Title: The effect of childhood multilingualism and bilectalism on implicature understanding

Short title: Multilingualism, bilectalism, and pragmatics

Authors: Kyriakos Antoniou[1, 2] and Napoleon Katsos[2]

[1]Département de Langues et Lettres/LaDisco, Université libre de Bruxelles

[2]Department of Theoretical and Applied Linguistics, University of Cambridge

Abstract

The present study compares the performance of multilingual children speaking Cypriot Greek, Standard Modern Greek, and English (and sometimes an additional language), bilectal children speakers of Cypriot Greek and Standard Modern Greek, and Standard Modern Greek-speaking monolingual children on a task that measures the comprehension of different types of implicature. Despite lower scores in language ability in the target language, multilingual and bilectal children performed at rates comparable to the monolinguals with implicature. Regression analyses indicated a positive correlation between implicature, language proficiency, and age (but not executive control), albeit language ability did not affect implicature within multilinguals. We suggest an interpretation according to which multilingual, bilectal, and monolingual children maintain a comparable level of implicature understanding, but they do so by relying on different resources. Finally, a principal component analysis on different implicature types revealed a single factor of implicature performance. This outcome has implications for pragmatic theory.

1.1 Introduction

Previous research has systematically explored the cognitive correlates of childhood bilingualism with many studies reporting positive effects on aspects of non-verbal cognitive functioning and negative effects on measures of language proficiency (e.g. Ahtar & Menjivar, 2012; Barac, Bialystok, Castro, & Sanchez, 2014; Bialystok, Craik, Green, & Gollan, 2009; Bialystok, Luk, Peets, & Yang, 2010; Oller & Elliers, 2002). Regarding language, research comparing bilingual and monolingual children has shown delays for bilingual children in vocabulary development – when each of their languages is considered separately- and in the acquisition of certain complex aspects of morpho-syntax (e.g. Bialystok et al., 2010; Nicoladis, Palnmer, & Marentette, 2007; Oller & Elliers, 2002; Oller, Pearson, & Cobo-Lewis, 2007; Paradis, 2010).

At the same time, many studies have reported a beneficial effect of bilingualism on executive control (henceforth, *EC*) (see e.g. Adesope, Lavin, Thompson, & Ungerleider, 2010; Akhtar & Menjivar, 2012; Barac et al., 2014; De Abreu, Cruz-Santos, Tourinho, Martin, & Bialystok, 2012). EC is a domain-general cognitive system which, according to an influential account proposed by Miyake, Friedman, Emerson, Witzki, Howerter, and Wager (2000), comprises three major processes: switching (the ability to flexibly switch attention between rules, representations, or perspectives), working memory (the ability to simultaneously maintain and manipulate task-relevant information in mind), and inhibition (the ability to suppress dominant, automatic, prepotent responses and resolve conflict by suppressing irrelevant information).

It should be noted, however, that there has been a recent spread of studies that failed to replicate the bilingual EC advantage, casting doubt on the very validity of the finding (see e.g. Antón, Duñabeitia, Estévez, Hernández, Castillo, et al., 2014; Duñabeitia, Hernández, Antón,

Macizo, Estévez, et al., 2014; Gathercole, Thomas, Jones, Guasch, Young, & Hughes, 2014; Klein, 2015; Paap, Johnson, & Sawi, 2014). It is yet not clear, though, why the effect is (or is not) found in some studies (see, among others, Bialystok, Craik, & Luk, 2012; Bialystok, Poarch, Luo, & Craik, 2014; Blom, Küntay, Messer, Verhagen, & Leseman, 2014; Green & Abutalebi, 2013; Kroll & Bialystok, 2013; Paap, 2014; Paap, Johnson, & Sawi, 2015; Paap & Sawi, 2014; Valian, 2014).

Although language and executive functions have been in the focus of research on bilingualism for many years, other areas of bilingual cognitive development, particularly pragmatic-communicative skills, have so far received little attention. Furthermore, most of previous research has focused on bilingual children and little is known about how bi-dialectalism or, rather, (to use the term introduced by Rowe & Grohmann, 2013) *bilectalism* affects children's cognitive skills. Bilectal speakers –that is, speakers of two linguistic varieties which are typologically close, show a high degree of structural and lexical similarity, and are considered to be dialects of the same language- remain a vastly under-researched population, even though bilectalism is prominent worldwide, and particularly so in Europe (see e.g. Auer, 2005).

The aim of this study is threefold: the first goal is to establish whether multilingualism confers an advantage in a specific aspect of pragmatics, the ability to understand pragmatically implied meanings (henceforth, *implicatures*). Second, we investigate the effect of bilectalism on this pragmatic-communicative skill by focusing on bilectal children speaking Cypriot Greek (henceforth, *CG*) and Standard Modern Greek (henceforth, *SMG*) in the diglossic sociolinguistic landscape of Greek-speaking Cyprus (see Appendix A for a description of the sociolinguistic situation in Greek-speaking Cyprus). Third, we aim to examine the linguistic and non-linguistic cognitive factors that underpin implicature understanding in children.

In the next sections, we sketch the theoretical framework on which our investigation of children's pragmatic-communicative development is based -Grice's theory of implicature- (section 1.2), we discuss studies that previously examined aspects of bilingual children's pragmatic-communicative competence and other socio-cognitive skills (1.3), and examine theoretical and empirical reasons that suggest a link between implicature understanding and EC (1.4).

1.2 Grice's theory of implicature

Philosopher Paul Grice (1975; 1989) argued for a division between the contents of an utterance which depend on the literal meaning of words and the structural relationships between them (syntax and semantics), and the aspects of an utterance that depend on context-sensitive inferential processes about speakers' intentions (pragmatics).

Grice (1975) suggested that a good deal of what is communicated in everyday conversation relies on the appreciation of certain conversational expectations. According to his account, speakers design their utterances with respect to the cooperative principle and maxims and listeners expect their interlocutors to adhere to these conversational principles. Grice's (1975, p. 45) maxims enjoin communicators to be no less and no more informative than is required for the purpose of the talk exchange (maxims of quantity I and II), to tell the truth and avoid statements for which they have not adequate evidence (maxim of quality), to be relevant (maxim of relation), and to be brief, orderly and avoid ambiguity and obscurity (maxim of manner) (Grice (1975), pp. 44-45).

In many communicative situations consideration of the maxims prompts interlocutors to draw inferences through which they attribute to speakers an implicit meaning, which goes

beyond what they literally said. These inferences are what Grice (1975, pp. 45) called *conversational implicatures*. For example, consider the following mini-discourse:

(1)   A: Did all of his students fail the exam?

         B: Some of his students failed the exam.

(2)   Not all of his students failed the exam

B's utterance in (1) will be understood as implying (2) even though this has not been stated explicitly. This implied proposition is an inference known as a *scalar implicature* (henceforth, *SI*). According to Grice's account, this implicature is derived through a reasoning process about the speaker's intentions, which involves taking into account a rich array of information: (a) what the speaker explicitly said, (b) the linguistic and non-linguistic context, (c) the assumption that B is cooperative, (d) sensitivity to the maxim of quantity I – that is, sensitivity to the fact that there exists a more informative proposition with the term *all* that could have been used but wasn't, (e) the assumption that the speaker is knowledgeable of the situation and that s/he would assert the more informative proposition with *all* if s/he knew it to be true, and (g) the assumption that all the above information is available to both interlocutors and that both interlocutors assume this to be the case.

     Grice (1975, p. 56) also introduced a distinction between particularised and generalised conversational implicatures. For instance, given a different question *Is he a promising teacher?*, B's answer in (1) would still imply (2), but will also imply the proposition in (3) or a related one. The implied proposition in (2) represents a case of a *generalised conversational implicature* (henceforth, *GCI*) and (3) is a case of a *particularised conversational implicature*

(henceforth, *PCI*).

(3) It is not certain that he is a promising teacher.

GCIs are associated with a specific form of words and seem to be stable across contexts. PCIs, on the other hand, seem to be more context dependent and are not associated with specific linguistic items. The example above illustrates the difference between the two types of implicature: while the GCI in (2) arises in both linguistic contexts, the PCI in (3) emerges only in the second as a result of the specific question asked.

For Grice (1989) and other so-called contextual theories of implicature (e.g. Geurts, 2010; Sperber & Wilson, 1986/1995), the generalised-particularised distinction is not necessarily of particular theoretical importance. Other theories of implicature (Chierchia, 2006; 2009; Levinson, 2000), however, suggest that GCIs are distinguished from PCIs in that the former arise by default mechanisms, which are distinct from the processes involved in PCIs. According to Levinson's (2000) view, GCIs are part of the lexical meaning of words and are generated automatically without reference to the speaker's intentions or the communicative context. Chierchia's (2006; 2009) proposal, on the other hand, suggests that GCIs are derived by the grammar, through a covert focus operator **O** which has similar properties to the word *only*. This silent grammatical operator **O** takes scope over the scalar term *some* and this leads to the negation of the alternative proposition with *all* and, hence, to the computation of the SI.

1.3 Bilingualism and pragmatic-communicative skills

Three studies conducted by Michael Siegal and his colleagues (2007; 2009; 2010) explored the

extent to which bilingualism confers an advantage on preschool children's competence with pragmatics. Siegal, Matsuo, Pond, and Otsu (2007) sought to determine whether four- to six-year-old English-Japanese bilingual children would outperform their monolingual (English or Japanese) peers in their ability to understand SIs. They used a task where a puppet described a situation and the children were asked to judge whether the puppet could have described the event better. In critical trials, the descriptions were under-informative utterances (e.g. *Teddy put some of the hoops on the pole*, as a description of a situation where Teddy had in fact put every hoop on the pole) and rejection of these utterances was taken as evidence that the children had generated a SI. The results of the Siegal et al. (2007) study showed that, despite their lower vocabulary scores in the language of testing, bilingual children were more advanced than their monolingual peers in rejecting the critical utterances and, hence, in computing SIs.

In two subsequent studies, Siegal, Iozzi, and Surian (2009) and Siegal, Surian, Matsuo, Geraci, Iozzi, Okumura, and Itakura (2010) employed a Conversational Violations Test (henceforth, *CVT*) in order to examine whether preschool-aged bilingual children also excel in their ability to detect utterances that violate Grice's conversational maxims (maxims of quantity I and II, and maxims of quality, relation, manner, and politeness[1]). In the CVT, children were presented with two doll speakers who answered a question. One answer violated a conversational maxim while the other was pragmatically appropriate. For example, in one of the items for the maxim of quality, the question was *Have you seen my dog?*, the pragmatically inappropriate answer was *Yes, he's in the sky*, and the pragmatically felicitous answer was *Yes, he's in the garden*. Children were asked to indicate the doll that said something silly or rude. Siegal et al. (2009, experiment 2) found that, although comparatively delayed in their vocabulary in the language of testing, bilingual children performed significantly better than their monolingual counterparts in the CVT. Similar results were obtained in their 2010 study.

Siegal et al. (2009) propose two possible explanations for the observed bilingual advantages in their experiments. According to the first explanation, bilingual children's precocious pragmatic development is linked to their enhanced EC skills, which allow them to "consider simultaneously the appropriateness of alternative responses to questions" (Siegal et al. (2009, p. 121)). The second explanation is that bilingual children develop heightened pragmatic abilities as compensation for their weaker knowledge of core language.

Other studies have examined additional aspects of bilingual children's wider pragmatic-communicative and socio-cognitive development in comparison to monolinguals. Genesee, Tucker, and Lambert (1975) reported that very young bilinguals were more sensitive to their listeners' communication needs when providing verbal information. Two studies by Yow and Markman (2011a; 2011b) showed that preschool-aged bilingual children were better able to understand referential gestures (such as eye gaze) in order to infer a person's referential intent and to use paralinguistic cues to judge the speaker's emotional state. Both of these advantages were evident in situations where conflicting information was available. Finally, a few studies have compared bilingual and monolingual children's performance in various Theory of Mind tasks (see Akhtar & Menjivar (2012) and Barac et al. (2014) for reviews). All these have consistently reported a bilingual advantage. Once more, advanced bilingual performance in Theory of Mind (henceforth, *ToM*) tasks has been attributed to bilingual children's enhanced EC skills by some researchers (Bialystok & Senman, 2004; Kovacs, 2008).

1.4 The relation between implicature understanding and executive control

Several theoretical and empirical factors suggest that EC might be positively associated to implicature understanding. First, a processing instantiation of Grice's account of implicature

(as outlined in section 1.2) suggests that implicature understanding is a process that involves taking into account information from various linguistic and other contextual sources. This requirement to coordinate linguistic and non-linguistic information while interpreting language raises the likelihood that understanding implicatures is an EC demanding process (see Bergen and Grodner (2012), Breheny, Katsos, and Williams (2006), Breheny, Ferguson, and Katsos (2013b), Grodner and Sedivy (2005), Huang and Snedeker (2009a), Tomlinson, Bailey, and Bott (2013) for some experimental evidence that, in deriving implicatures, interlocutors indeed consider some of the types of information proposed by Grice).

