

Measuring knowledge of multiple word meanings in children with English as a first and an additional language and the relationship to reading comprehension

BOOTON, Sophie A.¹

HODGKISS, Alex¹

MATHERS, Sandra¹

MURPHY, Victoria A.¹

¹Department of Education, University of Oxford.

Address for correspondence: Department of Education, University of Oxford, 15 Norham Gardens, Oxford, OX2 6PY.

Acknowledgements

The authors would like to thank the children, teachers and parents who helped with this research. We would also like to thank the other members of our research team, Rebecca Eynon, Fiona Jelley, for their guidance during the project, and Ferrero International for funding our work.

Key words: polysemy, second language, vocabulary.

Abstract

Polysemy, or the property of words having multiple meanings, is a prevalent feature of vocabulary. In this study we validated a new measure of polysemy knowledge for children with English as an additional language (EAL) and a first language (EL1) and examined the relationship between polysemy knowledge and age, language status, and reading comprehension. Participants were 112 British children aged 5 to 6 (n=61) or 8 to 9 years (n=51), 37% of whom had EAL (n=41). Participants completed the new measure of knowledge of polysemes, along with other measures of language, literacy and cognitive ability. The new measure was reliable and valid with EAL and EL1 children. Age and language status predicted children's polyseme knowledge. Polyseme knowledge uniquely contributed to reading comprehension after controlling for age, language status, non-verbal intelligence, time reading in English, and breadth of vocabulary. This research underscores the importance of polysemy for children's linguistic development.

Vocabulary is fundamental to children's development and academic achievement (Bleses, Makransky, Dale, Højen, & Ari, 2016; Schuth, Köhne, & Weinert, 2017). However, research on vocabulary has tended to focus solely on vocabulary as mapping a single word to a single meaning. Many or most words have multiple meanings (i.e. are polysemous; Brysbaert & Biemiller, 2017) and this is likely to impact children's language comprehension, particularly for children learning English as an additional language (EAL). Yet this dimension of vocabulary has largely been overlooked by research, partly due to a lack of appropriate measures of polysemous word knowledge for younger children. In the present study, a receptive assessment of children's vocabulary knowledge for polysemous words was developed and validated that reduces cognitive demands for children with EAL and younger children. The relation between polysemy knowledge and age, language status, and reading comprehension is also demonstrated.

Vocabulary Development

Vocabulary is a foundational aspect of children's development. Vocabulary forms the building blocks for language and communication. It underpins literacy, being a significant predictor of reading and writing skills both concurrently and longitudinally (e.g. Duff, Reen, Plunkett, & Nation, 2015; Ouellette, 2006). Due to the central role played by communication and literacy in education systems, vocabulary is a strong predictor of overall academic achievement (Bleses et al., 2016; Schuth et al., 2017). Furthermore, vocabulary is highly amenable to intervention (Marulis & Neuman, 2010), unlike other factors affecting academic achievement, such as socioeconomic status.

Vocabulary knowledge is critical for all children, but perhaps especially for children learning a second language. A large proportion of the world's population speaks more than one language, with for example 46% of Europeans being bilingual or multilingual (European Commission, 2012), and the number of bilinguals is rising in countries including the United

States (Ryan, 2013) and England (Department for Education, 2020). In England, the proportion of children defined as having EAL – defined as being exposed to a language at home other than English – has increased steadily since 2006, with 21.3% of primary pupils classified as speaking EAL (Department for Education, 2020). These pupils represent a diverse group with respect to first languages, volume of exposure to English at home, and onset of learning English. Whilst in educational contexts, EAL is defined as simply being exposed to a language other than English, in research contexts it has been useful to distinguish between first versus second language learners (Murphy, 2014) because of differences in their language learning contexts. Thus, in this paper we define EAL more specifically as having a first language other than English (i.e. being a sequential bilingual with a language other than English as a first language). By contrast, children with English as a first language (EL1) are those for whom English is a first language, which includes both English monolingual speakers (i.e. use English only) and bilinguals for whom English is one of their first languages.

Evidence has long suggested that bilinguals have less vocabulary knowledge within their individual languages than monolinguals (e.g. Pearson, 1993; Bialystok, Luk, Peets & Yang, 2010; Farnia & Geva, 2011). For example, Bialystok et al. (2010) found that bilingual children between 3 and 10 years had lower receptive vocabularies in their individual lexicons relative to monolingual peers, regardless of age or language pairings. They do, however, tend to have an overall vocabulary (including lexis from both languages) comparable to monolingual children (Costa, 2020). Critically, children with EAL, even with several years' exposure to English through a school setting, can struggle to catch up with their peers who have English as a first language (EL1) in vocabulary more so than other linguistic skills, such as phonological awareness and listening comprehension (Geva & Massey-Garrison, 2013). Indeed, duration and intensity of exposure to English are significant predictors of EAL children's vocabulary (e.g. Paradis, 2011; Paradis, Tulpar & Arppe, 2016; Paradis & Jia, 2017),

including exposure through digital media (Arndt & Woore, 2018; Sundqvist, 2019). As children with EAL tend to have had less exposure to English than their EL1 peers, vocabulary is a consistent area of challenge.

Much of the research on word learning to date focuses on vocabulary narrowly defined as mapping a single lexical unit to a single meaning. Many studies employ only measures of vocabulary breadth, or the number of words children know. For example, a commonly used standardised test is the British Picture Vocabulary Scale (BPVS; Dunn & Dunn, 2009), in which children are asked to select one image from four that depicts one meaning of a given word. Such measures do not tell us anything about the *depth* of children's word knowledge. Vocabulary depth refers to how well a word is known, including its semantic associates, use in collocations or idioms, and multiple senses of the word (Nation, 2001). Vocabulary depth reflects the fact that vocabulary is multi-dimensional: as such it cannot be measured adequately with only one type of test. Furthermore, considerable evidence suggests that depth of vocabulary knowledge is important for reading comprehension, for both children in their first language (Babayigit & Stainthorp, 2014; Catts, Adlof, & Weismer, 2006; Nation & Snowling, 2004) and additional languages (Geva & Farnia, 2012; Geva & Massey-Garrison, 2013; Lesaux, Crosson, Kieffer, & Pierce, 2010; Nakamoto, Lindsey, & Manis, 2007; Proctor, August, Carlo, & Snow, 2005).

Children with EAL lag behind their EL1 peers in vocabulary depth as well as breadth, and this can have consequences for their broader literacy skills. They tend to show poorer depth of understanding in relation to opaque multi-word phrases (e.g. *pay attention*, Smith & Murphy, 2015); metaphorical uses of words (Hessel & Murphy, 2019); and application of appropriate vocabulary in their writing (Cameron & Besser, 2004). Such differences in vocabulary depth could be a consequence of children's smaller general vocabularies or may represent a distinct relative weakness for EAL children. Either way, this knowledge partially accounts for

differences in reading comprehension between children's first and additional languages (Lervåg & Aukrust, 2010). Overall, children with EAL tend to have less well-developed vocabulary knowledge relative to their EL1 peers, and this has significant consequences for their other linguistic competences.

Polysemy

A crucial, but under-researched, aspect of vocabulary depth is polysemy. A large proportion of lexical forms are polysemous (Brysbaert & Biemiller, 2017). Polysemous words share the same written and/or spoken form but have multiple meanings, which may vary from being completely unrelated (e.g. homonyms like *bank*, which can mean a financial institution or the side of a river) to highly related (e.g. the verb and noun forms of *brush*). Estimates suggest that polysemous words are frequent in English and other languages. For example, approximately 4% of words across languages are homophones (Dautriche, 2015), that is, they have the same sound but a different written form to a word with another meaning (e.g. *I* and *eye*). Estimates for polysemy are more subjective and vary between 30% and 80% (Brysbaert & Biemiller, 2017; Clemmons, 2008; Rodd, Gaskell, & Marslen-Wilson, 2002). Regardless of the precise estimates, a significant proportion of the vocabulary children need to acquire is polysemous.

Despite the high prevalence of polysemes, and the significance of vocabulary depth for children's linguistic and academic achievement, few validated vocabulary tests exist which assess children's understanding of multiple meanings of the same word. Given that many or even most words have multiple senses, it is surprising that almost all standardised vocabulary measures assess understanding of just one meaning per word item. For example, the BPVS (Dunn & Dunn, 2009) asks children to select one image from four that depicts a given word, despite many of the words included in the test having multiple senses (e.g. fire, ring, ruler). This lack of available measures has significantly hampered research on polysemy. To date the only studies conducted have used a handful of bespoke measures, developed by individual

research teams. The conclusions of these studies – and gaps in knowledge arising from current limitations in the measurement of polysemy – are outlined below.

Polysemy Knowledge in L1

Research with children with EL1 suggests that understanding of polysemy shows an extended developmental trajectory. Children begin using multiple senses of words as young as two years old (Lippeveld & Oshima-Takane, 2015). By five years, children show some sensitivity to factors that determine which interpretation of a homonym to take, such as contextual plausibility (Rabagliati, Pytkänen, & Marcus, 2013; Srinivasan & Snedeker, 2011). Despite some early aptitude for understanding polysemy in their first language, older children (Doherty, 2004) and even adults (Degani & Tokowicz, 2010) have some difficulty in acquiring new meanings of polysemes. For example, Doherty (2004) showed that when five to nine-year-old children were exposed to a pseudo-homonym through a story ('cake' to refer to a novel animal), they struggled to recognise the correct meaning in a subsequent receptive test. Specifically, five-year-olds performed at or below chance level and whilst there was some improvement with age, even nine-year-olds only achieved 55% accuracy. This difficulty in learning words was specific to homonyms as children scored at ceiling when a nonsense word was used instead. The data suggested that children found attaching a second meaning to a known word form challenging due to interference from the known meaning: when the known meaning was not an option in the receptive test (e.g. there was no picture of a birthday cake for 'cake'), accuracy at all ages was much higher. Thus, while children can learn multiple meanings for homonyms at an early age, learning homonyms is more challenging than learning semantically unambiguous words and children may confuse the secondary with the primary meaning.

Research also implies that children's ability to provide multiple definitions of polysemes continues to grow in middle childhood, as indicated by two standardisation studies of polyseme knowledge tests (Richard & Hanner, 2005; Wiig & Secord, 1991) and grade of

acquisition norms (Brysbaert and Biemiller, 2017). For example, in the Multiple Meanings subtest of The Language Processing Test 3 – Elementary (Richard & Hanner, 1995), children were given a common polysemous word (e.g. *spring*) in three sentence contexts that disambiguated three different meanings of that word. They were then asked to provide a definition of the word in that context. In a sample of 1,126 US children (most of whom had EL1), the number of words for which children could provide at least two correct definitions increased overall from six years (32%) to eleven years (86%), with the greatest jump in accuracy on this test from six years (32%) to seven years (56%).

