX database now called Wordbank: Jørgensen, Dale, Bleses, & Fenson, 2009; Frank, Braginsky, Yurovsky, & Marchman, 2017) and for bilingual studies (e.g., Armon-Lotem, & Ohana, 2017; Cattani et al., 2014; Gatt, 2017; Pearson, Fernandez, & Oller, 1993; O'Toole et al., 2017). The reliability and validity of CDIs is long established for use in research (Mancilla-Martinez, Gamez, Vagh, & Lesaux, 2016; Marchman, Thal, Dale, & Reznick, 2006) and for clinical assessment (e.g., Charman, Drew, Baird, & Baird, 2003; Heilmann, Weismer, Evans, & Hollar, 2005; Thal, DesJardin, & Eisenberg, 2007).

To produce vocabularies for the 13 target Additional Languages, we could have used the words listed in each language-specific CDI as a proxy of toddler vocabulary for each Additional Language. However, these CDI vocabularies would largely reflect the cultural and physical environment in which the language was predominantly spoken. As our bilinguals all live in the UK it is likely that both their British English and Additional Language vocabulary would reflect a British English environment, rather than the environment of the monolingual Additional Language CDI. As such, we believe that the bilingual Additional Language lexicon would be better represented by translation equivalents of the words of the Oxford CDI. This approach also has the advantage that it is unaffected by variations in the methodologies used to construct Additional Language CDIs, which can result in wide differences in CDI word counts. Details of these toddler-centric metrics can be found in the method section (Chapter 2). Note that the use of a common CDI inventory across all language groups was restricted to the calculation of a metric of phonological overlap, and not to collect vocabulary knowledge from toddlers, which was performed with language-specific inventories.

With this first metric based on lexical phonological overlap, we expected phonological similarity between languages to facilitate the acquisition of words in each language. This hypothesis is supported by results from Bosch and Ramon-Casas (2014) showing that 18-month-old Catalan-Spanish bilinguals were more likely to produce highly similar cognates than less similar ones.

In addition to measures of linguistic distance based on lexical overlap, we also ranked language pairs on measures of grammatical distance, namely word order typology and morphological complexity. Languages can be broadly distinguished on the basis of the relative order of the main verb (V) and its object (O) (Dryer, 1991; Greenberg, 1963), with VO order as in English and Spanish, OV as in Bengali and Hindi/Urdu and mixed OV/VO for German and Dutch. Children's very first combinations of words show knowledge of this basic word order (e.g., Bates et al., 1984; Brown, 1973), demonstrated even earlier in comprehension (e.g., Höhle, Weissenborn, Schmitz, & Ischebeck, 2001), although the full knowledge of native language word order patterns takes several years to develop (e.g., Abbot-Smith, Lieven, & Tomasello, 2008; Akhtar & Tomasello, 1997; Guasti, 2002). To retrieve information about their language's syntactic typology, monolingual infants have been found to rely on the relative order of function and content words in their first year of life (Gervain, Nespor, Mazuka, Horie, & Mehler, 2008), as well as on prosodic correlates of word order (Bernard & Gervain, 2012; Christophe, Nespor, Guasti, & Van Ooyen, 2003). In bilingual infants, Gervain and Werker (2013) recently showed that when learning languages with opposite word orders, such as English and Hindi, 7-month-old bilinguals would also exploit the appropriate prosodic information, revealing an early sensitivity to cues that can be used to acquire basic word order in their two languages. Once acquired, knowledge of word order would logically boost word segmentation and grammatical parsing, and therefore learning two languages with a common word order might facilitate these processes. We examined whether vocabulary scores in English and in the Additional Language at the age of 2 could be predicted from the similarity of British English and the Additional Language in terms of word order typology.

Finally, languages can also be ranked on a continuum of morphological complexity (Brown, 2010; Comrie, 1989; Greenberg, 1960), with analytic or isolating languages on one end of the continuum (lowest ratio of morphemes to words, as in Cantonese, and to a lesser extent, in English), inflecting/fusional (Russian, Italian, French) as well as agglutinating (Finnish, Basque, Turkish, Hungarian) languages in the middle, and polysynthetic languages on the other end (highest ratio of morphemes to words, such as in Yupik). The most intuitive expectation is that speed of acquisition would vary as a function of the morphological complexity of the to-be-acquired language (see Caselli, Casadio, & Bates, 1999, for this suggestion). Xanthos et al. (2011) found, by contrast, that the more morphologically complex the ambient language, the faster children will acquire morphological rules (see also Leonard, 2000). However, it is possible that the nature of the cognitive resources engaged in language acquisition vary as a function of its morphological complexity, with analytic languages making greater demands on memory than synthetic languages (Fortescue & Lennert Olsen, 1992). In the bilingual situation, acquisition could be boosted when confronted with morphologically close pairs of languages as compared to more distant ones, simply because of a better alignment of cognitive demands. In support for this hypothesis is the study by Paradis (2011) showing that 6-year-old second language learners from immigrant families in Canada had a better command of English verb morphology if their language at home was closer to English in terms of verb morphology typology (marking tense/agreement). Here we evaluated whether vocabulary scores in English and in the Additional Language could be predicted from the proximity of languages in terms of morphological complexity.

Language community and culture. In addition to linguistic distance, each language pairing also comes with a range of cultural and social idiosyncrasies, which are impossible to explore exhaustively and often difficult to disentangle from linguistic factors. For example, cultural differences can relate to the degree to which parents communicate with infants

(Bavin, 1992; Brown, 2001), the characteristics of infant-directed speech (Fernald et al., 1989), the degree of focus on objects in conversations (Tardif, Shatz, & Naigles, 1997), parents' attitude towards bilingualism (De Houwer, 1999), or the level of involvement of siblings in early interactions (Super & Harkness, 2013). Studies which have attempted to examine the combined or separate effects of cultural and linguistic variations in early language development are, to our knowledge, very scarce. One such study by Barac and Bialystok (2012) compared three groups of 6-year-olds on English verbal tasks for receptive vocabulary, grammatical knowledge and metalinguistic awareness, testing Spanish-English, French-English and Chinese-English bilinguals, all attending similar schools in Canada. When controlling for SES and the amount of exposure to each language at home, it was found that Spanish-English bilinguals outperformed French-English and Chinese-English children on verbal tasks. The poorer performance of French-English children was attributed to the fact they were they were the only group whose schooling was in French – therefore resulting in less exposure to English - whereas the other groups received English schooling. More pertinently, the higher performance of Spanish-English bilinguals as compared to the Chinese-English bilinguals was accounted for by the greater linguistic similarity between Spanish and English than Chinese and English; this was, however, a speculative explanation since the authors acknowledged an unclear understanding of what would constitute cultural differences between these communities. Similarly, Bialystok, Majumber and Martin (2003) showed an advantage for metalinguistic awareness in Spanish-English bilinguals at 6-7 years of age, as compared to English monolinguals, with Chinese-English bilinguals performing worse than both other groups. The authors attributed these differences to English and Spanish being closer in terms of consonant-vowel alternation, as compared to the phonological and tonal structure of Chinese, and also to an advantage of Spanish itself provided by its greater phonological transparency. Once again, however, a combination of cultural and linguistic factors might explain the differences, for example if one community engaged more in communication overall, or made a greater use of intra-sentential code-mixing.

Disentangling the effect of linguistic distance from that of language community. It is impossible to account for all variation associated with each language, and a fortiori, for each language pairing. In the current study we examined a large number of language pairs (13 pairs) with British English common across all pairs. Our sample provides variation in terms of linguistic distance and cultural background, as a subsample of all possible pairings. These data were analysed with linear mixed models which include a factor for random variation due to language community, but also linguistic distance as a “between-language” fixed factor. If linguistic distance has an effect then it can be used to usefully explain some of the variance in scores that would otherwise simply be apportioned to language community.

In sum, we discussed how linguistic distance between a bilingual’s two languages can be estimated on three dimensions: phonological overlap, subject-verb order typology and morphological complexity. We also reviewed how current views on the structure of the early bilingual language system can accommodate an effect of linguistic distance on vocabulary size. This was the rationale for our first study, to quantify the amount of variation due to linguistic distance in the development of vocabulary at age two. In what follows we discuss our second, applied aim, which was to build a model of the bilingual early lexicon able to account for the variance due to situational factors related to bilingual experience, linguistic distance and language community, providing norms of vocabulary development for assessment purposes (the UKBTAT).

Study 2: Providing Norms of Vocabulary Development for Bilingual Toddlers

It is well documented that bilinguals generally command a smaller vocabulary in each language than monolinguals (e.g., Bialystok, Luk, Peets, & Yang, 2010; Hoff et al., 2012; Miękisz et al., 2016; Oller & Eilers, 2002; Perani et al., 2003; Portocarrero, Burright, &

Donovick, 2007; Smithson, Paradis, & Nicoladis, 2014). This fact is critical for assessing children's language development because vocabulary size is a central measure of progress in both the oral and literate forms of language (Muter, Hulme, Snowling, & Stevenson, 2004). Indeed, vocabulary size correlates highly with grammatical development (Conboy & Thal, 2006; Hoff, Quinn, & Giguere, 2017; Thal, Bates, Goodman, & Jahn-Samilo, 1997), and is strongly predictive of later language impairment (Conti-Ramsden & Botting, 1999; Dale, Price, Bishop, & Plomin, 2003) and reading comprehension skills (Duff, Reen, Plunkett, & Nation, 2015), making it a reasonable proxy of language development achievements in young toddlers (see also Cattani et al., 2014).

Getting an accurate estimate of the "bilingual difference" in vocabulary size for each language is a prerequisite to adapting existing assessment tools for the evaluation of bilingual toddlers' language achievements. Furthermore, language disorders occur with similar prevalence in both the monolingual and bilingual populations (Kohnert, 2010), with all children having a 7-15% chance of experiencing delayed language acquisition due, for example, to pervasive developmental disorders such as autism (e.g., 2.6%, Kim et al., 2011), hearing impairment (e.g., 1% in Fortnum et al., 2001) or Developmental Language Disorder (e.g., 7%, Tomblin et al., 1997). Current language assessment methods are based upon expectations for monolingual learners, and take no account of comparative delays seen in typical bilingual lexical development (Gathercole, 2007) or grammar (Bedore & Peña, 2008; Bialystok, 2009; Kohnert, 2010). Thus, according to circumstances, bilinguals, who are in reality typically-developing, may be diagnosed with spurious acquisition problems or have genuine problems ignored (e.g., Crutchley, 2000; Salameh, Nettelbladt, Håkansson, & Gullberg, 2002, for evidence of under-referral in bilingual children). Moreover, whilst recommendations to practitioners are that proficiency in both languages should be assessed (American Speech-Language-Hearing Association, 1999; Royal College of Speech and

Language Therapists, 2007), this is in practice complicated by the diversity of language pairs (Cattani et al., 2014; Thordardottir, Rothenberg, Rivard, & Naves, 2006).

An approach advocated by Pearson, Fernandez and Oller (1993) for Spanish-English and Junker and Stockman (2002) for German-English bilinguals is to obtain for each child her total vocabulary in the two languages, either by counting all tokens (Total Vocabulary or TV; e.g., *dog* and its French equivalent *chien* would count as one entry each) or counting two known translation equivalents as one (Total Conceptual Vocabulary or TCV; e.g., *dog* and *chien* would count as one entry). Based on findings that bilingual children score similar to monolinguals when using their TV or TCV measures, it was proposed that bilingual norms might not be necessary, as long as vocabulary could be estimated in both languages. However, Thordardottir et al. (2006) reported that the use of TCV/TV measures was not appropriate for balanced bilinguals, that is, those hearing equal amounts of each language on a regular basis, or children whose language dominance is not clear. They attributed this to the large overlap in knowledge from their two languages, which would modulate the relationship between measures using one language (monolingual norms) or two languages (TV/TCV). In addition and most importantly, they found that the comparison of TCV/TV to monolingual norms was highly dependent on which monolingual group is used for comparison, as vocabulary growth can vary substantially between languages (Thordardottir, 2005), for example when presented with a particularly complex vowel system as in Danish (Bleses et al., 2008). Core, Hoff, Rumiche and Señor (2013) also found that between 22 and 30 months, TCV scores placed significantly more bilinguals below the 25th percentile on monolingual norms than single-language scores did for monolinguals (see also Gross, Buac, & Kaushanskaya, 2014).

Therefore, measuring bilingual vocabulary appears to be an impractical task given the variety of factors that might shape lexical growth in these children, from the variation between language pairs to situational factors such as amount and mode of exposure (e.g., Hoff et al., 2012; Place & Hoff, 2011). The aim of our second study was to provide a new functional approach to the evaluation of vocabulary knowledge in bilingual children, addressing the diversity of bilingual children's situations, in particular, variation in the linguistic distance between each bilingual child's two languages. To our knowledge, this is the first time that data were collected from a large cohort of bilingual children learning a variety of language pairs, in an attempt to capture empirically the effect of language community and linguistic distance on other factors known to modulate vocabulary growth in bilingual children. In addition, by focusing on children who share one language (here, British English), we avoid having to rely on different standardised tools that may vary dramatically across languages (Thordardottir, 2005). Whilst the resulting norms of vocabulary growth are specific to children learning British English as one of their languages, our rationale and methods are, we believe, generalisable to any new population of bilingual toddlers.

Assessing language development: The case of the UK. In the historical context of the British Empire and more recent European Union expansion, the UK bilingual population is characterised by a great diversity of language backgrounds, with no predominant group such as Spanish-English in the US (with the exception of Welsh in North Wales). As such, it can be taken as a representative case study for the widespread situation where the professionals who assess young children have no easy access to an appropriate standardised monolingual assessment tool, let alone a bilingual tool. In what follows we will review briefly the current practices of screening and assessment of bilingual children in the UK.

Recently, the Dynamic Assessment of Preschoolers' Proficiency in Learning English (DAPPLE) was developed in the UK to respond to the clinical need to distinguish between a disorder and the bilingual difference, using a mixed group of bilinguals, that is, children learning English and a variety of Additional Languages (Hasson, Camilleri, Jones, Smith, & Dodd, 2013). This assessment examines the children's ability to learn vocabulary, sentence structure and phonology. This battery of language skills assessments sounds promising as a pre-diagnostic tool but is designed for children aged 42 months; in addition, it has been criticised for issues regarding inter-rater reliability (Hasson & Joffe, 2007) and is usually very time-consuming (De Lamo White & Jin, 2011).

Since 2013 the professionals who assess young children use general developmental questionnaires at age 2 which include language components, such as the revised Ages and Stages Questionnaires (ASQ: Squires & Bricker, 2009). However, each of these tools is designed and normed for monolingual development which, along with the lack of adequate language assessment tools for that age range, means that bilingual 24-month-olds continue to be at risk of under-referral (Crutchley, 2000; Salameh et al., 2002).

The linguistic heterogeneity of the UK regarding its bilingual population, and the clear clinical need, motivated the second and third studies reported in this paper: from the estimation of the impact of language community and linguistic distance on bilingual lexical development obtained in Study 1, we developed a full model of the bilingual lexicon at age 2 which provided norms of development, the UKBTAT (Study 2), and showed that it could be used for bilingual toddlers learning British English and a new Additional Language not included in the original target set (Study 3). The UKBTAT, designed to address issues faced by childhood professionals in a range of bilingual situations, has the following characteristics: (1) it is targeted for 24-month old children, a milestone age easy to remember for parents and practitioners, and relevant for the UK policy of assessing children from this age; (2) it is usable for any child learning British English and any other Additional Language from our 13 target languages – and can also provide useful information regarding English development for children learning another Additional Language; (3) it is user-friendly and easy to administer, relying on short parental questionnaires about the child’s vocabulary knowledge, and a 10-minute interview with the parent/carer (which can be done on the phone) to collect critical data on language exposure and demographics; (4) it provides interpretable results even if parents or carers estimate their child's knowledge of English only, although the added information about the Additional Language, when available, enriches the outcome; and (5) it can be used by a non-language-specialist practitioner, it is freely accessible online on a dedicated website and easily printable if needed.

From the cohort of toddlers tested in this project, we collected detailed information about family composition and characteristics, and level/mode of exposure to English and the Additional Languages. Many studies have examined situational factors that could impact the rate of language development in bilingual children, including socio-economic status (e.g., Calvo & Bialystok, 2014; Gross et al., 2014), relative amount of exposure to each language (e.g., Hoff et al., 2014), mode of exposure, such as number of speakers per language (e.g., Gollan, Starr & Ferreira, 2015), daycare attendance (e.g., Hansen & Hawkes, 2009), or language mixing (e.g., Byers-Heinlein, 2013). Such studies typically focus on no more than three factors at a time, making it difficult to quantify the relative contributions and interactions of all factors on a single measure of language achievement. A recent exception is the 250-children study by O'Toole et al. (2017), who compared vocabulary in six groups of bilingual children aged 24 to 36 months growing up in different countries (Maltese-English, Polish-English, French-Portuguese, Hebrew-English, Irish-English, Turkish-German) using adaptations of CDIs, and examining the contribution of a range of situational factors to TCV measures. They reported some large, unexplained variations in TCV measures across language groups, possibly due to linguistic distance, with for example Maltese-English and Polish-English groups scoring lower than the other groups. However because they tested children from a range of linguistic, cultural and geographical backgrounds, it is impossible to disentangle variance due to any of these factors from that due to linguistic distance.

The current study is the first extensive study of the effects and interactions of the potentially critical factors for language development in bilingual toddlers learning one common language: relative exposure to the two languages, family demographics, mode of exposure (which uncovers a range of factors, as described below) and, innovatively, language community (i.e., the specific Additional Language being spoken) and linguistic distance between English and the Additional Language. Quantifying the contributions of these factors to the trajectory of vocabulary development in bilingual toddlers will be critical to provide an accurate picture of expected language outcomes at 2 years. In what follows we review the potential factors that influence vocabulary knowledge at age 2 in bilingual toddlers.

Relative amount of exposure to languages. It is firmly established that the relative exposure to each language strongly influences bilingual children's rate of development in those languages (Welsh: Gathercole & Thomas, 2009; Spanish: Hoff et al., 2012; Pearson, Fernandez, Lewedeg, & Oller, 1997; French: Thordardottir, 2011; Cantonese and Mandarin: Law & So, 2006); indeed, relative exposure has been advocated as a proxy for language dominance (Unsworth, 2012). Relative exposure predicts development of phonology (Law & So, 2006), lexicon (Cattani et al., 2014) and grammar (Gathercole, 2002a; 2002b; 2002c; Nicholls, Eadie, & Reilly, 2011; Oller & Eilers, 2002). How to measure the exposure to each language varies from one study to the next or from one lab to the next, for example asking parents to use a prospective language diary (e.g., De Houwer & Bornstein, 2003; Place & Hoff, 2011), or using a detailed questionnaire or interview about regular exposure to each language (e.g., ALEQ: Paradis, 2011; Language Exposure Questionnaire: Bosch & Sebastián-Gallés, 1997; Thordardottir, 2011; Hoff et al., 2012). Because these questionnaires tend to be long and complex to administer, we developed our own Plymouth Language Exposure Questionnaire tool (Cattani et al., 2014), consisting of a 5 to 10 min interview with the parent (face-to-face or on the phone), comprising 10 to 12 simple questions about a child's typical week. In a group of 35 mixed bilinguals aged 30 months, Cattani et al. (2014) showed that the amount of exposure as measured by the Plymouth Language Exposure Questionnaire predicted vocabulary scores in comprehension and production as measured by the Oxford CDI (Hamilton, Plunkett, & Schafer, 2000): specifically, the more exposure to English (relative to the Additional Language) children experienced in a typical week, the more words they understood and used in English. Recently, Abdelwahab, Stone, Slee, Cattani and Floccia (2016) also showed a strong correlation between scores from the Plymouth Language Exposure Questionnaire and three other widely-used questionnaires (ALEQ: Paradis, 2011; MLEQ: Yang, Blume, & Lust, 2006; Language Exposure Questionnaire by Bosch & Sebastián-Gallés, 1997), with correlations ranging from .62 to .79.

The Plymouth Language Exposure Questionnaire primarily collects information about speech directed towards children, as the consensus is that children's language development benefits from joint attention situations (e.g., Tomasello & Farrar, 1986) and infant directed speech (Weisleder & Fernald, 2013). However, in many cultures children are not directly addressed (Lieven, 1994), and recent studies have showed that word learning at 18 months, for example, can be elicited from overheard speech (Floor & Akhtar, 2006). Furthermore, bilingual infants as young as 3.5 months benefit from overheard speech in their ability to discriminate between their two native language(s) (Molnar, Gervain, & Carreiras, 2014). In fact, in many tools used to quantify the relative amount of exposure to each language, both direct (speech to the child) and indirect (speech between adults) sources are taken into account in the calculations (e.g., Paradis, 2011; Yang et al., 2006). To complement the Plymouth Language Exposure Questionnaire measure and ensure that we quantify all possible sources of English/Additional Language input, the proportion of British English versus Additional Language used in overheard speech between parents was also evaluated in the current study.

Mode of exposure. Mode of exposure is probably the most complex factor to estimate, as it includes a range of variables such as the *source* of each language (e.g., presence of siblings, number of speakers per language, social context of exposure, e.g., crèche versus home), the *status* of the languages in the environment (minority language, such as Mandarin in a Plymouth family, or predominant cultural bilingualism as Welsh and British English in Bangor), and the *properties* of the input (language mixing; native versus non-native input). It is a current matter of debate whether these variables have a significant effect on bilingual language development, especially in toddlers.

Source of each language. The bilingual child's relative proficiency between languages appears to be modulated by the source of exposure, that is, who is speaking to them in each language. For example, Barrena, Ezeizabarrena and Garcia (2008) reported that Basque-Spanish young bilinguals knew more words in the Additional Language (Basque) when both parents were Additional Language speakers as compared to when only one was an Additional Language speaker - although this factor could be confounded with the amount of exposure in this study. It is also possible that there may also be differences depending on *which* parent is the source of Additional Language (please note that at this stage, the wording in the various questionnaires follows the heterosexual family model). Fathers generally direct less verbal output to their children than mothers, as they spend a greater proportion of their time interacting through play activities, especially physical play, which reduces the density of their speech (e.g., Pancsofar & Vernon-Feagans, 2006).

More generally, the effect of the number of speakers per language has recently been studied (e.g., Gollan et al., 2015), given theoretical proposals that variability in speech input supports the construction of phonological categories in early infancy (e.g., Rost & McMurray, 2009; Singh, 2008). Place and Hoff (2011) report that 25-month-olds Spanish-English bilinguals knew more words in English if they interacted with a larger number of speakers in that language, once corrected for the overall amount of exposure to English (see also Gollan et al., 2015). However, this finding was only partially replicated in Place and Hoff (2016) with a larger sample of 90 30-month-olds learning Spanish and English, where modest speaker number effects were found, predictive of knowledge in Spanish only. In the current study, we therefore examined the effect of vocabulary scores on the number of speakers per input language.

The impact of daycare attendance on children's development is a longstanding question in child development research, with mixed data (e.g., Brooks-Gunn, Han & Waldfogel, 2002; Ruhm, 2004). One of the few consistent overall results is that daycare attendance tends to benefit children from low SES. For example, using data from about 13,000 children in the British Millennium cohort, Côté, Doyle, Petitclerc and Timmins (2013) report a cognitive advantage at age 3 for children who have attended daycare, but only for those from low SES, and only below age 5. However, Hansen and Hawkes (2009), using the same cohort data, show that vocabulary outcomes are not significantly affected by daycare, with the exception of grandparent care which benefit mainly children from higher SES families. We measured how vocabulary outcomes are modulated by daycare attendance and language spoken within that environment.

The presence of older siblings is another potential source of variation. In monolingual homes, first-born children tend to acquire language faster than later born siblings (e.g., Fenson et al., 1994, 2007, for production; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 2010), presumably because they get more joint attention opportunities with adults and because child-directed speech produced by adults is of better quality than that produced by children. For bilingual children, this should result in a larger vocabulary in both languages for first-borns than later children. However, Bridges and Hoff (2014) found that North American bilingual toddlers with older siblings were more proficient in English compared to other bilingual children with no older siblings, presumably because older bilingual siblings were more likely to use English when addressing toddlers than other members of the family. Therefore, although input from older siblings may contain more ungrammatical structures and less diverse vocabulary, it nonetheless tends to contain a higher proportion of English, thus leading to higher levels of English vocabulary in the target children. Based on these findings we examined whether the number of older siblings at home had any effect on vocabulary development and especially English production skills.

Properties of the input. We use "properties of input" to refer to whether children interact with speakers who are consistent in their language use, and whether or not the input is produced by a native speaker. Within-speaker consistency can be measured as the degree of code-switching or language mixing (i.e., including elements of each language at the sentence level), which is estimated to occur significantly often for many bilingual children (e.g., over 20% in Marathi-English bilinguals: Tare & Gelman, 2011; between 2 and 10% in Brazilian-Portuguese English bilinguals: Nicoladis & Secco, 2000). It can also be measured more broadly as the adherence to the one-parent-one-language principle, that is, the extent to which children hear the two languages from the same person. Within-speaker inconsistency is potentially seen as a delaying factor in language development (especially code-switching at the sentence level), as language acquisition theorists usually argue that language separation early in development should optimise learning mechanisms (e.g., Curtin, Byers-Heinlein, & Werker, 2011). Indeed, Byers-Heinlein (2013) found that intra-sentential code-switching was detrimental at 18 months, although only marginally at 24 months. Evidence for the importance of the one-parent-one-language principle is less straightforward (e.g., de Houwer, 2007). Place and Hoff (2016) tested ninety 30-month-old Spanish-English bilingual children (see also Place & Hoff, 2011), and after controlling for gender, maternal education, and the child's relative language exposure, they reported no robust relations between the frequency of mothers' use of the two languages (measured at the discourse level, not within sentences) and measures of their children's English or Spanish skills. They concluded that the negative effect of the adherence to the one-parent-one-language principle may be minimal on children's language development or perhaps influential only at the very early stages of language development. Here we assessed the effect of consistency of parental language use on our target population of bilinguals. One possibility, which we examined here, is that the negative impact of parental code-switching on vocabulary development (Byers-Heinlein, 2013) might be mainly found for pairs of languages with a minimal phonological overlap, as Byers-Heinlein primarily examined distant language learners whereas Place and Hoff (2016) – who used the broader measure of adherence to the one-parent-one-language principle - looked at close language learners. The rationale behind this hypothesis is that frequent switches between distant languages would require a greater flexibility (and therefore cognitive resources) to navigate from one set of representations over another, as compared to close languages which activate overlapping representations.