Similarly, according to Relevance theory (Sperber & Wilson, 1986/1995), the interpretation of implicatures involves extra cognitive effort (as opposed to literal meaning). Sperber and Wilson (1986/1995) do not explicitly characterise the specific cognitive-psychological nature of this additional cognitive effort. However, several researchers have interpreted it either in terms of extra processing time (e.g. Bott & Noveck, 2004; Breheny, Katsos, & Williams, 2006) or in terms of employing additional cognitive resources (such as working memory) (e.g. De Neys & Schaeken (2007); see also Huang and Snedeker, (2009b) and Siegal and Surian, (2007), who explicitly suggest a link between executive functions and the ability to generate implicatures in children).

Second, several experimental investigations with adults have documented that the time course of SIs is associated with an additional processing cost relative to conventional meaning (see e.g. Bergen & Grodner (2012), Bott & Noveck (2004), Bott, Bailey, & Grodner (2011), Breheny, Katsos, & Williams (2006), Huang & Snedeker (2009a), Panizza, Chierchia, & Clifton (2009), Tomlinson, Bailey, & Bott (2013); but see Breheny, Ferguson, & Katsos (2013a); Degen & Tanenhaus, (2015); Grodner, Klein, Carbary, & Tanenhaus (2010)). This, in turn, suggests that computing SIs and, perhaps, implicatures in general is a non-automatic controlled process,

which possibly relies on EC resources. Third, empirical data in studies with adults documents that comprehending SIs specifically involves working memory (henceforth, *WM*) resources. De Neys and Schaeken (2007), Dieussaert, Verkerk, Gillard, and Schaeken (2011), and Marty and Chemla (2013) have shown that burdening adults' WM with a secondary task before judging under-informative utterances leads to a decrease in their SI responses.

It should be noted, however, that the three experimental studies by Siegal et al. (2007; 2009; 2010) directly tested for a link between pragmatic ability and EC in children, but failed to find any supporting evidence for a positive relation. Some methodological concerns, though, related to both the pragmatic and the EC tasks used by Siegal and colleagues do not allow for strong conclusions to be extracted. First, the three studies were limited in that they employed only two measures of EC (the Card Sort task and the Day/Night task) and did not include any measures of WM. Second, the selection of the EC tasks in their studies was not eventually felicitous in several respects. For the Card Sort task, Siegal and colleagues (2007; 2009) reported a near ceiling level of performance in both studies in which this test was used. The Day/Night task, on the other hand, is generally considered to measure a type of inhibition (response inhibition) for which bilingual advantages are often not reported (Carlson & Meltzoff, 2008; Martin-Rhee & Bialystok, 2008). Indeed, the choice of these specific tasks in Siegal et al.'s (2007; 2009; 2010) studies might have actually masked possible EC differences between the language groups and potential relations between pragmatic ability and EC.

Finally, the CVT employed in the Siegal et al. (2009; 2010) studies, is a test of fairly simple pragmatic abilities –it only required children to detect whether an utterance violated a conversational maxim or not. In this respect, it is likely that it didn't pose any significant burden on children's EC resources. Implicature understanding, on the other hand, might be considered

a more difficult, complex, and resource-demanding pragmatic ability which involves, but is not exhausted by, the skill to detect maxim violations (see section 1.2).

1.5 The present study

In the following study a multilingual, a bilectal, and a monolingual group of Greek-speaking children were administered a novel test on the ability to understand several types of implicature. Various other EC and vocabulary tests were administered as part of another study (Antoniou, Grohmann, Kambanaros, & Katsos, 2016).

For the pragmatics task, each implicature type was based on one of four of Grice's (1975; 1989) conversational maxims (relevance, quality, manner, and quantity I). Specifically, the task included one sub-test on relevance implicatures (testing whether children could comprehend e.g. the utterance *You are ill* as meaning *No, you can't buy an ice-cream because you are ill*), one on quality implicatures (metaphors) (testing whether children could understand e.g. the metaphorical utterance *George's father was a melting snowman* as meaning that *George's father was feeling sad*), one on manner implicatures (examining whether children could make the inference from e.g. *The man made the door open* to *The man opened the door in an atypical way*), and two sub-tests on quantity I (scalar) implicatures (examining whether children could make the inference from e.g. *There are stars on some of the cards* to *There are stars on not all of the cards*). In sum, there were five implicature sub-tests in the pragmatics task with three critical items each (where the generation of an implicature was required for accurate responding).

The implicatures included in the test can be further categorised in terms of the generalised-particularised distinction. Specifically, for the researchers who subscribe to this distinction, the

manner and quantity-I (scalar) implicatures can be considered cases of the generalised type, while the relevance and quality (metaphor) implicatures are exemplars of the particularised type.

The comparison between multilingual and monolingual children will provide evidence regarding the effect of mutilingualism on implicature understanding. We also decided to include a bilectal group of children and contrast their performance to that of monolinguals for various reasons. First, this comparison provides the opportunity to understand the cognitive effects of childhood bilectalism *per se.* To date, there is little research on bilectal children's linguistic and non-linguistic cognitive skills (at least from a comparative perspective that contrasts their performance to that of monolinguals and bilinguals) (but see e.g. Antoniou et al., 2016; Kambanaros, Grohmann, Michaelides, & Theodorou, 2012; Theodorou, Kambanaros, & Grohmann, 2013) and, to our knowledge, there is no work on bilectal children's pragmatic-communicative abilities. Second, the linguistic profile of bilectals as speakers of two different but minimally distant linguistic varieties allows to examine the issue of close language similarity (between the language pairs spoken by bilinguals) and how it modulates the outcomes of bilingualism. Third, it enables to examine the interplay between language proficiency and implicature understanding. It has been argued, for instance, that bilingual children learning two very similar languages may show less of a (or no) delay in their vocabulary skills because their languages likely share many cognates (e.g. Bialystok & Feng, 2011). To the extent that this prediction is met in our study, the comparison between bilectal and monolingual children provides the opportunity to examine implicature comprehension in a group of children who have better-developed language skills than bilingual children. Finally, the contrast between bilectals and monolinguals controls for various background factors that are often confounded with bilingualism: immigration status (the bilectal and monolingual groups came from local indigenous families in Cyprus and Greece, respectively; in contrast, the multilingual group was

13

more heterogeneous with many children having at least one immigrant parent), ethnicity and culture (Greek Cypriots and Hellenic Greeks differ minimally in these respects; multilinguals, on the other hand, were exposed to two cultures and many of them had a parent of a different nationality), and language of education (both bilectal and monolingual children were educated entirely in SMG; multilinguals, on the other hand, were educated in English with only a few hours of instruction in SMG).

Bilectal children were also tested in both their dialects (CG and SMG). This allowed to examine how their pragmatic performance compared to that of monolinguals in both their dominant, native variety (CG) and their second, non-native dialect (SMG).

Furthermore, the study offers an opportunity for a more careful investigation of the relation between implicature understanding and EC. First, as described in Antoniou et al. (2016), a wider range of EC tasks was administered measuring all components of EC. Moreover, Antoniou et al. (2016) reported that all EC measures revealed adequate variability in children's performance. Third, they found that multilingual and bilectal children outperformed their monolingual peers in EC performance after accounting for their lower language abilities in Greek. Fourth, the pragmatics test in the current study focuses on a pragmatic-communicative ability -implicature understanding- that from both a theoretical and an experimental perspective has been associated to EC skills.

Finally, this study may also provide data with respect to the psychological validity of the distinction between GCIs and PCIs.

2 Method

2.1 Participants

Participants included 64 bilectal children (speakers of CG and SMG; 32 boys and 32 girls aged 4;5–12;2, mean age 7;8, SD 1;6 years), 47 multilingual children (bilectal in CG and SMG, also speakers of English and in some cases an additional language; 24 boys and 23 girls; ages 5;0–11;5, mean age 7;8, SD 1;8 years), and 25 monolingual children (speakers of SMG only; 15 boys and 10 girls aged 6;2–9, mean age 7;4, SD 0;9 years). Four additional children were excluded from the bilectal group because they were exposed to a second language apart from the two dialects of Greek at home.

Multilingual children attended English-instruction private schools in the Republic of Cyprus. In these schools, they were taught SMG as a separate subject for six hours per week. Bilectal children were recruited from three different schools (two private, one public) where SMG was the language of instruction.   None of the children included in the analyses had any exposure to or use of a language other than the two Greek dialects at home.

All the monolingual children were recruited from a private primary school in Athens, Greece. They were instructed exclusively in SMG and their parents reported speaking only SMG to them at home.

2.2 Materials and procedure

All children were tested in two sessions taking approximately 50-60 minutes each. Seventeen of the bilectal children received additional tests in a third session, which took place at least a week after the second session.

In the first session, all children were administered the implicatures test (bilectals and multilinguals in CG and monolinguals in SMG), the Word Finding expressive Vocabulary Test-Greek, the WASI matrix reasoning test, the Simon task, and the Backward Digit Span Task. The

second session included the Colour-Shape task, Corsi Blocks test, Soccer task, and the Peabody Picture receptive Vocabulary Test-Greek (henceforth, *PPVT-Greek*). The PPVT-Greek was given only to the monolingual children and a group of 17 bilectal children and was administered in SMG to the former group and in CG to the latter group. The third session, taken only by the same subset of 17 bilectal children who also received the PPVT-Greek, included the implicatures test and the PPVT, both administered in SMG. Thus, 17 bilectals took the implicatures test and the PPVT in both their dialects.

From the EC tasks, the following measures were used (the task from which each measure was taken is given in parentheses): Simon effect (Simon task), switch cost (Colour-Shape task), Stop-Signal Reaction Time (Soccer task), and number of correctly recalled trials (in the Backward Digit Span Task, the forward, and backward versions of the Corsi Blocks task) (see Antoniou et al. (2016) for a detailed description of the EC tests and how these measures were calculated).

Implicatures test

This was a novel task testing children's comprehension of four types of implicatures: quantity I (scalar), relevance, manner, and quality (metaphors) implicatures. The quality, manner, and relevance implicature sub-tests had a picture-selection format where children had to respond by choosing one of two pictures. For SIs, two sub-tests were used -an action-based task (where children had to make a display match a target utterance) and a binary judgment task (where they had to judge whether an utterance was a correct or an incorrect description of a picture). In total, there were 15 implicature items (three per implicature sub-test), 48 filler items, and one practice item.

As already noted above, there were two language versions (CG and SMG) of the implicatures test. The same items were used in each language version, but recorded in a different dialect (CG or SMG) by native speakers. All multilingual and bilectal children took the test in CG, and monolinguals in SMG. A subset of 17 bilectal children received the test in both CG and SMG. Multilingual and bilectal children were tested in CG because this was the variety spoken in the community. This ensured that the children had adequate exposure in and knowledge of the language of testing (multilingual children had minimal exposure to SMG, mainly through a few hours of instruction at school and possibly through other sources, such as the media).

Moreover, there were two task versions of the pragmatics test. Each child was randomly given one of the two task versions so that an approximately equal number of children (overall and across language groups) was tested in each version. For the sub-tests that had a picture-selection design, two task versions were created in order to counterbalance various factors across the two versions - attractiveness of pictures provided as potential responses (picture one or two), target emotion expressed by the metaphors (sadness or anger), polarity of response in relevance implicatures (yes or no), markedness in manner implicatures (marked or unmarked) (see description of each sub-test below for more details). This ensured that level of performance in these tasks would reflect children's pragmatic understanding of the target utterances and not irrelevant (to pragmatics) factors (as listed above). For the two SI sub-tests, the same items were used in both task versions.

As described in Antoniou et al. (2016), children's performance in the 48 filler items of the implicatures test was considered a measure of their language comprehension skills in the language of testing (*language comprehension score*). A total comprehension score was calculated for each child by transforming the child's scores in the filler items of each implicature

sub-test into z scores and then averaging the three z scores.

All implicature parts, except for the binary judgment task on SIs, were designed and administered using Microsoft PowerPoint software. The binary judgment task on SIs was implemented using the E-Prime psychology software. All implicature sub-tests are described in detail in the following sections. The complete list of linguistic material used in each part of the implicatures test can be found in Appendix B.

Practice and relevance implicature items

For the practice and relevance implicature items, children were instructed that they would hear stories about a young male character named George and his mother, and that at the end of each story they had to point to a picture that showed how the story ended. All items were based on a previous study with children conducted by Bernicot, Laval, and Chaminaud (2007). A practice item did not require the generation of an implicature for accurate performance.

Each item was composed of two slides. In the first slide, children were shown a picture that visually established the story's interaction setting, and the target story was heard. Target stories in the three critical items were of the following format: George asked his mother a question (e.g. *Mom, can I buy an ice cream?*) and his mother replied with an utterance that implied either a negative or a positive answer (*You are ill* or *I've got money in my wallet*, respectively). The same was true for the practice item but in this case the reply was either explicitly negative or explicitly positive. In the second slide, the experimenter asked *What happened at the end of the story*? and introduced two pictures as possible endings (e.g. by saying *George bought an ice cream* or *George did not buy an ice cream* and pointing to the relevant pictures). One of the pictures depicted a situation compatible with a positive response (e.g. George eating an ice

cream) and the other showed a situation compatible with a negative response (e.g. George doing something else).