Grade of acquisition norms developed by Dale and O'Rourke (1981) and extended by Brysbaert and Biemiller (2017) suggest an extended trajectory from age nine to twenty years. They measured children's knowledge of 31,000 words through a written, three-alternative multiple-choice test. Children recognised definitions of multiple senses of some words before age nine to ten years (e.g. *show* – to point out, and *show* – performance). However, the majority of additional meanings of polysemes were learnt after age nine to ten years: children both learnt new word forms which are polysemous and acquired new meanings for word forms they already knew. This implies that children with a smaller overall vocabulary are also likely to know fewer polysemous words overall, and particularly secondary meanings of these words. These data suggest that children's ability to provide accurate definitions of polysemous words develops significantly across middle and late childhood.

Whilst these norming studies suggest that children improve in their ability to define polysemous words during childhood, there are significant limitations with the measures used that restrict our understanding of the development of polysemy knowledge. The existing tests of polysemy knowledge present incidental cognitive and linguistic demands for children. Even in the test normed for the youngest age group, six-year-olds (LPT-3 Multiple Meanings subtest), children must recognise the sense of the word indicated by the context and formulate

a coherent definition for at least two meanings. Furthermore, children are given no examples, practice items, or feedback, and no prompts to elaborate if they provide vague answers. This means that children must simply infer that the test requires elaborated definitions. The cognitive demands of these tests mean they likely underestimate children's knowledge of the specific words tested and are dependent on cognitive abilities that are not central to polysemy vocabulary. This makes them especially unsuitable for use with younger children and children who are learning English as an additional language.

Polysemy Knowledge in L2

Learning polysemes in an L2 is likely to be a challenge for children, given the differences in breadth and depth of vocabulary between bilinguals and monolinguals. Research with bilingual *adults* suggests that knowledge and processing of polysemous words is different compared to monolinguals. For instance, Arabic-English bilingual adults were able to identify fewer homophones in English than monolinguals (Gathercole & Moawad, 2010). Bilingual adults have also been shown to transfer homonyms between languages inappropriately, commonly from L1 to L2 (Elston-Güttler & Williams, 2008) but also from L2 to L1 in early bilinguals (Gathercole & Moawad, 2010). Once homonyms are learnt, they may still be processed differently in L2 for adults (Elston-Güttler & Friederici, 2005). For example, when shown homonyms in a disambiguating sentential context in a lexical decision task, advanced German learners of English were found to continue to process the inappropriate meaning despite the disambiguating context. Thus, bilingual adults know fewer homonyms and process them less effectively in their L2. We might assume, therefore, that children with EAL would equally struggle with these aspects of vocabulary knowledge in English.

While there is very limited evidence relating to children's knowledge of polysemes in their second language, there is one notable exception. One intervention study suggested that children with EAL may have poorer knowledge of English polysemous meanings of words

than monolingual English speakers (Carlo et al., 2004). Ten to eleven-year-old Spanish-speaking children with EAL were asked to generate sentences conveying different meanings of polysemous words (e.g. ring, settle), both before and after a daily, 15-week explicit vocabulary intervention. Children with EAL scored poorer at both pre- and post-test than monolingual English-speaking children but improved more than their monolingual peers in their knowledge of polysemous words following the vocabulary intervention. However, the authors acknowledged that the two language groups were drawn largely from different testing sites, with the monolingual children primarily from middle-class socioeconomic backgrounds while the children with EAL were primarily from working-class backgrounds. Additionally, the analyses did not control for vocabulary breadth, and the measure of polyseme knowledge used required additional cognitive skills (generating meanings in context, constructing sentences) which might underestimate the performance of children with EAL. Thus, it remains unclear whether children with EAL children truly differed in their knowledge of polysemous words. Because of the single age group used, no light is shed on the possible developmental trajectory of knowledge of polysemous words for children with EAL.

Polysemy and Reading Comprehension

Given the potential significance of polysemy knowledge for children's linguistic comprehension, the lack of valid measures of polysemy knowledge is highly problematic. Polysemy entails semantic ambiguity, and thus presents potential for misinterpretation of intended word meanings. Indeed, children often take a known, but incorrect, interpretation of a word, even when the context does not support it (Doherty, 2004). Knowledge of secondary word meanings should support comprehension of texts containing those senses of the words. For instance, children who know that *bat* can mean an animal will show better comprehension of texts about an animal bat than children who only know *bat* as a piece of sports equipment. Furthermore, knowledge of polysemy could have an impact on reading comprehension more

generally (i.e. not just comprehension for texts containing the polysemes tested). For example, children who know more polysemes may have greater metalinguistic awareness of the variety of word meanings. This may trigger them to more effectively infer word meanings from the context, whether they know the secondary word meaning at the outset or not. One study has addressed this question (Logan & Kieffer, 2017). The authors found that Spanish-speaking adolescents with EAL who could identify the definitions of secondary meanings of words in English also had better general reading comprehension in English. This was true after controlling for knowledge of primary meanings, decoding skill, and vocabulary breadth. This is compelling evidence, albeit from only one study, that polysemy knowledge predicts reading comprehension in adolescents with EAL. However, further data is needed to determine if this finding replicates with younger children and those with EL1.

The Present Study

There is a lack of data pertaining to three key aspects of children's knowledge of polysemy which the present study aims to address. Firstly, there is a dearth of developmental data on children's understanding of polysemous words, particularly those with EAL and younger children, due to a lack of tests of this critical dimension of vocabulary without extraneous cognitive demands. Thus, the present study aimed firstly to develop and validate a measure of knowledge of polysemy for children in the early primary school years; the Receptive Polyseme Vocabulary Test (RPVT). To reduce incidental demands (e.g. on reading, memory, and verbal expression) compared to the few previously available measures, a pictorial multiple-choice format was adopted. The test was assessed for reliability (test-retest and internal consistency) and convergent validity with other vocabulary measures. Thus, the first research question was: is the RPVT a valid and reliable measure of children's knowledge of polysemous words?

Secondly, there is a gap in the literature in terms of identifying whether there are differences between EAL children and their EL1 peers in their understanding of polysemy, and

what the nature of those differences might be. Therefore, the second aim of the study was to examine some candidate factors which might influence children's knowledge of polysemous words, as measured by the RPVT. Consequently, individual difference factors (the child's age and language status: EAL or EL1), and language exposure, including time spent learning English and use in communication and reading, were included in our investigation. Therefore, the second research question was: do age, language status, and language exposure affect children's knowledge of polysemous words? It was expected that older children, children with EL1, and children with greater English language exposure would demonstrate greater knowledge of polysemous words.

Thirdly, it remains unclear from the existing literature whether understanding polysemy plays a significant role in younger children's, and native English speakers', comprehension of text. As such, the third aim was to explore the relation between knowledge of polysemes and a concurrent standardised measure of reading comprehension. Therefore, the third research question was: does knowledge of polysemous words predict unique variance in reading comprehension?

Method

Design

A 2x2 between-subjects design was used with factors of year group (Year 1 and Year 4) and language status (EAL or EL1). Children completed a battery of cognitive measures, including the new polysemy test twice, at least one week apart, to assess test-retest reliability. All participants completed all tasks in the same fixed order. The methods and analyses were pre-registered on the Open Science Framework (<https://osf.io/ym3hu>) and any exploratory deviations from this are noted.

Participants

Participants were 112 children (59 female), recruited through 5 state schools in southern England which varied in their deprivation indices (from 3rd to 8th decile, child $M = 5.57$, $SD = 1.86$; Department For Communities and Local Government, 2019). All schools had a higher than average intake of EAL children for the country (range: 24.6 to 51.4%; Department for Education, 2012) and EAL and EL1 children were recruited from the same classes and schools. Children were from two school year groups: 61 Year 1 children (aged 5 to 6 years), and 51 Year 4 children (aged 8 to 9 years). Details of the participants in each group are shown in Table 1 and further details in Appendix 1.

Children were classified as having English as an additional language if English was not their first language (i.e. they were sequential bilinguals with a language other than English as a first language). English was an additional language for 41 children (37%). All other children spoke English as a first language, and so were classified as EL1. A proportion of the 71 EL1 children – though English was their first language – were exposed to other languages in the home, with 25 of the EL1 children (35%) using another language with any degree of fluency more than once a week (i.e. were bilingual). As is common in the UK context, the EAL children had a range of first languages: 17 different languages were used in total. The age of English onset was not significantly different for children with EAL in Year 1 ($M = 3.75$, $SD = 1.12$) and Year 4 ($M = 4.11$, $SD = 1.85$; $t(33) = 0.66$, $p = .513$): on average children began learning English at school onset in the UK aged 4 years (range: 2 to 8 years old). The number of years learning English was higher for Year 4 ($M = 4.94$, $SD = 1.74$) than Year 1 children ($M = 2.38$, $SD = 1.18$; $t(33) = 5.02$, $p < .001$) by approximately 3 years ($M = 2.56$, $SD = 0.51$). However, the sample did include some children who had only recently begun learning English (range: 0 to 7 years).

Table 1. Details of participants in each group in the sample.

Year	Language status	N	Female %	Age (SD)	Languages used
Year 1	EAL	20	40.0	6.14 (0.31)	Polish (7), Tetum (3), Arabic (2), Punjabi (2) Kurdish (1), Malayalam (1), Nepalese (1), Pashto (1), Portuguese (1), Romanian (1), Tamil (1).
	EL1	41	46.3	6.06 (0.29)	Urdu (5), Punjabi (2), Arabic (1), Latvian (1), Malayalam (1), Nepalese (1), Persian (1), Russian (1).
	All Year 1	61	44.3	6.09 (0.29)	
Year 4	EAL	21	52.3	9.07 (0.36)	Polish (5), Tetum (2), Albanian (1), Arabic (1), French (1), Kurdish (1), Lithuanian (1), Nepalese (1), Persian (1), Portuguese (1), Romanian (1), Spanish (1), Vietnamese (1).
	EL1	30	50.0	9.04 (0.22)	Malayalam (3), Punjabi (2), Albanian (1), Kiyaranda (1), Nepalese (1), Portuguese (1), Somali (1), Tagalog (1), Tamil (1).
	All Year 4	51	51.0	9.05 (0.29)	

Note: EAL=English as an Additional Language; EL1= English as a first language.

Materials

Receptive Polysemy Vocabulary Test






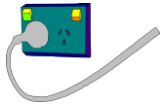





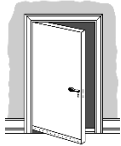
Task Format. The Receptive Polysemy Vocabulary (RPVT) test was developed for this study and aimed to measure children's knowledge of the meanings of polysemes. It uses a visual, multiple choice format like other receptive vocabulary measures (such as the BPVS; Dunn & Dunn, 2009). During the test, children heard a prompt word (e.g. *groom*) and had to select two pictures that they thought showed two different meanings of the prompt word. The format of questions was six-alternative forced choice, with 2 targets and 4 distractors. There were 30 items in total.