Another qualitative aspect of the input that might modulate bilingual children's vocabulary growth, is the nativeness of the adult speaker in each language. The hypothesis that native speakers provide more supportive input in their language than non-native speakers is a matter of controversy. Fernald (2006) suggested that being presented with both native and non-native speech in their two languages could lead infants to have more difficulties in discriminating between them, and subsequently impair development of phonological categories (young bilinguals do learn some phonological contrasts later than monolinguals: Bosch & Sebastián-Gallés, 2003; Burns, Werker, & McVie, 2003). It could be also that non-native speakers tend to use less varied vocabulary (Hoff, Coard, & Señor, 2013). However, Paradis (2011) examined 4- to 7-year-old immigrant children in Canada and found that mothers' proficiency in English was not a significant predictor of children's English vocabulary - although children's exposure to native speakers through media or friends was a predictor. In contrast, Hammer et al. (2012) reported that for Spanish-English 59-month-old bilinguals, English proficiency of mothers was a fair predictor of their children's scores in English tests, although they did not control for the relative exposure to English/Spanish. Place and Hoff (2011) reported that in 29 Spanish-English bilinguals aged 25 months, English vocabulary was positively correlated with the amount of English produced by native as opposed to non-native speakers, controlling for overall exposure. More recently, Place and Hoff (2016), with a larger sample of 90 30-month-old Spanish-English bilinguals, also reported a small but positive influence of the proportion of native-speaker input on English knowledge when measured with standardised tests (PLS-4: Zimmerman, Steiner & Pond, 2002; EOWPVT: Brownell, 2001). They found no effect on CDI measures, however. Based on these findings, we examined if the proportion of parental native versus non-native speech has an effect on children's vocabulary size in each language.

Status of the Additional Language. Whether children grow up bilingual in a monolingual or bilingual society may have consequences for their degree of achievement in both languages. Bilingual children from minority populations such as recent immigrants (e.g., Cantonese/English speakers in the UK) tend to have lower academic outcomes in mainstream education than monolinguals (Prevoo Malda, Mesman, & van IJzendoorn, 2016), likely driven by poorer English reading comprehension skills. This disadvantage is particularly acute if English is not their dominant language (Strand, Malmberg, & Hall, 2015). In contrast, when bilingualism is the norm within a particular society, such as in certain Welsh communities, the cognitive and academic achievements of bilingual children can be equal (Rhys & Thomas, 2013) or even better (i Trueta, Barrachina, & Pascual, 2012) than their monolingual peers. Thus, in a bilingual society children's achievements in both their languages might be advantaged, as compared to children learning English and an Additional

Language in a monolingual community. We explicitly tested this hypothesis through the inclusion of a target cohort of Welsh-English toddlers selected from a bilingual community in North Wales.

Demographic factors (SES and gender). Monolingual children from lower socio-economic groups tend to have poorer language skills than those from higher SES (Deutsch, 1965; Hoff-Ginsberg, 1998; Rack, Snowling, & Olson, 1992), perhaps because of the characteristics of maternal input (Hoff, 2003) and/or the home attitude towards literacy (Payne, Whitehurst, & Angell, 1994), which includes low frequency of shared reading activities at home (Britto, Fuligni, & Brooks-Gunn, 2002; see Tomalski et al., 2013, for showing that SES shapes brain activity in early infancy). Unsurprisingly, this SES related language advantage also extends to bilingual children (e.g., Calvo & Bialystok, 2014; Gathercole, Kennedy, & Thomas, 2016; Ollers & Eilers, 2002; Paradis, 2009). It has similarly been proposed that this effect is partially due to the higher quality of language provided by mothers with a high education background, improving their children's acquisition of the maternal language, and transferring to an advantage in the Additional Language even if the mother does not use it (Goldberg, Prause, Lucas-Thompson, & Himsel., 2008; Paradis, 2009).

It is important to note however, that in monolingual toddlers the effect of SES on vocabulary size as measured by CDIs is heavily dependent on the presence of children from *very low* SES backgrounds in the sample. Where such children are not systematically recruited there are negligible or null effects: for example, in the latest cohort tested with the MacArthur CDI (Fenson et al., 2007; N = 2007), no effect of SES (as measured by maternal education) was found between 13 and 20 months, and significant but very small effects were reported in production from the age of 21 months onwards. The original MacArthur CDI cohort (Fenson et al., 1994; N = 1130) found a very small correlation between SES and vocabulary production in 16-30 months old toddlers (r = .05). Similarly, in the Oxford CDI, Hamilton et al. (2000) did not find any correlation between SES and vocabulary scores in production or comprehension. In contrast, the studies of Fernald, Marchman and Weisleder (2013) and Arriaga, Fenson, Cronan and Pethick (1998), which sampled extensively from low SES families, demonstrated significant if modest SES effects.

This pattern of findings suggests that the effect of SES may be limited to the lower thresholds of SES indexing. Note that this is not to be necessarily expected from other measures of vocabulary, such as naturalistic recordings as used by Hoff (2003), who found differences between mid and high-SES children.

Gender is also a well-documented factor in vocabulary growth, with girls usually producing more words than boys (Huttenlocher et al., 1991), without necessarily showing better comprehension scores than boys (girl advantage in Fenson et al., 2007; equal levels in Eriksson et al., 2012). In the original MacArthur CDI cohort of 8- to 30-month-olds (Fenson et al., 1994), gender was found to account overall for 1 to 2% of the variance, and more in production than comprehension. This production advantage for girls was found to be consistent across linguistic communities as demonstrated with CDI data collected from 10 large non-English groups (Eriksson et al., 2012), pointing to a common neurophysiological explanation rather than sociological/cultural causes - or perhaps to widely shared conventions of encouraging more communication with girls.

To summarise, we have reviewed the situational factors which might shape vocabulary knowledge in bilingual toddlers: relative amount of exposure, mode of exposure (an umbrella term for a range of factors related to the source of the Additional Language, the properties of the input, and the Additional Language status), and demographics. These factors, together with language community and metrics of linguistic distance, were used to build a model of the bilingual lexicon at age 2 for assessment purposes, the UKBTAT (Study 2). In what follows, we discuss the rationale and objectives of the third study, which was to extend this model to bilingual toddlers speaking British English and any Additional Language different from those used to generate the lexical model in Study 2.

Study 3: Generalising Language Assessment to any UK-Raised Bilingual

As mentioned above, professional bodies such as the Royal College of Speech and Language Therapists recommend that bilinguals should be assessed in their two languages, as the variety of factors modulating bilingual development prevents a direct comparison of their achievements with corresponding monolingual norms (Kohnert, 2010). Our aim in Study 2 was to accommodate these factors to provide norms of development in English and 13 target Additional Languages and create the UKBTAT assessment tool; in Study 3 we aimed at establishing whether the estimates of English vocabulary generated for the UKBTAT would allow identification of possible language delays for any incoming non-target Additional Language (avoiding the requirement for normed data on all possible language pairs, which is practically impossible).

The possibility of an English-only screening tool is based on two theoretical premises. First, in bilinguals, Developmental Language Disorder, as well as language-based learning disabilities, affect both languages (Håkansson, Salameh, & Nettelbladt, 2003), insofar as such disorders are underpinned by a genotype affecting neurological functioning (e. g. Leonard, 1987; Schwarz, 2010). Therefore, any domain of impairment, for example morphosyntax – the most documented so far (e. g. Rice, 2004) – will be observable in the two languages. However, due to relative delays in growing the two languages, and to the language-specificity of the complexity of morphological rules, difficulties in language learning could appear more severe in one language as compared to the other. But our point remains that a delay or a difficulty is expected in both languages, when compared to norms of development.

The second premise is that, although not all late talkers in early childhood are going to develop a primary language impairment (e.g., Rescorla, 2005), almost all children diagnosed with DLD later in childhood had a history of initial language delay, in the form of protracted vocabulary development in comprehension and production (Conti-Ramsden et al., 1997). Hence, Dale et al. (2003) reported that monolingual 2-year-olds who scored below the 10th percentile on the MacArthur–Bates Short Form CDI in production were significantly more likely to be diagnosed as DLD when reaching the age of 3 to 4 years. Out of the 6,500 children who scored above the 10th percentile for vocabulary production at 2 years, only a tiny fraction exhibited grammatical difficulties one or two years later. Early receptive language difficulties have an even higher predictive validity of later language disorder diagnosis (e.g., Beitchman et al., 1994; Chiat & Roy, 2008).

The corollary of these two premises is as follows: if we could detect the late bilingual talkers at age 2 years in English only, when all other sources of variation have been factored in, we would have identified the main cohort from whom language impaired children will be diagnosed one or two years later. While we agree that collecting information about a bilingual child's language development in both languages is still relevant and informative for clinical and research purposes, practitioners or researchers might not always be familiar with both languages. Thus the ability to work with only one of the two languages of a bilingual is important from a theoretical as well as a practical point of view.

In order to achieve this goal, we examined whether the model of the bilingual lexicon generated in Study 2 using a sample of 13 target bilingual groups could reliably predict the English scores of a new sample of children learning British English and a non-target Additional Language.

Summary of Research Questions and Key Predictions

We analysed vocabulary data, situational factors, linguistic distance and language community in a sample of two-year-olds learning British English plus either one of 13 target Additional Languages (N = 372) or one non-target Additional Language (N = 58). Three studies were performed with the same data set: investigating the effect of linguistic distance on vocabulary development (Study 1, just using the target Additional Languages), building a model of the bilingual lexicon at age 2 using a range of situational variables and linguistic distance for the target Additional Languages learners, resulting in the UKBTAT (Study 2), and testing its generalisability for the non-target learners (Study 3). To fulfil the objectives of Study 1 and estimate the impact of linguistic distance on vocabulary development, it was necessary to first get a full estimate of the effects of all situational factors that were known or suspected to shape bilingual development. The following predictions were formulated:

- Given the solid evidence of the importance of the amount of language exposure in expressive and receptive vocabulary growth (e.g., Hoff, 2003), we expected this factor to be a robust predictor of vocabulary knowledge. In particular, we predicted that children with more exposure to English would have higher scores in both receptive and expressive vocabulary scores in English - and conversely, lower vocabulary scores in the Additional Language. We predicted that these effects would be strongest in our measure of direct exposure, obtained through the Plymouth Language Exposure Questionnaire, but might also be found in our measure of indirect exposure via overheard conversation between parents (referred to as Overheard speech).
- We predicted an effect of gender, with girls outperforming boys (Eriksson et al. 2012), especially in expressive vocabulary (Fenson et al., 2007).
- SES was expected to modulate language outcomes, with smaller expressive and receptive vocabularies for children from lower SES families (Gathercole et al., 2016).

- Regarding the effect of those factors related to the mode of exposure to each language (source, properties and status), one objective of this first study was to assess whether effects described in the literature for specific English-Additional Language pairings are robust over a range of languages.
- The effect of language community was an unknown variable, and was expected to generate a large amount of unexplained variance, due to a variety of cultural and linguistic factors. We hypothesised that part of the variance attributed to language community could be accounted for by linguistic distance between British English and the Additional Language, as measured by the phonological overlap between translation equivalents, the word order typology and morphological complexity.

Study 2 did not generate predictions per se, as its objectives were to build an assessment tool based on findings from Study 1. However, for Study 3, we predicted the following:

- In our models of vocabulary developed in Studies 1 and 2, Language Community was treated as a random factor to acknowledge that our 13 target languages comprise a non-exhaustive sampling of possible language pairs. The success with which the random effects modelling could produce a model that generalises beyond these target languages was assessed by establishing how well our models could generalise to a test set of 58 children whose Additional Language was not part of our 13 target languages, and therefore not represented in the model.

Chapter 2 - Methods

The methods described in this chapter cover the cohort data collection, which constituted the common data set for all three studies reported in this paper. A sample of 430 bilingual toddlers, learning British English and one of 13 target Additional Languages (N = 372), or any other Additional Language (N = 58), were identified over a 2-year period. To increase variation in both English/Additional Language pairs and in the situational factors outlined above (language exposure, mode of exposure and demographic factors), data was collected through trained research assistants recruited in the six universities involved in this project (Bangor, Birmingham, Kent, Liverpool, Oxford and Plymouth), as well as in Bristol and Leicester, each having access to multilingual populations to various degrees. However, since the testing platform was remotely accessible, the final sample comprised families from all areas of the UK, apart from Scotland and Northern Ireland. Bangor had the additional advantage of being located in a region with 75% bilingual Welsh-English children, providing a unique opportunity to compare language skills in bilinguals growing in a region with predominant bilingualism, to those whose bilingualism is linked to immigration.

When the child approached her second birthday, volunteer parents were contacted via the website UKBilingualToddlers, and the following data were collected in this order: English expressive and receptive vocabulary as measured through a bespoke Oxford Short Form CDI (Hamilton et al., 2000); Additional Language vocabulary as measured through the corresponding version of the CDI, when available; a family questionnaire with detailed questions about demographics (developed for the UK-CDI standardisation project, Alcock et al., in prep); and the Plymouth Language Exposure Questionnaire (Cattani et al., 2014) which provided the LEQ measure of relative exposure to each language.

To sum up, the current study measured four parent-assessed outcome variables at 24 months: receptive English vocabulary, expressive English vocabulary, receptive Additional Language vocabulary and expressive Additional Language vocabulary (through CDIs). For each of these outcome variables, we investigated the influence of the following factors: 1) gender, 2) SES (as assessed via parental income and educational level), 3) proportion of child-directed speech in English (LEQ), 4) proportion of overheard parental speech in English, 5) factors related to the source of each language (whether two parents were native Additional Language speakers or only one, number of sources of English, number of sources of the Additional Language, time in daycare in each language, number of older siblings), 6) factors related to the properties of the input (degree of language use consistency in parents' input, number of native and non-native speakers in each language), 7) status of the Additional Language (societal vs. minority), 8) the particular language community (i.e. which of the 13 additional languages the child was exposed to) and 9) the linguistic distance between English and the Additional Language as measured by a) phonological distance, b) morphological distance and c) syntactic distance (see Table 1 for a summary of these variables).

INSERT TABLE 1 HERE

Method

Participants. Data were collected for a total of 430 children between February 2014 and July 2016. The data of an additional 31 children were discarded as they had hearing problems (N = 7), had a diagnosed developmental delay (as reported by parents; N = 6), were too young or too old (N = 17), or had incomplete records (N = 1). The data of another 41 children could not be included as parents did not complete the study. Out of the remaining final sample of 430 children (aged 23.89 months, SD 0.39, from 23.0 to 25.0; 193 girls and 237 boys), 372 were learning English and one of the 13 target Additional Languages: Bengali, Cantonese, Dutch, French, German, Greek, Hindi/Urdu, Italian, Mandarin, Polish, Portuguese, Spanish, and Welsh. Following King (2001), spoken Hindi and Urdu were classified as two varieties of the same language. The remaining 58 were learning English and one non-target Additional Language (see Table 2). The proportion of children born in the UK was 94.1% for the 372 children learning a target Additional Language, and 93.1% for the 58 non-target Additional Language learners. Out of these 430 children, the information for family income (an optional field) was not supplied for 15 children (13 in the target language community and 2 in the non-target language community). See Table 2 for a full description of the sample.

INSERT TABLE 2 HERE

Procedure and instruments. The data collection was initiated when the children reached 23.5 months old. When signing the online consent form on the UKBT database, parents were notified that there were four tasks to complete for the study: none of these tasks involved testing the children, allowing for remote data collection. Specifically, the CDIs and the family questionnaire were completed on the online platform by the parents, and the Plymouth Language Exposure Questionnaire was completed by the research assistants during a final telephone interview. A paper copy of the questionnaires was sent to parents who were unable to access the internet. For some families who did not feel confident in English, the research assistants met with the parent(s) to help them go through the various questionnaires. When signing up, contact information and identification of the language(s) being spoken at home triggered the selection of the appropriate Additional Language CDI when available.

Metrics of linguistic distance. To create a toddler-centric representation of language distance, each of the 406 non-onomatopoeic words from the Oxford CDI (Hamilton et al., 2000), as well as their translation equivalents across the 13 target Additional Languages, were transcribed into broad phonological representations. These were produced by trained phoneticians, each of whom was a native speaker of the language they were asked to transcribe. Our metric of language distance was then calculated as the overlap between the phonological representation of a word in British English and its translation equivalent in the Additional Language. This overlap was based upon the Levenshtein distance, that is, the minimal number of insertions, deletions and translations that are required to get from the British English phonological representation to that of the Additional Language. To produce a proportional measure of overlap this distance was subtracted from the length of the longest phonological sequence in British English or Additional Language, and then divided by the same number. This produces a measure of phonological overlap for each word, between 0 (no overlap) and 1 (perfect cognate), that preserves sequence order and is proportional to the length of the word.

$$Overlap = \frac{Max(BE\ length,\ AL\ length) - Levenshtein\ distance}{Max(BE\ length,\ AL\ length)}$$

An example of a calculation for the British English word “lamp” and its Italian translation equivalent “lampada” is shown below:

BE lamp /l.æ.m.p/ : Sequence length = 4

Italian lampada /l.a.m.p.a.d.a/ : Sequence length = 7

Levenshtein distance (l.æ.m.p, l.a.m.p.a.d.a) = 4 (1 translation + 3 insertions)

$$Overlap = \frac{Max(4,7) - 4}{Max(4,7)} = 0.43$$

The language level phonological overlap between British English and each of the 13 Additional Languages is shown in Table 1, calculated as the average overlap across all 406 words.

For the measure of word order typology, the Additional Language was assigned a 1 if it had a VO order like British English, a 2 if it had a mixed VO/OV order, and a 3 for a OV order (see Table 1). Finally, morphological complexity was assessed on a 3 point scale, with analytic/isolating languages (Mandarin, Cantonese) being ranked closer to English (value 1), followed by fusional languages such as French and German (value 2) and agglutinative languages such as Hindi/Urdu and Bengali (value 3) (see Table 3). To illustrate, in analytic Mandarin number is not marked on nouns, as in 天 *yī tiān* "one day", 三天 *sān tiān* (lit.) "three day". In fusional French, the verbal suffix relates to grammatical mood, tense, aspect, person and number, as in *mangeais* "ate" (indicative, past, imperfective, second person singular) and *mangerions* "would eat" (conditional, present, perfective, first person plural). In agglutinative Bengali, nominative case for the word "river" is *nodi*, and the accusative *nodike*.

INSERT TABLE 3

Collecting demographic data. Demographic data were collected through the family questionnaire developed by Alcock et al. (in prep). This contains questions regarding (i) the health and development of the child, (ii) the child's family history, (iii) parental information (e.g., parents' educational level, income and postcode), and (iv) childcare arrangements. Some of these questions were repeated in the Plymouth Language Exposure Questionnaire (see below), but we tolerated overlap in order to retain each questionnaire's integrity. Following Arriaga et al. (1998), we focused on household income and educational levels when measuring SES, as typical indices of SES are highly correlated. Income was divided in four bands (variable Income), and education was measured on a seven point band that correspond to English qualification classifications, from no qualifications to a postgraduate degree (variables MumEd and DadEd; see appendix 2). Education was chosen as it is generally used as a proxy for SES (e.g., Bornstein, Hahn, Suwalsky, & Haynes, 2003; Fenson et al., 2007), and it is usually a better predictor of language development than income (e.g., Hoff, 2003); in addition we estimated that in the case of immigrant families, educational level might better reflect the child's learning environment than mere economic circumstances. The educational status of both parents was used since the correlation between these two predictors was not large (r = .29).

As with the Fenson et al. (2007) and Hamilton et al. (2000) studies, the current study had an under-representation of low SES children within our bilingual cohort. This may be representative of SES distribution across the national population of bilingual children: Dustmann and Frattini (2011), using a variety of large scale British and international sources collected between 1993 and 2009, observed that immigrant populations in the UK tend to leave the education system later and have higher wages than their native peers. It is also likely that this under-representation stems from sampling, with low SES bilingual families reluctant to take part in research, especially in cases when they are not confident in English.

Evaluating amount of exposure to each language. The Plymouth Language Exposure Questionnaire (Cattani et al., 2014) was used to obtain the percentage of direct language exposure received by the child in English and the Additional Language in a typical week based on a unique 5 to 10 minute phone interview (variable LEQ). The questionnaire (available at http://www.psy.plymouth.ac.uk/leq/) requested information about (i) the average number of hours spent by the child in nursery/with a childminder in each language environment (variables EngDaycare and ALDaycare); (ii) the language(s) spoken by each parent at home and the relative frequency of use of the two languages (variables MumPropEng and DadPropEng, measured on a 5-point scale); (iii) the number of hours spent by the child alone with each parent; (iv) whether the parents spoke equally with their child when both parents present; and (v) the number of hours of the child's sleep in a typical day (to evaluate the number of possible contact time during a week). The detail of these variables and calculations leading to the proportion of English vs the Additional Language in a typical week (variable LEQ) is found in appendix 3.

To obtain the proportion of English/Additional Language in overheard speech (variable referred to as Overheard speech), an added question (5-point scale) was inserted after the original Plymouth Language Exposure Questionnaire (see appendix 2). See Table 4 for a summary of the results per language group.

INSERT TABLE 4 HERE

Evaluating the mode of exposure (source, properties and status). Measures of the various factors underpinning the *source of each language* were derived from questions which were part of the initial sign-up sheet, the family questionnaire and the Plymouth Language Exposure Questionnaire (see Table 5). Straightforward measures based on individual questions were the identification of the type of family (binary score for two parents native Additional Language speakers or only one; variable FamLang), the number of hours per typical week in English or Additional Language daycare (EngDaycare and ALDaycare), and the number of older siblings living in the house (until the age of 18 years; variable Siblings). Regarding the number of speakers in each language, a score of 1 was given to each native speaker parent, each older sibling, and attendance to a form of daycare (variable SourcesEng and SourcesAL, with an observed range of 0 to 6; see appendix 2).

INSERT TABLE 5 HERE

Regarding the *properties of the input* (see Table 6), the degree of language use consistency from each parent was obtained through the questions in the Plymouth Language Exposure Questionnaire asking parents to quantify on a 5-point scale their relative use of English and Additional Language. Specifically, a parent would obtain a 1 for always speaking Additional Language, 2 for usually speaking Additional Language, 3 for English and Additional Language half of the time, 4 for usually speaking English and 5 for always speaking English (variable MumPropEng and DadPropEng). Then the degree of consistency would be recoded as a minimum of 1 if the answer to the above was a 1 or a 5; a 2 if the answer to the above was a 2 or a 4; and a maximum of 3 if the answer to the above was a 3 (variables MumConsistency and DadConsistency, averaged as Consistency).

The proportion of native/non-native speech produced by parents was calculated from the same question, in conjunction with whether the parent was a native speaker of Additional Language or not. That is, the number of hours spent with each parent during a typical week was calculated as: 168 (number of hours in a week) - total sleeping time - hours in daycare - hours alone with the other parent (variable A). Then, each parent's score on their respective PropEng variable (1 to 5) was re-expressed as a proportion from 0 to 1 (1 = 0, 2 = .25, 3 = .5, 4 = .75, 5 = 1) to obtain the proportion of English in their speech (variable B). The resulting amount of English in this parent's input was obtained by multiplying A by B. If this parent was a British English native speaker, then AB would correspond to the amount of native input, and if the parent was an Additional Language native speaker, AB would be the amount of non-native input. The final proportion of native English input across both parents (the variable PropEngN), was obtained by dividing the total amount of native English by the sum of native and non-native English. The proportion of native Additional Language input (the variable PropALN), was calculated with a similar logic (see appendix 2).

INSERT TABLE 6 HERE

Finally, regarding *status of the Additional Language*, Welsh-English children growing up in Wales were coded as societal bilinguals, all others not.

Measuring vocabulary. To measure children's vocabulary achievements in English and in their Additional Language for the 13 target languages, we used Communicative Developmental Inventories in each language. For the English CDI, we developed a 100-word version of the existing Oxford CDI (Hamilton et al., 2000), referred to as the Oxford Short Form CDI, by selecting words from the original 416 words which would (1) be representative of the words known and produced by 24-month-old monolinguals in the original norms that cover the same range of frequencies, and (2) contain the same distribution of syntactic categories (nouns, verbs, pronouns, etc). We selected 10 words understood and produced by 100% of 2-year-old monolingual toddlers as provided by the Oxford CDI database, then 10 words understood and produced by 90% of the same children, etc. Then we adjusted these words to include a proportion of nouns, verbs and function words similar to those found in the Oxford CDI (see appendix 5 for the full list). To verify the validity of this Oxford Short Form CDI, the parents of 134 monolingual children from the Plymouth area (including 72 girls) aged 10 to 26 months (mean age 17.9 months) completed both the short and the long CDI within a week (mean number of days between completions: 4.3 days, SD 5.5). Their mean score on the long Oxford CDI was 160.2 words in comprehension (out of 416; SD 119.7) and 80.3 in production (SD 107.9); their mean score on the Oxford Short Form CDI was 43.5 words in comprehension (SD 28.0) and 23.0 in production (SD 28.2). Children's scores in the two CDIs were highly correlated in comprehension ($r = .95$, $p < .0001$) and in production ($r = .86$, $p < .0001$).

We also compared the scores directly for the 100 words that were present on both the long and the 100-word versions of the Oxford CDI: monolingual children's parents reported higher scores on the Oxford Short Form CDI, both for comprehension (t (133) = 5.71, $p < .0001$, mean Oxford CDI score = 39.2%; mean Oxford Short Form CDI: 43.5%) and production (t (133) = 5.40, $p < .0001$, mean Oxford CDI score = 20.4%; mean Oxford Short Form CDI: 23.0%). This difference is likely due to a fatigue or attentional effect when having to fill in a CDI four times as long as the 100 word Oxford Short Form CDI. Across the two completions, parents reported the same outcome (known or unknown) for 85.8% of words in comprehension, and 92.6% in production. Correlations between the short and long CDIs for the 100 words were $r = .95$ and $r = .98$ for comprehension and production respectively ($p < .0001$), indicating excellent validity for the Oxford Short Form CDI.