There were two task versions for each story. One version of the story ended with an utterance that implied a negative response (or that explicitly stated a negative response for the practice item) and the other with an utterance that implied a positive response.

Quality implicatures (metaphors)

The sub-test was designed based on a previous study by Waggoner and Palermo (1989). Children were told that they would hear stories about George and his father and that at the end of each story they should point to a picture that showed how George's father felt.

They heard three stories ending in metaphors describing either the emotion of sadness (e.g. *When George's father returned home, he was a melting snowman*) or anger (e.g. *When George's father returned home, he was a thundering cannon*). Again, each critical trial was composed of two slides. In the first slide, children were presented with a picture relevant to the story and the target story was heard. In the second slide, the experimenter asked *How did George's father feel at the end of the story?* and presented two pictures -a picture of a sad man and a picture of an angry man. All metaphorical expressions were embedded in contexts that introduced the two emotions, but did not give away which of the two emotions was expressed by the metaphorical sentence. This was achieved by including a text such as the following in all story contexts (where X was a reason for feeling sad and Y was a reason for feeling angry): *He [George] wondered how his father would feel when coming back home. He didn't know whether his father would feel sad because X or whether he would feel angry because Y.*

Finally, only novel and apt metaphors were used. Fifteen metaphors generated by the

authors were rated by 28 adult native speakers of CG and 36 native speakers of SMG on a five-point scale. All metaphors included in the test received a mean rating of over 2.5 out of 5 for both novelty and aptness and by both Greek Cypriots and Hellenic Greeks.

Manner implicatures

The general design of this sub-test was a sentence-to-picture matching task. Participants were informed that they would hear George describing a picture from a book and that they had to point to a picture that matched George's description. There were six items in total, of which three were critical and three filler items.

Critical items were causatives for which a lexicalised and an opposed periphrastic alternative are available (e.g. *Opened the door* as opposed to *Made the door open*). Lexicalised causatives are associated with a direct, normal, more frequent, stereotypical causation, while their periphrastic alternatives are associated with an indirect, non-normal, less frequent, non-stereotypical causation (Levinson, 2000). Manner implicatures associated with causatives arise as a result of Grice's sub-maxim of manner *Be brief, avoid unnecessary prolixity* (Levinson, 2000). For instance, the utterance *The man made the door open* implies that *The man opened the door in an atypical way*. The additional meaning *in an atypical way* comes about because the speaker used a more prolix and complex sentence instead of the simpler statement *The man opened the door*. Filler items were literal expressions for which the correct picture could be identified on the basis of their explicit content.

Again, each item was composed of two slides. In the first slide, George was displayed holding a book and the target sentence was heard. The second slide featured two pictures as possible matches to the description. In the critical items, the two pictures contrasted an

unmarked, stereotypical way of causation (e.g. a picture depicting a man opening a door using the handle) with a marked, non-stereotypical way of causation (e.g. a picture displaying a man opening a door by kicking it). There were two task versions for each item. In one version, the unmarked, lexicalised expression was used and in the other, its corresponding marked, periphrastic expression was employed.

Scalar implicatures act-out task

This sub-test was a PowerPoint version of the action-based task used by Pouscoulous, Noveck, Politzer, and Baside (2007). Participants were presented with slides depicting five boxes and a selection of animals (five elephants, five turtles, five dolphins, and five hippopotamuses). There were three scenarios: in the 5/5 scenario, all boxes contained the same kind of animal, in the 2/5 scenario, two of the five boxes contained the same kind of animal, and in the 0/5 scenario, none of the boxes had any animals. For each scenario, children heard statements constructed with the quantifiers *all*, *some*, and *none* and three types of animals (elephants, turtles, and dolphins). This resulted in a total of 27 test items. There were also three practice items using numerals.

Children were told that they would hear George describing the display and that they had to make the display match the description either by putting animals inside the boxes, or by taking animals out of the boxes, or by leaving the display as it was without making any modifications (using the mouse as appropriate). There were three critical items. Critical items were statements with the quantifier *some* (e.g. *There are turtles in some of the boxes*) in the 5/5 scenario (e.g. in which all five boxes contained a turtle).

Scalar implicatures binary judgment task

In each trial of this test, the participant saw a depiction of five cards face down. An auditory stimulus was then played, *There are <X> on <Q> of the cards*, where X was the item type (rings, hearts, or stars) and Q the quantifier (*all*, *some*, or *none*). When the auditory stimulus ended, the cards were immediately 'turned over' to reveal the items. Participants were instructed to press a green-labelled key if the utterance were true and a red-labelled key if it were false, responding as quickly and accurately as possible.

The task included 23 trials comprising a practice block with two items and three test blocks of seven trials. There were three critical under-informative cases using the quantifier *some*, each presented in one of three test blocks. The rest of the items comprised an equal number of true and informative, and semantically false utterances with the quantifiers *some*, *all*, and *none*.

Socioeconomic status and language background questionnaire

This questionnaire was formed based on three questionnaires designed by other researchers: the Alberta Language Environment Questionnaire (Paradis, 2011), the Alberta Language Development Questionnaire (Paradis, Emmerzael, & Duncan, 2010), and the Family Affluence Scale (Currie, Molcho, Boyce, Holstein, Torsheim, & Richter, 2008).

Besides the items on the child's language use and exposure, the questionnaire also required information regarding the child's date of birth, gender, and the child's and parents' places of birth, among other topics.

Finally, three scores of the family's socioeconomic status could be extracted from the questionnaire: the family's wealth as measured by the Family Affluence Scale (henceforth, *FAS*)

and the parents' levels of education. Regarding maternal and paternal level of education, completion of junior high school was scored as *1*, senior high school as *2*, professional training in addition to completing senior high school as *3*, and higher education as *4*.

Language measures

Peabody Picture Vocabulary Test-Revised (Dunn & Dunn, 1981) (henceforth, PPVT)

A SMG version of the test adapted from English was administered as a measure of receptive vocabulary. The test was further adapted to and recorded in CG by the first author, who is a native CG speaker. The adaptations in the CG version were minimal and included phonetic changes in the pronunciation of words and also the substitution of SMG words with the corresponding CG words in cases where a more widely-used CG word existed. This test was administered only to the monolingual children and a subset of 17 bilectal children. The former group received the task in SMG while the latter in both CG and SMG.

Word Finding Vocabulary Test (henceforth, WFVT) (Renfrew, 1995)

The standardised SMG version (Vogindroukas, Protopapas, & Sideridis, 2009) of the test was administered to assess expressive vocabulary. For the bilectal and multilingual children, words in both CG and SMG were accepted as correct.

Non-verbal fluid intelligence test

The WASI Matrix Reasoning test (Wechsler, 1999) (henceforth, IQ)

In this test, participants were presented with pictures depicting a matrix from which a section was missing. They were required to complete the matrix by pointing at or stating the number corresponding to the correct response from five possible choices.

3 Results

3.1 Preliminary analyses

Principal component analyses

A principal component analysis (henceforth, *PCA*) was conducted on the implicature measures extracted from the various parts of the implicatures test (number of accurate responses in each implicature sub-test). These scores indicated performance in the CG language version of the test for bilectals and multilinguals, and in the SMG language version for monolinguals. However, because there was a ceiling level of performance for all children in the relevance implicature sub-test, scores from this part were not included in further analyses. Correlations between the various implicature measures are presented in table 1.

INSERT TABLE 1 AROUND HERE

The PCA on the four implicature indicators revealed only one factor with an eigenvalue above Kaiser's criterion of 1. This factor explained 36.8% of the variance. Table 2 summarises the PCA results.

INSERT TABLE 2 AROUND HERE

Composite scores

A composite score was computed for overall performance in the implicatures test (*overall implicature score-Greek*). The overall score was calculated by averaging participants' scores in each sub-test of the implicature task except for the relevance part in which ceiling performance was observed for all groups. The overall implicature-Greek score reflected performance in CG for bilectals and multilinguals and in SMG for monolinguals. Furthermore, for 17 of the bilectal children, a second overall implicature composite score was calculated based on their performance in the SMG version of the test (*overall implicature score-SMG*).

The following composite scores were also used: *general language ability-1*, *general language ability-2*, *general language ability-SMG*, and *SES*. General language ability-1 and general language ability-2 were measures of language proficiency in CG for bilectals and multilinguals and in SMG for monolinguals. The former was computed by transforming into z scores and then averaging the WFVT score-Greek[2] and the language comprehension score-Greek (indicating performance in the filler items of the implicatures test taken in CG by bilectals and multilinguals and in SMG by monolinguals). The latter was calculated for the sample of children who also took the PPVT, by collapsing the PPVT score-Greek (indicating performance in the CG version of the PPVT for bilectals and in the SMG version for monolinguals) and the general language ability score-1 into a new single indicator of language ability in Greek. Moreover, the general language ability-SMG score was created for the bilectal and monolingual children who also took the SMG version of the PPVT by averaging participants' z transformed scores in the WFVT, the filler items of the implicatures test (taken in SMG), and the PPVT (also taken in SMG). Finally, maternal level of education, paternal level of education, and FAS score

were also collapsed into a single score (in the same way as above) indicating socioeconomic status (henceforth, *SES*).

3.2 Main analyses

Between-group analyses on implicature understanding were performed in two stages. In the first step, the performance of the three groups of children was compared to each other. To this end, the three groups were matched in age by excluding from the analyses all multilingual and bilectal children who were above nine years or below six years of age.

In the second stage, the performance of a subset of 17 bilectal children was contrasted to that of the monolingual group. This second comparison provided the opportunity to (1) contrast the pragmatic performance of the two groups when more reliably controlling for language proficiency in the language of testing (CG or SMG) (since both groups in this comparison were further given the PPVT), (2) compare the performance of bilectal and monolingual children in each of the bilectals' varieties (since the bilectal children in this comparison further took the implicatures test in SMG), and (3) obtain a more reliable and robust test of potential group differences in implicature performance (since the two groups were comparable in several important background factors such as age, gender, IQ, SES, language comprehension in the language of testing and in SMG, ethnicity/culture, language of education, immigration history).

Where between-group Analyses of Covariance (henceforth, *ANCOVAs*) were conducted, the following variables were included as covariates: (1) any background measures for which statistically significant differences were found between the groups compared and (2) background variables that significantly correlated with the outcome variable. This allowed to partial out the effect of any of the background measures on the dependent variables and thus to

obtain a purer measure of the effect of interest (language group) and also ensured that any group differences in the background variables were not responsible for the presence or absence of group effects on the outcome variables[3].

Task Version (one versus two) for the implicatures test was also included as a between-subjects factor in the analyses in the cases where a significant correlation was found with implicature performance.

Comparison 1: Multilinguals versus bilectals versus monolinguals (matched in age)

Participants

In the following analyses the performance of 44 bilectal children (21 boys and 23 girls; ages 6;3–9, mean age 7;6, SD 0;9 years), 26 multilinguals (15 boys and 11 girls; ages 6;2–9, mean age 7;7, SD 0;9 years), and 25 monolinguals (15 boys and 10 girls; ages 6;2–9, mean age 7;4, SD 0;9 years) was compared.

Background measures

Background information for the three language groups in this comparison is presented in table 3. Language characteristics of the multilingual group as based on the Language Background Questionnaire are summarised in table 4.

INSERT TABLE 3 AROUND HERE

Bilectal, multilingual, and monolingual children did not statistically differ in age ($F(2, 92)=0.696$, $p>.05$), gender ($F(2, 92)=0.587$, $p>.05$), or language comprehension–Greek (in CG for bilectals and multilinguals, and in SMG for monolinguals; $F(2, 92)=0.319$, $p>.05$). Nevertheless, there were significant differences in SES ($F(2, 89)=9.622$, $p<.05$), IQ ($F(2, 92)=3.377$, $p<.05$, partial $\eta^2=.07$), and expressive vocabulary ($F(2, 92)=44.183$, $p<.05$, partial $\eta^2=.5$).