Word Stimuli. The full list of prompt words is shown in Appendix 1. All words were homonyms of 3 to 6 letters in length (1 or 2 syllables) and of medium to high frequency. Frequency counts obtained from the SUBTLEX norms for each word ranged from 3.65 to 5.70 ($M=4.57$, $SD=0.48$), which is approximately equivalent to between 10 to 1,000 instances per

324 million words. (Frequency counts for the two separate meanings of the words were not
325 available.) Stimuli were selected by piloting with children ($N=63$) and adults ($N=34$). More
326 details about stimulus selection are available in the online supplementary materials.

327 **Response Options.** The response options consisted of coloured drawings. These were
328 images sourced from Google Images and licensed for non-commercial reuse with modification.
329 Each item had two target response options, which were images of the two meanings of the word
330 (see Table 2). Each item also had 4 distractor images, one of each of 4 types of distractor. The
331 first distractor type was a semantic associate of the target word for meaning A. The second
332 distractor type was a semantic associate of the target word for the meaning B. The third
333 distractor type was a phonologically similar word to the prompt word. The fourth distractor
334 type was another image that was not a semantic associate of either word meaning.

Table 2. Examples of stimuli from the RPVT, including prompt words, target types and distractor types.

Word	Targets		Distractors			
	Meaning A	Meaning B	Meaning A associate	Meaning B associate	Phonological associate	Other word
train					 "drain"	
groom					 "game"	

Administration. The test was administered digitally using the Qualtrics app on a 10"

Android tablet. Children were introduced to the test as a quiz about words which have more than one meaning and given a simple example (*watch*). Children were then instructed to pick two pictures that showed two different meanings of the given word. An example using the word *watch* was worked through, in which the experimenter explained the reasons for selecting the targets and not the distractors. The experimenter explained that it was important to look carefully at all pictures before deciding, as some pictures could look similar. Children were also told to guess if they were not sure and to select two responses for each question. Two practice items were then given (*cold* and *play*) and children were given feedback to either reinforce the accuracy of their responses or correct their responses: 32% of children in session 1 and 10% in session 2 required corrective feedback. Children then completed the 30 test items, with a break half-way. Questions were presented in a fixed, random order. The task was self-paced: children took between 5 to 20 minutes to complete all items. The experimenter read out

the word for each item, and children used the touch screen to select two responses and press the continue button at the end of each question. Children were able to change their answers before pressing continue.

Scoring. Two types of accuracy scores were obtained. The first was the total percentage score, calculated as percentage accuracy on both meanings of all items. The second type was three frequency scores, calculated as the percentage of questions on which children scored none, one, or both meanings correct.

Child Language Background

A short survey (see Appendix 2) was conducted to ascertain the language background of the participants. The survey was completed by parents and contained six questions covering the child's first language(s), and other language(s); the age at which the child began learning English; frequency of use and proficiency in their home language; and time spent reading and using other media in English and their home language. Where there were contradictions in parental reporting of the child's first language (for example, marking English as a first language but stating the age of first learning English as 4 years old), the decision was referred to the child's class teacher. There were such conflicts for 25% of children. Analyses were rerun excluding these children. Since the pattern of results was unchanged, the results of the full sample are reported.

Children also reported on the dominance of English versus their home language in their day-to-day communication. Specifically, children were asked to what extent they spoke or were spoken to in English or their home language when interacting with parents; other family members; and friends, (6 questions in total). The scale for each question ran from 0 (completely in English) to 3 (completely in the home language), generating a total score from 0 (English completely dominant) through 9 (balanced) to 18 (home language completely dominant).

376 ***Homonym Knowledge***

377 To measure children's knowledge of homonyms, the Multiple Meanings subtest from the
 378 Elementary Language Processing Test 3 - Elementary (LPT-MM; Richard & Hanner, 2005)
 379 was used. In this task, children are presented with a stimulus word in three distinct sentence
 380 contexts. Children are asked to provide a definition for the word in at least two of the three
 381 sentences presented. For example, children are asked to define spring in the context 'Spring
 382 comes after winter', 'The horse stopped at the spring for a drink', and 'The weeds seem to
 383 spring up overnight'. Responses were scored according to the LPT-3 manual. Children received
 384 one point for each word in which they provided at least 2 accurate definitions, out of a total of
 385 12.

386 This subtest was normed with American children aged 6;0 to 11;11. Although the current
 387 study included some children under 6;0, the LPT3 was the only standardised measure of
 388 homonym vocabulary available for children under 8 years. The test contains four definitions of
 389 words which are based in the original American-English dialect. These were *fly* as in the
 390 baseball term *pop fly*; and *trunk* as in the storage compartment of a car; *cut* in the phrase "*cut*
 391 *it out!*"; and *check* as in the symbol (the latter three known more commonly in British English
 392 as *boot*, "*stop it*", and *tick* respectively). It was possible for a child to score full marks despite
 393 not knowing these terms, and indeed children were no less accurate on the words containing
 394 these 4 items than the other 8 words ($t(109) = 0.37$, $p = .713$). Some words and phrases
 395 particular to American English were used in the contextual sentences. These were adapted to
 396 British English to increase clarity. (Specifically, *math* was changed to *maths*; *field trip* to *school*
 397 *trip*; *groceries* to *shopping*; *grocery store* to *supermarket*; and *soccer* to *football*.)

398 ***Vocabulary Breadth***

399 To measure vocabulary breadth, the British Picture Vocabulary Scale version 3 (BPVS3; Dunn
 400 & Dunn, 2009) was administered according to the manual's instructions. Raw scores, which

reflect the number of items completed correctly, were taken as the key outcome measure (range 0 to 168). Standardised scores were also calculated, and where scores were “below 70” ($n = 9$), they were truncated at 70.

Reading Comprehension

The York Assessment of Reading for Comprehension (YARC; Snowling et al., 2009) passage reading subtest form A was used to measure reading comprehension. In this task, children read one or two passages of text, and are asked comprehension questions requiring: extraction of literal information; providing word definitions; and inference-making. Raw data on accuracy on comprehension questions, number of decoding errors, and speed of reading are converted to ability scores to control for the difficulty of the text read. For the purposes of this study, ability score on comprehension questions was the outcome measure (range: 0 to 85). Standardized scores are also reported only for descriptive purposes.

Non-Verbal Intelligence

To assess non-verbal intelligence, the Matrix Reasoning subtest from the Weschler Intelligence Scale for Children fifth edition (WISC-MR; Wechsler, 2016) was used. In this task, children see an array (a sequence or matrix) of shapes with one element missing and select the shape that matches the array from five response options. Raw scores (the number of items answered correctly) were taken as the outcome measure (range: 0 to 32).

Procedure

Consent was obtained from parents and verbal assent from children prior to testing. Testing was conducted in a quiet area in the child’s school. Children completed the battery of tasks across three sessions, in a fixed order. Sessions took 20 to 30 minutes to complete. In the first testing session, children completed the language dominance self-report, followed by the first RPVT session, followed by the BPVS. In the second testing session, children completed the LPT Multiple Meanings subtest, followed by the YARC passage reading test. In the third

testing session, which was conducted an average of 9 days after the first ($SD=3.8$, range: 6 to 29 days), children completed the RPVT for a second time.

Results

Data Analysis

The analyses conducted were pre-registered (<https://osf.io/swq49/>) and will be reported as planned. The pre-registered analyses comprise four parts: reliability and validity statistics; an ANCOVA examining the impact of age and language status on polysemy knowledge; a regression examining the relation of polysemy knowledge to reading comprehension; and correlations between language background factors and polysemy knowledge for children with EAL. One further exploratory analysis (i.e. ANCOVAs on the alternative frequency score from the RPVT) will also be reported.

Data Cleaning

The frequency scores on the RPVT were skewed, where ‘none correct’ and ‘both correct’ were positively skewed, whilst ‘one correct’ was negatively skewed. Because the skewness was in opposite directions, a square root transform was applied for the former and a square transform was applied to the latter (Osborne, 2010). The BPVS score was more strongly positively skewed for some groups, so a natural logarithm transformation was applied. The LPT-MM was strongly skewed for Year 1 groups, such that the minimum score (0) was the most common: because this cannot be corrected with transformation, non-parametric tests were used. Raw values are reported in all descriptives for ease of interpretation.

Two participants were missing all data on the LPT-MM because they did not provide any answers. Five participants were unable to complete the foundation level passage of the YARC; hence they received no score. Additionally, some parent-report data was missing from the language background questionnaire for the following questions: age of English (and years of learning English) ($n=6$); frequency of use of other language ($n=7$); ability in other language

($n = 3$); time spent reading in English ($n = 12$); time spent reading in other language ($n = 17$); time spent using other media in English ($n = 11$); time spent using other media in other language ($n = 16$). Listwise exclusion was used in all analyses.

Pre-Registered Analyses

RPVT Reliability and Validity

The RPVT was developed for this study, therefore reliability and validity of the test are examined first for the two outcome measures. Statistics on reliability and validity are reported in Table 3. Internal consistency was excellent in the full sample and was either good or acceptable in smaller sample subgroups. Test-retest reliability was examined using a Pearson's correlation between the two administration points revealing high stability over time in the full sample and high stability in all subgroups (all $ps < .001$). Scores for time 2 were slightly higher than time 1 ($t(111) = 4.50, p < .001$), indicating a practice effect. Total scores were 2.7% (1.6 points) higher at the second test ($M = 68.07, SD = 15.26$) than the first ($M = 65.39, SD = 14.79$).

Table 3. Reliability and validity statistics for the RPVT total and polysemy scores.

	Year 1			Year 4			Full sample
	EAL	EL1	All	EAL	EL1	All	
Cronbach's alpha	.70	.85	.84	.88	.79	.86	.92
Test-retest reliability	.78**	.82**	.83**	.83**	.85**	.85**	.91**
Correlation with LPT-MM^a	.39	.61**	.58**	.79**	.33	.62**	.84**
Correlation with BPVS	.63**	.59**	.62**	.87**	.77**	.85**	.76**

Note: * $p < .05$ ** $p < .01$. EAL=English as an Additional Language; EL1= English as a first language; LPT-MM=Language Processing Test – Multiple Meanings; BPVS = British Picture Vocabulary Scale. ^aLPT-MM correlations used Spearman's correlation; all other correlations were Pearson's.