For the Additional Languages, we used the adaptations of CDIs for 12 Additional Languages with the authors' permission (see list in the appendix 1), selecting the form adapted for the age of 24 months when multiple versions were available. Additional Language CDIs had lengths varying from 654 words in Greek (Kati, personal communication) to 62 in Bengali (Hamadani et al., 2010; see appendix 1 for references of CDIs).

We developed a new CDI for Hindi/Urdu as none were available. For simplicity we treated these two languages as dialects of the same language using different graphemic systems, so we developed the same version, written in the two alphabets. Following the method by Kern (2007), after a translation of the Oxford CDI, two focus groups of native Urdu speakers agreed on a cultural adaptation of the word list. Native Hindi speakers were consulted to check its adaptation to Hindi.

All parents were first asked to complete the Oxford Short Form CDI, assessing receptive and productive vocabulary separately (for each word, they were to assess whether the word was understood but not produced, or produced). If they felt unable to do so because, for example, they never spoke English at home and therefore could not estimate their child's English knowledge, a proficient English speaking caregiver would complete a printed version of the CDI (e.g., a childminder). Parents were asked to complete the appropriate Additional Language CDI within a week of the completion of the Oxford Short Form CDI.

Chapter 3 – Analyses and Results for Study 1: Estimating the Effect of Linguistic Distance on Vocabulary Development

To estimate the impact of linguistic distance on vocabulary development, which was the aim of Study 1, we needed to account for the effects of all situational factors that were known or suspected to shape bilingual development. This was achieved through a two-step analyses of the data from the 372 children whose Additional Language was one of our 13 target languages.

Plan of Analyses

In the first step, analyses were conducted on variables already established within the literature as strong predictors of vocabulary size (relative amount of exposure to each language in child-directed speech and overheard speech, gender and SES). Analyses were conducted initially in ANCOVAs (to include continuous variables such as LEQ) and then subjected to confirmation in linear mixed models, with variables entered as fixed effects predictors only if they reached significance in the ANCOVAs.

In the second step, analyses were then conducted on less well-established variables (factors relative to source, properties and status, and measures of linguistic distance), in models containing verified predictors from the initial stage. Again, ANCOVAs were followed by linear mixed models with the same logic as in the first step.

The reasons behind this two-step process extend beyond the aims of Study 1, as explained below. The effects from the ANCOVAs would hold for a population with the same breakdown of language communities as those in our sample (e.g., which was 17.4% German, 12.8% French etc.). The linear mixed models do not make that assumption and are thus strongly preferred if conclusions are to be generalised to all bilinguals. The purpose of the linear mixed models was indeed to provide the best possible test for the importance of candidate predictor variables for bilinguals generally, not merely those whose Additional Language was one of those 13 used in this paper (and therefore preparing for Study 3). For this it was imperative that Language Community be modelled as a random, not a fixed effect, as the ANCOVAs would do.

In contrast, the value of the ANCOVAs is that they allow a straightforward test of the significance of simultaneously entered predictor variables, something that is problematic with linear mixed models due to a lack of consensus on how to compute the degrees of freedom for each predictor (Baayen, Davidson, & Bates, 2008). Our procedure was thus to perform a preliminary selection of significant predictors from ANCOVAs, subject to confirmation in linear mixed models, before finally being included in predictive models.

Following steps 1 and 2, predictive linear mixed models of expressive and receptive vocabulary for the UKBTAT tool (Study 2) were calculated with predictors retained only if their effect size in the ANCOVAs was larger than $\eta^2 = 0.02$ (Cohen, 1988), which is considered a threshold for small effects in ANOVAs and multiple regressions. These final models do not include measures of linguistic distance, since we aimed at developing norms which could be applied to any bilingual children learning British English, and measures of linguistic distance would not be available for Additional Languages that are not amongst our 13 target Additional Languages. Predictive models for the UKBTAT (Study 2) and the test of their generalisation beyond the 13 target Additional Languages (Study 3) are presented in Chapter 4.

Predictor Variables

Step 1 predictors were language exposure scores (LEQ and Overheard speech), Gender, and SES. Language Community (which Additional Language is spoken by the child) was included as a 13 level dummy variable simply to control for its effects. Step 2 predictors were the mode of exposure variables, and three Linguistic Distance variables. Mode of exposure variables belonged to three categories: *source* of each language (whether two parents are native Additional Language speakers or only one, total number of English speakers, total number of Additional Language speakers, number of siblings, time spent in English speaking daycare, time spent in Additional Language speaking daycare), *properties of the input* (degree of language use consistency in parents' speech, proportion of native English, proportion of native Additional Language), and *status of the Additional Language* (Welsh group vs. all other Additional Languages). Linguistic Distance variables were phonological overlap, word order typology and morphological complexity.

Test language, or TestLang (English/Additional Language), was included as a repeated measures factor to examine the differential impact of predictors on English and Additional Language. See Table 1 for a summary of these variables.

Vocabulary Measures

Dependent variables were CDI counts of receptive and expressive vocabulary in English and the Additional Language. We conducted analyses on two different versions of the CDIs. Our starting point was a 30 word CDI made up of those 30 words present in the Oxford Short Form CDI and all 13 Additional Language CDIs (see Table A5-1 in appendix 5). This 30-word CDI had the advantage of holding items constant across a child's two test languages, thus controlling, amongst other things, for word frequency (although frequencies between translation equivalents differed, correlations were in the order of .8 over the 13 English-Additional Language pairings). The disadvantage of this approach is that, of the words common across all CDIs, a disproportionate number were high frequency words. This results in a ceiling effect with, for example, over a third of children scoring 100% on English comprehension. Our second CDI used the full 100 words of the Oxford Short Form and the full Additional Language CDI re-represented as a percentage, to accommodate the fact that the standardised CDIs varied considerably in length for each language. While these data suffer from no ceiling effects and maximise the sampling of vocabulary, they are the least satisfactory in that there is no control of word frequency between all Additional Language CDIs, that is, obtaining a 40% score in the German CDI is not necessarily equivalent to a 40% in the Portuguese CDI. The main analyses reported here were performed on the 30-word CDI, with a specific section added for the data from the full CDIs. Given that the results were essentially similar when using the 30-word CDI or the full CDIs, please note that the final equations in the UKBTAT (Chapter 4) are calculated from the 100 words of the Oxford Short Form for English (for increasing representativeness of the model's coefficients), and the 30 words for Additional Languages when relevant.

Descriptive Statistics

Predictors. Tables 2 to 6 presented summary data of all predictor variables broken down by Language Community. Because of the strong associations between Language Community and predictor variables, it was important to use models that included Language Community to avoid attributing to predictor variables the explanatory power that was actually simply due to variability over language communities.

Vocabulary measures. All data for vocabulary measures in English (for the full cohort) are reported in Table 7, and in Table 8 for each Additional Language (for the 372 target Additional Languages learners only). On the Oxford Short Form CDI, children understood on average 67.9 words and produced 41.2 words (variables CDI100Comp and CDI100Prod). On the Additional Language CDIs, which varied in length, children overall understood 54.9% of Additional Language words and produced 24.2% (PropALcomp and PropALprod). When restricting the analysis to the 30 words common to all CDIs, children understood on average 24.4 English words and produced 17.0 (variables CDI30comp and CDI30prod). For the target Additional Language, children understood on average 21.7 words and produced 11.2 (ALCDI30Comp and ALCDI30Prod), which was significantly less than in

English (comprehension: paired *t* (371) = 6.25, *p* < .0001; production *t* (371) = 11.51, *p* < .0001). As noted before, all main analyses provided below were run on the 30-word CDIs, with analyses on the full form CDIs provided in a specific section.

Of interest is the comparison of the bilingual scores to monolinguals. Based on the Oxford CDI database for 125 monolingual children aged 23.0 to 25.0, 24-month-olds understand 73.6% of the 416-word CDI (SD = 16.9) and produce 48.3% (SD = 25.8). To compare these scores to those of the Oxford Short Form CDI, we applied a correction ratio computed from the comparison between the long and short CDI described in the "Measuring vocabulary" section (see Methods). A score in the long CDI divided by 0.90 provides an equivalent score on the short CDI. That means that the 24-month-old monolinguals are estimated to understand 81.8% of words and produce 53.7% if assessed with the Oxford Short Form CDI. In contrast, the cohort of 430 bilinguals understood 67.9% of the Oxford Short Form CDI (SD = 25.0) and produced 41.2% (SD = 26.0), which is significantly less than the monolinguals (comprehension: *t* (553) = 5.84, p = .0001; production: *t* (553) = 4.74, p = .0001).

INSERT TABLE 7

INSERT TABLE 8

Results

Step 1 - Predictors firmly established in the literature. We looked at four predictors: LEQ (proportion of English in child-directed speech), Overheard speech (proportion of English spoken between parents), SES and Gender. Three indices (income, maternal education and paternal education) were initially selected as potential predictors for

SES, but due to high correlations between income and parental education, and because income had the widest observed range, the latter was selected. Results are very similar if parental education is used. Thus our step 1 ANCOVA consisted of the between-subjects predictors of LEQ, Overheard speech, Income and Gender, with the within-subjects predictor of TestLang (Additional Language/English). Language Community was included as a between-subjects factor, since we wished to ascertain in Study 3 (as aforementioned) the degree to which all other predictors are generalisable to bilingual 24-month-olds regardless of the particular Additional Language she or he is learning. Separate ANCOVAs were run for production and comprehension scores. All ANCOVAs used as dependent variables the 30-words CDIs (see Table 9 and 10 for full results in comprehension and production respectively).

INSERT TABLE 9 HERE

INSERT TABLE 10 HERE

For comprehension there was no main effect of LEQ, but an interaction of LEQ and TestLang ($F_{1,342}$ = 75.07, p < .001, η^2 =.18). Analysis of the effect of LEQ on each test language separately revealed that it significantly reduced Additional Language scores ($F_{1,342}$ = 18.81, p < .001, η^2 =.05) and increased English scores ($F_{1,343}$ = 21.46, p < .001, η^2 =.06): the more child-directed English children heard, the more English they understood, and the less Additional Language they understood (see Figure 1). Overheard speech (the proportion of English vs the Additional Language spoken between the parents when the child was present) significantly increased comprehension scores overall ($F_{1,342}$ = 10.55, p = .001, η^2 =.03), and showed an interaction with TestLang ($F_{1,342}$ = 38.72, p < .001, η^2 =.10). Breaking down the effect of Overheard speech for the two test languages, while no effect was seen for Additional Language ($F < 1$), a beneficial effect was seen for English ($F_{1,342}$ = 32.42, $p < .001$, η^2 =.09). The more English spoken between the parents, the more beneficial effect on English comprehension (see Figure 2). A main effect of Income was found ($F_{1,342}$= 11.97, p =.001, η^2 =.03) and no interaction with TestLang. There was no main effect of Gender ($F_{1,342}$ = 1.39, p = .24) or interaction with TestLang ($F_{1,342}$ = .05, p = .82).

INSERT FIGURE 1 HERE

For production no main effect of LEQ was observed but again an interaction with TestLang was seen ($F_{1,342}$ = 91.58, $p < .001$, η^2 =.21). As in comprehension, individual analyses on each of the test languages revealed a negative effect of LEQ on Additional Language scores ($F_{1,342}$ = 17.09, $p < .001$, η^2 =.05) and a positive effect of LEQ on English scores ($F_{1,342}$ = 25.94, $p < .001$, η^2 =.07) (see Figure 1). Overheard speech showed no main effect but did show an interaction with TestLang ($F_{1,342}$ = 34.08, $p < .001$, η^2 =.09). Breaking down the effect for the two test languages showed a significant detrimental effect on Additional Language ($F_{1,342}$= 3.91, p = .049, η^2 =.01) and a beneficial effect on English ($F_{1,342}$ = 143.30, $p < .001$, η^2 =.04) (see Figure 2).

INSERT FIGURE 2 HERE

Income did not have a significant effect ($F_{1,342}$ = 3.39, p =.07) and there was no interaction with TestLang ($F_{1,342}$ =.55, p =.46). There was a main effect of Gender ($F_{1,342}$= 21.00, $p < .001$, η^2=.06), with girls outperforming boys and no interaction with TestLang ($F_{1,342}$= .14, $p <= .71$) (see Figure 3).

INSERT FIGURE 3 HERE

Linear mixed models were then carried out with fixed effects only for those predictors that reached significance in the aforementioned ANCOVAs, with random slopes and intercept for Language Community and random intercepts for participants. Separate models were conducted for each fixed effect variable, with significance assessed by comparing each model against a null in which the fixed effect was absent. Linear mixed models were calculated using the lme4 package (Bates, Maechler, Bolker, & Walker, 2014) in the R environment (R Development Core Team, 2006, version 0.99.896). The coefficients for each model are given in Tables 11 (comprehension) and 12 (production). Note that the effect that is tested in these comparisons is the combined main effect and interaction with TestLang if one was indicated by the ANCOVAs.

INSERT TABLE 11

INSERT TABLE 12

For comprehension there was a significant effect of LEQ ($\chi^2(2) = 18.02$, $p < .001$) and Overheard speech ($\chi^2(2) = 18.62$, $p < .001$) but no effect of Income ($\chi^2(1) = 1.60$, $p = .21$). For production there was a significant effect of LEQ ($\chi^2(2) = 18.75$, $p < .001$), Overheard speech ($\chi^2(2) = 23.59$, $p < .001$) and Gender ($\chi^2(1) = 13.26$, $p < .001$). When these analyses were conducted again with each significant fixed effect entered into a model already containing the other significant fixed effects (Table 13), the last entered fixed effect retained its significance ($p < .006$) in each case. Because we consider the linear mixed models to be the more appropriate significance test for a model that generalises to all bilinguals (Study 3), Income was discarded as a predictor for subsequent analyses. The remaining predictors at the end of Step 1 were the relative amount of exposure to English in child-directed speech (LEQ), the proportion of English in parental overheard speech (Overheard speech), and Gender. This first step allowed us to confirm the robustness of predictors from the literature for the building of a model of the child's lexicon (Study 2), based on data from the 13 target Additional Languages learners.

INSERT TABLE 13 HERE

Step 2 - Secondary predictors. Secondary variables were then added to ANCOVAs containing those predictors shown to be significant in the Step 1 analysis above (LEQ, Overheard speech and Gender), with the 30 words common to all CDIs as dependent variables. These predictors were all the mode of exposure variables, and the three Linguistic Distance variables.

Mode of exposure variables were assessed by adding them individually to ANCOVAs containing LEQ, Overheard speech and, in the case of production, Gender. Societal status (variable Status), degree of language use consistency in parents' input (variable Consistency), presence of siblings (variable Siblings), and number of parental native Additional Language speakers (one or two; variable FamLang) were added to a model containing TestLang as a within–subjects factor. Models with only Additional Language or English test scores omitted the factor of TestLang but included factors describing the native input of test language (variables PropEngN and PropALN), the number of sources of test language (SourcesEng and SourcesAL), and the amount of day care provided in the test language (EngDaycare and ALDaycare).

Only two variables achieved significance: Consistency and PropEngN (see table A7-1 in appendix 7). Consistency interacted significantly with TestLang in determining production scores ($F_{1,355}$ = 3.94, p =. 047, η^2 = .01), due to English vocabulary being boosted by a decreasing consistency in parents' use of the two languages ($F_{1,355}$= 6.07, p = .014, η^2= .017): the more parents used a mix of English and the Additional Language, and the more English vocabulary was produced. The proportion of parental native English spoken (PropEngN) significantly improved English production scores ($F_{1,296}$ = 4.12, p = .043, η^2 = .01).

INSERT TABLE 14 HERE

The effect of Consistency on English production scores was confirmed with a linear mixed model (χ^2(1) = 5.79, p = .016), however, owing to the very small effect size, this variable was not included in the UKBTAT predictive models. The effect of proportion of native English spoken on English production scores was not supported by a linear mixed model (χ^2(1) < 1), and was not retained in the UKBTAT equations.

Finally, the three Linguistic Distance variables were assessed, phonological overlap, word order typology and morphological complexity. Because these showed a perfect to very high association with Language Community, these factors could not be added to ANCOVAs and were assessed in linear mixed models only (see Tables 15 to 17). These revealed a significant effect of Phonological Overlap on Additional Language production (χ^2(1) = 4.61, p = .032), a significant effect of Word Order typology on Additional Language comprehension (χ^2(1) = 6.02, p = .014), and a significant effect of Morphological Complexity on Additional Language comprehension (χ^2(1) = 4.80, p = .028). In all three cases, an advantage was found for children learning an Additional Language close to English. No effects on English were seen.

INSERT TABLE 15 HERE

INSERT TABLE 16 HERE

INSERT TABLE 17 HERE

In summary, all variables but two (Consistency and PropEngN) from Step 2 analyses were excluded at the ANCOVA stage, due to lack of significance. Consistency did have a significant effect in the subsequent linear mixed models but its effect size was too small to warrant integration in the UKBTAT predictive models (Study 2, next Chapter). PropEngN (proportion of English that is native) failed to reach significance in the linear mixed models, and therefore will not be included in the UKBTAT models. All measures of Linguistic Distance were found to be significant in the linear mixed models, fulfilling the predictions of Study 1. It must be kept in mind however that all results obtained at Step 2 should be taken in the context of multiple comparisons and viewed as subject to confirmation in future studies.

Metrics of linguistic distance will not be included in the UKBTAT models (Study 2) because our aim was to build a predictive model for any bilingual child learning British English (Study 3), and measures of linguistic distance will not be available for any Additional Language different from our 13 target languages.

Comparison with full CDI. The effects found in the Step 1 analysis were checked in the full CDI data (the proportion of words in the 100-word Oxford Short Form CDI and in the original Additional Language CDIs). The pattern of significance was identical with effect sizes highly comparable. In particular, linear mixed models once again showed no effect of Income on comprehension ($\chi^2(1) = 1.36$, $p = .24$).

On a final note, in the analyses provided in this paper, we have deliberately ignored item-level analyses as they are beyond our current scope (but see Table A5-2 in appendix 5 for a breakdown of comprehension and production of each English word in the 30-word CDI, as a function of exposure). However, further analyses at this level would provide a privileged insight of the internal organisation of the bilingual lexicon, complementing pioneering studies regarding the processes by which new words are added to the lexicon in monolinguals (e.g., Hills, Maouene, Maouene, Sheya, & Smith, 2009) or bilinguals (Bilson, Yoshida, Tran, Woods, & Hills, 2015). By comparison, where Bilson et al. (2015) collected data in 181 children spanning the age of 6 months to 78 months from eight different English-Additional Language pairs, using a version of the MCDI Toddler form designed for children aged 16 to 30 months of age (Fenson et al., 2007), our dataset includes data for both English and Additional Language (when available) from 430 24-month-olds, collected using age-appropriate tools. One application of this data now currently being conducted is to examine whether phonological overlap modulates the 2-year-old bilingual lexicon in terms of associative relationships and frequency for translation equivalents.

Key Findings

The aim of Study 1 was to establish whether measures of linguistic distance could predict vocabulary size at age 2 in bilingual toddlers. To achieve this, we first showed that vocabulary size was modulated by a set of robust factors from the literature: relative amount of exposure to English versus the Additional Language (English and Additional Language comprehension and production), proportion of English in overheard speech (English comprehension and production) and gender (English and Additional Language production). Two other, less established predictors emerged: first, English production was boosted when each parent used a mix of languages as compared to using one language only – but due to its very small effect size this variable will not be included in the subsequent UKBTAT equations.

Second, the proportion of native English spoken by parents significantly improved English production scores – but this effect did not survive in linear mixed models and therefore will not be included in the UKBTAT. No other variables reached significance, in particular not SES.

Importantly, we found that the three measures of linguistic distance predicted vocabulary knowledge in two-year-olds learning British English and one of 13 target Additional Languages, above and beyond variance due to Language Community: phonological overlap, SVO order typology and morphological complexity. Specifically, we found that children learning a language phonologically close to English had a larger production vocabulary in their Additional Language; similarly, those learning typologically or morphologically close languages to English had a larger Additional Language comprehension vocabulary.

Chapter 4 - Results for Studies 2 and 3: The UKBTAT Model and its Application to Non-Target Additional Language Learners

The aim of Study 2 was to develop the UKBTAT (UK Bilingual Toddlers Assessment Tool), the first screening tool for assessing the vocabulary size of bilingual 2-year-olds, in this case learning British English and one of the 13 target Additional Language. The aim of Study 3 was to establish the reliability of UKBTAT for bilingual children learning British English and any other non-target Additional Language. Norms for bilingual vocabulary in English were obtained through regression equations in linear mixed models, using variables shown to be predictive of comprehension and production in the previous analyses (Chapter 3). Similarly, norms for the Additional Language vocabulary were calculated for children learning one of the 12 target Additional Languages (Spanish is sadly absent from the final list of target languages due to some disagreement with the Spanish CDI editors, TEA Ediciones).

Study 2: Characteristics of the UKBTAT

To be made freely accessible online at www.psy.plymouth.ac.uk/UKBTAT for professionals working with young children and academics, the UKBTAT is similar to the platform used for data collection in this project, but with modifications suited to an applied setting. Firstly, access is secured for practitioners or academics through a personal account, allowing confidential storage of patient test results. Practitioners can use the system to send a link to parents requesting the completion of tests, and have full access to all responses if required. Alternatively tests can be printed and used offline with parents. The tests are still presented in the same order as in this study, with English 100-word Oxford Short Form CDI, full Additional Language CDI (when Additional Language is one of the supported target languages), finishing with the Plymouth Language Exposure Questionnaire. The Plymouth Language Exposure Questionnaire is still the last component that must be filled in by the practitioner, either on the phone or in a live interview with the parent(s). Many of the questions from the family questionnaire and the Plymouth Language Exposure Questionnaire have been merged in an abbreviated Plymouth Language Exposure Questionnaire, which retains only the questions relevant to the significant predictors (amount of exposure, gender, overheard speech).

In the UKBTAT all children are assessed in English, with those whose Additional Language is one of the 12 target languages also assessed in their Additional Language. This data is used to calculate a percentile score for the child's position in their cohort for expressive and receptive vocabulary. For children whose Additional Language is assessed, separate ratings are provided for each language, otherwise only a single rating is provided for English.

Study 2: UKBTAT Predictive Equations

In order to be included in the UKBTAT model, predictors were required to reach significance in the ANCOVAs and the subsequent linear mixed models, and have an effect size of at least η^2 =.02 in the ANCOVAs (see Chapter 3). With these criteria, predictors that made it through Step 1 were the relative amount of English child-directed speech (LEQ), the proportion of English in parental overheard speech (Overheard speech) and Gender. No predictor relating to SES, the source of each language, the properties of the input and the status of the Additional Language survived Step 2 analyses. From the outset, we ruled out including Language Distance predictors in the equations since our aim was to provide models applicable to any Additional Language (Study 3).

Coefficients were obtained from the final mixed models shown in Table 18 . In the case of English, these were obtained from a model run on the full cohort of 430 children, and on the 100-word Oxford Short Form CDI data rather than the 30-word CDI, in order to improve representativeness for the UKBTAT implementation (as mentioned before, the 30-

word CDI shows a ceiling effect for a third of children in English comprehension). For the Additional Languages, the coefficients were obtained from models run on the 372 children who provided Additional Language data, and on the 30-word CDIs.

INSERT TABLE 18 HERE

Altogether, these equations provide predicted scores for a bilingual of unspecified Additional Language in English, and in the Additional Language if it is part of our 12 target languages (Table 19).

INSERT TABLE 19 HERE

For example, an Italian-English girl has 50% exposure to English (LEQ = 50) with parents speaking English and Italian equally often between themselves (Overheard speech = 3; this variable uses a 5-point scale, with 1 = parents always use the Additional Language when addressing one another, 2 = usually the Additional Language, 3 = English about half the time, 4 = usually English, 5 = always English; see Appendix 2). Using the coefficients from the row label “Predicted English 100 score” in Table 19, the equation to be used would be 37.62 + 0.23 x LEQ + 5.19 x Overheard speech. Replacing LEQ with 50 and Overheard speech with 3, the child should have a predicted English score of 64.7 in comprehension, meaning she should understand 65 words from the 100 word Oxford Short Form CDI. Similarly, she obtains a score of 45.9 in English production using the appropriate equation in Table 19, meaning she would produce 46 out of the 100 words from the English Short Form CDI. In the Additional Language she obtains 18.83 in comprehension and 12.02 in production, meaning she should understand 19 words in Italian from the 30-word Italian CDI and produce 12.

As a diagnostic tool, it is important to be able to interpret the difference between a child's predicted and observed scores as a percentile. A reasonable threshold for suspecting a language delay is a score within the 10th percentile (Fenson et al., 2007; Rescorla, 2002; Tomblin, Records, & Zhang, 1996), so access to these ratings will allow practitioners to make an informed decision as to whether a referral might be necessary in a near future, or whether a wait and see approach is more appropriate.

We therefore examined the distribution of observed - predicted residuals in the mixed models from which the coefficients above were generated. Standard deviations of these residuals were as follows: English comprehension 21.19, English production 22.28, Additional Language comprehension 6.04, Additional Language production 7.60. These provide a basis for converting an observed - predicted difference scored in items into a percentile, which is what UKBTAT reports for a child screened by this tool. Thus in English comprehension, an observed - predicted decrement of 21.96 items places a child at exactly the 15th percentile, and a decrement of 27.16 items at exactly the tenth percentile. In the example above, if the Italian-English girl who was predicted an English comprehension score of 64.7 words scored in reality less than 42.7 (that is, 64.7-21.96), she would be in the 15th percentile (see Table 8).

Figure 4 provides the percentiles of word comprehension and production in English, illustrating the gender effect in production, and the well-documented difference between bilingual scores and monolingual data.