Regarding SES, there was a significant difference between bilectals and multilinguals, in that bilectal children were of a lower SES than multilingual children ($p<.05$, multiple comparisons with Bonferroni correction applied). For IQ, a multilingual advantage over monolinguals was found ($p<.05$, Bonferroni correction for multiple comparisons applied). Finally, with respect to expressive vocabulary, monolingual children had a significantly higher expressive vocabulary score than both bilectal and multilingual children, and bilectal children were significantly better than multilinguals (all p$s<.05$, Bonferroni correction applied). Within the multilingual group, CG-dominant children exhibited better vocabulary skills than their English-dominant multilingual peers, although their scores were still significantly lower than bilectal and monolingual participants (all ps$<.05$, Bonferroni correction applied). Further analyses indicated that the scores of both multilingual and bilectal children were within the normal monolingual range (bilectal scores were significantly higher than the 20th percentile ($t(26)=5.698$, $p<.05$) and multilingual scores did not significantly differ from the 20th percentile ($t(14)=1.024$, $p>.05$)). The same was true for both the CG-dominant ($t(9)=1.612$, $p>.05$) and English-dominant ($t(4)=-1.3600$, $p>.05$) multilinguals' expressive vocabulary performance[4]. In general, the results on expressive vocabulary are in line with existing research (e.g. Bialystok et al., 2010; Oller & Eilers, 2002; Oller et al., 2007) and validate accounts that attribute lower vocabulary or grammatical performance in bilinguals to environmental factors related to

bilingual children's language input rather than to deficits of any kind (cognitive, perceptual, or other) (see e.g. Ahtar & Menjivar, 2012; Hoff et al., 2012; Oller et al., 2007; Thordardottir, 2011). They also corroborate previous suggestions and experimental evidence in the bilingualism literature that high language proximity between the language pairs spoken by bilinguals (or bilectals) closes the gap in their vocabulary development relative to monolinguals (e.g. Barac & Bialystok, 2012; Bialystok & Feng, 2011).

INSERT TABLE 4 AROUND HERE

Implicatures test-Greek

Implicature scores included in the following analyses reflected performance in CG for bilectals and multilinguals and in SMG for monolinguals. Mean scores and SDs for the various measures from the implicatures test by language group are reported in table 5. Correlations between implicature scores and background variables can be found in table 1.

Overall Implicature composite score

An ANCOVA was conducted with Group (multilinguals versus bilectals versus monolinguals) and Task Version (one versus two) as between-subjects factors, and IQ, General language ability-1, Age, and SES as covariates. The results of this analysis were as follows. First, the effect of Group was not significant ($F(2, 82)=0.808$, $p>.05$, partial $\eta^2=.02$). Moreover, there was a significant effect of Task Version ($F(1, 82)=11.678$, $p<.05$) in that version one was easier than version two. Finally, a significant Group x Task Version interaction emerged ($F(2, 82)=3.111$, $p=.05$, partial $\eta^2=.07$).

In order to further investigate this interaction, two follow-up ANCOVAs (with the same covariates as above) were conducted separately for each version of the implicatures test. These ANCOVAs showed a clearly non-significant effect of Group in version two of the test (F(2, 38)=0.59, p>.05) and a marginally non-significant effect of Group in version one of the test (F(2, 40)=3.15, p=.054, partial η^2=.14). The marginally non-significant effect of Group in version one was due to a numerical advantage for multilinguals and monolinguals over bilectals. However, post-hoc pair-wise comparisons with Bonferroni correction applied showed that neither the multilingual nor the monolingual advantage was statistically significant (both ps>.09).

To determine whether multilingual and bilectal children exhibited equivalent to monolinguals implicature performance despite their lower core language proficiency, a similar ANCOVA as above was conducted with the Language comprehension score-Greek instead of the General language ability-1 score included as a covariate. Remember that, contrary to expressive vocabulary, the three groups did not differ in terms of the Language comprehension score-Greek. The Language comprehension score, though, significantly and positively correlated with overall implicature performance (Spearman's rho=.51, p(two-tailed)<.05) and for this reason it was covaried in the analysis. Results were the same as in the previous ANCOVA (for the effect of Group: F(2, 82)=0.340, p>.05), apart from the Group x Task Version interaction, which was marginally non-significant in this analysis (F(2, 82)=3.03, p=.054).

Finally, to establish whether amount of exposure to/use of the language of testing (CG) had an effect on implicature performance, a further analysis was conducted on the implicature composite score with the multilingual group divided in terms of language dominance. A 4 (Group: CG-dominant multilinguals versus English-dominant multilinguals versus bilectals versus monolinguals) x 2 (Task Version: one versus two) ANCOVA with IQ, Age, SES, and

Language ability-1 as covariates, returned a non-significant effect of Group ($F(3, 79)=1.029$, $p>.05$, partial $\eta^2=.04$) and a significant effect of Task Version ($F(1, 79)=10.64$, $p>.05$).

INSERT TABLE 5 AROUND HERE

Comparison 2: bilectals versus monolinguals

We start by reporting the analyses comparing bilectal and monolingual children's implicature performance in their native dialect (CG for bilectals and SMG for monolinguals), when a more reliable control over children's language ability was performed. Then, we describe how bilectals' and monolinguals' implicature performance compared when both groups were tested in SMG. We note that preliminary analyses comparing bilectal children's pragmatic performance in CG and SMG indicated no significant differences ($t(16)=-1.509$, $p>.05$).

Participants

17 bilectal (10 boys and 7 girls; ages 6;3–8;9, mean age 7;6, SD 0;9 years) and 25 monolingual children (15 boys and 10 girls; ages 6;2–9, mean age 7;4, SD 0;9 years) were included in the following analyses.

Background measures

Table 6 outlines background information about the two groups. The two groups did not differ in age ($t(40)=0.868$, $p>.05$), gender ($t(40)=-.074$, $p>.05$), language comprehension-Greek

(t(40)=-0.48, p>.05), language comprehension-SMG (t(40)=0.77, p>.05), IQ (t(40)=1.246, p>.05), or SES (t(38)=-1.373, p>.05).

However, monolingual children significantly outperformed bilectal children in both expressive (WFVT) (t(40)=-4.365, p<.05) and receptive vocabulary (PPVT) (when each group took the test in their native variety: (t(40)=-2.212, p<.05, r=.3); and when both groups took the test in SMG: (t(40)=-1.907, p=.06, r=.3).

INSERT TABLE 6 AROUND HERE

Implicatures test-Greek

Descriptive statistics for all measures from the implicatures test by language group for this comparison can be found in table 7.

Overall implicature composite score (after more reliably controlling for group differences in language proficiency)

An ANCOVA on the overall implicature measure, with Group (bilectals versus monolinguals) and Task Version (one versus two) as between-subjects factors, Language ability-2, Age, and IQ as covariates indicated a non-significant effect of Group (F(1, 35)=0.52, p>.05, partial η^2=.02) and a significant Group x Task Version interaction (F(1, 35)=7.38, p<.05, partial η^2=.17).

Follow-up ANCOVAs on overall implicature performance for each version separately (with the same covariates as in the previous analysis) revealed a clearly non-significant effect of Group in version one of the test (t(17)=0.17, p>.05) and a significant effect of Group in version two of the test (t(15)=2.33, p<.05, r=.5) indicating a bilectal advantage.

Implicatures test-SMG

Mean scores and SDs for children's performance in the SMG version of the implicatures test are reported by language group in table 8.

Overall Implicature composite score

An ANCOVA on the implicature-SMG score with Group as a between-subjects factor (bilectals versus monolinguals) and General language ability-SMG as a covariate showed no significant differences between the two groups (t(39)=1.61, p=.116, r=.26). Results were the same when including the Language comprehension score-SMG instead of the General language ability-SMG score as a covariate (t(39)=0.268, p>.05, r=.04).

Were the children included in the between-group analyses too competent with implicature for significant differences between the three language groups to emerge?

In order to examine this possibility, we divided the whole sample of children tested as part of this study (including the children not included in the between-group analyses above; see section 2.1., p. 15 and section 3.2, p. 26) into three age groups: (1) very young children (n=19, mean age 5;2, SD 0;4, age range 4;3-5;8), (2) middle-aged children (n=97, mean age 7;6, SD 0;9, age

range 6;2-9), and (3) older children (n=23, mean age 10;1, SD 0;9, age range 9;01-12;2). Middle-aged children were the participants included in the between-group analyses on implicature performance reported in the previous sections. We then conducted a between-group analysis on overall implicature performance in Greek with age group as a between-subjects factor. This analysis indicated a significant effect of age group ($F(2, 136)=25.182$, $p<.05$). Planned contrasts indicated that the younger age group (mean performance 1.33, SD 0.7) performed significantly worse than the middle-aged group (mean performance 2.22, SD 0.5) ($p<.05$) and that the middle-aged group performed significantly worse than the older group (mean performance 2.49, SD 0.5) ($p<.05$). These results indicate a developmental trend in children's implicature understanding skills. More importantly, they show that the children included in the between-group analyses on implicature performance (the middle-aged group) were yet not at the peak of their implicature understanding skills. This suggests that there was still room in children's pragmatic skills for further improvement, if indeed bilingualism had a positive effect on children's implicature understanding performance.

Power of statistical analyses for the effect of Language Group

To examine the power of our statistical comparisons to detect an effect of language group on overall implicature performance, we first calculated the likely magnitude of the bilingualism effect on pragmatic ability based on the three previous studies by Siegal et al. (2007; 2009; 2010), which tested similar skills in preschool-aged children. To do so, we conducted a basic meta-analysis on SPSS using the syntax files provided by Field and Gillett (2010). Table 10 shows the included studies, the effect sizes from each study (by comparison), and the sample size for each comparison.

INSERT TABLE 10 HERE

The results of the meta-analysis showed that the mean effect size based on Hedges and Vevea's (1998) random-effects model was r=.39; the 95% confidence interval was .27 (lower) and .49 (upper), which had a significant associated z score (z=6.24, p<.001). This represents a medium-to-large effect size according to Cohen (1988). We thus calculated the power of our between-group statistical tests to detect an effect of the magnitude found by the Siegal and colleagues studies (2007; 2009; 2010) in combination (see O'Keefe, 2007; Sun et al., 2011). We focus here on the first comparison that included the larger samples from each group, since both comparisons yielded similar results.

A power calculation using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) indicated that the first comparison had a power of .96 for the main effect of Group (for an ANOVA with three groups and a total sample of 95 children), which is well above Cohen's (1988) recommended level of .8[5]. We further calculated the power of our statistical comparison to detect a Group effect of moderate size (f=.25) (to allow for a more conservative estimate of the effect size than that expected based on the studies by Siegal and colleagues). This power calculation indicated that the first comparison had a power of .56 for the main effect of Group, which is well below the recommended level of .8 (Cohen, 1988). Nevertheless, these power calculations do not take into account two aspects of our statistical analyses that are known to have a positive effect on power. The first is the use of ANCOVA, which allows removing the effect of factors that correlate with the dependent measure of interest (and, hence, reduces error variance and improves the relationship between Group and the dependent variable). The second is the measurement of implicature understanding skill through multiple indicators that leads to more reliable estimates of this ability (Cohen, 1988:535-542; Rushton, Brainerd, & Pressley, 1983).

Regarding the first aspect, in our between-group comparisons we included three factors

that correlated with pragmatic performance (see table 1): Language ability, IQ, and Age. A regression analysis of Language ability–1, IQ, and Age on overall implicature performance showed that these three variables accounted collectively for 32% of the total variance in the dependent measure. Using the formula in Cohen (1988:540-541), this means that an ANCOVA comparison with these three factors as covariates would need only 68% of participants to achieve a power of .8 compared to a comparison without these variables covaried. In essence, this means a gain of 50 participants, which brings the power level of the Group comparison in this study to .76, just below the required level of .8.

We would expect, however, that the above estimates of power are even higher. When multiple measures of a target construct are tested and these measures show some degree of convergent validity and are combined into a single indicator of that component, variance idiosyncratic to each task is averaged out and general variance accumulates (see Carlson, 2003; Rushton et al., 1983). This leaves a less biased, more stable, more reliable, and purer estimator of the relevant component (Rushton et al., 1983) and achieves "more psychometric precision" (Carlson, 2003:142). This reduction in error in the dependent measure and consequent increase in reliability is also salutary for power as noted by Cohen, (1988:537; see also references therein).

The relation between implicature performance, children's language proficiency, and EC

In order to explore the verbal and non-verbal cognitive skills that potentially predict children's implicature performance, correlational analyses were conducted. These analyses were performed on the overall implicature composite score that reflected implicature performance in CG for multilinguals and bilectals, and in SMG for monolinguals.

Antoniou et al. (2016) reported that a PCA on six indicators of EC ability revealed two

distinct components of EC, which they interpreted as representing Inhibition and WM (with the Simon effect, Stop-Signal Reaction Time, and the Switching cost loading on the first factor and the rest on the second factor). Following the PCA, Antoniou et al. (2016) formed composite scores from the measures that loaded on each factor by transforming into z scores and then averaging the relevant scores. Thus, the correlational analyses reported here were conducted with these two EC composite scores –Inhibition and WM.

For all groups, measures of theoretical interest included the two EC composite scores. Age and language ability are also two factors that have been found in some studies to affect children's implicature understanding skills and for this reason correlations with these aspects were also considered (General language ability-1 was included as a measure of children's language proficiency in the language of testing) (see e.g. Johnson (1991), Norbury (2005), Rundblad & Annaz, (2010), and references therein for effects of language ability and Guasti, Chierchia, Crain, Foppolo, Gualmini, and Meroni (2005), Noveck (2001), Pouscoulous et al. (2007), Winner, Levy, Kaplan, & Rosenblatt (1988), Winner (1997) for effects of age). Correlations between these measures and implicature performance are presented in table 1 for the whole sample of children in this experimental study. Independent variables that significantly correlated with implicature performance were included in the regression analyses as predictors.