Validity was examined by correlating RPVT scores with a measure of children's definition of homonyms (LPT-MM), and a measure of vocabulary breadth (BPVS). Partial correlations controlling for age revealed a strong and significant positive relationship between the RPVT score and BPVS for the full sample and all subgroups (see Table 3). Spearman's correlations showed a similar pattern of positive correlations between the RPVT and LPT-MM, although these did not reach significance for two of the subgroups. This is possibly due to limited variability in the LPT-MM for these groups (see Table 5). Overall, these results suggest that the RPVT correlates with other measures tapping into similar constructs.

Furthermore, we wanted to see whether the items from our test converged with existing grade of acquisition norms collected with a different test (Brysbaert & Biemiller, 2017). To do this, we examined the Pearson's correlation between the grade of acquisition and score for the 30 test items: grade of acquisition (from 2-12) was averaged across the two meanings for each item. This revealed a significant correlation ($r = -.50, p = .005$), such that children tended to score lower on items with a higher estimated grade of acquisition. This suggests that the RPVT converged with existing norms, but that these existing norms did not completely align with test performance.

The Effects of Age and Language Status on Polysemy Knowledge

Descriptive statistics for RPVT score at both time points for the 4 groups of children are presented in Table 4. Note that all of the groups performed above a chance score of 33.3% in the RPVT according to one-sample t-tests ($ts > 9.79, ps < .001$).

Table 4. Descriptive statistics for the RPVT by the 4 groups of participants.

Year	Language status	RPVT time 1				RPVT time 2			
		<i>M</i>	<i>SD</i>	Min	Max	<i>M</i>	<i>SD</i>	Min	Max
Year 1	EAL	50.75	7.95	33.33	75.00	51.67	10.97	33.33	81.67
	EL1	58.73	11.15	35.00	86.67	61.63	10.35	41.67	83.33
	All Year 1	56.12	10.83	33.33	86.67	58.37	11.48	33.33	83.33
Year 4	EAL	71.75	11.77	50.00	93.33	76.43	11.93	46.67	93.33
	EL1	79.78	8.67	63.33	95.00	81.95	8.52	61.67	95.00
	All Year 4	76.47	10.72	50.00	95.00	79.67	10.32	46.67	95.00
Full sample		65.39	14.79	33.33	95.00	68.07	15.26	33.333	95.00

Note: Chance score = 33.3. EAL=English as an Additional Language; EL1= English as a first language; RPVT = Receptive Polysemy Vocabulary Test.

The means and standard deviation of scores on the WISC-MR and LPT-MM, and raw, ability and standardised scores for the BPVS and YARC reading comprehension by the 4 groups of children are presented in Table 5.

Table 5. Means and standard deviations for other tests by the 4 groups of participants.

Year	Language status	Stat	LPT-MM	WISC-MR	BPVS raw	BPVS stand.	YARC ability	YARC stand.
Year 1	EAL	<i>M</i>	0.79	10.45	69.70	84.30	25.84	93.00
		<i>SD</i>	1.23	4.19	18.14	11.34	12.74	11.52
	EL1	<i>M</i>	1.85	9.29	85.37	95.66	39.49	106.35
		<i>SD</i>	2.63	3.41	15.40	12.22	11.36	11.34
	All Year 1	<i>M</i>	1.51	9.67	80.23	91.93	34.86	101.82
		<i>SD</i>	2.31	2.69	17.81	13.01	13.15	12.97
Year 4	EAL	<i>M</i>	5.81	16.48	105.19	85.76	50.29	94.10
		<i>SD</i>	3.87	3.33	19.06	15.88	9.50	10.65
	EL1	<i>M</i>	7.60	14.33	117.80	93.33	59.23	105.00
		<i>SD</i>	1.92	3.17	14.98	13.85	8.90	11.41
	All Year 4	<i>M</i>	6.86	15.22	112.61	90.22	55.55	100.51
		<i>SD</i>	2.64	3.37	17.74	15.04	10.09	12.26
Full sample		<i>M</i>	3.99	12.20	94.97	91.15	44.72	101.20
		<i>SD</i>	3.64	4.49	23.99	13.93	15.79	12.59
		<i>N</i>	110	112	112	112	107	107

Note: EAL=English as an Additional Language; EL1= English as a first language; LPT-MM=Language Processing Test – Multiple Meanings; WISC-MR = Weschler Intelligence Scale for Children – Matrix Reasoning; BPVS = British Picture Vocabulary Scale; YARC = York Assessment of Reading for Comprehension – Comprehension score.

Before conducting the key analysis, two 2 (year: Year 1 and Year 4) by 2 (language status: EAL and EL1) ANOVAs were performed to determine whether the groups were similar in terms of vocabulary breadth and non-verbal intelligence. The two year-groups did not differ in terms of BPVS standardised scores ($F(1,108) = 0.027, p = .87$), suggesting that the two age groups were of equivalent general linguistic ability for their age. As predicted given previous research, there was a main effect of language status ($F(1,108) = 13.09, p < .001, \eta_p^2 = .108$) in favour of EL1 children, and there was no significant interaction ($F(1,108) = 0.52, p = .47$). The four groups did differ in non-verbal intelligence according to raw scores on the WISC Matrix Reasoning test: there was not only an expected effect of year ($F(1,108) = 65.03, p <$

.001, $\eta_p^2 = .376$), but also of language status ($F(1,108) = 5.78, p = .018, \eta_p^2 = .051$). This was such that the children with EAL had higher non-verbal intelligence ($M = 13.46, SD = 4.81$) than those with EL1 ($M = 11.81, SD = 4.13$). The interaction was not significant ($F(1,108) = 0.52, p = .47$). Thus, in our sample the children with EAL have higher non-verbal intelligence, and thus this will be controlled in further analyses.

A 2 (year) x 2 (language status) between-subjects ANCOVA was performed on RPVT total score, with non-verbal intelligence as a covariate. The results are shown in Figure 1. There was a significant and large main effect of year ($F(1,107) = 39.99, p < .001, \eta_p^2 = .272$), such that children in Year 4 had higher scores ($M = 72.20, SD = 11.28$) than those in Year 1 ($M = 57.33, SD = 11.08$). There was also a significant and medium-large main effect of language status ($F(1,107) = 26.69, p < .001, \eta_p^2 = .200$), where EL1 children had higher scores ($M = 69.68, SD = 9.58$) than children with EAL ($M = 59.83, SD = 9.65$). There was no interaction between age and language status ($F(1,107) = 0.09, p = .76$), indicating that the effect of language status remained consistent between year groups.

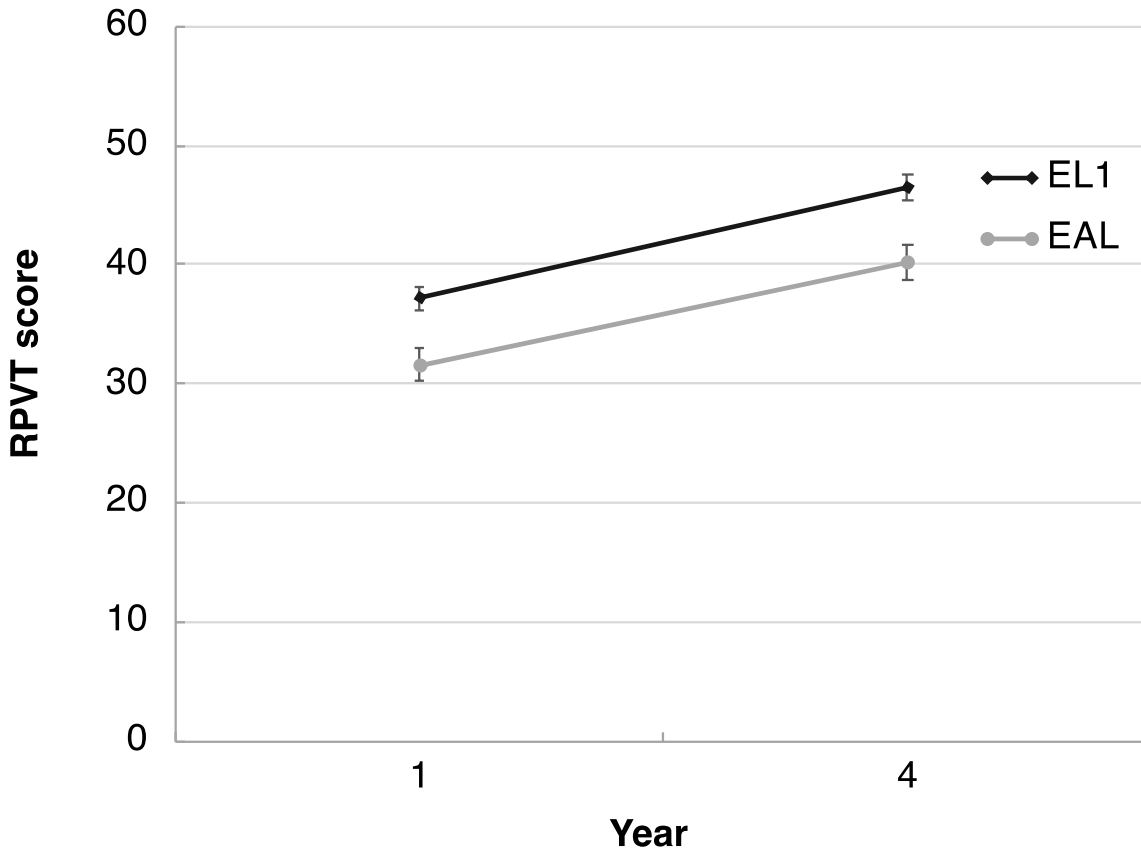


Figure 1. Graph of mean and standard errors of performance on RPVT by year group and language status.

Note: Estimated marginal means with the covariate of WISC-MR = 12.20. Maximum score = 60. Chance score = 20.

To determine whether the effects of language status and year on polysemy knowledge could be accounted for by vocabulary breadth, an additional 2 x 2 between-subjects ANCOVA was conducted on total score. This ANCOVA followed the same model as the one reported above except that we also included BPVS raw score as a covariate. The significant effect of year remained ($F(1,106) = 18.32, p < .001, \eta_p^2 = .148$), as did the lack of interaction ($F(1,106) = 1.08, p = .301$). Importantly, the effect of language status remained significant ($F(1,106) = 6.72, p = .011, \eta_p^2 = .060$), although the effect size was reduced: children with EL1 performed better ($M = 67.63, SD = 8.68$) than those with EAL ($M = 62.88, SD = 8.88$). This finding

suggests that the effect of language status on polysemy knowledge cannot be entirely explained by breadth of vocabulary differences.