INSERT FIGURE 4 HERE

It is worth pointing out that a common practice recommended by Rescorla (1989), known as the Delay 3 cutoff, is to refer for further assessment any (American English) monolingual 2-year-old child who produces fewer than 50 words from the LDS (which contains 310 words, therefore 16%), which identifies about 15% of children. Our findings clearly point to the infeasibility of this "one size fits all" approach for bilingual children: the English vocabulary production score that would be needed to be in the 15th percentile or below varies between 0 and 42.3 (out of 100), depending on the extreme values of the predictors. That is, a boy with the minimum exposure to English, and whose parents would always speak the Additional Language between them, should produce on average 14.0 words out of 100 on the English CDI, and a score of 0 would put him on the 15th percentile. In contrast, a girl hearing 100% English as measured by the LEQ and whose parents always speak English between them, should produce 65.6 on the English CDI, and a score of 42.3 would place her on the 15th percentile.

Study 3: Testing the Predictive Model

We tested the validity of the equations above by seeing how accurately they could be used to predict novel data, namely those of the 58 non-target Additional Language children, which was the aim of Study 3. Predicted scores for these children were generated using the equations (recalculated for the 372 children learning a target Additional Language) and then correlated with the observed scores. This was done for English scores only since non-target children had supplied no Additional Language data. For both comprehension and production, strong correlations were seen between predicted and observed scores (comprehension: $r = .60$; production: $r = .59$). Importantly, there was not any systematic under-prediction or over-prediction of scores in these novel data, as established by t-tests of the means of observed and predicted scores ($t < 1$).

We also calculated the number of children who would be identified as having a delayed language acquisition, by applying the conservative criterion of scoring less than 1 SD (16th percentile) below the mean of the overall distribution (Conti-Ramsden, Botting, & Faragher, 2001). For each child in the 58 non-target Additional Language learners, using the standard deviations of residuals calculated above (UKBT predictive equations derived from the 372 children leaning the target Additional Languages), we converted the difference between each predicted score and the observed score as a percentile. Out of the 58 children, 9 were at or below the 16th percentile (15.5%) in comprehension and 11 in production (19.0%). Five out of the 9 children with low comprehension scores had production scores at or below the 16th percentile, with the other 4 scoring also relatively low in production. Two children scoring very low on production had normal comprehension scores. Given the prevalence of 7-15% of experiencing delayed language acquisition (Kohnert, 2010) in monolingual children, these results suggest a very satisfactory sensitivity for the UBTAT equations.

This demonstrates that the model was not simply fitted to UKBT data a posteriori, but also has predictive validity, not only to fit a new set of data, but also to identify children with potential language delays. Furthermore, as the key test of the validity of our norms was carried out with data from children outside of the range of Additional Languages used to develop the model, it shows that the model is predictive of *general* bilingual vocabulary, whatever Additional Language is spoken by the British-English learning bilingual. We therefore conclude that we have developed bilingual norms for 24-month-olds learning British English and any Additional Language, which have clear cut-offs for referral once the proportions of English in child-directed input and parental overheard speech have been determined, fulfilling the objectives of Studies 2 and 3.

Chapter 5 - General Discussion

Ours is the first study to directly measure the relative contribution of linguistic distance to the acquisition of a bilingual toddler's two languages, which was the aim of Study 1. To address this question we tested 372 24-month-olds learning British English and one of 13 target Additional Languages. We found that a higher phonological overlap between these Additional Languages and British English led to higher levels of Additional Language CDI vocabulary production. Similar effects in comprehension were found for our other measures of linguistic distance, namely degree of similar in morphological complexity and word order. Importantly, linguistic distance contributed unique variance even when other key factors (proportion of English in child-directed input, proportion of English in parental overheard speech and gender) were entered into the same model.

Ours is also the first study to develop bilingual norms specified for the proportion of a particular language that a child hears in the input, which was the objective of Study 2. To establish the reliability of our norms, we applied the model developed with the 372 24-month-olds learning British English and one of 13 target Additional Languages, to a cohort of 58 24-month-olds learning British English and an Additional Language which was not part of our 13 target languages. We found that the English vocabulary scores derived from the model were highly predictive of the vocabulary scores of these 58 children, showing a strong validity of our model for estimating word knowledge in any bilingual toddler growing up in the UK, which was the aim of Study 3. This demonstrates the feasibility of assessing bilingual toddlers in the majority language only, providing professionals working with young children with a practical solution to a long-standing problem in societies where bilinguals come from heterogeneous backgrounds.

Study 1: Linguistic Distance and the Bilingual Lexicon

To explore the effect of linguistic distance on vocabulary outcome, we were faced with the problem of disentangling variance due to linguistic distance from that due to cultural diversity amongst language communities. In our initial analyses we treated our 13 target language pairs as a random factor. This provided an excellent fit between predicted and observed values from bilinguals learning an Additional Language outside of our 13 target languages, indicating that this factor should be included in any future study using heterogeneous sampling of bilinguals. In our second analysis we replaced the random effect of language with a linear predictor of linguistic distance, as measured by the phonological overlap between English words and their translation equivalents in the Additional Language. We found that children's production of Additional Language words was improved when this language was phonologically close to English (such as Dutch, Welsh, German) as compared to more distant languages (such as Cantonese, Polish and Greek). Testing other measures of linguistic distance, we found that children have a receptive vocabulary boost in their Additional Language if they learn a language with the same word order typology as British English (such as Polish or Portuguese) and/or a morphologically close language (such as Cantonese or Mandarin).

Phonological overlap distance. In the literature the possibility that bilinguals do not face the same learning challenges depending on the language pairs being acquired is rarely acknowledged (but see Argyri & Sorace, 2007). Here we established that linguistic distance between a bilingual child's languages shapes their word learning at the age of two. This finding is perhaps less surprising given the recent growing evidence from monolingual research showing that young learners follow slightly different paths in speech perception (e.g., Mani & Plunkett, 2010; Nazzi, 2005), segmentation (e.g., Höhle et al., 2009) and lexical growth (e.g., Bleses et al., 2008) depending on the language they acquire. It is also supported by research examining adult bilingual lexical access which has investigated the effect of lexical overlap of translation equivalents on lexical access (e.g., Costa, Caramazza, & Sebastián-Gallés, 2000; Strijkers, Costa, & Thierry, 2009). Cognates, translation equivalents with form overlap, such as *bed* and *Bett* in German, are known to provide an advantage to production, similar to the linear measure of phonological overlap used in the current study. In word production, cognates are produced faster (e.g., Costa et al., 2000; Hoshino & Kroll, 2008), elicit a different brain signature (e.g., Strijkers et al., 2009), and produce higher levels of activation in priming tasks (e.g., Colomé & Miozzo, 2010) than non-cognates. The cognate advantage in picture naming is also found in young proficient bilinguals as early as 4 years (Sheng, Lam, Cruz, & Fulton, 2016; see also Poarch & van Hell, 2012).

To explain this advantage it has been proposed that cognates have different representations than non-cognates in the lexicon, perhaps because of a larger conceptual overlap (van Hell & De Groot, 1998), a shared morphological representation (Sánchez-Casas & García-Albea, 2005), or because they might have been learned earlier (Costa, Pannunzi, Deco, & Pickering, 2016). Alternatively, it has been proposed that the cognate advantage is the by-product of the dynamic interactions between the lexical and the phonological (and orthographical) levels of processing within the lexicon (Costa, Sanesteban, & Caño, 2005). Altogether, these findings with cognates show good equivalence with our own findings, where phonological overlap between translation equivalents increased Additional Language expressive vocabulary. This finding also provides support to the proposal that the cognate advantage is due to cognates being acquired before non-cognates in early childhood (Costa et al., 2016), leading to an ease of processing later in life.

Finally, the cognate advantage is an emerging property of recent computational models of the bilingual lexicon (BLINCS: Shook & Marian, 2013; DEV-LEX II: Zhao & Li, 2010): the activation of a concept in a common semantic lexicon during word production would generate parallel top-down activation of phonological representations from lexical representations in both languages. Any segments that overlapped between language representations would receive top-down activation from both of the languages lexical representations. This would provide an advantage to cognate or more overlapped translation equivalents, as they would become activated faster than those with less or non-overlapped representations (see also Costa et al., 2005).

One aspect of our results worth noting is that, although we found an increase in Additional Language vocabulary with phonological overlap in production, we found no effect in comprehension. This disparity also has some parallels in prior literature, where the evidence for a cognate advantage in bilingual spoken word recognition is mixed (see a review in Lagrou, Hartsuiker, & Duyck, 2011; see the review by Sánchez-Casas & García-Albea, 2005, for the cognate effect in *visual* word recognition). Although there is evidence for non-selective lexical access in word recognition as is found in production (Lagrou et al., 2011), that is, for online activation of words in both language, Shook, Goldrick, Engstler and Marian (2015) did not find any cognate advantage in a word recognition task in spoken English sentences in German-English bilinguals. When using an eye-tracking word recognition task for sets of pictures, Blumenfeld and Marian (2007) reported a cognate advantage dependent on the listener's proficiency in the target language. That is, whereas a cognate and non-cognate equally activate words in a highly proficient language, only cognates boost the activation of words in a less proficient language. In development all prior studies of cognate advantage in spoken word recognition have been with Spanish-English bilinguals. Again, findings are mixed, with some studies not finding any advantage in receptive vocabulary recognition (in first graders: Umbel, Pearson, Fernandez, & Oller, 1992; first to sixth graders: Umbel & Oller, 1994), while others did find an advantage (fifth and sixth graders:

Cunningham & Graham, 2000; 8 to 13 years old: Kelley & Kohnert, 2012; kindergarten and first graders: Pérez, Peña, & Bedore, 2010). Kelley and Kohnert (2012) actually directly compared the cognate effect in production and comprehension using standardised tests, and found a very small advantage in comprehension, and a small to medium effect in production. In comprehension a large proportion of the variance in the cognate advantage was due to age, with older children showing a greater advantage than younger ones.

In sum, our findings are the first to demonstrate that phonological overlap at the lexical level appears to boost the acquisition of expressive vocabulary in bilingual toddlers, which fits nicely with past demonstrations of a cognate advantage in older children and adults.

Syntactic and morphological distance. Our findings of a facilitatory effect of word order typology distance on receptive vocabulary can be explained in three, non-exclusive ways. First, children learning two languages with the same word order, such as British English and Polish, can probably use similar phrase-level prosodic cues for segmentation of syntactic constituents: indeed VO languages (such as English and French) primarily use duration to express prosodic prominence of a content word as compared to a functor, while OV languages (such as Bengali and Hindi) tend to use pitch/intensity (Nespor et al., 2008). The search for similar cues in the speech signal would lead to a single mechanism for prosodic-driven segmentation, boosting the retrieval of words and the assignment of syntactic categories. An additional explanation of this facilitatory effect is that, instead of activating both languages, bilinguals transfer the structure of their native language processing to that of the new language (Costa et al., 2016). Although this proposal applies to lexical processing in sequential bilinguals, the idea that a processing structure for one language can be 'carried over' to the other language also has currency for word order computation. Finally, some studies have proposed that word order variations amongst the world's languages are constrained by computational or learnability limitations, leading to a proliferation of easy-to-learn orders (e.g., Ferrer-i-Cancho, 2015; Lupyan & Christiansen, 2002). In this perspective, pairs of languages that differ on the word order dimension will add to the learnability issue by adding computational complexity.

In addition to word order typology we also found a facilitative effect of morphological similarity, with children learning Additional Languages morphologically close to English, such as isolating Cantonese or Mandarin, having better Additional Language comprehension vocabulary than those with a distant Additional Language, such as the agglutinative Bengali. In monolingual research, it has been argued that complex morphological systems may hinder language acquisition (e.g., Slobin, 1973) - although complex systems can be learnt quickly provided that morphological rules are regular and obligatory (e.g., Devescovi et al., 2005; Kim, McGregor, & Thompson, 2000). Our findings suggest that bilinguals learning languages with similar morphology may benefit from the training of cognitive resources engaged in supporting the learning of one kind of morphology over another, for example, memory for isolating languages, versus rule-based learning for synthetic languages, as suggested by Fortescue and Lennert Olsen (1992). Some aspects of this cognitive training might be driving the bilingual cognitive advantage (Bialystok, 2009), raising the interesting possibility that linguistic distance could be partially responsible for the elusive nature of this cognitive advantage (Paap & Greenberg, 2013).

On a final note, when considering these findings it should be noted that language distance was found to modulate Additional Language scores only, with English scores being more resistant to facilitatory effects from the Additional Language proximity. We interpret this as showing that English acquisition benefits from the overarching effect of the English-speaking environment, whereas the Additional Language acquisition relies, in most cases, solely on one or two parents' input. While measures of exposure do not quantify the weight of everyday social interactions (shopping, visit to the doctor, media exposure, etc), a telling outcome that can be derived from Figure 1 is that in order for a bilingual child from our corpus to know an equal proportion of words in English and Additional Language, her/his exposure to English as measured by the Plymouth Language Exposure Questionnaire only needs to be at 30% of English in comprehension (they would then know 23 out the 30 common words in English and the Additional Language) and 20% in production (they would then produce 14 out of the 30 common words in English and the Additional Language). These low English exposure values clearly fail to capture the overwhelming influence of the language of the surrounding community, and point to the relative vulnerability of the growth of the Additional Language.

In this study we have focussed on three measures of linguistic distance, phonological overlap of the child lexicon, syntax typology and morphological complexity, that reflect the core aspects of language development in the second year of life. However, 'linguistic distance' is fundamentally complex, with a multidimensional representation that goes beyond the factors examined thus far. Our database provides an opportunity to continue the exploration of how additional measures of linguistic distance can account for variations in bilingual vocabulary, for example, through examination of the effect of prosody on early word development (e.g., rhythmic families, final word lengthening, etc; White et al., in prep), or the effect of cross-linguistic differences in infant-directed speech styles (e.g., Fernald et al., 1989).

In conclusion, the finding that linguistic distance shapes vocabulary knowledge in the early years strongly supports the idea of independent (Genesee et al., 1995), yet interfering language systems from the onset of development (DeAnda et al., 2016, Von Holzen & Mani, 2012). It is unknown at this stage whether the facilitation effects we found for phonologically or structurally close languages are due to the sharing of information processing mechanisms, or to the duplication of processing structures from one language to the next, following the idea of a transfer proposed by Costa et al. (2016). Further research will be needed to discover the mechanisms underlying language transfers in early childhood, in order to shed some light on how the brain organises information sharing when two systems compete for resources.

Study 2: Predicting Vocabulary Scores in Bilingual Toddlers

The second key aim of the current paper was to develop a model of English and Additional Language vocabulary to be used with UK-raised bilinguals learning one of the target Additional Languages. To this aim we investigated which predictors should be included in this model, out of an extensive inventory which we summarise below (see also Table 1).

Relative amount of exposure to languages. As expected, the most robust predictor of English and Additional Language vocabulary was found to be the relative amount of exposure to child-directed English versus the Additional Language (e.g., Hoff, 2003). This was predictive of both comprehension and production, with greater exposure to English increasing vocabularies in English, and reducing them in the Additional Language. Similarly, the proportion of English/Additional Language spoken *between* parents was also strongly predictive of comprehension and production in English, in line with previous studies showing that children encode information from overheard speech (Akhtar, 2005; Shneidman, Arroyo, Levine, & Goldin-Meadow, 2013).

Mode of Exposure. We explored the predictive value of factors related to the source, properties, and status of the mode of exposure to English and the Additional Language. Only the two predictors related to the properties of the input survived our analytical process - yet statistical criteria prevented their inclusion in the final predictive models: the proportion of native English spoken to the child by parents, and the degree of language use consistency in parents' speech.

Source of each language. First we examined whether the fact that two parents were native Additional Language speakers, or only one, had an impact on vocabulary knowledge, as was suggested by findings by Barrena et al. (2008). This factor did not have any significant impact in our data, perhaps because much of the variance associated to this variable was already apportioned to the relative amount of exposure to each language.

Second we looked at whether the total number of native English speakers, or the total number of Additional Language speakers around the child would affect word knowledge in each language. This was previously found by Gollan et al (2015), testing 8-year-old Hebrew-English bilinguals in a picture naming task, who reported that children interacting with more Hebrew speakers could name more pictures in this language (see also Place & Hoff, 2011). However Place and Hoff (2016) only reported modest effects of the number of speakers in Spanish-English 30-month-olds, and only in Spanish. Here we did not find any significant advantage for being surrounded by many native speakers. Further research is needed to clarify the impact of this factor on vocabulary growth, as it is of major theoretical interest to delineate the role of input variability in phonological development (Rost & McMurray, 2009).

Thirdly we found no effect of the time spent in daycare (in each language) on vocabulary development. Given that daycare attendance in monolingual children tends to benefit only those from low SES (Côté et al., 2013), and given that our sampling failed to capture a significant portion of low SES families, it is perhaps not surprising that we did not report any effect of this factor.

Finally, regarding the effect of the number of older siblings, predictions from the literature were mixed, with monolingual data pointing to larger vocabularies overall in first-born children (Huttenlocher et al., 2010), and bilingual studies suggesting larger English vocabularies for North American bilingual toddlers with older siblings (Bridges & Hoff, 2014). Here we did not report any significant effect of this factor, which perhaps is related to the fact that in our data collection we did not distinguish between toddlers who had school-aged siblings and toddlers whose siblings were still too young. Indeed Bridges and Hoff (2014) reported that it was school-aged siblings who were mainly responsible for an increase in the proportion of English spoken at home and the resulting increase in toddlers' English vocabulary.

Properties of the input. We found that English production increased with the proportion of native English parental input, in line with previous findings that native speakers provide more supportive linguistic input than non-natives (Hammer et al., 2012; Place & Hoff, 2011, 2016; although see Paradis, 2011, who found no effect). Fernald (2006) suggested that non-native speech might hinder the development of phonological categories, but this appears contradicted by the fact that we did not report any *negative* effect of non-native input on vocabulary. Our findings may be better accounted for by non-native speakers using less varied vocabulary than native speakers, as suggested by Hoff et al. (2013).

Finally, we found that the degree of language use consistency in parents' speech was positively associated with English production, so that children whose parents used a mix of the two languages knew more words in English, perhaps because more code-switching increases the relative frequency of the majority language. However, previous findings on the impact of within-speaker consistency (measured at the sentence level as in code-switching, or at the discourse level as in the degree of language use consistency) suggest mixed outcomes: Byers-Heinlein (2013) reported a detrimental effect of code-switching at 18 months, but only marginally at 24 months, whereas Place and Hoff (2016) reported no robust relations between the degree of language use consistency in mothers' speech and language development at 30 months. In addition to an age-related explanation, we argued that the discrepancy between these two prior studies would be due to language distance, as Byers-Heinlein tested mainly distant languages learners whilst Place and Hoff examined close languages learners. To examine whether the positive effect found in the current study was different for close and distant language learners, we ran separate post-hoc regression analyses on the median split of children learning close (N = 240) and distant (N = 132) languages (based on phonological overlap). For distant language learners, the degree of language use consistency in parental speech did not predict comprehension nor production, once corrected for exposure to English and the proportion of English in overheard speech. However, for children learning close languages, the degree of language use consistency in parental speech contributed significantly to improve English production scores (main model: F (4, 239) = 11.34, $p < .001$; standardised *ß* for language mixing = 0.15, $p = .016$; no effect for comprehension, or for Additional Language scores). These results do provide some preliminary support for our original hypothesis: distant language learners perform poorly with language mixing input measured through code-switching (Byers-Heinlein, 2013) or show no effect (this study), while close language learners either benefit from language mixing measured through language use consistency (this study) or show no effect (Place & Hoff, 2016). The explanation behind these findings would be that frequent switches from one set of representations to another in the case of distant languages impinge on cognitive resources, whereas close languages activate overlapping representations and therefore do not necessitate an increased cognitive flexibility.

Status of the Additional Language (societal vs. minority). We hypothesized that the societal status of the Additional Language would have consequences for children's achievements in the two languages, because of people's more positive attitude towards bilingualism, and also because of a more balanced exposure to the two languages in everyday life. However, Welsh-English toddlers in Wales did not have significantly higher vocabulary scores in their two languages than the other bilinguals, whether these scores were corrected by situational factors or not. In the literature, Welsh-English bilinguals' vocabulary in

English tends to be lower than English monolinguals' (Rhys & Thomas, 2013) at the ages of 7 and 11 years, with a strong effect of language dominance (as defined by exposure and family language) and a complex relation between home language exposure, SES, and age (Gathercole, et al., 2016). Our study shows that this "bilingual difference", previously found in Welsh, is similar to that of any other bilingual minority group in the UK, at least at the age of 24 months. This point contributes to strengthen the generality of the UKBTAT model for predicting language outcomes in any incoming bilingual toddler growing in the UK. Importantly, the fact that there was no effect for the societal status of the Additional Language being learned implies that our findings apply both to bilingual toddlers learning a minority language and also to those learning two languages which have more equivalent societal status, such as perhaps English and French in Canada.

Demographic factors (gender and SES). As expected, gender was found to be a reliable predictor of production vocabulary in both languages, with girls producing more words than boys (Huttenlocher et al., 1991); no effect was found on comprehension, which is also in line with previous findings (Eriksson et al., 2012).

Contrary to our initial expectations, SES effects were absent in our data, in keeping with prior studies which also had much reduced sampling at the low end of the SES spectrum (Fenson et al., 2007; Hamilton et al., 2000). The effects of these missing low SES children on the accuracy of UKBTAT norms are unclear. If they had been fully represented then we might have expected reduced vocabulary scores for low SES children, meaning that the current UKBTAT model would over-refer low SES children, and under-refer high SES children. There are a number of dangers with this assumption however, beyond those inherent in assuming that SES would have reached significance if provided with a fully balanced spectrum of samples. Fernald et al. (2013) reported that their SES effect, one of the strongest found in the literature, was stronger at the lower end of the SES spectrum. We confirmed this with a re-analysis of data points retrieved from Fernald et al's published scatterplot using the application Plot Digitizer (www.plotdigitizer.sourceforge.net), first reproducing the originally reported correlation of .34. We then went on to split this data at median SES and analyse high and low SES data separately, finding that the correlation of SES with CDI comprehension was significant in low SES children (r = .42, *p* = .001), but not in high SES children (r = .27, *p* > .05). The nonlinearity of this effect means that it would be poorly modelled by linear equations, such as those used in the UKBTAT.

In summary, the three factors that were identified to be the key predictors of vocabulary knowledge in 2-year-old bilinguals learning English and one target Additional Languages were the relative amount of exposure to English versus the Additional Language in child-directed input (English and Additional Language production and comprehension) and parental overheard speech (English production and comprehension), and gender (Additional Language and English production). The resulting models were built into the UKBTAT assessment tool, which fulfilled the objectives of Study 2. The UKBTAT uses measures of English vocabulary on the 100-word Oxford Short Form CDI, and if the child learns one of the target Additional Languages, the full Additional Language CDI is administered, from which the 30 words common to all CDIs are extracted to derive predictions. In addition, the features of the UKBTAT (online, based on parental report and interview) are particularly appealing for an initial screening at such an early age, where language assessments are notoriously challenging and yet very desirable to plan appropriate intervention.

Application of the UKBTAT with Bilingual Children with Non-Target Additional Languages

We created the UKBTAT norms for assessing the vocabulary of any bilingual 24-month-old learning British English and an Additional Language, regardless of which Additional Language it would be. These norms are available in British English, and in the

Additional Language when this is part of our target languages. The recommendations of professional bodies are that bilinguals should be assessed in both their languages as there is a, quite justified, risk that an individual child's situational factors may render an assessment performed on only one language unrepresentative of their overall linguistic capability. Unfortunately, for pragmatic reasons due to the heterogeneity of additional languages spoken in the world - and in the UK in particular - this recommendation can only be followed in a relatively small minority of cases, with bilinguals generally only assessed in the majority language. Therefore, in this project we have sought to identify the situational factors that affect bilingual acquisition and quantify their contribution to improve the quality of English only testing, in the case of UK-raised toddlers. We have shown that only the relative amount of English exposure in child-directed input and overheard speech interacted with test language in the determination of vocabulary scores, thus we have no evidence that situational factors beyond these apply differently to the two test languages. This means there is no reason to suppose that Additional Language scores provide useful information beyond that of an English score, once corrected for these factors. Therefore, for clinical purposes, information provided by English vocabulary scores alone appears to provide an effective proxy for overall linguistic attainment, and can be safely used as normative data for any new incoming bilingual toddler. It is our hope that this pragmatic approach to bilingual screening will be generalised to other countries who are, like the UK, facing a growing number of bilingual infants from an increasingly heterogeneous background of languages.

Limitations and Future Directions

An important limitation of this study pertains to the relatively small proportion of children with a low level of exposure to English, either because of SES characteristics, date of arrival in the country or simply family characteristics. Regarding the poor representation of low SES families in our sample (only 4 children out of the cohort scored on the lowest income band): as noted earlier, the current UKBTAT model would probably over-refer low SES children and, possibly, under-refer high SES children; only further research will allow to address this issue satisfyingly. Regarding date of arrival in the country, it must be noted that we did not include age of acquisition in the analyses, because the vast majority of toddlers were born in the UK (see participants section), which we took as a reasonable proxy for simultaneous acquisition, but it would be unreasonable to use the UKBTAT for a child whose age of acquisition of British English is only a few months prior assessment: not only is the effect of age of acquisition a very established predictor of bilingual development (Flege, Yeni-Komshian, Liu, 1999), but the Plymouth Language Exposure Questionnaire, which is implemented in the UKBTAT, does not provision (yet) for variation in the age of acquisition. Finally, the characteristics of the families included in the survey were such that only 7% of the 430 children had less than 10% exposure to English (as measured by the Plymouth Language Exposure Questionnaire). This probably relates again to the poor sampling of low SES families, and calls for caution in the use of the UKBTAT when assessing toddlers with very low English exposure.