The first regression analysis was conducted on the whole sample of participants in this study with Age, General language ability-1, WM, IQ, and Task Version as predictors. Overall, this regression model was significant (F (5, 127)=15.465, p<.05) accounting for 38% of the variance in the dependent measure. When looking at the coefficients, only Age, General language ability-1, and Task Version significantly predicted overall implicature performance, the association with the first two measures being positive and the association with Version being negative (t(127)=3.400, p<0.05, t(127)=2.745, p<.05, t(127)=-3.793, p<.05,

respectively). Results of this linear regression analysis are presented in table 9.

INSERT TABLE 9 AROUND HERE

Similar regression analyses were conducted for each group of participants separately (monolinguals, bilectals, and multilinguals). The regression analysis for the monolingual group included General language ability-1 and Task Version as predictors and revealed only a significant negative effect of Version (t(22)=-2.60, p<.05) and a positive effect of General language ability-1 (t(22)=2.69, p<.05). The regression analysis for the bilectal group included Age, General language ability-1, WM, Inhibition, and IQ as predictors. When looking at the individual predictors, there was only a significant positive effect of General language ability-1 (t(54)=2.93, p<.05). Finally, the regression analysis for the multilingual group included Age, General language ability-1, WM, IQ, and Task Version and revealed only a significant positive effect of Age (t(39)=3.38, p<.05) and a significant negative effect of Version (t(37)=-4.589, p<.05).

4 Discussion

The present research investigated how growing up with two languages or two different varieties of the same language affects the ability to understand implicatures in children. We were also interested in potential relations between pragmatic language understanding and EC. There were several interesting findings of this study and these are discussed in the following sections.

4.1 The components of implicature understanding skill

A Principal Component Analysis on children's scores from four implicature sub-tests (on

metaphor, manner, and scalar implicature comprehension) revealed only a single factor of implicature performance. This was not a trivial result given that the scores for each implicature type were extracted from different sub-tests that varied methodologically (e.g. in terms of instructions, method of response, verbal demands). In terms of the generalised-particularised implicature debate, this finding provides support to pragmatic theories (e.g. Geurts, 2010; Grice, 1975; Sperber & Wilson, 1986/1995) that treat all types of pragmatically inferred meanings as the outcome of a single pragmatic interpretation process that involves uncovering the speaker's intentions behind an utterance. On the contrary, this result is not easily accommodated by theories that postulate a categorical distinction in the mechanisms involved in deriving PCIs and GCIs (Chierchia, 2004; Levinson, 2000).

4.2 The effects of childhood multilingualism and bilectalism on implicature understanding

Our study did not provide any hard evidence for differences in multilingual, bilectal, and monolingual children's implicature understanding skills. When considering the more reliable indicators of pragmatic ability –the overall implicature composite scores–, some differences emerged, but these differences were not robust and not in the direction of one group of children consistently out-performing the others. First, comparison one indicated better multilingual and monolingual performance relative to bilectals. However, this advantage was evident in only one of the two task versions of the implicatures test and was not significant in the follow-up analyses and post-hoc pairwise comparisons between the three groups. Furthermore, there were some indications for a bilectal advantage relative to monolinguals in overall implicature performance when children's language proficiency was more reliably controlled in comparison two. However, again, this difference in performance favoring bilectals was evident only in one

version of the implicatures task when bilectals and monolinguals were tested in their native dialects (bilectals in CG and monolinguals in SMG) and was not significant when bilectal and monolingual children's implicature performance in SMG was considered. Thus, it seems fair to conclude that there was no strong evidential basis in this study on which to maintain that the three groups differed in their implicature understanding skills.

Furthermore, multilingual children exhibited monolingual-like implicature understanding whether considering children who were dominant in the language of testing (CG) or children who were dominant in English. Similarly, bilectals performed comparably to monolinguals whether they were tested in their dominant, native variety (CG) or in their second, non-native variety (SMG). Finally, multilingual and bilectal children exhibited comparable to monolinguals implicature performance even when no statistical control was performed on their lower formal language proficiency.

These results in combination show that multilingual and bilectal children exhibit comparable to monolinguals level of implicature understanding irrespective of factors such as level of language proficiency or language dominance.

Towards reconciling the null multilingual effect on implicature understanding with previous positive reports

The above conclusion, however, stands in contrast to the findings of the studies by Siegal and colleagues who reported precocious bilingual development either in similar (2007) or closely related pragmatic-communicative skills (2009; 2010), and even in the face of bilingual children's lower language abilities. Here, we consider several explanations that might help reconcile the discrepancy in the findings.

To start with, Siegal et al.'s (2007) finding of a bilingual advantage in generating SIs is a

single finding in the literature and it's possible that it is not replicable (a type I error)[6]. The present study provides some evidence in support to this position since it included a similar (albeit not fully identical) task, but obtained no evidence for enhanced multilingual performance. Both the Siegal et al. (2007) task and the binary judgment task in the current study required children to judge the accuracy of sentences using the quantifiers *some* and *all* as descriptions of various scenarios, and both tasks required the generation of a SI in order to reject the critical under-informative sentences. Yet the two studies obtained different results. Furthermore, the present study did not document a multilingual advantage in understanding SIs in a second task that tested the same type of implicature, but had a different format (action-based task). This observation makes it unlikely that the present study gave rise to a type II error in two tasks compared to the likelihood of a type I error in Siegal et al. (2007) who used just one task.

Two additional explanations are related to the particular test employed in the Siegal et al. (2009; 2010) studies. First, Siegal et al.'s (2009; 2010) Conversational Violations Test examined the ability to detect utterances that violated conversational rules, while the pragmatic test in the present study investigated the skill to interpret implicatures. It is likely that the documented bilingual advantage in detecting violations of Gricean maxims does not extend to the ability to understand implicatures. We can see only one reason why this might not be the case. As described earlier, experimental evidence shows that bilingual children excel in almost all components that have been theoretically and/or experimentally linked to the process of implicature understanding (e.g. EC, detecting violations of Gricean maxims, ToM) but lag behind in core language knowledge. It is conceivable that superior implicature performance was not obtained in this study because the pragmatic task used posed more demands on children's language proficiency than the test employed in Siegal et al. (2009; 2010). This might have neutralised any potential multilingual advantages in other implicature-related facets.

Indeed, in support to this view, the regression analyses conducted on the whole sample of children indicated that core language ability was the primary factor affecting children's implicature understanding performance.

Second, Katsos, Roqueta, Estevan, and Cummins (2011) have questioned the CVT's validity as a test of pragmatic ability. Specifically, they argued that the CVT falls short of being a real test of pragmatic sensitivity, at least as far as the maxims of quality, relation, and politeness are concerned. As Katsos et al. (2011, p. 44) note, given the question *Have you seen my dog?*, children can reject an answer such as *Yes, he's in the sky* on the grounds of it being an implausible statement, just by drawing on their encyclopaedic knowledge about the world and the likelihood of dogs being in the sky, and not because they were sensitive to the maxim of quality. Katsos et al. (2011) express similar concerns about the statements used to investigate children's sensitivity to the maxim of relation (e.g. *I know your name* as a response to the question *What game do you like?*). Finally, they contend that identification of rude utterances such as the ones used by Siegal et al. (2009, 2010) to test children's sensitivity to the maxim of politeness (*No, it's disgusting* as a response to the question *Do you like my dress?*) depends on culture-specific social norms rather than on "culture-independent considerations of communicative efficiency that underpin the maxims of quantity, relation and manner" (Katsos et al. (2011, p. 44)). Nevertheless, it should be noted that Siegal and colleagues did report a bilingual advantage for the items on the maxims of manner (Siegal et al. 2009; 2010) and quantity II (Siegal et al., 2009; 2010) that unequivocally tested pragmatic sensitivity.

Another explanation for the contrast in the results might be that this study included children of an older age (school-aged children between 6–9 years old) than the children tested in Siegal et al. (2007; 2009; 2010; pre-schoolers between 4–6 year olds). It is possible that the impact of bilingualism is more profound during the preschool years, a period during which pragmatic understanding is at the very early stages of development and children present

difficulties in a range of pragmatic tasks (see e.g. Guasti et al., 2005; Noveck, 2001; Pouscoulous et al., 2007; Vosniadou, 1987; Winner 1997). Alternatively, one might argue that monolingual children catch up with their bilingual peers as they become older and accumulate more experience in interpreting pragmatic language. Thus, even though our data shows that a multilingual/bilectal benefit is not found in early school-aged children, it cannot exclude the possibility that such an advantage might be found when testing younger children.

That said, it should be clarified that the lack of significant differences between the language groups cannot be attributed to lack of variability in children's pragmatic performance or to the fact that the children in our study were too competent with implicature for significant differences to emerge. Overall implicature performance across the three groups was at 73% in the current study, which is far from ceiling. Furthermore, this level of performance is lower than the 79.5% and 74.4% overall performance obtained in the CVT by Siegal and colleagues (2009, experiments 1 and 2, respectively) and the 80% and 79% obtained by Siegal et al. (2010, experiments 1 and 2, respectively). This shows that our task had more variability in children's performance than the CVT in Siegal et al. (2009; 2010). Furthermore, the fact that significant differences between bilinguals and monolinguals were found in the experiments by Siegal and collaborators but not in the present one, suggests that the null effect in our experiment cannot be attributed to high level of performance or lack of sufficient variability. Finally, the implicature understanding skills of the children included in the between-group analyses on implicature comprehension (aged 6;2-9;0 years) were not yet at a peak level, but were still developing, as indicated by the fact that the implicature performance of these children was significantly lower than that of a group of older children (aged 9;01-12;2, whose data were discarded in the between-group analyses due to the matching in age).

4.3 The cognitive foundations of implicature understanding in children

Another purpose of this study was to investigate the cognitive factors that affect implicature understanding in children. A regression analysis on the data from the whole sample of children indicated that language proficiency in the language of testing, as well as age, were the critical predictors of implicature understanding. The findings of this analysis suggest that implicature understanding is a pragmatic-communicative skill that largely depends on children's language abilities.

Further regression analyses, however, on each of the three groups of children individually revealed an interesting divide. While for monolingual and bilectal children general language ability emerged as the decisive factor, for multilingual children there was a clearly non-significant effect of language proficiency and only a significant positive contribution of age.

How is it possible to account for this pattern of results? One interpretation might be that multilingual, bilectal, and monolingual children achieve the same ends when interpreting pragmatic language but through a different route. The findings of the present study indicate that, in general, implicature understanding is a pragmatic-communicative ability that primarily draws on children's language proficiency. This of course might be particularly true for children with better-developed language skills such as monolingual and bilectal children. The well-documented delays in bilingual children's language development, however, suggest that bilingual children might be less able to rely on their core language knowledge when comprehending implicatures. In this respect, they need to accomplish the implicature interpretation process in a different way. This possibility is further reinforced by the fact that age was a significant predictor of pragmatic performance only within the multilingual group, suggesting that other non-linguistic factors, besides EC, that develop with age are possibly involved in and sustain the process of computing implicatures in multilinguals.

Exactly what these other non-linguistic cognitive resources or this alternative way of understanding implicatures might be is beyond the reach of this study and such a discussion can only be speculative. One possibility is that bilingual children compensate for their disadvantage in language ability by being more attentive and sensitive to contextual aspects of an utterance, as suggested by Siegal et al (2010). Another option might be that bilinguals engage to a greater degree their (potentially better-developed) ToM skills.

4.5 The role of confounding factors

In the current study we controlled for several factors that correlate with bilingualism and/or that affect implicature performance by matching the groups in the relevant variables and/or by measuring these variables and excluding them as potential confounds through the use of ANCOVA. Ethnicity, culture and immigration status were also controlled by default given the groups of children compared (see section 1.5).

Language proficiency was carefully measured via two measures in comparison one and three measures in comparison two. Thus, we can be fairly confident that the language proficiency composite scores calculated from these measures were reliable indicators of the target skill and that the ANCOVA adequately adjusted for differences between the groups. No significant differences in overall implicature performance emerged even when comparing bilectal and monolingual children (who showed smaller differences in language proficiency as seen in the analyses on expressive vocabulary in comparison one) or even when more robustly controlling for differences in language proficiency by covarying a composite score of three language measures in comparison two.

Similarly, for SES, three different indicators were aggregated into a single composite measure, which was covaried in the between-group analyses. Again, it seems reasonable to assume that this composite score precisely and reliably measured this socio-demographic facet

and that the ANCOVA led to appropriate adjustments of differences between the three groups. Furthermore, SES cannot confound the results of the between-group analyses because it did not significantly correlate with implicature performance in this study. In addition, significant differences in SES were found only between the multilingual and bilectal children. Thus, in the critical comparisons between multilingual and monolingual children, and between bilectal and monolingual children that tested for the effects of main interest the groups were of comparable SES.