Error types were also checked to see if these differed by group. The number of each of the four possible error types on the RPVT was compared by conducting four 2 (year) by 2 (language status) ANCOVAs, controlling for non-verbal intelligence. (Separate analyses were performed rather than a repeated measures ANCOVA due to the non-independence of the four error rates.) The means are reported in Appendix 3 and the ANCOVA results in Appendix 4. All analyses yielded a main effect of year and a main effect of language status, with small to medium effect sizes. This suggests that older children and children with EL1 are less likely to select all types of distractors. There was no significant interaction between year and language status. It is notable that for the main effect of age, semantic associates showed the strongest effect sizes, whereas for the main effect of language status, the phonological associate showed the strongest effect size.

Polysemy Knowledge and Reading Comprehension

The third research question was to examine whether knowledge of polysemous words is related to reading comprehension. Correlations between the RPVT and reading comprehension, and covariates of breadth of vocabulary, non-verbal intelligence, and parent-reported minutes per day spent reading in English are shown in Table 6. There were significant positive correlations between all cognitive tests, both before and after controlling for age and language status. Reading in English was correlated with all cognitive tests, but after controlling for age and language status, only the relation to the WISC-MR and YARC comprehension score remained significant.

Table 6. Bivariate and partial correlations between RPVT and other test scores.

		RPVT total	BPVS	WISC- MR	YARC ability	Reading in English
RPVT total	<i>r</i>		.78**	.58**	.80**	.31**

Polysemy knowledge and reading comprehension

	<i>N</i>	112	112	107	91
BPVS	<i>r</i>	.59**	.58**	.82**	.35**
	<i>N</i>	108	112	107	91
WISC-MR	<i>r</i>	.34**	.39**	.53**	.32**
	<i>N</i>	108	108	107	91
YARC ability	<i>r</i>	.54**	.66*	.32**	.44**
	<i>N</i>	103	103	103	87
Reading in English	<i>r</i>	.10	.18	.23*	.30**
	<i>N</i>	87	87	87	83

Note: ** $p < .01$ * $p < .05$. Correlations above the diagonal are bivariate, below the diagonal are after partialling out age and language status.

We used hierarchical multiple linear regression to examine whether polysemy knowledge was a significant predictor of reading comprehension. The results of the three models are shown in Table 7. In the first step, age, language status and non-verbal intelligence (WISC-MR) were entered as predictors. This model was significant ($R^2 = .622$, $F(3,83) = 45.51$, $p < .001$). In the next step, BPVS score and time spent reading in English were entered. This predicted an additional 14.2% of the variance, which was a significant change ($R^2_{change} = 0.142$, $F_{change}(2, 81) = 24.22$, $p < .001$). In the final model, RPVT score was added. This predicted an additional 1.2% of the variance, which was a significant change ($R^2_{change} = .012$, $F_{change}(1,80) = 4.45$, $p = .038$).

Table 7. Regression analysis models predicting reading comprehension.

Model	Variable	<i>B</i>	β	<i>t</i>	<i>p</i>
1	(Constant)	-17.62		-3.06	.003
$R^2 = .621$	Age	5.70	.545	5.64	<.001
	Language status	13.29	.393	5.69	<.001
	WISC	0.92	.242	2.47	.016
2	(Constant)	-146.64		-6.82	<.001
$R^2 = .763$	Age	2.45	.234	2.58	.012
	Language status	5.43	.161	2.48	.015
	WISC	0.31	.081	0.99	.325
	Time reading in English	0.15	.169	2.90	.005

	BPVS	35.47	.527	6.13	<.001
3	(Constant)	-112.64		-4.24	<.001
	Age	2.02	.193	2.13	.037
	Language status	4.35	.129	1.97	.052
$R^2=.776$	WISC	0.22	.058	0.72	.476
	Time reading in English	0.16	.176	3.08	.003
	BPVS	25.29	.376	3.39	<.001
	RPVT	0.43	.231	2.11	.038

584

585 ***The Effects of Language Exposure on Polysemy Knowledge for EAL Children***

586 We also planned to examine whether aspects of language exposure affected RPVT performance
587 in EAL children by factors of the child's language background and the RPVT total score for
588 children with EAL. These correlations, controlling for age are reported in Appendix 5,
589 alongside correlations with the BPVS for comparison. After controlling for age, RPVT scores
590 were significantly negatively correlated with the age of onset for learning English ($r = -.53$, p
591 $= .001$), and significantly positively correlated with years spent learning English ($r = .48$, p
592 $= .004$). Longer exposure to English seems then to be related to better knowledge of polysemes
593 for children with EAL. No other variables showed a significant correlation. The pattern was
594 the same for BPVS scores.

595 **Unregistered Exploratory Analyses**

596 To further explore where differences lay between year and language groups on the RPVT, the
597 percentage of questions in which children obtained no responses correct, one correct, or both
598 correct were calculated for each group. This is of interest because getting both meanings correct
599 signifies that children recognise both the primary and secondary meanings of the word. Figure
600 2 shows the frequency of each score for the 4 subgroups. A separate 2 (year) x 2 (language
601 status) between-subjects ANCOVA, with non-verbal intelligence as the covariate, was
602 performed on each of the three scores (none correct, one correct, and both correct).

Polysemy knowledge and reading comprehension

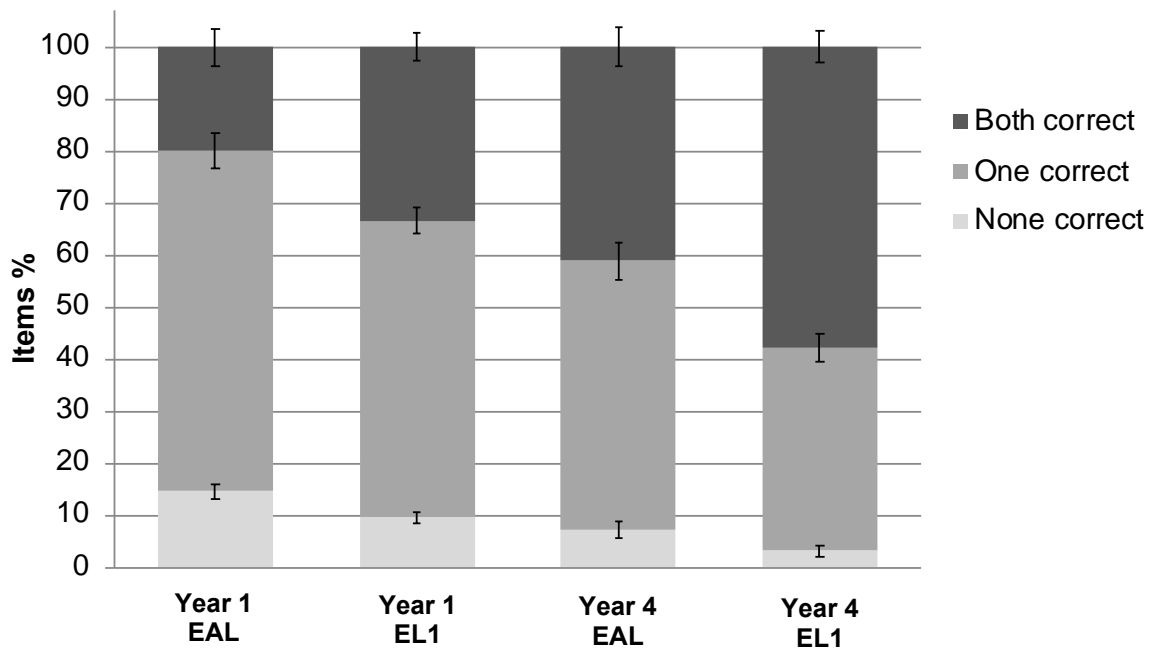


Figure 2. Percentage of items with neither, one, or both answers correct on RPVT with standard error. Note: Estimated marginal means with the covariate of WISC-MR = 12.20.

Neither correct (no knowledge). Children in Year 1 more often obtained neither word meaning correct ($M = 12.05$, $SD = 7.46$) than those in Year 4 ($M = 5.14$, $SD = 7.57$) after adjusting for non-verbal intelligence ($F(1,107) = 31.84$, $p < .001$, $\eta_p^2 = .229$) with a large effect size. Furthermore, children with EAL ($M = 10.87$, $SD = 6.47$) more often obtained neither word meaning correct than those with EL1 ($M = 6.33$, $SD = 6.44$; $F(1,107) = 15.98$, $p < .001$, $\eta_p^2 = .130$) with a medium effect size. There was no interaction between age and language status ($F = 0.31$, $p = .579$). Thus, younger children and those with EAL were more likely to get neither the primary nor secondary meaning of a word in the RPVT relative to older children and children with EL1.

One correct (partial knowledge). Children in Year 1 more often knew only one word meaning ($M = 61.24$, $SD = 16.87$) than those in Year 4 ($M = 45.33$, $SD = 17.21$) after adjusting for non-verbal intelligence ($F(1,107) = 20.10$, $p < .001$, $\eta_p^2 = .158$). Furthermore, children with EAL ($M = 58.59$, $SD = 14.66$) more often got only one meaning correct than those with EL1

($M = 47.98$, $SD = 14.58$; $F(1,107) = 13.18$, $p < .001$, $\eta_p^2 = .110$). Both effect sizes were medium. There was no interaction between age and language status ($F(1,107) = 0.01$, $p = .924$). Again, younger children and those with EAL were more likely to display partial knowledge of a word's meanings in the RPVT relative to older and children with EL1.

Both correct (primary and secondary meaning knowledge). Children in Year 1 less often demonstrated knowledge of both word meanings ($M = 26.71$, $SD = 18.20$) than those in Year 4 ($M = 49.53$, $SD = 18.56$) after adjusting for non-verbal intelligence ($F(1,107) = 33.36$, $p < .001$, $\eta_p^2 = .238$). Furthermore, children with EAL ($M = 30.54$, $SD = 15.88$) less often demonstrated knowledge of both meanings than those with EL1 ($M = 45.70$, $SD = 15.76$; $F(1,107) = 28.00$, $p < .001$, $\eta_p^2 = .207$). Both effect sizes were medium-large. There was no interaction between age and language status ($F(1,107) = 0.08$, $p = .782$). Thus, younger children and those with EAL were less likely to manifest knowledge of both meanings of a word in the RPVT relative to older and children with EL1.

Discussion

The first aim of this study was to validate the RPVT, a new receptive measure of children's understanding of the multiple meanings of words. Our results indicate that this new measure is both reliable and valid with a sample of British children between five to nine years old. The RPVT showed high reliability in terms of both internal consistency and temporal stability. It also showed convergent validity with existing vocabulary measures. Performance on the RPVT was correlated with an existing measure of children's ability to define meanings of polysemes based on their context (the LPT Multiple Meanings subtest), and with grade of acquisition norms for recognising definitions of polysemes (Brysbaert & Biemiller, 2017). This was the case despite the differences in format, response modality and cognitive demands between the tests. The RPVT also correlated with an existing receptive measure of vocabulary breadth, the

BPVS. These findings suggest that the test is a valid measure of polyseme vocabulary knowledge.