In the same vein, another limitation relates to the mode of calculation of exposure to each language, which was here exclusively focused on the *relative* exposure (as most exposure questionnaires do: e.g., De Houwer & Bornstein, 2003; Paradis, 2011), and not on the *quantity* of exposure. That is, a child who scored 60% of English on our measure of exposure could experience a few hours of cumulated English speech per week or a few dozen, depending on the communication style of her caregivers. Measures of quantity (obtained through naturalistic recordings) are known to predict language growth in young monolinguals (Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010), and thus a full measure of exposure should encompass both relative and absolute measures of language input.

Additional research will be needed to evaluate if quantity of bilingual exposure explains an additional portion of the variance in explaining vocabulary growth.

Another limiting factor relates to the age of toddlers, which we selected to be two years old, as an important milestone for parents, researchers and practitioners, in time for preschool referral if needed. Whether our key findings would hold for younger and older children remain to an empirical question, but we anticipate that measures of exposure and gender would remain central predictors of bilingual development throughout development (see Hyde & Linn, 1988 for effects of gender through development). What is more interesting is how linguistic distance metrics would predict language development before and after the age of two. Generally speaking, it is likely that the type of distance measures that would modulate bilingual language growth would depend both on the age of the child and to-be-assessed language skills, with perhaps phonological overlap metrics becoming less influential as grammar and morphology grow in the child. We are currently exploring these claims as a substantial subset of the cohort tested in this paper were re-assessed at age 3 with the same tools.

Finally, the language-specificity of our findings implies an obvious caveat, as the UKBTAT is currently valid only for UK-raised toddlers who learn British English as one of their languages. However there is no reason to think that the key results – that linguistic distance is predictive of vocabulary growth, and that a useful model of the early lexicon needs only a handful of predictors – would be language-specific. Of particular interest would be to replicate this cohort study in North America, as minimal adjustments to the 100-word Oxford Short Form CDI would be required, allowing the problem of assessing a growing number or bilinguals from heterogeneous populations (Statistics Canada, 2011; United States Census 2020) to be addressed.

Conclusion

The current research included three key strands of research questions. In the first study we established that linguistic distance between the two languages plays a role in predicting bilingual toddlers' vocabulary at age 2, a result which moved the field forward. Given the interweaving of word-learning abilities and cognitive processes in development (Baddeley, Gathercole, Papagno, 1998), the impact of linguistic distance might extend beyond language acquisition, and could constitute a key factor in the study of the debated bilingual cognitive advantage (Bialystock, 2009; Paap & Greenberg, 2013).

In Study 2, which had an applied orientation, we identified the amount of exposure to each language in child-directed and overheard speech as key predictors of vocabulary development in bilingual toddlers, along with infant gender, and developed a model of the bilingual lexicon usable for assessment purposes, the UKBTAT. Finally, in Study 3, we explored the feasibility of assessing a bilingual toddler's vocabulary in the majority language only, in order to provide professionals working with young children with a practical solution when encountering bilingual children from heterogeneous backgrounds, as is the case is most of the world. The validity of this approach was successfully demonstrated when the UKBTAT model was applied to a new cohort of bilinguals learning a different Additional Language to our original target languages. It is our hope that the UKBTAT will enhance the early detection of children at risk of language delays in the growing UK bilingual population, and perhaps in other countries where the same approach could be applied.

References

Abbot-Smith, K., Lieven, E., & Tomasello, M. (2008). Graded representations in the acquisition of English and German transitive constructions. *Cognitive Development*, *23*(1), 48-66.

Abdelwahab, A., Stone, C., Slee, C., Cattani, A., & Floccia, C. (2016). Evaluating existing language exposure questionnaires for the study of language development in bilingual children between 12- 24 months Poster presented at the *The Young Language Learners (YLL) Symposium 2016*, Department of Education, Oxford University, Oxford, UK, 6-7 July.

Adam, G., & Bat-El, O. (2009). When do universal preferences emerge in language development? The acquisition of Hebrew stress. *Brill's Journal of Afroasiatic Languages and Linguistics*, *1*(1), 255-282.

Akhtar, N. (2005). The robustness of learning through overhearing. *Developmental Science, 8*(2), 199-209.

Akhtar, N., & Tomasello, M. (1997). Young children's productivity with word order and verb morphology. *Developmental Psychology*, *33*(6), 952.

Alcock, K. J., Meints, K., Rowland, C. F., Christopher, A., Just, J., & Brelsford, J. (in prep). The UK Communicative Development Inventory: Words and Gestures.

American Speech-Language-Hearing Association (ASHA) (1999). *Guidelines for the Roles and Responsibilities of School-Based Speech–Language Pathologists.* Rockville, MD: American Speech–Language–Hearing Association.

Argyri, E., & Sorace, A. (2007). Crosslinguistic influence and language dominance in older bilingual children. *Bilingualism: Language and Cognition*, *10*(01), 79-99.

Armon-Lotem, S., & Ohana, O. (2017). A CDI study of bilingual English-Hebrew children – frequency of exposure as a major source of variation. *International Journal of Bilingual Education and Bilingualism, 20(2)*, 201-217.

Arriaga, R. J., Fenson, L., Cronan, T., & Pethick, S. J. (1998). Scores on the MacArthur Communicative Development Inventory of children from low- and middle-income families. *Applied Psycholinguistics*, *19*, 209-223.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390-412.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*(1), 158.

Barac, R., & Bialystok, E. (2012). Bilingual effects on cognitive and linguistic development: Role of language, cultural background, and education. *Child Development*, *83*(2), 413-422.

Barrena, A., Ezeizabarrena, J., & Garcia, I. (2008). Influence of the linguistic environment on the development of the lexicon and grammar of Basque bilingual children. *A Portrait of the Young in the New Multilingual Spain*, *9*, 86.

Bates, E., MacWhinney, B., Caselli, C., Devescovi, A., Natale, F., & Venza, V. (1984). A cross-linguistic study of the development of sentence interpretation strategies. *Child Development*, *55*(2), 341-354.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. *R package version*, *1*(7).

Bavin, E. (1992). The acquisition of Warlpiri. In D. T. Slobin (Ed.), *The crosslinguistic study of language acquisition*, *3*, 309-372. Hillsdale, NJ: Lawrence Erlbaum.

Bedore, L. M., & Peña, E. D. (2008). Assessment of bilingual children for identification of language impairment: Current findings and implications for practice. *International Journal of Bilingual Education and Bilingualism*, *11*(1), 1-29.

Beitchman, J. H., Brownlie, E. B., Inglis, A., Wild, J., Mathews, R., Schachter, D., ... & Lancee, W. (1994). Seven-year follow-up of speech/language-impaired and control children: Speech/language stability and outcome. *Journal of the American Academy of Child & Adolescent Psychiatry*, *33*(9), 1322-1330.

Bergmann, C., & Cristia, A. (2016). Development of infants' segmentation of words from native speech: A meta-analytic approach. *Developmental Science*, *19*(6), 901-917.

Bernard, C., & Gervain, J. (2012). Seven-month-old infants expect converging cues to word order: The joint use of word frequency and phrasal prosody. *Frontiers in Psychology, 3*:451.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition, 12*(1), 3-11.

Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, *13*(04), 525-531.

Bialystok, E., Majumder, S., & Martin, M. M. (2003). Developing phonological awareness: is there a bilingual advantage? *Applied Psycholinguistics*, *24*, 27–44.

Bleses, D., Vach, W., Slott, M., Wehberg, S., Thomsen, P. I. A., 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.

Bilson, S., Yoshida, H., Tran, C. D., Woods, E. A., & Hills, T. T. (2015). Semantic facilitation in bilingual first language acquisition. *Cognition*, *140*, 122-134.

Blumenfeld, H. K., & Marian, V. (2007). Constraints on parallel activation in bilingual spoken language processing: Examining proficiency and lexical status using eye-tracking. *Language and Cognitive Processes*, *22*(5), 633-660.

Bornstein, M. H., Hahn, C.-S., Suwalsky, J. T. D., & Haynes, O. M. (2003). Socioeconomic status, parenting, and child development: The Hollingshead four-factor index of social status and the socioeconomic index of occupation*s*. In M. H. Bornstein & R. H. Bradley (Eds.), *Socioeconomic status, parenting, and child development* (pp.29–82). Mahwah, NJ: Erlbaum.

Bouchon, C., Floccia, C., Fux, T., Adda-Decker, M., & Nazzi, T. (2015). Call me Alix, not Elix: Vowels are more important than consonants in own-name recognition at 5 months. *Developmental Science*, *18*(4), 587-598.

Bosch, L., & Ramon-Casas, M. (2014). First translation equivalents in bilingual toddlers' expressive vocabulary: Does form similarity matter?. *International Journal of Behavioral Development, 38*(4), 317-322.

Bosch, L., & Sebastián-Gallés, N. (1997). Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition, 65*, 33-69.

Bosch, L., & Sebastián-Gallés, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy*, *2*(1), 29-49.

Bosch, L., & Sebastián-Gallés, N. (2003). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech*, *46*(2-3), 217-243.

Britto, P. B., Fuligni, A. S., & Brooks-Gunn, J. (2002). Reading, rhymes, and routines: American parents and their young children. In N. Halfon, K.T. McLearn, & M.A. Shuster, (Eds.), *Child rearing in America: Challenges facing parents with young children* (pp. 117-145). New York, NY, US: Cambridge University Press.

Bridges, K., & Hoff, E. (2014). Older sibling influences on the language environment and language development of toddlers in bilingual homes. *Applied Psycholinguistics*, *35*(2), 225–241.

Brooks-Gunn, J., Han, W., & Waldfogel, J. (2002). Maternal employment and child cognitive outcomes in the first three years of life: The NICHD Study of Early Child Care. *Child Development*, *73*(4),1052– 1072.

Brown, R. (1973). *A first language: The early stages*. London: George Allen & Unwin.

Brown, D. (2010). Morphological typology. In J. J. Song, (Ed.), *Handbook of linguistic typology* (pp. 487-503). Oxford: Oxford University Press.

Brownell, R. (2001). *Expressive one-word picture vocabulary test manual: Spanish-bilingual edition*. Novato, CA: Academic Therapy Publications.

Burns, T. C., Werker, J. F., & McVie, K. (2003). Development of phonetic categories in infants raised in bilingual and monolingual environments. In B. Beachley, A. Brown, & F. Conlin (Eds.), *Proceedings of the 27th Annual Boston University Conference on Language Development* (Vol. 1, pp. 173–184). Somerville, MA: Cascadilla Press

Byers-Heinlein, K. (2013). Parental language mixing: Its measurement and the relation of mixed input to young bilingual children's vocabulary size. *Bilingualism: Language and Cognition*, *16*(1), 32–48.

Calvo, A., & Bialystok, E. (2014). Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition, 130*(3), 278-288.

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, i-174.

Caselli, C., Casadio, P., & Bates, E. (1999). A comparison of the transition from first words to grammar in English and Italian. *Journal of Child Language*, *26*(01), 69-111.

Cattani, A., Abbot-Smith, K., Farag, R., Krott, A., Arreckx, F., Dennis, I., & Floccia, C. (2014). How much exposure to English is necessary for a bilingual toddler to perform like a monolingual peer in language tests? *International Journal of Language & Communication Disorders*, *49*(6), 649-671.

Charman, T., Drew, A., Baird, C., & Baird, G. (2003). Measuring early language development in preschool children with autism spectrum disorder using the MacArthur Communicative Development Inventory (Infant Form). *Journal of Child Language*, *30*(1), 213-236.

Chiat, S., & Roy, P. (2008). Early phonological and sociocognitive skills as predictors of later language and social communication outcomes. *Journal of Child Psychology and Psychiatry*, *49*(6), 635-645.

Chiswick, B. R., & Miller, P. W. (2005). Linguistic distance: A quantitative measure of the distance between English and other languages. *Journal of Multilingual and Multicultural Development*, *26*(1), 1-11.

Christophe, A., Nespor, M., Guasti, M. T. & Van Ooyen, B. (2003). Prosodic structure and syntactic acquisition: The case of the head-direction parameter. *Developmental Science*, *6*, 211–220.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge. ISBN 1-134-74270-3.

Colomé, À., & Miozzo, M. (2010). Which words are activated during bilingual word production? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(1), 96.

Comrie, B. (1989). *Language universals and linguistic typology*. Oxford: Blackwell.

Conboy, B. T., & Thal, D. J. (2006). Ties between the lexicon and grammar: Cross-sectional and longitudinal studies of bilingual toddlers. *Child Development*, *77*(3), 712-735.

Conti-Ramsden, G., & Botting, N. (1999). Classification of children with Specific Language Impairment: Longitudinal considerations. *Journal of Speech, Language, and Hearing Research*, *42*(5), 1195-1204.

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*(6), 741-748.

Conti-Ramsden, G., Crutchley, A., & Botting, N. (1997). The extent to which psychometric tests differentiate subgroups of children with SLI. *Journal of Speech, Language, and Hearing Research*, *40*(4), 765-777.

Conway, A. R., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, *7*(12), 547-552.

Core, C., Hoff, E., Rumiche, R., & Señor, M. (2013). Total and conceptual vocabulary in Spanish–English bilinguals from 22 to 30 months: Implications for assessment. *Journal of Speech, Language, and Hearing Research*, *56*(5), 1637-1649.

Costa, A., Caramazza, A., & Sebastián-Gallés, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1283.

Costa, A., Pannunzi, M., Deco, G., & Pickering, M. J. (2016). Do bilinguals automatically activate their native language when they are not using it? *Cognitive Science*. 10.1111/cogs.12434.

Costa, A., Santesteban, M., & Caño, A. (2005). On the facilitatory effects of cognate words in bilingual speech production. *Brain and Language*, *94*(1), 94-103.

Côté, S. M., Doyle, O., Petitclerc, A., & Timmins, L. (2013). Child care in infancy and cognitive performance until middle childhood in the millennium cohort study. *Child Development*, *84*(4), 1191-1208.

Crutchley, A., (2000). Bilingual children in language units: does having ‘well-informed parents’ make a difference? *International Journal of Language and Communication Disorders, 35*, 65–81.

Cunningham, T. H., & Graham, C. (2000). Increasing native English vocabulary recognition through Spanish immersion: Cognate transfer from foreign to first language. *Journal of Educational Psychology, 92*(1), 37.

Curtin, S., Byers-Heinlein, K., & Werker, J. F. (2011). Bilingual beginnings as a lens for theory development: PRIMIR in focus. *Journal of Phonetics, 39*(4), 492-504.

Dale, P. S., Price, T. S., Bishop, D. V. M., & 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*, *46*, 544-560.

DeAnda, S., Poulin-Dubois, D., Zesiger, P., & Friend, M. (2016). Lexical processing and organization in bilingual first language acquisition: Guiding future research. *Psychological Bulletin, 142*(6), 655-667.

De Houwer, A. (1999). Environmental factors in early bilingual development: The role of parental beliefs and attitudes. In G. Extra and L. Verhoeven (Eds.), *Bilingualism and migration: Studies on Language acquisition*, pp75-96. Berlin: Mouton de Gruyter.

De Houwer, A. (2007). Parental language input patterns and children's bilingual use. *Applied Psycholinguistics*, *28*(3), 411-424.

De Houwer, A., & Bornstein, M. (2003, April). Balancing on the tightrope: Language use patterns in bilingual families with young children. *Fourth International Symposium on Bilingualism, Tempe, Arizona*.

De Lamo White, C., & Jin, L. (2011). Evaluation of speech and language assessment approaches with bilingual children. *International Journal of Language & Communication Disorders*, *46*(6), 613-627.

Delle Luche, C., Floccia, C., Granjon, L., & Nazzi, T. (2016). Infants' first words are not phonetically specified: Own name recognition in British English-learning 5-month-olds. *Infancy, 22*(3), 362-388

Deutsch, M. (1965). The role of social class in language development and cognition. *American Journal of Orthopsychiatry, 35(1),* 78-88.

Devescovi, A., Caselli, M. C., Marchione, D., Pasqualetti, P., Reilly, J., & Bates, E. (2005). A crosslinguistic study of the relationship between grammar and lexical development. *Journal of Child Language*, *32*(04), 759-786.

Dijkstra, A. (2005). Bilingual visual word recognition and lexical access. In J. F. Kroll & A. De Groot, (Eds.). *Handbook of Bilingualism: Psycholinguistic Approaches* (pp. 178–201). Oxford University Press.

Dryer, M. S. (1991). SVO languages and the OV: VO typology. *Journal of Linguistics*, *27*(02), 443-482.

Duff, F. J., Reen, G., Plunkett, K., & Nation, K. (2015). Do infant vocabulary skills predict school-age language and literacy outcomes? *Journal of Child Psychology and Psychiatry*, *56*(8), 848-56.

Dustmann, C,. & Frattini, T. (2011). *The impact of migration on the provision of UK public services*. Final report to the Migration Advisory Committee. UK Home Office.

Dyen, I., Kruskal, J. B., & Black, P. (1992). An Indo-European classification: A lexicostatistical experiment. *Transactions of the American Philosophical Society*, 82(5), pp. 1-132. Philadelphia: The American Philosophical Society.

Ellison, M. & Kirby, S. (2006). Measuring language divergence by intra-lexical comparison. In *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the ACL*, pp. 273-280. Association for Computational Linguistics.

Eriksson, M., Marschik, P. B., Tulviste, T., Almgren, M., Pérez Pereira, M., Wehberg, S., . . . Gallego, C. (2012). Differences between girls and boys in emerging language skills: Evidence from 10 language communities. *British Journal of Developmental Psychology, 30*(2), 326-343.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5), 1-173; discussion 174-185.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur–Bates Communicative Development Inventories: User's guide and technical manual* (2nd edition). Baltimore, MD: Brookes Publishing.

Ferrer-i-Cancho, R. (2015). The placement of the head that minimizes online memory. *Language Dynamics and Change, 5*(1), 114-137.

Fernald, A. (2006). When infants hear two languages: Interpreting research on early speech perception by bilingual children. In P. D. McCardle & E. Hoff (Eds.), *Childhood bilingualism: Research on infancy through school age* (pp. 19-29). Clevedon, UK: Multilingual Matters.

Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science, 16*(2), 234–248.

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language, 16*(3), 477-501.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language, 41*(1), 78-104.

Floccia, C., Keren-Portnoy, T., DePaolis, R., Duffy, H., Delle Luche, C., Durrant, S., . . . Vihman, M. (2016). British English infants segment words only with exaggerated infant-directed speech stimuli. *Cognition, 148*, 1-9.

Floccia, C., Nazzi, T., Delle Luche, C., Poltrock, S., & Goslin, J. (2014). English-learning one-to two-year-olds do not show a consonant bias in word learning. *Journal of Child Language*, *41*(05), 1085-1114.

Floor, P., & Akhtar, N. (2006). Can 18-month-old infants learn words by listening in on conversations? *Infancy*, *9*(3), 327-339.

Fortescue, M., & Lennert Olsen, L. (1992). The acquisition of West Greenlandic. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition* (vol. 3, pp. 111– 220). Hillsdale, NJ: Lawrence Erlbaum.

Fortnum, H. M., Summerfield, A. Q., Marshall, D. H., Davis, A. C., Bamford, J. M., ... & Hind, S. (2001). Prevalence of permanent childhood hearing impairment in the United Kingdom and implications for universal neonatal hearing screening: questionnaire based ascertainment study. Commentary: Universal newborn hearing screening: implications for coordinating and developing services for deaf and hearing impaired children. *British Medical Journal*, *323*(7312), 536-539.

Francis, B., & Archer, L. (2005). British—Chinese pupils' and parents' constructions of the value of education. *British Educational Research Journal, 31*(1), 89-108.

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2017). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*, 3, 677-694.

French, R. M., & Jacquet, M. (2004). Understanding bilingual memory: Models and data. *Trends in Cognitive Sciences*, *8*(2), 87-93.

Gathercole, V. C. M. (2002a). Command of the mass/count distinction in bilingual and monolingual children: An English morphosyntactic distinction. In D. K. Oller, & R. E.

Eilers (Eds.), *Language and literacy in bilingual children* (pp. 175-206). Clevedon: Multilingual Matters.

Gathercole, V. C. M. (2002b). Grammatical gender in bilingual and monolingual children: A Spanish morphosyntactic distinction. In D. K. Oller, & R. E. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 207-219). Clevedon: Multilingual Matters.

Gathercole, V. C. M. (2002c). Monolingual and bilingual acquisition: Learning different treatments of *that*-trace phenomena in English and Spanish. In D. K. Oller, & R. E. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 220-254). Clevedon: Multilingual Matters.

Gathercole, V. C. M. (2007). Miami and North Wales, so far and yet so near: Constructivist account of morphosyntactic development in bilingual children. *International Journal of Bilingual Education and Bilingualism, 10*(3), 224–247.

Gathercole, V. C. M., Kennedy, I., & Thomas, E. (2016). Socioeconomic level and bilinguals' performance on language and cognitive measures. *Bilingualism: Language and Cognition*, *19*(5), 1057-1078. doi:10.1017/S1366728915000504

Gathercole, V. C. M., & Thomas, E. M. (2009). Bilingual first-language development: Dominant language takeover, threatened minority language take-up. *Bilingualism: Language and Cognition, 12*(02), 213-237.

Gatt, D. (2017). Bilingual vocabulary production in young children receiving Maltese-dominant exposure: individual differences and the influence of demographic and language exposure factors. *International Journal of Bilingual Education and Bilingualism, 20(2),* 163-182.

Genesee, F. (1989). Early bilingual development: one language or two?. *Journal of Child Language, 16*(1), 161-179.

Genesee, F., Geva, E., Dressler, C., & Kamil, M. (2006). Synthesis: Cross-linguistic relationships. *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth*, 153-174.

Genesee, F., Nicoladis, E., & Paradis, J. (1995). Language differentiation in early bilingual development. *Journal of Child Language*, *22*(3), 611-631.

Gervain, J., Nespor, M., Mazuka, R., Horie, R. & Mehler, J. (2008). Bootstrapping word order in prelexical infants: A Japanese-Italian cross-linguistic study. *Cognitive Psychology*, *57*, 56–74.

Gervain, J., & Werker, J. F. (2013). Prosody cues word order in 7-month-old bilingual infants. *Nature Communication.* 4:1490 doi: 10.1038/ncomms2430.

Gildersleeve-Neumann, C. E., Kester, E. S., Davis, B. L., & Pena, E. D. (2008). English speech sound development in preschool-aged children from bilingual English-Spanish environments. *Language, Speech and Hearing Services in Schools, 39*(3), 314-328.

Gleason, J. B., & Ratner, N. B. (2017). *The development of language. Ninth edition*. Pearson.

Goldberg, W. A., Prause, J. A., Lucas-Thompson, R., & Himsel, A. (2008). Maternal employment and children’s achievement in context: A meta-analysis of four decades of research. *Psychological Bulletin, 134*, 77– 108.

Gollan, T. H., Starr, J., & Ferreira, V. S. (2015). More than use it or lose it: The number-of-speakers effect on heritage language proficiency. *Psychonomic Bulletin & Review*, *22*(1), 147-155.

Greenberg, J. H. (1960). A quantitative approach to the morphological typology of language. *International Journal of American Linguistics, 26*(3), 178–194.

Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (Ed.), *Universals of human hanguage*, (pp. 73-113). Cambridge, Mass: MIT Press.

Gross, M., Buac, M., & Kaushanskaya, M. (2014). Conceptual scoring of receptive and expressive vocabulary measures in simultaneous and sequential bilingual children. *American Journal of Speech-Language Pathology*, *23*(4), 574-586.

Guasti, M. T. (2002). *Language acquisition : the growth of grammar*. Cambridge, Mass.: MIT Press.

Håkansson, G, Salameh E, & Nettelbladt U. (2003). Measuring language development in bilingual children: Swedish-Arabic children with and without language impairment. *Linguistics*, *41*, 255–288.

Hamilton, A., Plunkett, K., & Schafer, G. (2000). Infant vocabulary development assessed with a British communicative development inventory. *Journal of Child Language, 27*(3), 689-705.

Hammer, C. S., Komaroff, E., Rodriguez, B. L., Lopez, L. M., Scarpino, S. E., & Goldstein, B. (2012). Predicting Spanish–English bilingual children's language abilities. *Journal of Speech, Language, and Hearing Research*, *55*(5), 1251-1264.

Hansen, K., & Hawkes, D. (2009). Early childcare and child development. *Journal of Social Policy*, *38*(02), 211-239.

Harding, R. M., & Sokal, R. R. (1988). Classification of the European language families by genetic distance. *PNAS*, *85(23)*, 9370-9372.

Hasson, N., Camilleri, B., Jones, C., Smith, J., & Dodd, B. (2013). Discriminating disorder from difference using dynamic assessment with bilingual children. *Child Language Teaching and Therapy*, *29*(1), 57-75.

Hasson, N., & Joffe, V. (2007). The case for dynamic assessment in speech and language therapy. *Child Language Teaching and Therapy*, *23*(1), 9-25.

Heilmann, J., Weismer, S. E., Evans, J., & Hollar, C. (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.

Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. (2009). Longitudinal analysis of early semantic networks preferential attachment or preferential acquisition? *Psychological Science*, *20*(6), 729-739.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, *74*(5), 1368-1378.

Hoff, E. (2013). *Language development. Fifth edition*. Wadsworth: Cengage Learning.

Hoff, E., Coard, W., & Señor, M. (2013). Bilingual parents' child-directed speech in two languages. Poster presented to the meetings of the Society for Research in Child Development; Seattle, WA. April 18–20.

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*(01), 1-27.

Hoff, E., Quinn, J. M., & Giguere, D. (2017). What explains the correlation between growth in vocabulary and grammar? New evidence from latent change score analyses of simultaneous bilingual development. *Developmental Science*. 10.1111/desc.12536.

Hoff, E., Welsh, S., Place, S., Ribot, K., Grüter, T., & Paradis, J. (2014). Properties of dual language input that shape bilingual development and properties of environments that shape dual language input. In T. Grüter and J. Paradis (Eds.) *Input and experience in bilingual development* (pp. 119-140). Amsterdam, The Netherlands: John Benjamins.

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

Höhle, B., Bijeljac-Babic, R., Herold, B., Weissenborn, J., & Nazzi, T. (2009). Language specific prosodic preferences during the first half year of life: Evidence from German and French infants. *Infant Behavior and Development*, *32*(3), 262-274.