Between-group differences in non-verbal fluid intelligence were controlled for by covarying performance in the WASI matrix reasoning test in all the between-group analyses on overall implicature understanding skill. According to Miller and Chapman (2001:45), the application of ANCOVA is acceptable when the between-group differences on the covariate do not reflect true population differences, but can be attributed to chance. A review of the literature reveals that there is no expectation for a bilingual advantage in non-verbal fluid intelligence and the majority of studies that tested this aspect of cognition in bilingual and monolingual children reported a null bilingual effect (see Barac et al., 2014; but see Marzecova et al., 2013; Tao et al., 2011). This raises the possibility that the positive effect of multilingualism on IQ in this study was not a true effect and was due to chance.

But even if the multilingual advantage in IQ performance is real does this pose any challenges on our interpretation of the results in this study? Our answer is again *no* and several different arguments can be offered to ground this response. First, if between-group differences in IQ reflect a true effect, then the only error that could possibly be committed in our analyses is a type II error against monolinguals (i.e. that there is a monolingual advantage over multilinguals that was not detected due to the fact that the between-group difference in IQ was not adequately controlled). This by itself is a novel finding given that one would expect, not a monolingual advantage, but a medium-to-large multilingual advantage based on previous

research. However, we do not think that this scenario is plausible. First, the WASI matrix reasoning test is part of a widely-used cognitive battery whose psychometric characteristics have been investigated and established. For participants' scores in the matrix reasoning test, Wechsler (1999) reports split-half reliability of .96 and test–retest reliability of .72 for the age range in this sample. Second, we also calculated the reliability of this score based on the sample of children in our study. Cronbach's α for the IQ score is .956, which is quite high (see Field, 2013:715). Second, even though IQ significantly correlated with implicature understanding skill, the regression analyses indicated that it does not actually affect pragmatic performance and that the significant correlation can be explained in terms of third factors (age, EC, or language proficiency). Third, it is possible that the multilingual advantage in the WASI matrix reasoning test was due to multilinguals' enhanced EC skills. If this is true, then, this would justify treating IQ as an EC measure. Again, differences in EC cannot explain the between–group results as indicated by the non-significant effect of EC on implicature understanding in the regression analyses in section 3.2 (p. 32). Finally, in the second comparison of the current study, bilectal and monolingual sub-groups were used that did not statistically differ in terms of either SES or fluid intelligence performance. Yet, again, no significant differences between the two groups were obtained in overall implicature understanding.

Appendix A. The (socio-)linguistic situation in Greek-speaking Cyprus and the relation between Cypriot Greek and Standard Modern Greek

The sociolinguistic terrain in Greek-speaking Cyprus has been often described as diglossic with Cypriot Greek being the low variety and Standard Modern Greek the high variety (Arvaniti, 2010; Rowe & Grohmann, 2013; but see Auer, 2005; Karyolemou, 2006). CG and SMG are both dialects of Greek. CG is natively acquired and used in everyday communication. SMG is acquired sequentially, mainly through formal education (from kindergarten onwards, around the age of five), it is the language of education and it is also the language used in formal situations –particularly writing, public speaking, and in the media. It is also the official language of the Republic of Cyprus and the language spoken by Hellenic Greeks in the Republic of Greece.

CG and SMG exhibit a high degree of language similarity. To begin with, the two varieties are genetically related. According to the traditional comparative method, the two dialects are classified as parts of the Greek language family of Indo-European languages (Lewis, Gary, & Fennig, 2014, see *Classification* under *Greek Language* entry; see also Lyovin, 1997). Furthermore, the *Ethnologue* (Lewis et al., 2014, see *Dialects* under *Greek Language* entry) reports a lexical similarity of 84%-93% between the two varieties, with an 85% similarity being the cut-off point for two variants being dialects of the same language[7] (see also Terkourafi, (2005, p. 316); Newton, (1972, pp. 111-112)). Finally, there is some degree of intelligibility between the two dialects as spoken by native speakers among themselves with some varieties of CG being more easily understood by SMG native speakers than others.

Nevertheless, the two dialects also show differences in both grammar and the lexicon (see e.g. Terkourafi, 2005): morphology (for example, CG has verbs that are formed with the suffix –*isko*, as in CG [miˈnisko], *I stay*, instead of the SMG [ˈmeno]), syntax (e.g. CG uses a

clefting strategy for forming *wh*-questions, which does not exist in SMG, as in the CG [ˈinda mbu ˈkamnis?], meaning *what is it that you are doing*, opposed to the SMG [ti ˈkaneis] (Grohmann, Panagiotidis, & Tsiplakou, 2006)), phonetics and phonology (e.g. in CG the voiced fricatives [v], [ð], [ɣ] are sometimes deleted intervocalically, so the SMG [tiɣaˈnizo], meaning *I fry*, becomes [tiaˈnizo] in CG), semantics (e.g. the word [laˈlo] means *I crow* in SMG and *I say* in CG), pragmatics (Greek Cypriots and Hellenic Greeks often use different politeness strategies (e.g. Terkourafi, 1997)) and lexis (e.g. for the word *run*, SMG speakers use the word [ˈtrexo], while Greek Cypriot speakers frequently use the word [vuˈɾo]).

Appendix B. List of linguistic materials in the Implicatures test.

Item		Task Version	Correct Response
	List of linguistic material used in the picture-selection part of the Implicatures test.		
	Practice items		
1	George and his mother were in the living room. George asked his mother: "Mom, shall I open the window?" His mother replied: "Yes, it's very hot in here."	1	Picture 2
2	George and his mother were in the living room. George asked his mother: "Mom, shall I open the window?" His mother replied: "No, it's very cold outside."	2	Picture 1
	Relevance implicature items		
3	George and his mother were in the living room. George asked his mother: "Mom, can I go to the beach with my friends?" His mother replied: "You are ill."	1	Picture 1
4	George and his mother were in the living room. George asked his mother: "Mom, can I go to the beach with my friends?" His mother replied: "Don't forget to take some sunscreen with you."	2	Picture 2
5	George and his mother were in the living room. George asked his mother: "Mom, can I buy some ice cream?" His mother replied: "I've got money in my wallet."	1	Picture 1
6	George and his mother were in the living room. George asked his mother: "Mom, can I buy some ice cream?" His mother replied: "You are ill."	2	Picture 2
7	George and his mother were in the living room. George asked his mother: "Mom, shall I feed the dog?" His mother replied: "The dog chow is in the kitchen."	1	Picture 2
8	George and his mother were in the living room. George asked his mother: "Mom, shall I feed the dog?" His mother replied: "I've just fed the dog."	2	Picture 1
	Metaphor items		
9	Today, George's father's favourite football team was playing a game. George's father went to the stadium to watch the game. His team did not play well and lost. George heard the final score on the news. He wondered how his father would feel when coming back home. He didn't know whether his father would feel sad because his favourite team lost or angry because the team did not play well. When George's father returned home, he was a melting snowman.	1	Picture 2
10	Today, George's father's favourite football team was playing a game. George's father went to the stadium to watch the game. His team did not play well and lost. George heard the final score on the news. He wondered how his father would feel when coming back home. He didn't know whether his father would feel sad because his favourite team lost or angry because the team did not play well. When George's father returned home, he was a thundering cannon.	2	Picture 1
11	George's father had a dog that he loved dearly. The dog lived in a fenced area in the backyard of the house. Today, George's father was at work and George was told to feed the dog. George gave the dog some food but, when leaving, he accidentally left the fence	1	Picture 1

	door open. The dog escaped and George could not find him anywhere. George wondered how his father would feel when he told him about the dog. He didn't know whether his father would feel sad because his beloved dog had gone or angry because George had carelessly left the fence door open. When George told his father what had happened, George's father was a foaming sea.		
12	George's father had a dog which he loved dearly. The dog lived in a fenced area in the backyard of the house. Today, George's father was at work and George was told to feed the dog. George gave the dog some food, but when leaving, he accidentally left the fence door open. The dog escaped and George could not find him anywhere. George wondered how his father would feel when he told him about the dog. He didn't know whether his father would feel sad because his beloved dog had gone or angry because George had carelessly left the fence door open. When George told his father what had happened, George's father was a knocked down tower.	2	Picture 2
13	Today, George's father's favourite singer was in the town for a concert. The tickets for the concert were limited and they soon sold out. George's father, though, managed to buy two tickets - one for himself and one for George's mother. He kept the tickets on a table in the living room. At some point, however, George accidentally spilt water on the table and ruined the tickets. George wondered how his father would feel upon finding out. He didn't know whether his father would feel sad because he would miss his favourite singer's concert or angry because George had carelessly ruined the tickets. When George's father found out what had happened, he was a rumbling thunder.	1	Picture 1
14	Today, George's father's favourite singer was in the town for a concert. The tickets for the concert were limited and they soon sold out. George's father, though, managed to buy two tickets - one for himself and one for George's mother. He kept the tickets on a table in the living room. At some point, however, George accidentally spilt water on the table and ruined the tickets. George wondered how his father would feel upon finding out. He didn't know whether his father would feel sad because he would miss his favourite singer's concert or angry because George had carelessly ruined the tickets. When George's father found out what had happened, he was a sinking ship.	2	Picture 2
	Manner Implicature items		
15	In this picture, a man opened the door.	1	Picture 2
16	In this picture, a man made the door open.	2	Picture 1
17	In this picture, a man made the safe open.	1	Picture 2
18	In this picture, a man opened the safe.	2	Picture 1
19	In this picture, a man spilled the water on the floor.	1	Picture 1
20	In this picture, a man made the water spill on the floor.	2	Picture 2
	Filler items		
21	In this picture, a man is sleeping.	1	Picture 2
22	In this picture, a man turned off the alarm-clock.	2	Picture 1
23	In this picture, a man unlocked the lock.	1	Picture 1
24	In this picture, a man broke the lock.	2	Picture 2
25	In this picture, a boy is playing in the sand.	1	Picture 2
26	In this picture, a boy is swimming.	2	Picture 1
27	George and his mother were in the living room. George's friends	1	Picture 1

	were going to play football and George really wanted to join them. George asked his mother: "Mom, can I play football with my friends?" George's mother replied: "No, you have a test at school tomorrow."			
28	George and his mother were in the living room. George's friends were going to play football and George really wanted to join them. George asked his mother: "Mom, can I play football with my friends?" George's mother replied: "Yes, your football is in the kitchen."	2	Picture 3	
29	George and his mother were in the living room. George's favourite film was going to come on TV in a little while and George really wanted to watch it. George asked his mother: "Mom, can I watch TV?" George's mother replied: "Yes, I would like to watch TV, too."	1	Picture 3	
30	George and his mother were in the living room. George's favourite film was going to come on TV in a little while and George really wanted to watch it. George asked his mother: "Mom, can I watch TV?" George's mother replied: "No, the baby is sleeping."	2	Picture 2	
31	George and his mother were in the living room. His parents had just bought him a new bike and George really wanted to go out and ride it. George asked his mother: "Mom, can I go play with my bike?" George's mother replied: "No, you haven't finished your homework yet"	1	Picture 3	
32	George and his mother were in the living room. His parents had just bought him a new bike and George really wanted to go out and ride it. George asked his mother: "Mom, can I play with my bike?" George's mother replied: "Yes, but be careful not to get hurt"	2	Picture	

List of linguistic material used in the binary judgment part of the Implicatures test.				
Item	**Target utterance**	**Visual display**	**Utterance Status**	**Expected Response**
	Practice items			
33	There are hearts on none of the cards.	**5/5**	**False**	**Incorrect**
34	There are rings on all of the cards.	**5/5**	**True and informative**	**Correct**
	Test items			
35	**There are hearts on some of the cards.**	**5/5**	**True but under-informative**	**Correct**
				Incorrect
36	There are hearts on some of the cards.	3/5	True and informative	Correct
37	There are hearts on some of the cards.	0/5	False	Incorrect
38	There are hearts on all of the cards.	5/5	True and informative	Correct
39	There are hearts on all of the cards.	3/5	False	Incorrect

40	There are hearts on none of the cards.	3/5	False	Incorrect
41	There are hearts on none of the cards.	0/5	True and informative	Correct
42	**There are stars on some of the cards.**	**5/5**	**True but under-informative**	**Correct**
				Incorrect
43	There are stars on some of the cards.	3/5	True and informative	Correct
44	There are stars on some of the cards.	0/5	False	Incorrect
45	There are stars on all of the cards.	5/5	True and informative	Correct
46	There are stars on all of the cards.	3/5	False	Incorrect
47	There are stars on none of the cards.	3/5	False	Incorrect
48	There are stars on none of the cards.	0/5	True and informative	Correct
49	**There are rings on some of the cards.**	**5/5**	**True but under-informative**	**Correct**
				Incorrect
50	There are rings on some of the cards.	3/5	True and informative	Correct
51	There are rings on some of the cards.	0/5	False	Incorrect
52	There are rings on all of the cards.	5/5	True and informative	Correct
53	There are rings on all of the cards.	3/5	False	Incorrect
54	There are rings on none of the cards.	3/5	False	Incorrect
55	There are rings on none of the cards.	0/5	True and informative	Correct

Note: critical under-informative utterances appear in bold. The rest of the test utterances are filler items.