The results indicate that the RPVT is a valid and reliable measure for younger children and those with EAL. The test is suitable for children aged as young as five years, and therefore enables testing with a younger age group than was previously possible with existing measures (e.g. LPT-MM, Richard & Hanner, 1995). Unlike previous measures, it has also been demonstrated to be reliable and valid with children with English as an additional language, in a sample which represented a diverse range of first languages and levels of experience with English. Therefore, pictorial multiple-choice is an effective paradigm for measuring knowledge of polysemes with monolingual and multilingual children.

There are also some hints from the data that previous test formats may have underestimated children's knowledge of polysemy. During the first pilot, one of the items removed for being too easy for children was *fly*, for which 100% of children ($N=7$) selected both meanings. For the same item and meanings in the LP-MM, the accuracy was 32%. In addition, the younger children in our sample more often knew 2 meanings of words on the RPVT (23%) than the LPT-MM (13%), despite many items in the RPVT (e.g. *groom*, *pupil*, *mount*) being ostensibly more challenging than the LPT (e.g. *spring*, *match*, *set*). This was not the case for the older children. This suggests that the RPVT more sensitively captures younger children's developing understanding of polysemy than existing tests, perhaps due to the reduced extraneous task demands.

The second aim of the present research was to examine which factors are associated with polysemy knowledge. The first relevant factor was age: children aged eight to nine years performed better than those aged five to six years. This mirrors research showing that the ability to define polysemes improves gradually between six and twenty years of age (Dale & O'Rourke, 1981; Wiig & Secord, 1991; Richard & Hanner, 1995). The current study extends

existing knowledge by showing that children's *recognition* of meanings of polysemes improves between five and nine years, and importantly that this improvement cannot be completely accounted for by an increase in non-verbal intelligence or vocabulary breadth.

The results also showed that children with EAL had poorer knowledge of polysemes than children with EL1, with a large effect size, equivalent to 1 standard deviation. This is consistent with previous research suggesting that bilingual children have smaller general vocabularies in each of their languages than monolingual children (e.g. Farnia & Geva, 2011), and with bilingual adults suggesting that bilinguals knew fewer polysemes in one language than monolinguals (Gathercole & Moawad, 2010). This research goes further by demonstrating a specific difficulty with polysemy knowledge for children with EAL and by showing that poorer performance cannot be entirely explained by their smaller breadth of vocabulary in English. Exploratory analyses also hinted that children with EAL might find learning the secondary meanings of words particularly difficult. Children with EAL were less likely than children with EL1 to correctly identify both meanings of words in the RPVT, but more likely to identify only one meaning. This implies that children with EAL seem to be behind partly due to not knowing the secondary meanings of some words. This could suggest that children with EAL find the secondary meanings of words harder to learn than primary meanings. However, children with EAL showed a similar profile of polysemy knowledge to younger EL1 children, who also displayed specific challenges with polysemous vocabulary that could not be fully explained by their vocabulary breadth. If the challenges experienced by EAL children are typical of EL1 learners with lower general linguistic proficiency, we might consider them to be simply further behind on the same developmental trajectory – and that greater exposure to English will further their development.

In support of the benefits of English exposure to polysemy knowledge in English, the study found that children with EAL performed better on the RPVT if they had a greater duration

of exposure to English – a relationship which has previously been identified for vocabulary breadth (e.g. Paradis, 2011). However, the findings also indicate that simple exposure is not enough. The gap between children with EAL and EL1 in polysemy knowledge was consistent across age groups. Despite having on average an additional 3 years of exposure to English through school, and on average 5 years of exposure in total, EAL children in Year 4 were behind their EL1 peers to the same degree as children in Year 1. Thus, this cross-sectional sample of children with a comparable age of onset of English and age-standardised breadth of vocabulary scores suggests that children with EAL are not catching up to their peers over time in terms of their polysemy vocabulary. These findings confirm and extend previous longitudinal evidence showing that the vocabulary breadth of EAL students continue to lag behind their EL1 peers even after 6 years of schooling (Farnia & Geneva, 2011) by indicating that the same may also apply for vocabulary breadth, particularly for children with middle-to-lower socio-economic status and below average language skills (Paradis & Jia, 2017) – as was the case in our sample. Thus, whilst exposure to English through schooling contributes to polysemy knowledge, it may not be sufficient to enable EAL students to catch up with their EL1 peers. This suggests that children with EAL may need specific support to develop their polysemy knowledge, rather than relying on simple exposure.

There could be several possible explanations for why children with EAL have a specific challenge with polysemous words, and particularly their secondary meanings, compared to their EL1 peers. Children with EAL may experience interference from polysemous words in their L1 as adults do (Elston-Güttler & Williams, 2008), although given the diversity of languages used in our sample and the small number of words tested, this is unlikely to be the only explanation. It may be that acquiring secondary meanings of words is a specific skill which is more challenging for children with EAL. Learning secondary meanings of words requires children firstly to recognise that the primary meaning is not appropriate in the given context

(e.g. Doherty, 2004), and then to infer the appropriate meaning, a skill which depends on having greater reading comprehension and vocabulary knowledge (Cain, Lemmon, & Oakhill, 2004; Cain, Oakhill, & Elbro, 2003). Many children with EAL will still be developing these skills. This could also explain why younger children, whose inference skills are also less developed (Lepola, Lynch, Laakkonen, Silvén, & Niemi, 2012), struggled with polysemy knowledge in our study. It could also be that properties of polysemous words (such as the lower frequency of secondary meanings) make learning them more sensitive to levels of target language exposure, which tends to be reduced in children with EAL (e.g. Paradis & Jia, 2017). Further research would be needed to distinguish between these possible explanations for why children with EAL have reduced knowledge of words that are polysemous.

In line with the third aim of the present research, we also demonstrated that polyseme vocabulary is important for reading comprehension. Polyseme vocabulary predicted reading comprehension, and this effect remained after controlling for age, language status, non-verbal intelligence, time spent reading in English, and breadth of vocabulary. Polyseme knowledge added an additional 1.2% of variance, suggesting it makes a unique contribution to children's comprehension of what they read beyond the range of factors controlled for, including breadth of vocabulary. This is a particularly compelling finding given that the standardised measure of reading comprehension included a broad range of skills including locating factual information, defining words, and several types of inference. It is worth noting that only three of the words in the RPVT were used across the seven passages and questions in the YARC (hide, deck and, throw) and in none of these cases were the secondary meanings used. The conclusions of the current study thus extend previous research showing that adolescents who know the secondary meanings of academic words have better reading comprehension skills (Logan & Kieffer, 2017). Our findings show that knowledge of polysemes is associated with better reading comprehension using different measures and in younger children and those with EL1. There

are several possible reasons why knowledge of polysemes might relate to reading comprehension. It could be that knowing additional, less dominant meanings of words helps children to infer meaning in semantically ambiguous contexts whilst reading. However, it is also possible that the relationship is bidirectional: children who have better reading comprehension may also be more able to acquire the additional meanings of words through inference.

The findings of this study have significant methodological, theoretical, and educational implications. The study makes a major methodological contribution to the field, by presenting a novel receptive test for assessing children's understanding of the multiple meanings of words. The test ostensibly presents fewer demands on children's other language or memory abilities than previous measures (e.g. LPT MM) and is appropriate for children at least as young as age five, and those with EAL. Furthermore, the test requires no manual scoring on the part of the experimenter and so can be administered quickly and easily, and in contexts without an experimenter present (e.g. in a digital context). The test can be used to measure this important dimension of vocabulary alongside other measures of vocabulary breadth and depth to obtain a fuller picture of the child's vocabulary knowledge. Furthermore, the findings contribute to our theoretical understanding pertaining to children's knowledge of polysemes. It suggests that polysemy is an aspect of depth of vocabulary that children with EAL and younger children struggle with specifically, beyond what would be expected from their vocabulary size. In addition, it indicates that polysemy knowledge plays a role in reading comprehension. The results have further implications for practitioners in educational settings. They show that knowing the secondary meanings of words is important for reading comprehension, and therefore accessing the curriculum. Yet, these words can be more challenging to learn, and so may require specific targeting in classrooms. Interventions for vocabulary could be used to highlight the multiple senses of words and inferring the intended meaning from context. The

findings also suggest that specific provision may be needed to help EAL children to catch up with their EL1 peers in terms of polysemy knowledge.

The present research is limited in some respects. As a cross-sectional sample was used, we cannot infer developmental trends as clearly as with a longitudinal sample. Because our definition of EAL was more narrow than the definition usually used in education (Department for Education, 2020), the findings are not generalisable to all children defined by schools as having EAL (specifically, those who are bilingual with English as a first language). Using the school definition, differences between groups might not be as pronounced, because bilingual children with English as a first language likely perform better than those with English as a second language in terms of English vocabulary. Thus, interventions for polyseme vocabulary should perhaps be targeted more towards children for whom English is not a first language. Whilst we measured several factors that could explain variability in polyseme knowledge, there are likely other individual differences factors not accounted for (e.g. socioeconomic status). Features of the words themselves might also be important, as secondary meanings of polysemes may differ from words with singular meanings in terms of psycholinguistic factors such as frequency which make them more difficult to learn.

Future research should examine further which factors predict children's polyseme knowledge. For example, more studies could be conducted to measure the developmental trajectory of polysemous word knowledge longitudinally over a wider range of ages and years of exposure to English in children with EAL to determine to what extent children are able to catch up to their classmates over time. Furthermore, given the gap identified in polysemous word knowledge for EAL children, the question is raised as to how to support this group of children to learn additional word meanings. Vocabulary teaching strategies applied to single meanings of words might be extended to polysemous words, but perhaps better would be to

develop strategies which raise metalinguistic awareness of polysemy, as has been done with children in their first language (Zipke, 2011; Zipke, Ehri, & Cairns, 2009).

In conclusion, the present research demonstrates the viability of a new approach to measuring children's understanding of polysemes and the validity and reliability of the RPVT as a measure of this aspect of vocabulary for early primary-school age children. Further, it demonstrates that children with EAL struggle with this important aspect of vocabulary and do not appear to catch up with their peers through schooling. This is important because knowledge of polysemes was also shown to be significant for children's reading comprehension. Overall, these findings support the notion that polysemy is a distinct and important dimension of vocabulary for researchers to measure and for practitioners to teach.