Höhle, B., Weissenborn, J., Schmitz, M., & Ischebeck, A. (2001). Discovering word order regularities. In J. Weissenborn & B. Höhle (Eds.), *Approaches to bootstrapping: Phonological, lexical, syntactic and neurophysiological aspects of early language acquisition*, Vol. 1 (pp. 249-265).

Højen, A., & Nazzi, T. (2016). Vowel bias in Danish word-learning: Processing biases are language-specific. *Developmental Science*, *19*(1), 41-49.

Hoshino, N., & Kroll, J. F. (2008). Cognate effects in picture naming: Does cross-language activation survive a change of script? *Cognition*, *106*(1), 501-511.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*(2), 236.

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, *61*(4), 343-365.

Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin*, *104*(1):53–69.

i Trueta, M. S., Barrachina, L. A., & Pascual, E. S. (2012). *Results of the language model adopted by schools in Catalonia: The empirical evidence*. Universitat Oberta de Catalunya.

Iverson, J. M. (2010). Developing language in a developing body: The relationship between motor development and language development. *Journal of Child Language*, *37*(2), 229-261.

Jones, S. S., & Smith, L. B. (2005). Object name learning and object perception: A deficit in late talkers. *Journal of Child Language*, *32*(1), 223-240.

Jørgensen, R. N., Dale, P. S., Bleses, D., & Fenson, L. (2009). CLEX: A cross-linguistic lexical norms database. *Journal of Child Language*, *37(2),* 419-428.

Junker, D. A., & Stockman, I. J. (2002). Expressive vocabulary of German–English bilingual toddlers. *American Journal of Speech–Language Pathology, 11,* 381–395.

Kelley, A., & Kohnert, K. (2012). Is there a cognate advantage for typically developing Spanish-speaking English-language learners? *Language, Speech, and Hearing Services in Schools*, *43*(2), 191-204.

Kim, M., McGregor, K. K., & Thompson, C. K. (2000). Early lexical development in English-and Korean-speaking children: Language-general and language-specific patterns. *Journal of Child Language*, *27*(02), 225-254.

Kim, Y. S., Leventhal, B. L., Koh, Y. J., Fombonne, E., Laska, E., Lim, E. C., ... & Song, D. H. (2011). Prevalence of autism spectrum disorders in a total population sample. *American Journal of Psychiatry, 168*(9), 904-12.

Kohnert, K. (2010). Bilingual children with primary language impairment: Issues, evidence and implications for clinical actions. *Journal of Communication Disorders*, *43*(6), 456-473.

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor: University of Michigan Press.

Lagrou, E., Hartsuiker, R. J., & Duyck, W. (2011). Knowledge of a second language influences auditory word recognition in the native language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(4), 952-965.

Lauro, J., & Schwartz, A. I. (2017). Bilingual non-selective lexical access in sentence contexts: A meta-analytic review. *Journal of Memory and Language*, *92*, 217-233.

Law, N. C. W. & So, L. K. H. (2006). The relationship of phonological development and language dominance in bilingual Cantonese-Putonghua children. *International Journal of Bilingualism, 10*(4), 405–428.

Leonard, L. B. (1987). Is specific language impairment a useful construct? In S. Rosenberg (Ed.), *Advances in applied psycholinguistics: Disorders of first language development* (pp. 1-39). Cambridge, UK: Cambridge University Press.

Leonard, L. B. (2000). Specific language impairment across languages. In D. V. M. Bishop & L. B. Leonard (Eds.), *Speech and language impairments in children: Causes, characteristics, intervention and outcome* (pp. 115-129). Psychology Press.

Lieven, E. V. M. (1994). Crosslinguistic and crosscultural aspects of language addressed to children. In C. Gallaway & B. J. Richards (Eds.), *Input and interaction in language acquisition* (pp. 56-73). New York, NY, US: Cambridge University Press.

Lindgren, E., & Muñoz, C. (2012). The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension. *International Journal of Multilingualism*, *10*(1), 105-129.

Ling, L. E., Grabe, E., & Nolan, F. (2000). Quantitative characterizations of speech rhythm: Syllable-timing in Singapore English. *Language and Speech*, *43*(4), 377-401.

Lupyan, G., & Christiansen, M. H. (2002). Case, word order, and language learnability: Insights from connectionist modeling. In *Proceedings of the 24th Annual Conference of the Cognitive Science Society* (pp. 596-601). Mahwah, NJ: Lawrence Erlbaum.

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

Mani, N., & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language, 57*, 252-272.

Mani, N,. & Plunkett, K. (2010). In the infant's mind's ear: evidence for implicit naming in 18-month-olds. *Psychological Science, 21*(7), 908-913.

Marchman, V. A., Fernald, A., & Hurtado, N. (2010). How vocabulary size in two languages relates to efficiency in spoken word recognition by young Spanish–English bilinguals. *Journal of Child Language, 37*(04), 817-840.

Marchman, V. A., Thal, D. J., Dale, P. S., & Reznick, J. S. (2006). *MacArthur-Bates Communicative Development Inventories: user's guide and technical manual*. L. Fenson (Ed.). Brookes Publ.

McMahon, A., & McMahon, R. (2005). *Language classification by numbers*. Oxford University Press on Demand.

Mercier, J., Pivneva, I., & Titone, D. (2014). Individual differences in inhibitory control relate to bilingual spoken word processing. *Bilingualism: Language and Cognition, 17*(1), 89-117.

Miękisz, A., Haman, E., Łuniewska, M., Kuś, K., O'Toole, C., & Katsos, N. (2016). The impact of a first-generation immigrant environment on the heritage language: productive vocabularies of Polish toddlers living in the UK and Ireland. *International Journal of Bilingual Education and Bilingualism*, 1-18.

Molnar, M., Gervain, J., & Carreiras, M. (2014). Within-rhythm class native language discrimination abilities of Basque-Spanish monolingual and bilingual infants at 3.5 months of age. *Infancy, 19*(3), 326-337.

Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: evidence from a longitudinal study. *Developmental Psychology, 40*(5), 665.

NALDIC (2012), School Census January 2012, Department for Education, http://www.naldic.org.uk/eal-advocacy/eal-news-summary/210612

Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: Differences between consonants and vowels. *Cognition*, *98*, 13-30.

Nazzi, T., Floccia, C., Moquet, B., & Butler, J. (2009). Bias for consonantal information over vocalic information in 30-month-olds: Cross-linguistic evidence from French and English. *Journal of Experimental Child Psychology*, *102*(4), 522-537.

Nazzi, T., Mersad, K., Sundara, M., Iakimova, G., & Polka, L. (2014). Early word segmentation in infants acquiring Parisian French: Task-dependent and dialect-specific aspects. *Journal of Child Language*, *41*(03), 600-633.

Nerbonne, J., Heeringa, W., van den Hout, E., van der Kooi, P., Otten, S., & van de Vis, W. (1996). Phonetic distance between Dutch dialects. In G. Durieux, W. Daelemans & S. Gillis (Eds.), CLIN VI: *Proceedings from the Sixth CLIN Meeting* (pp. 185–202). Center for Dutch Language and Speech, University of Antwerpen (UIA), Antwerpen.

Nespor, M., Shukla, M., van de Vijver, R., Avesani, C., Schraudolf, H., & Donati, C. (2008). Different phrasal prominence realization in VO and OV languages. *Lingue e Linguaggio*, *7*(2), 1-28.

Nicholls, R. J., Eadie, P. A., & Reilly, S. (2011). Monolingual versus multilingual acquisition of English morphology: what can we expect at age 3? *International Journal of Language and Communication Disorders*, *46*(4), 449-463.

Nicoladis, E. & Secco, G. (2000). The role of a child's productive vocabulary in the language choice of a bilingual family. *First Language, 58*, 3-28.

Nishibayashi, L. L., & Nazzi, T. (2016). Vowels, then consonants: Early bias switch in recognizing segmented word forms. *Cognition*, *155*, 188-203.

Oller, D. K., & Eilers, R. E. (2002). *Language and literacy in bilingual children*. Clevedon: Multilingual Matters.

O'Toole, C., Gatt, D., Hickey, T. M., Miękisz, A., Haman, E., Armon-Lotem, S., ... & Kern, S. (2017). Parent report of early lexical production in bilingual children: A cross-linguistic CDI comparison. *International Journal of Bilingual Education and Bilingualism*, 1-22.

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, *66*(2), 232-258.

Pancsofar N., & Vernon-Feagans L. (2006). Mother and father language input to young children: Contributions to later language development. *Journal of Applied Developmental Psychology, 27*, 571–587.

Paradis, M. (2009). *Declarative and procedural determinants of second languages* (vol. 40). Amsterdam: John Benjamins.

Paradis, J. (2011). Individual differences in child English second language acquisition: Comparing child-internal and child-external factors. *Linguistic Approaches to Bilingualism*, Volume 1(3).

Paradis, J., Crago, M., Genesee, F., & Rice, M. (2003). French-English bilingual children with SLI: How do they compare with their monolingual peers? *Journal of Speech, Language, and Hearing Research*, *46*(1), 113-127.

Paradis, J. & Genesee, F. (1996). Syntactic acquisition in bilingual children: Autonomous or interdependent? *Studies in Second Language Acquisition, 18*, 1-25.

Payne, A. C., Whitehurst, G. J., Angell, A. L. (1994). The role of the home literacy environment in the development of language ability in preschool children from low-income families. *Early Childhood Research Quarterly, 9,* 427-440.

Pearson, B. Z., Fernández, S. C., Lewedeg, V., & Oller, D. K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics, 18,* 41-58.

Pearson, B. Z., Fernández, S. C., & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language Learning, 43,* 93-120.

Perani, D., Abutalebi, J., Paulesu, E., Brambati, S., Scifo, P., Cappa, S. F., & Fazio, F. (2003). The role of age of acquisition and language usage in early, high-proficient bilinguals: an fMRI study during verbal fluency. *Human Brain Mapping*, *19*(3), 170-182.

Pérez, A. M., Peña, E. D., & Bedore, L. M. (2010). Cognates facilitate word recognition in young Spanish-English bilinguals' test performance. *Early Childhood Services (San Diego)*, *4*(1), 55-67.

Pinker, S. (1995). *The language instinct: The new science of language and mind* (Vol. 7529). Penguin UK.

Place, S., & Hoff, E. (2011). Properties of dual language exposure that influence two-year-olds' bilingual proficiency. *Child Development, 82,* 1834–1849.

Place, S., & Hoff, E. (2016). Effects and noneffects of input in bilingual environments on dual language skills in 2 ½-year-olds. *Bilingualism: Language and Cognition*, *19*(5), 1023-1041.

Poarch, G. J., & van Hell, J. G. (2012). Cross-language activation in children's speech production: Evidence from second language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology*, *111*(3), 419-438.

Polka, L., Orena, A. J., Sundara, M., & Worrall, J. (2017). Segmenting words from fluent speech during infancy–challenges and opportunities in a bilingual context. *Developmental Science, 20*(1).

Portocarrero, J. S., Burright, R. G., & Donovick, P. J. (2007). Vocabulary and verbal fluency of bilingual and monolingual college students. *Archives of Clinical Neuropsychology*, *22*(3), 415-422.

Poulin-Dubois, D., Bialystok, E., Blaye, A., Polonia, A., & Yott, J. (2013). Lexical access and vocabulary development in very young bilinguals. *International Journal of Bilingualism*, *17*(1), 57-70.

Prevoo, M. J., Malda, M., Mesman, J., & van IJzendoorn, M. H. (2016). Within-and cross-language relations between oral language proficiency and school outcomes in bilingual children with an immigrant background: A meta-analytical study. *Review of Educational Research*, *86*, 237-276.

Rack, J. P., Snowling, M. J., & Olson, R. K. (1992). The nonword reading deficit in developmental dyslexia - A review. *Reading Research Quarterly*, *27*(1), 28-53.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, *73(3)*, 265-292.

Rescorla, L. (1989). The language development survey: A screening tool for delayed language in toddlers. *Journal of Speech and Hearing Disorders*, *54*(4), 587-599.

Rescorla, L. (2002). Language and reading outcomes to age 9 in late-talking toddlers. *Journal of Speech, Language, and Hearing Research*, *45*(2), 360-371.

Rescorla, L. (2005). Age 13 language and reading outcomes in late-talking toddlers. *Journal of Speech, Language, and Hearing Research*, *48*(2), 459-472.

Rhys, M., & Thomas, E. M. (2013). Bilingual Welsh–English children's acquisition of vocabulary and reading: implications for bilingual education. *International Journal of Bilingual Education and Bilingualism*, *16*(6), 633-656.

Rost, G. C., & McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Developmental Science*, *12*(2), 339-349.

Royal College of Speech and Language Therapists (RCSLT) (2007). Good Practice for Speech and Language Therapists Working with Clients from Linguistic Minority Communities. Guidelines Prepared by the RCSLT Specific Interest Group in

Bilingualism (London: RCSLT) (available at: https://www.rcslt.org/members/publications/publications2/linguistic_minorities) (accessed on 2 May 2017).

Ruhlen, M. (1987). *A guide to the world's languages*. Stanford, CA: Stanford Univ. Press, Vol. 1.

Ruhm C. J. (2004). Parental employment and child cognitive development. *Journal of Human Resources, 39*(1), 155–192.

Salameh, E. K., Nettelbladt, U., Håkansson, G., & Gullberg, B. (2002). Language impairment in Swedish bilingual children: a comparison between bilingual and monolingual children in Malmö. *Acta Paediatrica*, *91*(2), 229-234.

Sánchez-Casas, R., & García-Albea, J. E. (2005). The representation of cognate and noncognate words in bilingual memory. *Handbook of bilingualism: Psycholinguistic approaches*, 226-250.

Schwartz, R. G. (Ed.). (2010). *Handbook of child language disorders*. Psychology press.

Serva, M., & Petroni, F. (2008). Indo-European languages tree by Levenshtein distance. *EPL (Europhysics Letters), 81*(6), 68005(5pp).

Sheng, L., Lam, B. P. W., Cruz, D., & Fulton, A. (2016). A robust demonstration of the cognate facilitation effect in first-language and second-language naming. *Journal of Experimental Child Psychology*, *141*, 229-238.

Shneidman, L. A., Arroyo, M. E., Levine, S. C., & Goldin-Meadow, S. (2013). What counts as effective input for word learning? *Journal of Child Language*, *40*(03), 672-686.

Shook, A., Goldrick, M., Engstler, C., & Marian, V. (2015). Bilinguals show weaker lexical access during spoken sentence comprehension. *Journal of Psycholinguistic Research*, *44*(6), 789-802.

Shook, A., & Marian, V. (2013). The bilingual language interaction network for comprehension of speech. *Bilingualism: Language and Cognition, 16*(2), 304-324.

Singh, L. (2008). Influences of high and low variability on infant word recognition. *Cognition, 106*(2), 833-870.

Slobin, D. I. (1973). Cognitive prerequisites for the development of grammar. *Studies of Child Language Development, 1*, 75-208.

Smithson, L., Paradis, J., & Nicoladis, E. (2014). Bilingualism and receptive vocabulary achievement: Could sociocultural context make a difference? *Bilingualism: Language and Cognition, 17*(4), 810–821.

Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science*, *10*(3), 281-284.

Squires, J., & Bricker, D. (2009). *Ages & Stages Questionnaires,* Third Edition (ASQ-3). Baltimore, MD: Brookes.

Statistics Canada, 2011, http://www12.statcan.gc.ca/census-recensement/2011/as-sa/98-314-x/98-314-x2011001-eng.cfm.

Stevens, J. S., Gleitman, L. R., Trueswell, J. C., & Yang, C. (2017). The pursuit of word meanings. *Cognitive Science*, *41*(S4), 638-676.

Strand, S., Malmberg, L., & Hall, J. (2015). *English as an additional language and educational achievement in England: An analysis of the National Pupil Database*. London: Educational Endowment Fund.

Strijkers, K., Costa, A., & Thierry, G. (2009). Tracking lexical access in speech production: electrophysiological correlates of word frequency and cognate effects. *Cerebral Cortex*, *20* (4), 912-928.

Super, C. M., & Karkness, S. (2013). The development of affect in infancy and early childhood. In M. Woodhead, D. Faulkner and K. Littleton (Eds.), *Cultural worlds of early childhood*, pp 34-47. London and New York: Routledge.

Tardif, T., Gelman, S. A., & Xu, F. (1999). Putting the “noun bias” in context: A comparison of English and Mandarin. *Child Development, 70*(3), 620-635.

Tardif, T., Shatz, M., & Naigles, L. (1997). Caregiver speech and children's use of nouns versus verbs: A comparison of English, Italian, and Mandarin. *Journal of Child Language*, *24*(3), 535-565.

Tare, M., & Gelman, S. A. (2011). Bilingual parents' modeling of pragmatic language use in multiparty interactions. *Applied Psycholinguistics*, *32*(4), 761-780.

Thal, D., Bates, E., Goodman, J., & Jahn-Samilo, J. (1997). Continuity of language abilities in late- and early-talking toddlers. In D. Thal & J. Reilly (Eds.), *Special issue: Origins of communicative disorders. Developmental Neuropsychology*, *13* (3), 239–273.

Thal, D., DesJardin, J. L., & Eisenberg, L. S. (2007). Validity of the MacArthur–Bates Communicative Development Inventories for measuring language abilities in children with cochlear implants. *American Journal of Speech-Language Pathology*, *16*(1), 54-64.

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, *7*(1), 53-71.

Thordardottir, E. (2005). Early lexical and syntactic development in Quebec French and English: Implications for cross-linguistic and bilingual assessment. *International Journal of Language and Communication Disorders*, *40*, 243 – 278.

Thordardottir, E. (2011). The relationship between bilingual exposure and vocabulary development. *International Journal of Bilingualism*, *15*(4), 426-445.

Thordardottir, E., Rothenberg, A., Rivard, M. -E., & Naves, R. (2006). Bilingual assessment: Can overall proficiency be estimated from separate measurement of two languages? *Journal of Multilingual Communication Disorders*, *4*(1), 1-21.

Tomalski, P., Moore, D. G., Ribeiro, H., Axelsson, E. L., Murphy, E., Karmiloff-Smith, A., ... & Kushnerenko, E. (2013). Socioeconomic status and functional brain development–associations in early infancy. *Developmental Science*, *16*(5), 676-687.

Tomasello, M. (2009). *Constructing a language*. Harvard University Press.

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child Development*, *57*(6), 1454-1463.

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*, *40*(6), 1245-1260.

Tomblin, J. B., Records, N. L., & Zhang, X. (1996). A system for the diagnosis of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*, *39*(6), 1284-1294.

Umbel, V. M., & Oller, D. (1994). Developmental changes in receptive vocabulary in Hispanic bilingual school children. *Language Learning*, *44*(2), 221-242.

Umbel, V. M., Pearson, B. Z., Fernandez, M. C., & Oller, D. K. (1992). Measuring bilingual children's receptive vocabularies. *Child Development*, *63*(4), 1012-1020.

United States Census 2020, https://www.census.gov/newsroom/press-releases/2015/cb15-185.html, retrieved 13/07/2017

Von Holzen, K., & Mani, N. (2012). Language nonselective lexical access in bilingual toddlers. *Journal of Experimental Child Psychology*, *113*(4), 569-586.

van Hell, J. G., & De Groot, A. M. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition*, *1*(03), 193-211.

Volterra, V., & Taeschner, T. (1978). The acquisition and development of language by bilingual children. *Journal of Child Language*, *5*(2), 311-326.

Weisleder, A., & Fernald, A. (2013). Talking to children matters early language experience strengthens processing and builds vocabulary. *Psychological Science*, *24*(11), 2143-2152.

Werker, J. F., Yeung, H. H., & Yoshida, K. A. (2012). How do infants become experts at native-speech perception?. *Current Directions in Psychological Science*, *21*(4), 221-226.

White, L., Delle Luche, C., Kwok, R. K. W., Goslin, J., Plunkett, J., & Floccia, C. (in prep). Prosodic distance in production and perception.

Xanthos, A., Laaha, S., Gillis, S., Stephany, U., Aksu-Koç, A., Christofidou, A., ... & Korecky-Kröll, K. (2011). On the role of morphological richness in the early development of noun and verb inflection. *First Language*, *31*(4), 461-479.

Yang, S., Blume, M., & Lust, B. (2006). *Virtual Linguistic Lab (VLL) Child Multilingualism Questionnaire.* Cornell Language Acquisition Laboratory, Cornell University.

Zhao, X., & Li, P. (2010). Bilingual lexical interactions in an unsupervised neural network model. *International Journal of Bilingual Education and Bilingualism*, *13*(5), 505–524.

Zimmerman , I. L., Steiner, V. G., & Pond, R. E. (2002). *Preschool language scale: Examiner's manual (4th ed.)*. San Antonio, Texas: Harcourt Assessment.

Acknowledgements

	Description	Codename	Measure
Step 1 predictors and variables	Gender	Gender	male/female
	Income (SES)	Income	4-point scale
	Maternal education (SES)	MumStudy	7-point scale
	Paternal education (SES)	DadStudy	7-point scale
	Relative amount of exposure to English in child directed speech	LEQ	proportion
	Proportion of English/AL in overheard speech	Overheard speech	5-point scale
	Language community	Language community	dummy variable from 1 to 13 (Spanish, etc)
	Which language is measured (within-subject)	TestLang	English vs AL
Step 2 predictors	Number of parental AL speakers (source)	FamLang	1 or 2
	Total number of English speakers (source)	SourcesEng	from 0 upwards
	Total number of AL speakers (source)	SourcesAL	from 0 upwards
	Number of older siblings (source)	SIB	from 0 upwards
	Time in English daycare (source)	EngDaycare	hours
	Time in AL daycare (source)	ALDaycare	hours
	Degree of language use consistency in parents' speech (properties)	Consistency	3-point scale
	Proportion of native/non-native in English input (quality)	PropEngN	proportion
	Proportion of native/non-native in AL input (quality)	PropALN	proportion
	Societal status of the AL (status)	Status	Welsh vs all other ALs
	Linguistic distance measured through phonological overlap	Overlap	from 0 upwards
	Linguistic distance measured through morphological complexity	Morph	3-point scale
	Linguistic distance measured through word order typology	WordOrder	3-point scale
Dependent variables	English comprehension on the Oxford Short Form CDI	CDI100comp	0 to 100
	English production on the Oxford Short Form CDI	CDI100prod	0 to 100
	English comprehension on the 30-word CDI	CDI30comp	0 to 30
	English production on the 30-word CDI	CDI30prod	0 to 30
	AL comprehension on the AL full CDI	PropALComp	proportion
	AL production on the AL full CDI	PropALProd	proportion
	AL comprehension on the 30-word CDI	ALCDI30comp	0 to 30
	AL production on the 30-word CDI	ALCDI30prod	0 to 30

Table 1. List of all variables and predictors used in the analyses. More details on calculations are provided in appendices 2 and 3 (for LEQ). AL stands for Additional Language.

	Number of children	Number of children per gender		Age in months		Income bracket		Maternal education (MumEd)		Paternal education (DadEd)	
	total	girls	boys	mean	std	mean	std	mean	std	mean	std
Bengali	13	7	6	24.02	0.38	3.45	0.69	5.15	1.95	5.77	1.64
Cantonese Chinese	8	6	2	23.90	0.46	3.88	0.35	6.50	0.53	6.13	0.35
Dutch	14	6	8	23.96	0.42	3.71	0.47	6.43	0.65	6.00	1.24
French	55	28	27	23.96	0.34	3.74	0.48	6.69	0.60	6.24	1.09
German	75	31	44	23.87	0.38	3.60	0.62	6.36	1.01	6.21	1.18
Greek	12	6	6	23.85	0.37	3.67	0.89	6.75	0.62	6.33	1.15
Hindi/Urdu	9	2	7	23.99	0.30	3.67	0.50	6.44	1.01	6.78	0.44
Italian	33	13	20	23.84	0.36	3.70	0.47	6.55	0.56	5.97	1.24
Mandarin Chinese	8	6	2	24.09	0.48	3.25	1.16	6.63	0.52	6.00	1.41
Polish	29	17	12	23.92	0.43	3.26	0.71	6.38	1.01	5.38	1.50
Portuguese	10	4	6	23.93	0.34	3.33	0.87	6.30	0.67	6.10	0.99
Spanish	43	18	25	23.79	0.30	3.59	0.74	6.42	0.73	5.81	1.22
Welsh	63	20	43	23.95	0.44	3.61	0.71	6.05	1.07	5.70	1.27
Other	58	29	29	23.80	0.42	3.54	0.71	6.33	0.87	5.81	1.53
Grand total/mean	**430**	**193**	**237**	**23.89**	**0.39**	**3.59**	**0.66**	**6.36**	**0.94**	**5.96**	**1.28**

Table 2. Demographic characteristics of participants per Additional Language community: gender, mean age in months (when completion of Oxford Short Form CDI), mean household income (4-point scale scale from 1 to 4), maternal and paternal education (7-point scale scale from 1 to 7). Details for non-target Additional Language children are provided under the row label "Other". See appendix 2 for details of the calculations of Income, MumEd and DadEd.

Additional Language	**Phonological Overlap**	**Word order typology**	**Morphological complexity**
Dutch	0.2214	2	2
Welsh	0.2163	1	2
German	0.1975	2	2
Italian	0.1076	1	2
French	0.1034	1	2
Bengali	0.0941	3	3
Hindi	0.0899	3	3
Spanish	0.0874	1	2
Polish	0.0828	1	2
Greek	0.0807	1	2
Portuguese	0.0801	1	2
Cantonese	0.0422	1	1
Mandarin	0.0197	1	1

Table 3. Average phonological overlap between 406 British English words of the Oxford CDI and their translation equivalents for the 13 Additional Languages. The higher the number, the closer the languages. Word order typology distance to British English (1 = VO language like British English; 2 = VO/OV language; 3 = OV language). Morphological complexity distance to British English (1 = analytic/isolating language like British English; 2 = fusional; 3 = agglutinative).