List of linguistic material used in the act-out part of the Implicatures test.				
Item	**Target utterance**	**Visual display**	**Utterance Status**	**Expected Response**
	Practice items			
56	Two of the boxes have dolphins.	0/5	False	Add one dolphin in each one of two empty boxes
57	Two of the boxes have turtles.	2/5	True	No action
58	Three of the boxes have elephants.	5/5	False	Remove two elephants from the boxes
	Test items			
59	**Some of the boxes have turtles.**	**5/5**	**True but under-informative**	**Remove at least one and no more than 4 turtles from the boxes**
				No action
60	Some of the boxes have turtles.	2/5	True and informative	No action or add one turtle in at most two of the three empty boxes

61	Some of the boxes have turtles.	0/5	False	Add at least one and no more than 4 turtles in the boxes
62	All of the boxes have turtles.	5/5	True and informative	No action
63	All of the boxes have turtles.	2/5	False	Add one turtle in each one of the three empty boxes
64	All of the boxes have turtles	0/5	False	Add one turtle in each one of the five empty boxes
65	None of the boxes have turtles	2/5	False	Remove the two turtles from the boxes
66	None of the boxes have turtles.	0/5	True and informative	No action
67	None of the boxes have turtles.	5/5	False	Remove all turtles from the boxes
68	**Some of the boxes have elephants.**	**5/5**	**True but under-informative**	**Remove at least one and no more than four of the elephants from the boxes**
				No action
69	Some of the boxes have elephants.	2/5	True and informative	No action or add one elephant in at most two of the three empty boxes
70	Some of the boxes have elephants.	0/5	False	Add at least one and no more than four elephants in the boxes
71	All of the boxes have elephants.	5/5	True and informative	No action
72	All of the boxes have elephants.	2/5	False	Add one elephant in each one of the three empty boxes
73	All of the boxes have elephants	0/5	False	Add one elephant in each one of the five empty boxes
74	None of the boxes have elephants.	2/5	False	Remove the two elephants from the boxes
75	None of the boxes have elephants.	0/5	True and informative	No action
76	None of the boxes have elephants.	5/5	False	Remove all elephants from the boxes
77	**Some of the boxes have dolphins.**	**5/5**	**True but under-informative**	**Remove at least one and no more than four of the dolphins from the boxes**
				No action
78	Some of the boxes have dolphins.	2/5	True and informative	No action or add one dolphin in at most two of the three empty boxes
79	Some of the boxes have dolphins.	0/5	False	Add at least one and no more than four dolphins in the boxes

80	All of the boxes have dolphins.	5/5	True and informative	No action
81	All of the boxes have dolphins.	2/5	False	Add one dolphin in each one of the three empty boxes
82	All of the boxes have dolphins	0/5	False	Add one dolphin in each one of the five empty boxes
83	None of the boxes have dolphins.	2/5	False	Remove the two dolphins from the boxes
84	None of the boxes have dolphins.	0/5	True and informative	No action
85	None of the boxes have dolphins.	5/5	False	Remove all dolphins from the boxes.
Note: critical under-informative utterances appear in bold. The rest of the test utterances are filler items.				

Acknowledgements

[1]Grice (1975; 1989) himself did not propose a maxim of politeness (*be polite*), although he did discuss the possibility that such a maxim might be necessary to provide a full account of pragmatic meaning (Grice (1989, p. 28)).

[2]We collapse the distinction between CG and SMG and call the score, *WFVT score-Greek* because, in the expressive vocabulary test, words coming from both varieties were accepted as correct when testing multilingual and bilectal children.

[3]Field (2013:486) and Huitema (2011) note that the independence of the covariate and the experimental effect is not a statistical requirement. Zinbarg Suzuki, Uliaszek, and Lewis (2010, p. 3) also note that it is consensually meaningful to use ANCOVA when differences on the covariate are not a central, inherent, conceptual part of the Group variable (e.g. in the case of lower language proficiency in bilinguals, one can include language proficiency as a covariate, if there exist bilinguals in real life who do not differ from monolinguals in language proficiency). Differences between the comparison groups on the covariate, however, might bias the ANCOVA and lead to an error. This bias in ANCOVA might arise when the groups differ on the covariate but the covariate is not reliably measured (i.e. includes measurement error) (see e.g. Reichardt, 1979, who also describes other sources of bias in ANCOVA, which are not relevant here). This situation will lead to an under-adjustment of differences between the comparison groups. Nevertheless, when bias in the ANCOVA is due to the unreliability of the covariate, including the covariate into the analysis is generally better than not including it at all. The inclusion of the covariate will lead to some partial —but possibly not complete— adjustment of differences between the two groups on the covariate (see also Reichardt, 1979; Miller & Chapman, 2001; Huitema, 2011; Zinbarg et al., 2010, for elaborate discussions of these issues). We return to this briefly in the *Discussion*, where we discuss the issue of confounds in our experiment.

[4]Vogindroukas et al. (2009) provide norms only for monolingual children aged five to eight years old. For this reason, only scores from bilectal and multilingual children within this age range were converted into percentiles.

[5]Note that we first had to convert the effect size *r* to an effect size *f* (the effect size used by G*Power) using the formulas reported in Cohen (1988). The conversion indicated that an effect size r=.39 equals an effect size f=.42.

[6]Remember that Siegal and colleagues (2009; 2010) did not test implicature comprehension in their studies but a related pragmatic skill -the ability to detect violations of conversational maxims (see section 1.3). We note, though, that when considering performance in the items on the maxim of quantity I (which is the maxim on which computation of SIs is based), Siegal et al. (2009) found no significant differences between bilinguals and monolinguals either in their experiment 1 (where the bilingual advantage was found in overall performance and not in a specific maxim) or in experiment 2 (where a significant Language Group x Maxim interaction was reported and the bilingual advantage was found only in the items for the maxim of relation, quality, manner, and quantity II). Siegal et al. (2010) did not include any items on quantity I in their experiments.

[7]According to *Ethnologue* "The percentage of lexical similarity between two linguistic varieties is determined by comparing a set of standardised wordlists and counting those forms that show similarity in both form and meaning. Percentages higher than 85% usually indicate a speech variant that is likely a dialect of the language with which it is being compared. Unlike intelligibility, lexical similarity is bidirectional or reciprocal." (Lewis et al. (2014, see *Dialects* under *Language Information* section)). It is not clear, however, why *Ethnologue* suggests this percentage of lexical similarity as the cut-off point for distinguishing between dialects and languages.

References

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research, 80*(2), 207-245.

Akhtar, N., & Menjivar, J. A. (2012). Cognitive and linguistic correlates of early exposure to more than one language. *Advances in Child Development & Behavior, 42*, 41–78.

Antón, E., Duñabeitia, J. A., Estévez, A., Hernández, J. A., Castillo, A., Fuentes, L. J., & Carreiras, M. (2014). Is there a bilingual advantage in the ANT task? Evidence from children. Frontiers in Psychology, 5, 398.

Antoniou, K., Grohmann, K., Kambanaros, M., & Katsos, N. (2016). The effect of childhood bilectalism and multilingualism on executive control. *Cognition*, *149*, 18-30.

Arvaniti, A. (2010). Linguistic practices in Cyprus and the emergence of Cypriot Standard Greek. *Mediterranean Language Review 17*, 15-45.

Auer, P. (2005). Europe's sociolinguistic unity, or: A typology of European dialect/standard constellations. In N. Delbecque, J. van der Auwera & D. Geeraerts (Eds.), *Perspectives on Variation: Sociolinguistic, Historical, Comparative* (pp. 7-41). Berlin/New York: Mouton de Gruyter.

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.

Barac, R., Bialystok, E., Castro, D. C., & Sanchez, M. (2014). The cognitive development of young dual language learners: A critical review. *Early Childhood Research Quarterly*. doi.org/10.1016/j.ecresq.2013.08.004

Bernicot, J., Laval, V., & Chaminaud, S. (2007). Nonliteral language forms in children: In what order are they acquired in pragmatics and metapragmatics? *Journal of*

Pragmatics, 39, 2115–2132.

Bergen, L., & Grodner, D. J. (2012). Speaker knowledge influences the comprehension of
pragmatic inferences. *Journal of Experimental Psychology: Learning, Memory, and
Cognition, 38*(5), 1450-1460.

Bialystok, E., Craik, F. I., & Luk, G. (2012). Bilingualism: consequences for mind and brain.
Trends in cognitive sciences, 16(4), 240-250.

Bialystok, E., Craik, F. I., Green, D. W., & Gollan, T. H. (2009). Bilingual minds.
Psychological Science in the Public Interest, 10(3), 89-129.

Bialystok, E. & Feng, X. (2011). Language proficiency and its implications for monolingual
and bilingual children. In A. Y. Durgunoglu & C. Goldenberg (Eds.), *Language and
literacy development in bilingual settings*, (pp 121–138). New York: Guilford.

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., Poarch, G., Luo, L., & Craik, F. I. (2014). Effects of bilingualism and aging on
executive function and working memory. Psychology and aging, 29(3), 696.

Bialystok, E., & Senman, L. (2004). Executive processes in appearance–reality tasks: The
role of inhibition of attention and symbolic representation. *Child development, 75*(2),
562-579.

Blom, E., Küntay, A. C., Messer, M., Verhagen, J., & Leseman, P. (2014). The benefits of
being bilingual: Working memory in bilingual Turkish–Dutch children. Journal of
Experimental Child Psychology, 128(0), 105-119.

Bott, L., Bailey, T.M., and Grodner, D. (2011). Distinguishing speed from accuracy in scalar
implicatures. *Journal of Memory and Language. 66*(1), 123–142.

Bott, L. & Noveck, I. A. (2004). Some utterances are underinformative: The onset and time
course of scalar inferences. *Journal of Memory and Language, 51*(3), 437-57.

Breheny, R., Ferguson, H. J., & Katsos, N. (2013a). Investigating the timecourse of accessing conversational implicatures during incremental sentence interpretation. *Language and Cognitive Processes*, *28*(4), 443-467.

Breheny, R., Ferguson, H. J., & Katsos, N. (2013b). Taking the epistemic step: toward a model of on-line access to conversational implicatures. *Cognition*, *126*(3), 423-440.

Breheny, R., Katsos, N., and Williams, J. (2006). Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition*, *100*(3), 434–463.

Carlson, S. M. (2003). Executive function in context: Development, measurement, theory, and experience. Monographs of the Society for Research in Child Development, 68(3), 138-151.

Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental science, 11*(2), 282-298.

Chierchia, G. (2004). Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In A. Belletti (Ed.), *Structures and Beyond,* (pp. 39–103). Oxford: Oxford University Press.

Cohen, J. (1988). Statistical power analysis for the behavioural sciences. Hillsdale, NJ: Erlbaum.

Currie, C., Molcho, M., Boyce, W., Holstein, B., Torsheim, T., & Richter, M. (2008). Researching health inequalities in adolescents: the development of the Health Behaviour in School-Aged Children (HBSC) family affluence scale. *Social science & medicine, 66*(6), 1429-1436.

de Abreu, P. M. E., Cruz-Santos, A., Tourinho, C. J., Martin, R., & Bialystok, E. (2012). Bilingualism Enriches the Poor Enhanced Cognitive Control in Low-Income Minority Children. *Psychological Science, 23*(11), 1364-1371.

Degen, J., & Tanenhaus, M. K. (2015). Availability of Alternatives and the Processing of

   Scalar Implicatures: A Visual World Eye-Tracking Study. *Cognitive science*.

de Neys, W., & Schaeken, W. (2007). When people are more logical under cognitive load:

   dual task impact on scalar implicature. *Experimental Psychology, 54*(2), 128-33.

Dieussaert, K., Verkerk, S., Gillard, E., & Schaeken, W. (2011). Some effort for some:

   further evidence that scalar implicatures are effortful. *The Quarterly Journal of

   Experimental Psychology*, *64*(12), 2352-2367.

Dunn, L. & Dunn, D. (1981). *Peabody Picture Vocabulary Test-Revised*. Circle Pines, MN:

   American Guidance Service.

Duñabeitia, J. A., Hernández, J. A., Antón, E., Macizo, P., Estévez, A., Fuentes, L. J., &

   Carreiras, M. (2014). The inhibitory advantage in bilingual children revisited: Myth or

   reality? *Experimental Psychology, 61*(3), 234.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical

   power analysis program for the social, behavioral, and biomedical sciences. *Behavior

   Research Methods*, *39*, 175-191.

Field, A. (2013). *Discovering Statistics using SPSS* (4th edition). London: Sage.

Field, A. P., & Gillett, R. (2010). How to do a meta-analysis. *British Journal of Mathematical

   and Statistical Psychology*, *63*(3), 665-694.

Foppolo, F., Guasti, M. T., & Chierchia, G. (2012). Scalar implicatures in child language:

   Give children a chance. *Language learning and development*, *8*(4), 365-394.