References

- Babayiğit, S., & Stainthorp, R. (2014). Correlates of early reading comprehension skills: A componential analysis. *Educational Psychology*, 34(2), 185–207.
<https://doi.org/10.1080/01443410.2013.785045>
- Bleses, D., Makransky, G., Dale, P. S., Højen, A., & Ari, B. A. (2016). Early productive vocabulary predicts academic achievement 10 years later. *Applied Psycholinguistics*, 37(6), 1461–1476. <https://doi.org/10.1017/S0142716416000060>
- Brysbaert, M., & Biemiller, A. (2017). Test-based age-of-acquisition norms for 44 thousand English word meanings. *Behavior Research Methods*, 49(4), 1520–1523.
<https://doi.org/10.3758/s13428-016-0811-4>
- Cain, K., Lemmon, K., & Oakhill, J. (2004). Individual differences in the inference of word meanings from context: The influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of Educational Psychology*, 96(4), 671–681.
<https://doi.org/10.1037/0022-0663.96.4.671>
- Cain, K., Oakhill, J., & Elbro, C. (2003). The ability to learn new word meanings from context by school-age children with and without language comprehension difficulties. *Journal of Child Language*, 30(3), 681–694.
<https://doi.org/10.1017/s0305000903005713>
- Cameron, L., & Besser, S. (2004). *Writing in English as an additional language at Key Stage 2*. Retrieved from [https://www.naldic.org.uk/Resources/NALDIC/Research and Information/Documents/RR586.pdf](https://www.naldic.org.uk/Resources/NALDIC/Research%20and%20Information/Documents/RR586.pdf)
- Carlo, M. S., August, D., McLaughlin, B., Snow, C. E., Dressler, C., Lippman, D. N., ... White, C. E. (2004). Closing the gap: Addressing the vocabulary needs of English-language learners in bilingual and mainstream classrooms. *Reading Research Quarterly*, 39(2), 188–215. <https://doi.org/10.1598/RRQ.39.2.3>

- 828 Catts, H. W., Adlof, S. M., & Weismer, S. E. (2006). Language deficits in poor
 829 comprehenders: A case for the simple view of reading. *Journal of Speech, Language,*
 830 *and Hearing Research*, 49(2), 278–293. [https://doi.org/10.1044/1092-4388\(2006/023\)](https://doi.org/10.1044/1092-4388(2006/023))
- 831 Clemmons, K. (2008). *The problem of polysemy in the first thousand words of the general*
 832 *service list: A corpus study of secondary chemistry texts*. Retrieved from
 833 [http://login.proxy.library.vanderbilt.edu/login?url=http://search.proquest.com/docview/6](http://login.proxy.library.vanderbilt.edu/login?url=http://search.proquest.com/docview/621754893?accountid=14816%5Cnhttp://sfx.library.vanderbilt.edu/vu?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&genre=dissertations+%26+theses&sid=P)
 834 [21754893?accountid=14816%5Cnhttp://sfx.library.vanderbilt.edu/vu?url_ver=Z39.88-](http://login.proxy.library.vanderbilt.edu/login?url=http://search.proquest.com/docview/621754893?accountid=14816%5Cnhttp://sfx.library.vanderbilt.edu/vu?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&genre=dissertations+%26+theses&sid=P)
 835 [2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&genre=dissertations+%26+theses](http://login.proxy.library.vanderbilt.edu/login?url=http://search.proquest.com/docview/621754893?accountid=14816%5Cnhttp://sfx.library.vanderbilt.edu/vu?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&genre=dissertations+%26+theses&sid=P)
 836 [&sid=P](http://login.proxy.library.vanderbilt.edu/login?url=http://search.proquest.com/docview/621754893?accountid=14816%5Cnhttp://sfx.library.vanderbilt.edu/vu?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&genre=dissertations+%26+theses&sid=P)
- 837 Dale, E., & O'Rourke, J. (1981). *The living word vocabulary: A national vocabulary*
 838 *inventory*. Chicago: World Book.
- 839 Dautriche, I. (2015). *Weaving an ambiguous lexicon* (Université Paris Descartes).
 840 <https://doi.org/http://dx.doi.org/10.1016/j.cognition.2015.06.003>
- 841 Degani, T., & Tokowicz, N. (2010). Ambiguous words are harder to learn. *Bilingualism*,
 842 13(3), 299–314. <https://doi.org/10.1017/S1366728909990411>
- 843 Department For Communities and Local Government. (2019). *English indices of deprivation*.
 844 Retrieved from <http://imd-by-postcode.opendatacommunities.org/imd/2019>
- 845 Department for Education. (2012). *EAL pupils by school*. Retrieved from
 846 <https://www.naldic.org.uk/research-and-information/eal-statistics/eal-pupils/>
- 847 Department for Education. (2020). *Schools, pupils and their characteristics*. Retrieved from
 848 [https://explore-education-statistics.service.gov.uk/find-statistics/school-pupils-and-their-](https://explore-education-statistics.service.gov.uk/find-statistics/school-pupils-and-their-characteristics)
 849 [characteristics](https://explore-education-statistics.service.gov.uk/find-statistics/school-pupils-and-their-characteristics)
- 850 Doherty, M. J. (2004). Childrens difficulty in learning homonyms. *Journal of Child*
 851 *Language*, 31(1), 203–214. <https://doi.org/10.1017/S030500090300583X>
- 852 Duff, F. J., Reen, G., Plunkett, K., & Nation, K. (2015). Do infant vocabulary skills predict

- 853 school-age language and literacy outcomes? *Journal of Child Psychology and*
 854 *Psychiatry and Allied Disciplines*, 56(8), 848–856. <https://doi.org/10.1111/jcpp.12378>
- 855 Dunn, D. M., & Dunn, L. M. (2009). BPVS III. In *British Picture Vocabulary Scales* (Third
 856 Edit). London: GL Assessment.
- 857 Elston-Güttler, K. E., & Friederici, A. D. (2005). Native and L2 processing of homonyms in
 858 sentential context. *Journal of Memory and Language*, 52(2), 256–283.
 859 <https://doi.org/10.1016/j.jml.2004.11.002>
- 860 Elston-Güttler, K. E., & Williams, J. N. (2008). First language polysemy affects second
 861 language meaning interpretation : evidence for activation of first language concepts
 862 during second language reading. *Second Language Research*, 24(2), 167–187.
- 863 European Commission. (2012). *Special Eurobarometer 386: Europeans and their languages*.
 864 Retrieved from
 865 https://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs_386_en.pdf
- 866 Gathercole, V. C. M., & Moawad, R. A. (2010). Semantic interaction in early and late
 867 bilinguals: All words are not created equally. *Bilingualism*, 13(4), 385–408.
 868 <https://doi.org/10.1017/S1366728909990460>
- 869 Geva, E., & Farnia, F. (2012). Developmental changes in the nature of language proficiency
 870 and reading fluency paint a more complex view of reading comprehension in ELL and
 871 EL1. *Reading and Writing*, 25(8), 1819–1845. [https://doi.org/10.1007/s11145-011-](https://doi.org/10.1007/s11145-011-9333-8)
 872 9333-8
- 873 Geva, E., & Massey-Garrison, A. (2013). A Comparison of the Language Skills of ELLs and
 874 Monolinguals Who Are Poor Decoders, Poor Comprehenders, or Normal Readers.
 875 *Journal of Learning Disabilities*, 46(5), 387–401.
 876 <https://doi.org/10.1177/0022219412466651>
- 877 Hessel, A. K., & Murphy, V. A. (2019). Understanding how time flies and what it means to

- 878 be on cloud nine: English as an Additional Language (EAL) learners' metaphor
 879 comprehension. *Journal of Child Language*, 46(2), 265–291.
 880 <https://doi.org/10.1017/S0305000918000399>
- 881 Lepola, J., Lynch, J., Laakkonen, E., Silvén, M., & Niemi, P. (2012). The role of inference
 882 making and other language skills in the development of narrative listening
 883 comprehension in 4-6-year-old children. *Reading Research Quarterly*, 47(3), 259–282.
 884 <https://doi.org/10.1002/RRQ.020>
- 885 Lervåg, A., & Aukrust, V. G. (2010). Vocabulary knowledge is a critical determinant of the
 886 difference in reading comprehension growth between first and second language learners.
 887 *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 51(5), 612–620.
 888 <https://doi.org/10.1111/j.1469-7610.2009.02185.x>
- 889 Lesaux, N. K., Crosson, A. C., Kieffer, M. J., & Pierce, M. (2010). Uneven profiles:
 890 Language minority learners' word reading, vocabulary, and reading comprehension
 891 skills. *Journal of Applied Developmental Psychology*, 31(6), 475–483.
 892 <https://doi.org/10.1016/j.appdev.2010.09.004>
- 893 Lippeveld, M., & Oshima-Takane, Y. (2015). The Effect of Input on Children's Cross-
 894 Categorical Use of Polysemous Noun-Verb Pairs. *Language Acquisition*, 22(3), 209–
 895 239. <https://doi.org/10.1080/10489223.2014.943902>
- 896 Logan, K. J., & Kieffer, M. J. (2017). Evaluating the role of polysemous word knowledge in
 897 reading comprehension among bilingual adolescents. *Reading and Writing*, 30(8), 1687–
 898 1704. <https://doi.org/10.1007/s11145-017-9745-1>
- 899 Marulis, L. M., & Neuman, S. B. (2010). The Effects of Vocabulary Intervention on Young
 900 Children's Word Learning: A Meta-Analysis. *Review of Educational Research*, 80(3),
 901 300–335. <https://doi.org/10.3102/0034654310377087>
- 902 Murphy, V. A. (2014). Bilingual development in young children. In *Second Language*

- 903 *Learning in the Early School Years: Trends and Contexts-Oxford Applied Linguistics*
 904 (pp. 37–81). Oxford: Oxford University Press.
- 905 Nakamoto, J., Lindsey, K. A., & Manis, F. R. (2007). A longitudinal analysis of English
 906 language learners' word decoding and reading comprehension. *Reading and Writing*,
 907 20(7), 691–719. <https://doi.org/10.1007/s11145-006-9045-7>
- 908 Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*.
 909 [https://doi.org/10.1016/S0889-4906\(02\)00014-5](https://doi.org/10.1016/S0889-4906(02)00014-5)
- 910 Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills
 911 contribute to the development of reading. *Journal of Research in Reading*, 27(4), 342–
 912 356. <https://doi.org/10.1111/j.1467-9817.2004.00238.x>
- 913 Osborne, J. W. (2010). Improving your data transformations : Applying the Box-Cox
 914 transformation. *Practical Assessment, Research & Evaluation*, 15(12), 1–9.
- 915 Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word
 916 reading and reading comprehension. *Journal of Educational Psychology*, 98(3), 554–
 917 566. <https://doi.org/10.1037/0022-0663.98.3.554>
- 918 Paradis, J., & Jia, R. (2017). Bilingual children's long-term outcomes in English as a second
 919 language: language environment factors shape individual differences in catching up with
 920 monolinguals. *Developmental Science*, 20(1). <https://doi.org/10.1111/desc.12433>
- 921 Proctor, C. P., August, D., Carlo, M., & Snow, C. (2005). Native spanish-speaking children
 922 reading in english: Toward a model of comprehension. *Journal of Educational*
 923 *Psychology*, 97(2), 246–256. <https://doi.org/10.1037/0022-0663.97.2.246>
- 924 Rabagliati, H., Pylkkänen, L., & Marcus, G. F. (2013). Top-down influence in young
 925 children's linguistic ambiguity resolution. *Developmental Psychology*, 49(6), 1076–
 926 1089. <https://doi.org/10.1037/a0026918>
- 927 Richard, G. J., & Hanner, M. A. (2005). *Language processing test 3 Elementary*. Austin, TX:

- 928 Pro-Ed Inc.
- 929 Rodd, J. M., Gaskell, G., & Marslen-Wilson, W. (2002). Making sense of semantic
930 ambiguity: Semantic competition in lexical access. *Journal of Memory and Language*,
931 46(2), 245–266. <https://doi.org/10.1006/jmla.2001.2810>
- 932 Ryan, C. (2013). *American Community Survey Reports: Language Use in the United States*.
933 Retrieved from [https://www2.census.gov/library/publications/2013/acs/acs-22/acs-](https://www2.census.gov/library/publications/2013/acs/acs-22/acs-22.pdf)
934 22.pdf
- 935 Schuth, E., Köhne, J., & Weinert, S. (2017). The influence of academic vocabulary
936 knowledge on school performance. *Learning and Instruction*, 49, 157–165.
937 <https://doi.org/10.1016/j.learninstruc.2017.01.005>
- 938 Smith, S. A., & Murphy, V. A. (2015). Measuring productive elements of multi-word phrase
939 vocabulary knowledge among children with English as an additional or only language.
940 *Reading and Writing*, 28(3), 347–369. <https://doi.org/10.1007/s11145-014-9527-y>
- 941 Snowling, M. J., Stothard, S. E., Clarke, P., Bowyer-Crane, C., Harrington, A., Truelove, E.,
942 ... Hulme, C. (2009). *YARC York Assessment of Reading for Comprehension Passage*
943 *Reading*. (Second Edi). London: GL Assessment.
- 944 Srinivasan, M., & Snedeker, J. (2011). Judging a book by its cover and its contents: The
945 representation of polysemous and homophonous meanings in four-year-old children.
946 *Cognitive Psychology*, 62(4), 245–272. <https://doi.org/10.1016/j.cogpsych.2011.03.002>
- 947 Wechsler, D. (2016). *Wechsler intelligence scale for children—Fifth UK Edition*. London:
948 Harcourt Assessment.
- 949 Wiig, E. H., & Secord, W. (1991). *Test of word knowledge*. San Antonio, TX: Psychological
950 Corporation.
- 951 Zipke, M. (2011). First graders receive instruction in homonym detection and meaning
952 articulation: The effect of explicit metalinguistic awareness practice on beginning

- 953 readers. *Reading Psychology*, 32(4), 349–371.
- 954 <https://doi.org/10.1080/02702711.2010.495312>
- 955 Zipke, M., Ehri, L. C., & Cairns, H. S. (2009). Using Semantic Ambiguity Instruction to
- 956 Improve Third Graders' Metalinguistic Awareness and Reading Comprehension: An
- 957 Experimental Study. *Reading Research Quarterly*, 44(3), 300–321.
- 958 <https://doi.org/10.1598/RRQ.44.3.4>
- 959

960 **Appendix 1**961 **Language Exposure Descriptives**

	EAL			EL1		
	N	M	SD	N	M	SD
Reading in English (mins per day)	30	25.86	15.98	61	31.13	18.2
Media use in English (mins per day)	29	36.63	22.42	62	38.14	22.6
Dominance	39	8.97	3.38	23	7.087	2.59
Other language frequency	34	4.91	0.29	25	4.72	0.46
Listening skill (other language)	38	3.39	0.86	24	3.29	0.81
Speaking skill (other language)	38	3.11	1.06	24	2.5	1.02
Reading skill (other language)	37	1.84	0.93	23	1.61	0.89
Writing skill (other language)	37	1.76	0.96	23	1.39	0.66
Reading in other language (mins per day)	24	7.29	14.14	13	3.46	11.1
Media use in other language (mins per day)	25	21.92	19.2	12	8.13	13.5

962 Note: Aside from Reading in English and Media use in English, all variables reported for
 963 bilingual participants only. Media use in English & Media other language are averages per
 964 media reported by the parent (i.e. videos, music, websites, and games).

965 **Appendix 2**966 **Prompt Words Used in Polysemy Test with Definitions**

Item	Definition meaning 1	Definition meaning 2
box	(n) a (usually rectangular) container	(v) engage in a boxing match
cold*	(adj) having a low temperature	(n) an illness causing coughing and sneezing
crane	(n) lifts and moves heavy objects	(n) large long-necked wading bird
cycle	(v) ride a bicycle	(n) an interval during which a recurring sequence of events occurs
deck	(n) platforms built into a ship	(n) a pack of 52 playing cards
degree	(n) a unit of temperature	(n) an award conferred by a university
duck	(n) a water bird with short legs, a short neck, and a large flat beak	(v) to move (the head or body) quickly downwards or away
file	(n) a set of records (written or electronic) kept together	(n) a tool used for smoothing nails
fire	(n) the event of something burning	(v) cause a gun to go off
groom	(v) give a neat appearance to	(n) a male participant at his own wedding
hide	(v) prevent from being seen or discovered	(n) the dressed skin of an animal
horn	(n) an instrument which makes a loud noise	(n) bony outgrowths on the heads of animals
iron	(v) press and smooth with a heated iron	(n) a heavy metallic element
kid	(n) a young person of either sex	(n) young goat
litter	(n) rubbish carelessly dropped or left about	(n) the multiple offspring of one birth
mount	(v) get up on the back of (a horse)	(v) to fix an object firmly on something
nail	(n) covering at the end of fingers and toes	(n) a pointed piece of metal used as a fastener
order	(n) arranged so one thing follows another	(v) ask, command
pen	(n) a writing implement	(n) an enclosure for confining livestock
play*	(v) be engaged in playful activity	(n) a theatrical performance of a drama
pupil	(n) student	(n) the aperture in the centre of the eye
rare	(adj) especially valued for its uncommonness	(adj) (of meat) cooked quickly; still red inside
ring	(n) jewellery made of a circlet of precious metal	(v) sound loudly and sonorously
row	(v) propel a boat with oars	(n) an arrangement of objects in a line
scale	(n) flat pieces of skin that cover fish or reptiles	(n) a measuring instrument for weighing
school	(n) a building where children receive education	(n) a large group of fish
sign	(n) a public display of a message	(v) mark with one's signature
staff	(n) workers or employees	(n) a rod or stick
tank	(n) a large vessel for holding gases or liquids	(n) an enclosed armoured military vehicle
throw	(v) move your arm quickly to propel an object	(n) a light rug or cover for a sofa or bed
train	(n) transport made of carriages that run on a railway	(v) to teach; coach; educate
wave	(n) ridges of water that move across the sea	(v) signal with the hands
watch*	(n) a timepiece worn on the wrist	(v) observe

967 * Practice and example items.

Appendix 3

Language Background Survey for Parents

What is/are your child's first language(s)? ☐ English ☐ Other – please specify: _____

2. Does your child use any other language(s)? **Yes / No**

2a. **If Yes**, which other language(s)? _____

3. At what age did your child begin learning English? _____ years (If from birth, write 0.)

4. **If your child uses language(s) other than English**, how often do they use their strongest other language? (Circle one)

Daily	2-6 times a week	Once a week	1-3 times a month	Less than once a month
-------	------------------	-------------	-------------------	------------------------

5. **If your child uses language(s) other than English**, please rate your child's ability in listening, speaking, reading and writing, in their strongest language that is not English. (Tick ✓ one in each row.)

	Poor	Okay	Good	Excellent
Listening	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Speaking	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Reading	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Writing	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

6. Please indicate how many minutes your child spends on the following activities in an average day, in English and (if applicable) their strongest language that is not English.

	In English	In <u>strongest</u> other language (if applicable)
Reading	minutes per day	minutes per day
Watching programs or videos	minutes per day	minutes per day
Listening to music or radio	minutes per day	minutes per day
Using websites	minutes per day	minutes per day
Playing games	minutes per day	minutes per day

988

Appendix 4**989 Mean and Standard Deviations for Number of Each Error Type on the RPVT**

Year	Language status	Meaning A semantic		Meaning B semantic		Phonological		Other	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Year 1	EAL	11.75	3.84	6.95	2.52	7.30	2.85	3.55	2.16
	EL1	11.12	3.77	5.59	2.41	5.36	2.48	2.68	2.27
	Total	11.33	3.77	6.03	2.51	6.00	2.74	2.97	2.25
Year 4	EAL	6.33	3.95	3.90	1.86	5.09	2.36	1.62	1.02
	EL1	5.10	3.27	2.86	2.03	3.20	2.28	0.97	0.96
	Total	5.61	3.58	3.29	2.01	3.98	2.48	1.24	1.03
Full sample		8.72	4.65	4.79	2.67	5.08	2.80	2.18	2.00

990 Note: Maximum = 30, Chance = 10.

Appendix 5

Results of ANCOVA Analyses for Error Type on the RPVT.

	Main effect: age			Main effect: language status			Interaction: age * language status		
	<i>F</i>	<i>p</i>	η_p^2	<i>F</i>	<i>p</i>	η_p^2	<i>F</i>	<i>p</i>	η_p^2
Meaning A semantic	19.24	<.001	.152	4.39	.038	.039	0.46	.500	.004
Meaning B semantic	18.65	<.001	.148	8.82	.004	.076	0.08	.785	.001
Phonological	8.27	<.001	.072	16.56	.000	.134	0.00	.972	.000
Other	4.44	.037	.040	9.20	.003	.079	0.01	.942	.000

994 **Appendix 6**995 **Correlations Between Language Exposure and RPVT and BPVS Scores for Children**996 **with EAL**

		Age Eng.	Years Eng.	Freq. Other	Ability Other	Domin -ance	Reading Eng.	Reading Other	Media Use Eng.	Media Use Other
RPVT	<i>r</i>	-.53**	.48**	.24	.06	-.09	-.16	-.35	-.19	-.30
	<i>df</i>	32	32	31	35	36	26	21	26	21
BPVS	<i>r</i>	-.48**	.47**	.33	.03	-.10	-.09	-.40	-.07	-.28
	<i>df</i>	32	32	31	35	36	26	21	26	21

997 Note: Partial correlations controlling for age. **p<.01 *p<.05. Eng .= English, Freq. =

998 Frequency.