	Proportion of English vs. AL in child directed speech		Proportion of English in mother input		Proportion of English in father input		Proportion of English vs. AL between parents	
	LEQ		MumPropEng		DadPropEng		Overheard speech	
	mean	std	mean	std	mean	std	mean	std
Bengali	49.80	26.44	3.00	1.22	3.15	1.34	3.46	1.56
Cantonese Chinese	61.77	21.74	3.13	1.25	4.00	1.07	4.25	1.16
Dutch	47.01	25.44	2.14	1.35	3.57	1.74	4.36	1.28
French	56.02	22.37	2.25	1.40	3.67	1.55	4.36	1.08
German	46.27	21.64	1.96	1.21	4.07	1.48	4.33	1.33
Greek	46.53	25.91	2.08	1.62	3.17	1.85	3.58	1.83
Hindi/Urdu	52.14	29.77	2.89	1.36	2.78	1.20	2.56	1.59
Italian	48.04	25.08	2.06	1.37	2.88	1.71	3.24	1.77
Mandarin Chinese	48.06	15.05	1.88	0.64	3.63	1.69	3.50	1.85
Polish	45.25	22.50	1.86	1.06	3.55	1.72	4.00	1.67
Portuguese	57.14	24.81	2.40	1.35	2.50	1.84	2.70	1.89
Spanish	45.88	22.88	2.00	1.35	2.98	1.77	3.40	1.80
Welsh	46.79	28.02	2.35	1.48	3.22	1.67	4.03	1.47
Other	51.26	24.65	2.10	1.27	3.55	1.67	3.71	1.73
Mean	**49.13**	**24.14**	**2.17**	**1.33**	**3.45**	**1.65**	**3.87**	**1.59**

Table 4. Mean amount of exposure to English versus Additional Language (AL) in child directed speech (LEQ) for each Additional Language group; proportion of English vs the Additional Language in maternal (or paternal) speech on a 5-point scale from 1 to 5 (MumPropEng and DadPropEng); proportion of English vs the Additional Language spoken between parents in overheard speech (Overheard speech; 5-point scale from 1 to 5; 1 = only Additional Language to 5 = only English). Details for non-target Additional Language children are provided under the row label "Other". See appendix 3 for details of LEQ calculation, and Appendix 2 for details of calculation of the other variables.

	Number of families with one or both parents AL speakers		Total number of English speakers		Total number of AL speakers		Number of hours per week in English speaking		Number of hours per week in AL speaking		Number or older siblings	
	FamLang		SourcesEng		SourcesAL		EngDaycare		ALDaycare		Siblings	
	AL only	Mixed	mean	std	mean	std	mean	std	mean	std	mean	std
Bengali	11	2	3.31	1.18	3.31	1.55	8.31	10.04	7.50	8.34	0.92	1.32
Cantonese Chinese	3	5	2.75	0.46	2.38	0.74	25.75	14.41	15.81	16.33	0.00	0.00
Dutch	3	11	2.64	1.22	2.29	0.83	18.68	14.38	9.39	19.56	0.43	0.65
French	4	51	2.80	0.89	2.31	1.10	24.93	14.74	4.30	11.09	0.44	0.71
German	10	65	2.84	1.10	2.27	0.98	16.56	13.32	5.33	12.46	0.59	0.77
Greek	3	9	2.42	1.38	2.50	1.00	23.46	17.28	9.54	19.00	0.50	0.90
Hindi/Urdu	7	2	3.33	1.32	3.00	1.22	9.89	11.34	1.67	3.94	0.89	1.17
Italian	14	19	2.52	0.94	2.45	1.06	21.95	15.39	2.76	5.56	0.36	0.55
Mandarin Chinese	3	5	2.75	0.89	2.00	0.53	24.06	14.86	4.00	9.75	0.13	0.35
Polish	7	22	2.48	0.78	2.17	1.00	18.33	13.42	3.74	8.62	0.28	0.53
Portuguese	5	5	2.20	0.79	2.30	0.82	29.45	18.21	6.95	15.52	0.20	0.42
Spanish	14	29	2.23	0.72	2.23	0.90	21.47	16.98	3.92	8.39	0.23	0.43
Welsh	26	37	2.51	1.09	2.63	1.11	14.44	14.61	8.70	11.36	0.46	0.78
Other	20	38	2.55	1.13	2.09	0.90	18.03	15.38	1.68	5.28	0.36	0.58
Grand Total/mean	**130**	**300**	**2.63**	**1.03**	**2.36**	**1.03**	**19.03**	**15.19**	**5.20**	**10.99**	**0.43**	**0.71**

Table 5. Characteristics of participants per Additional Language community in terms of mode of exposure (sources). Details for non-target Additional Language children are provided under the row label "Other". Number of families where both parents are Additional Language speakers (Additional Language only) or only one (Mixed) (variable FamlLang); Total number of regular English (or Additional Language) speakers around the child (SourcesEng and SourcesAL; observed range 0-6); weekly number of hours in an English (or Additional Language) speaking daycare (EngDaycare and ALDaycare); number of older siblings in the household (Siblings). See appendix 2 for details of the calculations of these variables.

	Degree of language use consistency in parents' input		Proportion of native English in parents' input		Proportion of native AL in parents' input	
	Consistency		PropEngN		PropALN	
	mean	std	mean	std	mean	std
Bengali	2.08	0.64	0.031	0.19	1.000	0.00
Cantonese Chinese	1.94	0.68	0.450	0.41	0.969	0.09
Dutch	1.43	0.43	0.708	0.38	0.918	0.15
French	1.47	0.39	0.695	0.55	0.932	0.15
German	1.37	0.40	0.737	0.35	0.972	0.08
Greek	1.29	0.50	0.836	0.32	0.958	0.10
Hindi/Urdu	2.06	0.77	0.114	0.25	0.936	0.15
Italian	1.44	0.46	0.572	0.44	0.958	0.14
Mandarin Chinese	1.63	0.35	0.524	0.37	0.978	0.06
Polish	1.43	0.44	0.735	0.38	0.932	0.15
Portuguese	1.55	0.64	0.393	0.57	0.982	0.06
Spanish	1.42	0.56	0.697	0.42	0.905	0.24
Welsh	1.50	0.52	0.575	0.43	0.933	0.18
Other	1.47	0.49	0.761	0.98	0.967	0.10
Mean	**1.48**	**0.50**	**0.649**	**0.55**	**0.948**	**0.14**

Table 6. Characteristics of participants per Additional Language community in terms of mode of exposure (properties). Details for non-target Additional Language children are provided under the row label "Other". Degree of language use consistency in parents' input (Consistency; 3-point scale from 1 to 3); proportion of native vs non-native English input in parental speech (PropEngN); proportion of native vs non-native Additional Language input in parental speech (PropALN). See appendix 2 for details on the calculation of these variables.

	Oxford Short Form CDI				30 words common to all CDIs			
	Proportion of words understood		Proportion of words produced		Number of words understood		Number of words produced	
	CDI100Comp		CDI100Prod		CDI30comp		CDI30prod	
	Mean	std	Mean	std	Mean	std	Mean	std
Bengali	65.1	28.7	42.5	28.2	22.4	8.1	16.6	8.2
Cantonese Chinese	71.5	19.1	57.0	23.4	26.4	4.2	22.4	6.7
Dutch	76.9	16.1	44.7	18.5	27.2	3.9	18.7	7.1
French	76.9	19.5	45.0	24.3	26.9	4.2	19.0	8.1
German	68.3	23.8	38.4	24.2	24.7	5.7	16.2	8.8
Greek	71.4	15.7	46.5	23.1	26.7	2.9	19.4	8.3
Hindi/Urdu	61.6	36.4	32.6	33.4	20.9	10.6	12.6	11.2
Italian	71.7	23.6	34.8	26.5	25.8	5.0	15.3	9.8
Mandarin Chinese	65.5	23.8	47.1	28.3	23.9	5.7	19.1	8.2
Polish	60.6	28.1	37.1	27.2	22.5	8.6	15.3	10.1
Portuguese	65.0	24.1	46.1	24.6	24.4	6.4	20.3	7.2
Spanish	62.3	26.3	36.0	22.4	23.6	7.3	16.1	8.7
Welsh	65.2	29.0	43.9	30.6	22.7	8.2	16.3	9.4
Other	66.5	24.6	42.9	26.2	24.2	7.3	17.7	9.5
Mean	**67.9**	**25.0**	**41.2**	**26.0**	**24.4**	**6.7**	**17.0**	**9.0**

Table 7. English vocabulary measures for each Additional Language group. Details for non-target Additional Language children are provided under the row label "Other". CDI100comp and CDI100prod: mean number of words respectively understood and produced on the Oxford Short Version CDI; CDI30comp and CDI30prod: mean number of words understood and produced out of the 30 words common to all CDIs (30-word CDI).

	Additional Language CDIs				30 words common to all CDIs				AL CDI length
	Proportion of words understood		Proportion of words produced		Number of words understood		Number of words produced		
	PropALComp		PropALProd		ALCDI30Comp		ALCDI30Prod		
	Mean	std	Mean	std	Mean	std	Mean	std	
Bengali	50.6	25.9	32.3	25.0	6.2	2.2	4.2	2.5	62
Cantonese Chinese	40.7	25.7	20.7	13.6	15.5	9.0	9.1	7.1	389
Dutch	60.9	21.3	31.9	16.6	24.1	7.1	16.4	6.9	442
French	59.8	21.4	25.2	21.2	24.4	5.3	12.4	9.0	414
German	57.6	22.8	22.8	18.3	25.5	4.7	14.1	8.2	600
Greek	44.7	22.1	15.1	12.6	21.8	6.2	10.1	6.6	654
Hindi/Urdu	17.5	13.3	7.0	6.7	6.2	5.9	1.9	1.8	444
Italian	64.6	25.4	24.7	23.7	24.5	7.2	11.0	9.9	413
Mandarin Chinese	52.6	21.4	30.9	14.9	20.4	6.3	12.8	4.9	411
Polish	63.0	25.2	22.5	20.8	22.5	7.4	8.5	7.6	381
Portuguese	59.2	27.7	37.1	23.7	11.3	4.7	6.2	4.4	91
Spanish	49.3	21.9	20.1	17.1	23.1	7.1	12.2	9.4	591
Welsh	50.4	24.3	26.5	22.4	19.5	7.6	9.9	8.1	430
Mean	**54.9**	**24.3**	**24.2**	**20.3**	**21.7**	**8.0**	**11.2**	**8.5**	**459.9**

Table 8. Additional Language vocabulary measures for each Additional Language group. PropALComp and PropALProd: mean proportion of words respectively understood and produced in the Additional Language as measured against the corresponding CDI; ALCDI30comp and ALCDI30prod: mean number of words understood and produced out of the 30-word CDIs; AL CDI length: total number of words in each Additional Language CDI.

	Sum of squares	df	*F*	*p*	η^2
With TestLang included					
TestLang	15.31	1,342	41.77	.000	.109
LEQ	.21	1,342	.21	.64	.001
Gender	1.39	1,342	1.39	.24	.004
Income	11.97	1,342	11.97	.001	.034
Overheard speech	10.56	1,342	10.56	.001	.030
Language community	86.52	12,342	7.21	.000	.202
TestLang*LEQ	27.51	1,342	75.07	.000	.180
TestLang*Gender	.019	1,342	.052	.820	.000
TestLang*Income	.259	1,342	.707	.401	.002
TestLang*Overheard speech	14.19	1,342	38.72	.000	.102
TestLang*Language community	48.83	12,342	11.10	.000	.280
On English scores only					
LEQ	16.30	1,342	21.46	.000	.059
Gender	.87	1,342	1.14	.286	.003
Income	4.35	1,342	5.73	.017	.016
Overheard speech	24.61	1,342	32.42	.000	.087
Language community	16.198	12,342	1.78	.050	.059
On Additional Language scores only					
LEQ	11.43	1,342	18.81	.000	.052
Gender	.54	1,342	.89	.34	.003
Income	7.87	1,342	12.95	.000	.036
Overheard speech	.13	1,342	.22	.64	.001
Language community	119.14	12,342	16.33	.000	.364

Table 9. ANCOVA results for ***comprehension*** *on the 30-word CDIs (English: CDI30Comp; Additional Language: ALCDI30Comp), in* ***Step 1 analyses*** *(predictors firmly established in the literature). TestLang = English vs Additional Language. All variables are z-scored. N = 359 as income data was not provided for 13 children out of 372.*

	Sum of squares	df	*F*	*p*	η^2
With TestLang included					
TestLang	5.29	1,342	13.66	.000	.038
LEQ	.41	1,342	.32	.57	.001
Gender	26.82	1,342	21.00	.000	.058
Income	4.33	1,342	3.39	.066	.010
Overheard speech	1.19	1,342	.93	.34	.003
Language community	23.92	12,342	1.56	.101	.052
TestLang*LEQ	35.45	1,342	91.58	.000	.211
TestLang*Gender	.054	1,342	.14	.71	.000
TestLang*Income	.212	1,342	.55	.46	.002
TestLang*Overheard speech	13.19	1,342	34.08	.000	.091
TestLang*Language community	22.96	12,342	4.94	.000	.148
On English scores only					
LEQ	21.75	1,342	25.94	.000	.071
Gender	14.64	1,342	17.47	.000	.049
Income	1.31	1,342	1.57	.21	.005
Overheard speech	11.15	1,342	13.30	.000	.037
Language community	7.32	12,342	.73	.72	.025
On Additional Language scores only					
LEQ	14.11	1,342	17.09	.000	.048
Gender	12.23	1,342	14.81	.000	.042
Income	3.23	1,342	3.91	.049	.011
Overheard speech	3.29	1,342	3.91	.049	.011
Language community	39.56	12,342	3.99	.000	.123

Table 10. ANCOVA results for ***production*** *on the 30-word CDIs (English: CDI30Prod; Additional Language: ALCDI30Prod), in* ***Step 1 analyses*** *(predictors firmly established in the literature). TestLang = English vs Additional Language. All variables are z-scored. N = 359 as income data was not provided for 13 children out of 372.*

Fixed effect		Coef.	Std. Error	*t*	X^2	*p*
LEQ	Intercept	-0.63	0.40	-1.59		
	LEQ	-0.64	0.11	-5.79		
	TestLang	0.30	0.18	1.65		
	LEQ:TestLang	0.47	0.07	6.36	18.02	.000
Overheard speech	Intercept	-0.79	0.48	-1.66		
	Overheard speech	-0.56	0.12	-4.78		
	TestLang	0.43	0.24	1.76		
	Overheard speech:TestLang	0.47	0.08	6.04	18.62	.000
Income	Intercept	-0.68	0.42	-1.60		
	Income	0.09	0.06	1.48		
	TestLang	0.33	0.20	1.65	1.60	.21

Table 11. Coefficient estimates from the linear mixed models used in ***Step 1 analyses*** *after the ANCOVAs, to estimate the robustness of each predictor individually, in* ***comprehension,*** *on the 30-word CDIs (English: CDI30Comp; Additional Language: ALCDI30Comp). For each variable which survived the ANCOVAs, such as LEQ, tests of significance (x^2) to compare the full model and its null are provided in the last two columns. TestLang = English vs Additional Language. All variables are z-scored. N = 372 apart from the last model in which N=359 as income data was not provided for 13 children out of 372.*

Fixed effect		Coef.	Std. Error	*t*	X^2	*p*
LEQ	Intercept	-0.27	0.22	-1.24		
	LEQ	-0.73	0.10	-7.21		
	TestLang	0.13	0.11	1.22		
	LEQ:TestLang	0.52	0.07	7.64	18.75	.000
Overheard speech	Intercept	-0.39	0.26	-1.49		
	Overheard speech	-0.55	0.11	-5.01		
	TestLang	0.22	0.15	1.49		
	Overheard speech:TestLang	0.42	0.06	7.25	23.50	.000
Gender	Intercept	-0.28	0.22	-1.24		
	Gender	-0.20	0.04	-4.75		
	TestLang	0.14	0.12	1.20	13.26	.000

Table 12. Coefficient estimates from the linear mixed models used in ***Step 1 analyses*** *after the ANCOVAs, to estimate the robustness of each predictor individually, in* ***production,*** *on the 30-word CDIs (English: CDI30Prod; Additional Language: ALCDI30Prod). For each variable which survived the ANCOVAs, such as LEQ, tests of significance (x^2) to compare the full model and its null are provided in the last two columns. TestLang = English vs Additional Language. N = 372. All variables were z-scored.*

		Coef.	Std. Error	*t*	χ^2	*p*
Comprehension Full model						
	Intercept	-0.67	0.39	-1.69		
	LEQ	-0.58	0.11	-5.40		
	TestLang	0.40	0.19	2.10		
	Overheard speech	-0.31	0.08	-3.67		
	LEQ:TestLang	0.37	0.07	5.28		
	TestLang:Overheard speech	0.34	0.05	6.33		
Comparison full/null model for variable:	LEQ				13.60	.001
	Overheard speech				16.21	.000
Production Full model	Intercept	-0.24	0.18	-1.39		
	LEQ	-0.60	0.11	-5.60		
	TestLang	0.16	0.09	1.69		
	Overheard speech	-0.32	0.10	-3.24		
	Gender	-0.20	0.05	-4.38		
	LEQ:TestLang	0.41	0.08	4.85		
	TestLang:Overheard speech	0.27	0.07	4.11		
Comparison full/null model for variable:	LEQ				12.20	.002
	Overheard speech				10.93	.006
	Gender				10.28	.001

Table 13. Coefficient estimates from the linear mixed models used in ***Step 1 analyses*** *after the ANCOVAs, to estimate the robustness of each predictor when incremented onto models containing the other ones. Dependent variables are the scores on the 30-word CDIs in comprehension (English: CDI30Comp; Additional Language: ALCDI30Comp) and production (English: CDI30Prod; Additional Language: ALCDI30Prod). For variables which survived the ANCOVA stages and the initial linear mixed models with single predictors (tables 9 to 12), the coefficients for the full models are presented. Then for each variable of interest, such as LEQ, tests of significance (χ^2) to compare the full model and its null are provided in the last two columns. TestLang = English vs Additional Language. N = 372. All variables were z-scored.*

	Sum of squares	df	*F*	*p*	η^2
Effect of Mixing on production with TestLang included					
TestLang*LEQ	29.92	1,355	79.16	.000	.182
TestLang*Gender	.013	1,355	.035	.85	.000
TestLang*Overheard speech	14.35	1,355	37.95	.000	.097
TestLang*Consistency	1.49	1,355	3.94	.048	.011
TestLang*Language community	21.65	12,355	4.77	.000	.139
Effect of Mixing on English production					
LEQ	17.21	1,355	20.93	.000	.056
Gender	12.63	1,355	15.36	.000	.041
Overheard speech	14.40	1,355	17.52	.000	.047
Consistency	4.99	1,355	6.07	.014	.017
Language community	8.48	12,355	.86	.59	.028
Effect of the proportion of native English on English production					
LEQ	20.69	1,296	24.20	.000	.076
Gender	7.20	1,296	8.42	.004	.028
Overheard speech	9.67	1,296	11.31	.001	.037
EngPropN	3.53	1,296	4.12	.043	.014
Language community	8.39	12,296	.82	.63	.032

Table 14. ANCOVA results for ***Step 2 analyses*** *(surviving predictors from Step 1 plus less well established predictors), restricted to those predictors providing significant effects. The dependent variables are scores in* ***production*** *on the 30-word CDIs (English: CDI30Prod; Additional Language: ALCDI30Prod). TestLang = English vs Additional Language. All variables were z-scored. The effect of Consistency (degree of language use consistency in parental speech, see Appendix 2) is calculated with N = 372. The effect of the proportion of native vs non –native English in parental speech (PropEngN) is calculated for N=313 due to 59 children for whom this variable could not be computed (see appendix 2).*

		Coef.	Std. Error	t	X^2	p
English comprehension	Intercept	0.04	0.07	0.54		
	Phonological Overlap	-0.02	0.06	-0.37		
	LEQ	0.22	0.07	3.02		
	Overheard speech	0.31	0.06	5.47	0.11	.74
Additional Language comprehension	Intercept	-0.16	0.23	-0.70		
	Phonological Overlap	0.37	0.19	1.90		
	LEQ	-0.17	0.05	-3.38	2.41	.12
English production	Intercept	0.005	0.05	0.11		
	Phonological Overlap	-0.04	0.05	-0.73		
	LEQ	0.26	0.08	3.22		
	Overheard speech	0.21	0.06	3.68		
	Gender	-0.20	0.05	-4.19	0.53	.47
Additional Language production	Intercept	-0.03	0.10	-0.30		
	Phonological Overlap	0.24	0.09	2.61		
	LEQ	-0.22	0.05	-4.05		
	Gender	-0.18	0.06	-3.31	4.61	.032

Table 15. Coefficient estimates from the linear mixed models used in ***Step 2 analyses*** *after the ANCOVAs, to estimate the effect of* ***Phonological Overlap*** *on vocabulary in each language, when incremented onto models containing the predictors retained from Step 1 analyses. Dependent variables are scores on the 30-word CDIs in comprehension (English: CDI30Comp; Additional Language: ALCDI30Comp) and production (English: CDI30Prod; Additional Language: ALCDI30Prod). Tests of significance to compare each full model to its null (without the fixed effect of Phonological Overlap) are provided in the last two columns. N = 372. All variables were z-scored.*

		Coef.	Std. Error	t	X^2	p
English comprehension	Intercept	0.04	0.07	0.62		
	Word Order	-0.09	0.05	-1.62		
	LEQ	0.22	0.07	2.91		
	Overheard speech	0.30	0.052	5.87	2.29	.13
Additional Language comprehension	Intercept	-0.25	0.18	-1.42		
	Word Order	-0.39	0.14	-2.79		
	LEQ	-0.17	0.06	-3.02	6.02	.014
English production	Intercept	0.005	0.05	0.10		
	Word Order	-0.04	0.05	-0.92		
	LEQ	0.26	0.08	3.34		
	Overheard speech	0.21	0.06	3.62		
	Gender	-0.20	0.05	-4.30	0.84	.36
Additional Language production	Intercept	-0.10	0.11	-0.92		
	Word Order	-0.12	0.09	-1.37		
	LEQ	-0.21	0.06	-3.76		
	Gender	-0.19	0.049	-3.91	1.50	.22

Table 16. Coefficient estimates from the linear mixed models used in ***Step 2 analyses*** *after the ANCOVAs, to estimate the effect of* ***Word Order typology*** *on vocabulary in each language, when incremented onto models containing the predictors retained from Step 1 analyses. Dependent variables are scores on the 30-word CDIs in comprehension (English: CDI30Comp; Additional Language: ALCDI30Comp) and production (English: CDI30Prod; Additional Language: ALCDI30Prod). Tests of significance to compare each full model to its null (without the fixed effect of Word Order) are provided in the last two columns. N = 372. All variables were z-scored.*

		Coef.	Std. Error	*t*	X^2	*p*
English comprehension	Intercept	0.04	0.07	0.57		
	Morph	-0.07	0.05	-1.36		
	LEQ	0.22	0.07	3.11		
	Overheard speech	0.30	0.05	5.93	1.77	.18
Additional Language comprehension	Intercept	-0.32	0.19	-1.70		
	Morph	-0.28	0.11	-2.49		
	LEQ	-0.19	0.06	-3.27	4.80	.028
English production	Intercept	0.005	0.05	.10		
	Morph	-0.05	0.05	-1.07		
	LEQ	0.26	0.08	3.40		
	Overheard speech	0.20	0.06	3.55		
	Gender	-0.20	0.05	-4.21	1.13	.29
Additional Language production	Intercept	-0.10	0.09	-1.05		
	Morph	-0.14	0.06	-2.17		
	LEQ	-0.22	0.06	-3.96		
	Gender	-0.17	0.05	-3.46	3.80	.051

Table 17. Coefficient estimates from the linear mixed models used in ***Step 2 analyses*** *after the ANCOVAs, to estimate the effect of* ***Morphological complexity*** *(Morph) on vocabulary in each language, when incremented onto models containing the predictors retained from Step 1 analyses. Dependent variables are scores on the 30-word CDIs in comprehension (English: CDI30Comp; Additional Language: ALCDI30Comp) and production (English: CDI30Prod; Additional Language: ALCDI30Prod). Tests of significance to compare each full model to its null (without the fixed effect of Morph) are provided in the last two columns. N = 372. All variables were z-scored.*

		Coef.	Std. Error	*t*
English comprehension	Intercept	37.62	3.78	9.97
	LEQ	0.23	0.08	2.78
	Overheard speech	5.19	0.90	5.76
Additional Language comprehension	Intercept	21.83	1.87	11.70
	LEQ	-0.06	0.02	-3.09
English production	Intercept	34.48	5.76	5.99
	LEQ	0.24	0.09	2.68
	Overheard speech	3.85	1.14	3.39
	Gender	-12.16	2.44	-4.98
Additional Language production	Intercept	18.69	2.10	8.91
	LEQ	-0.07	0.02	-3.71
	Gender	-3.17	0.83	-3.80

Table 18. Coefficient estimates from the linear mixed models ***for the UKBTAT****, for comprehension and production, in each language (English and the Additional Language). Dependent variables are scores on the 100-word CDI in English (comprehension: CDI100comp; production: CDI100prod) and on the 30-word CDI in the Additional Language (comprehension: ALCDI30Comp; production: ALCDI30Prod). For English, the models were calculated with N = 430 children and with N = 372 for the Additional Language. Variables were not z-scored so that they could be directly applied to new raw scores.*

	Coefficients	Decrements

	Intercept	LEQ	Overheard speech	Gender	10th percentile decrement	15th percentile decrement
Comprehension						
Predicted English 100 score	37.62	0.23	5.19		27.16	21.96
Predicted AL 30 score	21.83	-0.06			28.55	23.09
Production						
Predicted English 100 score	34.48	0.24	3.85	-12.16	7.74	6.26
Predicted AL 30 score	18.69	-0.07		-3.17	9.74	7.88

Table 19. Coefficients from the linear mixed models of vocabulary knowledge in English and the Additional Language, in comprehension and production, and decrements derived from standard deviations of residuals. For example, the predicted English score in comprehension is: 37.62 + 0.23 x LEQ + 5.19 x Overheard speech, with LEQ ranging from 0 to 100 (proportion of exposure to English vs the Additional Language in child directed speech), and Overheard speech ranging between 1 and 5 (1 = parents always speak the AL between them; 2 = parents usually speak the Additional Language between them; 3 = parents speak the Additional Language and English half of the time; 4 = parents usually speak English between them; 5 = parents always speak English between them). Gender is assigned a value of 1 for girls and 2 for boys. Decrement is then applied to the predicted score, and compared to the observed score to determine if the child's score is below the 10th or the 15th percentile.

Figure Captions

Figure 1. Vocabulary scores measured on the 30-words CDI as a function of the amount of exposure to English in child directed speech as measured by the LEQ (%). Top left: Additional Language comprehension; top right: Additional Language production; bottom left: English comprehension; bottom right: English production.

Figure 2. English vocabulary scores in comprehension and production measured on the 30-words CDI as a function of the proportion of English spoken between parents (Overheard speech). Overheard speech ranges from 1 (only Additional Language) to 5 (only English).

Figure 3. English and Additional Language vocabulary in comprehension (CDIComp) and production (CDIProd) on the 30-words CDI as a function of child's gender.

Figure 4. Percentile values in English comprehension (left) and production (right) measured from the Oxford Short Form CDI, as function of gender, in the full cohort of bilingual toddlers (N = 430). For comparative purposes, monolingual data for comprehension and production are included on the same graphs (N = 125, taken from the Oxford CDI database, Hamilton et al., 2000). For example, in comprehension, children who understand a maximum of 50% of the CDI (Y axis) constitute about 75% of bilinguals, and 85% of monolinguals (X axis).

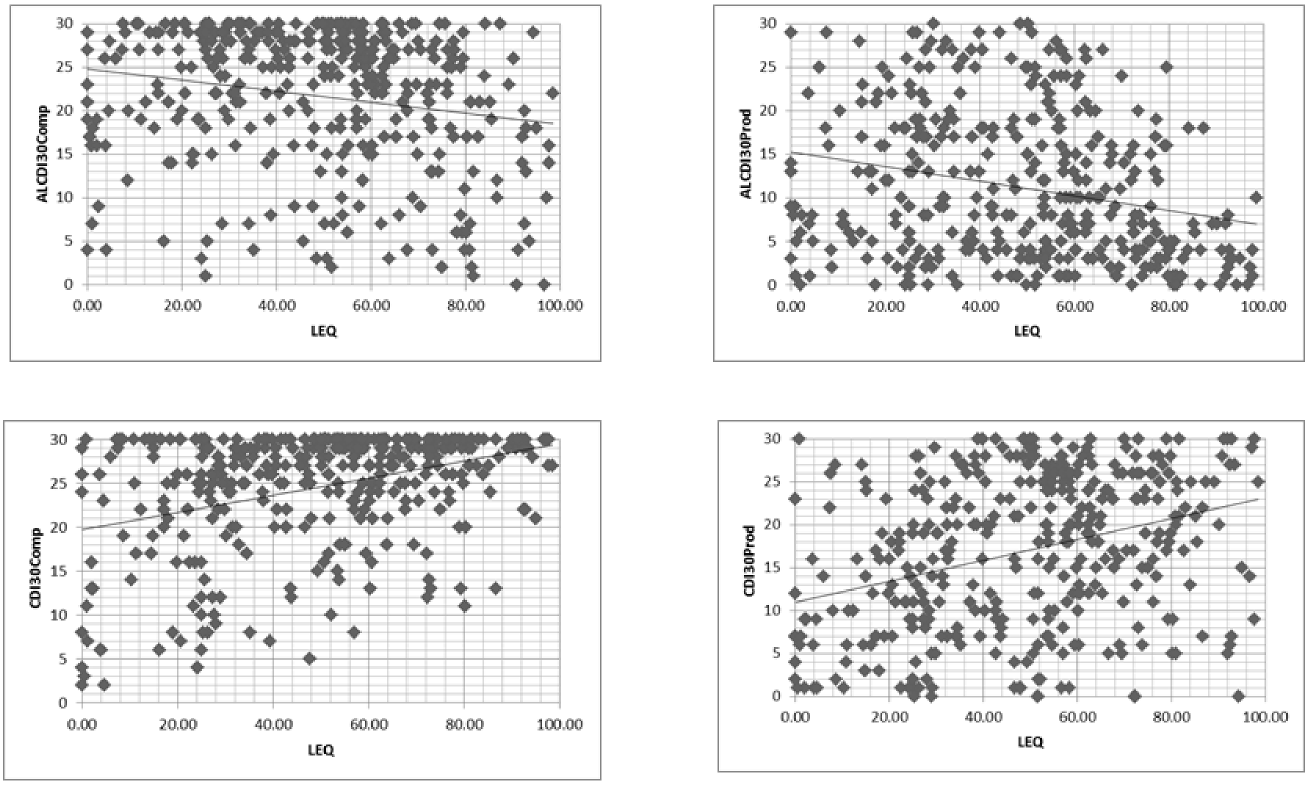
ALCDI30Comp
LEQ
ALCDI30Prod
LEQ
CDI30Comp
LEQ
CDI30Prod
LEQ

Figure 1

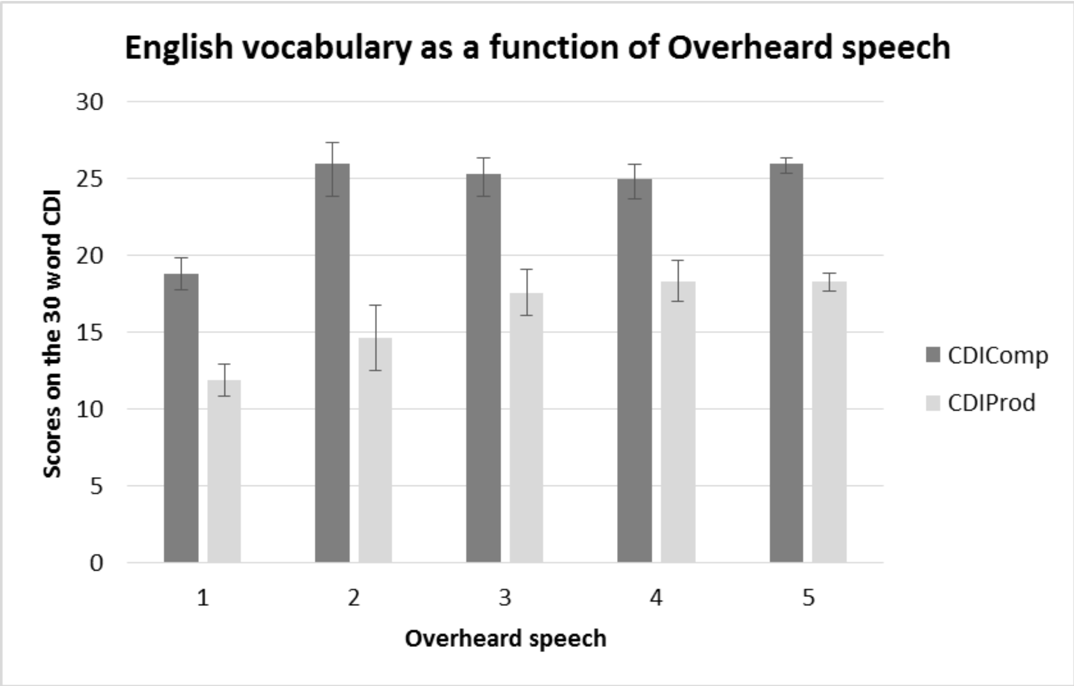
English vocabulary as a function of Overheard speech
Scores on the 30 word CDI
30
25
20
15
10
5
0
1
2
3
4
5
Overheard speech
CDIComp
CDIProd

Figure 2

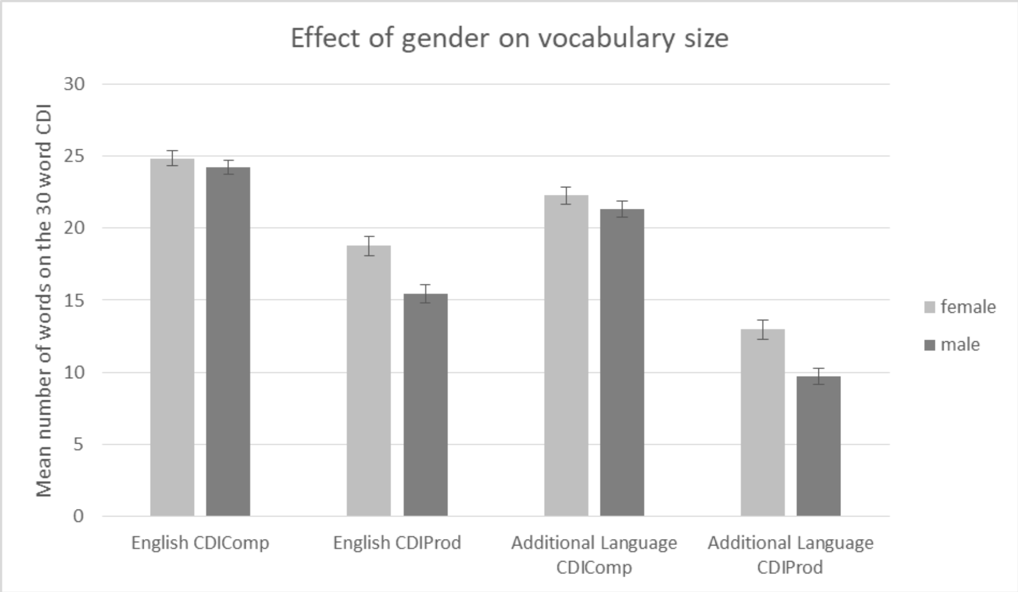
Effect of gender on vocabulary size
Mean number of words on the 30 word CDI
30
25
20
15
10
5
0
English CDIComp
English CDIProd
Additional Language CDIComp
Additional Language CDIProd
female
male

Figure 3

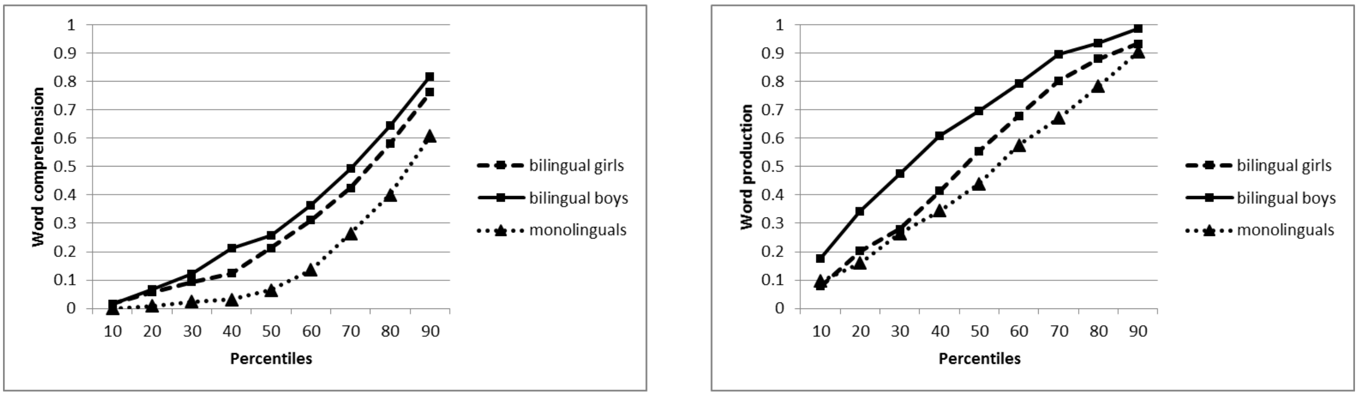
Word comprehension
Percentiles
bilingual girls
bilingual boys
monolinguals
Word production
Percentiles
bilingual girls
bilingual boys
monolinguals

Figure 4

Appendix 1 - Additional Language CDIs References

Bengali

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

Cantonese (Hong Kong) and Mandarin (Beijing)

Tardiff, T., & Fletcher, P. (2008). Chinese Communicative Development Inventories: User's guide and manual, Peking University Medical Press, Beijing, China.

Dutch

Zink, I, & Lejaegere, M. (2002). N-CDIs: Lijsten voor Communicatieve Ontwikkeling. Aanpassing en hernormering van de MacArthur CDIs van Fenson et al. Acco, Leuven (Belgium) / Leusden(Netherlands). (A CDI user's manual with normative and validity data).

French

Kern, S., & Gayraud, G. (2010), Inventaire Français du Développement Communicatif (IFDC), Grenoble, La Cigale, 978-2-912457-91-2

German

FRAKIS: Szagun, G., Stumper, B., & Schramm, A.S. (2009). Fragebogen zur frühkindlichen Sprachentwicklung (FRAKIS) und FRAKIS-K (Kurzform). Frankfurt: Pearson Assessment. http://www.pearsonassessment.de

Greek

Personal communication from Prof Demetra Kati, University of Athens, May 2014.

Italian

Caselli, M.C., & Casadio, P. (1995). Il primo vocabolario del bambino: Guida all'uso del questionario MacArthur. Milan, Italy: Franco Angeli.

Polish

Smoczyńska, M. (1999). Inwentarz Rozwoju Mowy i Komunikacji: Słowa i Zdania [Polish Adaptation of The MacArthur-Bates Communicative Development Inventory: Words and Sentences]. Unpublished material. Krakow: Jagiellonian University.

Portuguese

Frota, S., Butler, J., Correia, S., Severino, C., Vicente, S., & Vigário, M. (2016). Infant communicative development assessed with the European Portuguese MacArthur-Bates Communicative Development Inventories Short forms. *First Language*, *36*(5): 525-545. doi: 10.1177/0142723716648867

Spanish

López Ornat, S., Gallego, C., Gallo, P., Karousou, A., Mariscal, S., & Martínez, M. Evaluación de los niveles de lenguaje y comunicación de los niños pequeños. Inventario de desarrollo comunicativo de MacArthur. ISBN: 84-7174-820-7

Welsh

Mills, D., Gathercole, V., & Ebanks, N. (2013). The Bangor Welsh Communicative Development Inventory: Words and Gestures. Bangor University

Appendix 2 - Calculation of UKBT Variables

Income	£0-£14,000 1 £14,001-£24,000 2 £24,001-£42,000 3 £42,001 or more 4
Maternal education (MumEd)	No qualifications 1 Below standard for a pass on the school-leaving examination 2 O-levels (left school at 16) 3 A-levels (left school at 18) 4 Tertiary vocational qualifications 5 An undergraduate degree 6 A postgraduate degree 7
Paternal education (DadEd)	As MumEd
Number of parental AL speakers (FamLang)	Value of 1 if only one parent is a native AL speaker, and 2 if both are.
Total number of speakers of English (SourcesEng)	A value from 0 upwards. Possible sources include a maximum of one BE speaking parent (scores 1), English DayCare (see below) (scores 1) and older children in the home (1 for each). Observed range for SourcesEng was 0-6.
Total number of speakers of the AL (SourcesAL)	A value from 1 upwards. Possible sources include a minimum of one and a maximum of two AL speaking parents (scores 1 each), AL daycare (see below) (scores 1) and older children in the home (1 for each). Observed range for SourcesAL was 1-7.
Time in English speaking daycare (EngDaycare)	Number of hours a week on average the child spends in an English speaking environment (nursery / day care / preschool / childminder / relative or friend)
Time in AL speaking daycare (ALDaycare)	Number of hours a week on average the child spends in an Additional Language environment (nursery / day care / preschool / childminder / relative or friend)
Number of older siblings (Siblings)	How many other children (24 months – 18 years) live in the home

Proportion of English/AL in overheard speech (Overheard speech)	When both parents are together with the child, and they talk between the two of them, which language do they speak? 1. Always AL, 2. Usually AL, 3. English about half the time, 4. Usually English, 5. Always English
Proportion of English/AL in maternal speech (MumPropEng)	Does the mother speak 1. Always AL, 2. Usually AL, 3. English about half the time, 4. Usually English, 5. Always English
Proportion of English/AL in paternal speech (DadPropEng)	As Mother
Degree of language use consistency in mother's speech (MumConsistency)	Derived from MumPropEng as follows: 1 or 5 = 1, 2 or 4 = 2, 3 = 3
Degree of language use consistency in mother's speech (DadConsistency)	As MumConsistency
Degree of language use consistency in parents' speech (Consistency)	Average of MumConsistency and DadConsistency
Proportion of native/non-native English (PropEngN)	This variable concerns input from parents only. The hours spent with each parent was computed as 168 – 7*Sleep – EngDaycare – ALDaycare – Hours alone with other parent (for 59 children this value was negative and PropEngN and PropALN could not be computed). Parents' score on PropEng variable (1 to 5) was re-expressed as a proportion from 0 to 1 (1 = 0, 2 = .25, 3 = .5, 4 = .75, 5 = 1). This was multiplied by the hours the child spent with the parent to provide the hours of English input. If the parent was BE speaker this was native English input, if AL speaker this was non-native input. Both parents were assessed in this way. PropEngN was given by native English / (native + non-native Eng).
Proportion of native/non-native AL (PropALN)	As above, parents' score on PropEng was expressed as a proportion from 0 to 1 (1 = 0, 2 = .25, 3 = .5, 4 = .75, 5 = 1), with this value then subtracted from 1 to provide a proportion of time speaking AL. This was multiplied by the hours the child spent with the parent to provide the hours of AL input. If the parent was AL speaker this was native AL input, if BE speaker this was non-native AL input. Both parents were assessed in this way. PropALN was given by native AL / (native + non-native AL).

Appendix 3 - Details of the Calculation of Percentage of English Exposure in a Typical Week in the Year Preceding Testing (LEQ, adapted from Cattani et al., 2014)

A. Input from the parents:

Number of hours a week in English-speaking nursery/childminder/playgroup = EngDaycare

Number of hours a week in an Additional Language speaking nursery/relatives = ALDaycare

Number of sleeping hours per night = Sleep

Does the mother always speak the Additional Language (AL) to the Child, or usually, or equally often English and the AL, or usually English, or always English (5 possible responses) = MumPropEng

Does the father always speak the AL to the Child, or usually, or equally often English and the AL, or usually English, or always English (5 possible responses) = DadPropEng

When together, who speaks most to the child? Mother, father or both = Most

Number of hours per week spent with mother only = HM

Number of hours per week spent with father only = HF

B. Calculations

1. Assign a percentage to M and F, to estimate the proportion of English in each parent's input to the child.

If MumPropEng (or DadPropEng) = Always AL then ME (or FE) = 100

If MumPropEng (or DadPropEng) = Usually AL then ME (or FE) = 75

If MumPropEng (or DadPropEng) = Equally AL and English then ME (or FE) = 50

If MumPropEng (or DadPropEng) = usually English then ME (or FE) = 25

If MumPropEng (or DadPropEng) = always English then ME (or FE) = 0

2. Correct HM and HF to give more weight to the time spent with the mother, as it is found usually that fathers tend to produce less verbal output to their child, therefore

directly impacting on the amount of exposure in English and the Additional Language (e.g., Pancsofar & Vernon-Feagans, 2006).

Corrected time with mother = CHM = HM*4/3

Corrected time with father = CHF = HF*2/3

3. Assign a value (MI to Most), to give more weight to the mother's input. What is obtained corresponds to the percentage of the mother's input during the time when both parents are with the child.

If Most = Mother then MI = 90

If Most = Father then MI = 50

If Most = Both then MI = 70

4. Calculate the number of hours per week with both parents together

TBP = 7(24-Sleep) - EngDaycare - ALDaycare - CHM - CHF

5. Calculate the total number of hours of English exposure in a week (E) with the following formula:

E = English from mother when mother alone + English from father when father alone + English from mother when both parents together + English from father when both parents together + English from nursery or equivalent

$$E=\frac{\text{CHM}(100\text{-ME})}{100}+\frac{\text{CHF}(100\text{-FE})}{100}+\text{EngDaycare}+0.01*\text{TBP}*\frac{\text{MI}(100\text{-ME})}{100}+0.01*\frac{\text{TBP}(100\text{-MI})(100\text{-FE})}{100}$$

With English from mother when mother alone = CHM(100-ME)/100

English from father when father alone = CHF(100-FE)/100

English from mother when both parents together = 0.01*TBP*MI(100-ME)/100

English from father when both parents together = 0.01*TBP(100-MI)(100-FE)/100

6. Calculate the percentage of exposure to English

$$P=\frac{E}{7(24-Sleep)}$$

Appendix 4 - Breakdown of Languages for the Non-Target Additional Language Community

Arabic	6
Bosnian	1
Bulgarian	3
Catalan	1
Czech	3
Danish	3
Finnish	3
French (Quebec)	3
Hebrew	2
Hungarian	1
Japanese	5
Kannada	1
Latvian	1
Lithuanian	1
Norwegian	1
Punjabi	1
Romanian	2
Russian	1
Serbian	1
Slovak	8
Swedish	3
Tagalog	1
Tamil	1
Turkish	3
Ukrainian	1
Yoruba	1

Table A4. Number of bilingual children per language group (N = 58); they all learn British English and one of 26 Additional Languages which are not part of our 13 target Additional Languages.

Appendix 5 - List and Results for Individual English Words in the Oxford Short Form CDI

	Nouns	35	**door**		**Others**
1	donkey	36	**table**	70	bye bye
2	**elephant**	37	bowl	71	cockadoodledoo
3	**fish**	38	broom	72	dinner
4	goose	39	brush	73	nap
5	kitten	40	**cup**	74	peekaboo
6	**lion**	41	glass	75	yes
7	penguin	42	key	76	**big**
8	pig	43	lamp	77	clean
9	squirrel	44	light	78	**cold**
10	**aeroplane / plane**	45	money	79	**dirty**
11	**car**	46	**scissors**	80	fast
12	**ball**	47	**soap**	81	happy
13	balloon	48	watch	82	**hot**
14	block / brick	49	**flower**	83	old
15	**book**	50	outside	84	soft
16	pen	51	**sky**	85	wet
17	butter	52	swing	86	**what**
18	cake	53	**tree**	87	**where**
19	cereal	54	wall	88	why
20	meat	55	aunt	89	now
21	**milk**	56	**mummy**	90	today
22	tea		**Action words**	91	tomorrow
23	arm	57	call	92	back
24	mouth	58	carry	93	in
25	**nose**	59	catch	94	all
26	toe	60	drop	95	not
27	bib	61	**fall**	96	another
28	glasses / specs	62	finish	97	some
29	jacket	63	go	98	there
30	**shoe**	64	**play**	99	I
31	**sock**	65	splash	100	her
32	zip	66	swim		
33	**bed**	67	tickle		
34	**chair**	68	walk		
		69	want to		

Table A5-1. List of words in the Oxford Short Form CDI and the 30-word CDI (in bold) in their order of presentation to parents.

	Production				Comprehension			
LEQ (exposure)	0-25%	>25-50%	>50-75%	>75-100%	0-25%	>25-50%	>50-75%	>75-100%
	N=68	N=105	N=142	N=57	N=68	N=105	N=142	N=57
aeroplane/plane	38.2	65.7	69.0	70.2	66.2	86.7	91.5	91.2
ball	69.1	84.8	92.3	93.0	86.8	99.0	99.3	100.0
bed	38.2	56.2	71.8	77.2	73.5	85.7	93.0	98.2
big	22.1	41.0	51.4	56.1	47.1	65.7	71.8	75.4
book	54.4	75.2	83.8	89.5	83.8	93.3	97.2	100.0
car	72.1	83.8	91.5	89.5	91.2	96.2	97.9	96.5
chair	27.9	50.5	65.5	71.9	63.2	85.7	93.0	94.7
cold	19.1	47.6	50.0	59.6	45.6	78.1	79.6	82.5
cup	29.4	47.6	58.5	63.2	60.3	80.0	89.4	91.2
dirty	20.6	41.9	49.3	64.9	47.1	77.1	83.8	84.2
door	32.4	61.9	76.8	84.2	73.5	87.6	95.1	96.5
elephant	44.1	49.5	61.3	59.6	83.8	85.7	89.4	93.0
fall	10.3	29.5	43.0	50.9	45.6	67.6	73.9	80.7
fish	64.7	72.4	80.3	78.9	85.3	90.5	95.8	94.7
flower	29.4	52.4	63.4	64.9	60.3	81.0	90.8	96.5
hot	42.6	66.7	74.6	84.2	69.1	89.5	90.8	94.7
lion	32.4	56.2	59.2	64.9	73.5	85.7	90.8	94.7
milk	44.1	66.7	70.4	78.9	75.0	91.4	93.0	94.7
mummy	94.1	95.2	92.3	96.5	97.1	98.1	97.9	100.0
nose	47.1	77.1	80.3	84.2	82.4	95.2	97.9	96.5
play	25.0	45.7	54.2	56.1	63.2	78.1	90.1	89.5
scissors	5.9	14.3	16.2	24.6	30.9	39.0	50.0	59.6
shoes	66.2	83.8	88.0	91.2	89.7	97.1	99.3	98.2
sky	13.2	35.2	40.8	45.6	30.9	64.8	66.9	75.4
soap	11.8	24.8	28.2	36.8	41.2	45.7	64.1	70.2
sock	51.5	74.3	81.0	80.7	76.5	91.4	96.5	98.2
table	20.6	38.1	50.0	52.6	60.3	82.9	90.1	93.0
tree	30.9	48.6	64.8	68.4	57.4	77.1	88.0	94.7
what	19.1	31.4	33.8	36.8	42.6	53.3	64.1	68.4
where	14.7	34.3	33.8	38.6	61.8	73.3	79.6	84.2
Mean	36.4	55.1	62.5	67.1	65.5	80.8	86.7	89.6

Table A5-2. For words from the 30-word CDI, proportion of bilingual children (all 13 target Additional Languages collapsed) who produce and understand each word in English. Children's data are binned as a function of exposure *(as measured by the Plymouth Language Exposure Questionnaire). For example, the word 'ball' is produced by 69.1% of those children exposed to English between 0 and 25% of the time, and by 93.0% of those children exposed to English between 75 and 100%.*