Gathercole, V. C. M., Thomas, E. M., Kennedy, I., Prys, C., Young, N., Guasch, N. V.,

   Jones, L. (2014). Does language dominance affect cognitive performance in

   bilinguals? Lifespan evidence from preschoolers through older adults on card sorting,

   Simon, and metalinguistic tasks. Frontiers in Psychology, 5, 11. doi:

   10.3389/fpsyg.2014.00011

Genesee, F., Tucker, G. R., & Lambert, W. E. (1975). Communication skills of bilingual

    children. *Child development*, 1010-1014.

Geurts, B. (2010). *Quantity Implicatures*. Cambridge: Cambridge University Press.

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control

    hypothesis. Journal of Cognitive Psychology, 25(5), 515-530.

Grice, H.P. (1975). Logic and conversation. In P. Cole and J.L. Morgan (Eds.), *Syntax and*

    semantics, (Vol. 3, pp. 225-242). New York: Seminar Press.

Grice, H. P. (1989). *Studies in the Way of Words*. Cambridge, MA: Harvard University

    Press.

Grodner, D., Klein, N., Carbary, K. & Tanenhaus, M. (2010). "Some," and possibly all,

    scalar inferences are not delayed: Evidence for immediate pragmatic enrichment.

    Cognition, *116*, 42-55.

Grodner, D., & Sedivy, J. (2005). The effect of speaker-specific information on pragmatic

    inferences. In N. Pearlmutter, and E. Gibson (Eds.), *The processing and acquisition of*

    reference. Cambridge, MA: MIT Press.

Grohmann, K. K., Panagiotidis, P., & Tsiplakou, S. (2006). Properties of *wh*-question

    formation in Cypriot Greek. In M. Janse, B.D. Joseph, & A. Ralli (Eds.), *Proceedings*

    of the 2nd International Conference on Modern Greek Dialects and Linguistic Theory

    (pp. 83-98). Patras: University of Patras.

Guasti, T.M., Chierchia, G., Crain, S., Foppolo, F., Gualmini, A., & Meroni, L. (2005). Why

    children and adults sometimes (but not always) compute implicatures. *Language and*

    Cognitive Processes, 20(5), 667-696.

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.

Huang, Y., & Snedeker, J. (2009a). On-line interpretation of scalar quantifiers: insight into

 the semantics-pragmatics interface. *Cognitive Psychology, 58,* 376-415.

Huang, Y. T., & Snedeker, J. (2009b). Semantic meaning and pragmatic interpretation in 5-

 year-olds: Evidence from real-time spoken language comprehension. *Developmental*

 psychology, 45(6), 1723.

Huitema, B. (2011). *The Analysis of Covariance and Alternatives: Statistical Methods for*

 Experiments, Quasi-Experiments, and Single-Case Studies. Hoboken, NJ: John Wiley

 & Sons.

Johnson, J. (1991). Developmental versus language-based factors in metaphor interpretation.

 Journal of Educational Psychology, 83(4), 470.

Kambanaros, M., Grohmann, K., Michaelides, M., & Theodorou, E. (2012). Comparing

 multilingual children with SLI to their bilectal peers: Evidence from object and

 action picture naming. International Journal of Multilingualism 10. 60–81. DOI:

 10.1080/14790718.2012.705846

Karyolemou, M. (2006). Reproduction and innovation of communicative patterns in a

 former-"diglossic" community. In R. Muhr (Ed.), *Reproduction and Innovation in*

 Language and Communication in different Language Cultures/Reproduktionen und

 Innovationen in Sprache und Kommunikation verschiedener Sprachkulturen (pp. 39-

 56). Vienna: Peter Lang Verlag.

Katsos, N., Roqueta, C. A., Estevan, R. A. C., & Cummins, C. (2011). Are children with

 Specific Language Impairment competent with the pragmatics and logic of

 quantification? *Cognition, 119*(1), 43-57.

Kovács, Á. M. (2008). Early bilingualism enhances mechanisms of false-belief reasoning.

 Developmental Science, 12(1), 48-54.

Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for

     language processing and cognition. *Journal of Cognitive Psychology, 25*(5), 497-514.

Klein, R. M. (2015). Is there a benefit of bilingualism for executive functioning?.

     Bilingualism: Language and Cognition, 18(01), 29-31.

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational*

     implicature. Cambridge, MA: MIT Press.

Lewis, M. P., Gary F. S., & C. D., Fennig (Eds.). (2014). *Ethnologue: Languages of the*

     World, (17th ed.). Dallas, Texas: SIL International. Online version:

     http://www.ethnologue.com.

Lyovin, A. (1997). *An Introduction to the Languages of the World*. New York: Oxford

     University Press.

Martin-Rhee, M. M., & Bialystok, E. (2008). The development of two types of inhibitory

     control in monolingual and bilingual children. *Bilingualism Language and Cognition*

     11(1), 81.

Marty, P. P., & Chemla, E. (2013). Scalar implicatures: working memory and a comparison

     with only. *Frontiers in Psychology, 4*, 403. doi: 10.3389/fpsyg.2013.00403

Marzecova, A., Asanowicz, D., Krivá, L. U., & Wodniecka, Z. (2013). The effects of

     bilingualism on efficiency and lateralization of attentional networks. *Bilingualism:*

     Language and Cognition, 16(3), 608-623.

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of*

     Abnormal Psychology, 110(1), 40.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D.

     (2000). The Unity and Diversity of Executive Functions and Their Contributions to

     Complex "Frontal Lobe" Tasks: A Latent Variable Analysis. *Cognitive Psychology,*

     41(1), 49-100.

Newton, Brian (1972). *The Generative Interpretation of Dialect: A Study of Modern Greek Phonology*. Cambridge: Cambridge University Press.

Nicoladis, E., Palmer, A., & Marentette, P. (2007). The role of type and token frequency in using past tense morphemes correctly. *Developmental science, 10*(2), 237-254.

Noveck, I. A. (2001). When children are more logical than adults: investigations of scalar implicature. *Cognition*, *78*(2), 165-88.

Norbury, C. F. (2005). The relationship between theory of mind and metaphor: Evidence from children with language impairment and autistic spectrum disorder. *British Journal of Developmental Psychology, 23*(3), 383-399.

O'Keefe, D. J. (2007). Brief report: post hoc power, observed power, a priori power, retrospective power, prospective power, achieved power: sorting out appropriate uses of statistical power analyses. *Communication methods and measures*, *1*(4), 291-299.

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

Oller, D. K., Pearson, B. Z., & Cobo-Lewis, A. B. (2007). Profile effects in early bilingual language and literacy. *Applied Psycholinguistics, 28*(02), 191-230.

Paap, K. R. (2014). The role of componential analysis, categorical hypothesising, replicability and confirmation bias in testing for bilingual advantages in executive functioning. *Journal of Cognitive Psychology, 26*(3), 242-255.

Paap, K. R., Johnson, H. A., & Sawi, O. (2014). Are bilingual advantages dependent upon specific tasks or specific bilingual experiences?. Journal of Cognitive Psychology, 26(6), 615-639.

Paap, K. R., Johnson, H. A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. Cortex.

Paap, K. R., & Sawi, O. (2014). Bilingual advantages in executive functioning: Problems in convergent validity, discriminant validity, and the identification of the theoretical constructs. *Frontiers in Psychology: Language Sciences, 5*, 962. doi: 10.3389/fpsyg.2014.00962

Panizza, D., Chierchia, G. and Clifton, C., J. (2009). On the role of entailing patterns in the interpretation and processing of numerals and scalar quantifiers. *Journal of Memory and Language*, *61*, 503-518.

Papafragou, A., & Musolino, J. (2003). Scalar implicatures: Experiments at the semantics–pragmatics interface. *Cognition*, *86*(3), 253-282.

Paradis, J. (2010). Bilingual children's acquisition of English verb morphology: Effects of language exposure, structure complexity, and task type. *Language learning, 60*(3), 651-680.

Paradis, J. (2011). Individual differences in child English second language acquisition: Comparing child-internal and child-external factors. *Linguistic Approaches to Bilingualism, 1*(3), 213-237.

Paradis, J., Emmerzael, K., & Duncan, T. S. (2010). Assessment of English language learners: Using parent report on first language development. *Journal of communication disorders, 43*(6), 474-497.

Pouscoulous, N., Noveck, I., Politzer, G. and Bastide, A. (2007). A developmental investigation of processing costs in implicature production. *Language Acquisition*, *14,* 347-76.

Reichardt, C. S. (1979). The statistical analysis of data from non-equivalent group designs. In T. Cook & D. Campbell (Eds.), *Quasi-experimentation: Design & analysis issues for field settings* (pp. 147–205). Boston, MA: Houghton Mifflin.

Renfrew, C. (1995). *Word Finding Vocabulary Test*. 4th Edition. Oxon: Winslow.

Rowe, C., & Grohmann, K. K. (2013). Discrete bilectalism: Towards co-overt prestige and diglossic shift in Cyprus. *International Journal of the Sociology of Language, 2013*(224), 119-142.

Rundblad, G., & Annaz, D. (2010). Development of metaphor and metonymy comprehension: Receptive vocabulary and conceptual knowledge. *British Journal of Developmental Psychology, 28*(3), 547-563.

Rushton, J. P., Brainerd, C. J., & Pressley, M. (1983). Behavioral development and construct validity: The principle of aggregation. Psychological bulletin, 94(1), 18.

Siegal, M., Iozzi, L., & Surian, L. (2009). Bilingualism and conversational understanding in young children. *Cognition, 110*(1), 115-122.

Siegal, M., Matsuo, A., Pond, C., & Otsu, Y. (2007). Bilingualism and cognitive development: Evidence from scalar implicatures *Proceedings of the Eighth Tokyo Conference on Psycholinguistics* (pp. 265-280). Tokyo, Japan: Hituzi Syobo

Siegal, M., & Surian, L. (2007). Conversational understanding in young children. In E. Hoff & M. Shatz (Eds.), *Blackwell handbook of language development* (pp. 304–323). Oxford: Blackwell.

Siegal M, Surian L, Matsuo A, Geraci A, Iozzi L, et al. (2010) Bilingualism Accentuates Children's Conversational Understanding. PLoS ONE 5(2): e9004. doi:10.1371/journal.pone.0009004

Sperber, D., & Wilson, D. (1986/1995). *Relevance: Communication and cognition.* Oxford: Blackwell.

Sun, S., Pan, W., & Wang, L. L. (2011). Rethinking Observed Power. *Methodology*, *7*(3), 81−87.

Tao, L., Marzecová, A., Taft, M., Asanowicz, D., & Wodniecka, Z. (2011). The efficiency of attentional networks in early and late bilinguals: The role of age of acquisition.

Frontiers in Psychology, 2: 123, doi: 10.3389/fpsyg.2011.00123.

Terkourafi, Maria (1997). *The discourse functions of diminutives in the speech of Cypriot Greeks and Mainland Greeks.* Unpublished M.Phil. Dissertation, University of Cambridge. Cambridge, U.K.

Terkourafi, M. (2005). Understanding the present through the past: Processes of koineisation in Cyprus. *Diachronica, 22*(2), 309-372.

Theodorou, E., Kambanaros, M., & Grohmann, K. K. (2013). Specific language impairment in Cypriot Greek: diagnostic issues. Linguistic Variation, 13(2), 217-236.

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

Tomlinson Jr, J. M., Bailey, T. M., & Bott, L. (2013). Possibly all of that and then some: Scalar implicatures are understood in two steps. *Journal of memory and language, 69*(1), 18-35.

Valian, V. (2015). Bilingualism and cognition. *Bilingualism: Language and Cognition, 18*(01), 3-24.

Vogindroukas, I., Protopapas, A. & Sideridis, G. (2009). *Test of Expressive Vocabulary* [in Greek]. Chania: Glafki.

Vosniadou, S. (1987). Children and metaphors. *Child development*, 870-885.

Waggoner, J. E., & Palermo, D. S. (1989). Betty is a bouncing bubble: Children's comprehension of emotion-descriptive metaphors. *Developmental psychology, 25*(1), 152.

Waggoner, J. E., Palermo, D. S., & Kirsh, S. J. (1997). Bouncing bubbles can pop: Contextual sensitivity in children's metaphor comprehension. *Metaphor and Symbol, 12*(4), 217-229.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio: Pearson Assessment, Inc.

Winner, E. (1997). *The point of words: Children's understanding of metaphor and irony*. Cambridge, MA: Harvard University Press.

Winner, E., Levy, J., Kaplan, J., & Rosenblatt, E. (1988). Children's understanding of nonliteral language. *Journal of Aesthetic Education*, 51-63.

Yow, W. Q., & Markman, E. M. (2011a). Bilingualism and children's use of paralinguistic cues to interpret emotion in speech. *Bilingualism: Language and Cognition, 14*(04), 562-569.

Yow, W. Q., & Markman, E. M. (2011b). Young bilingual children's heightened sensitivity to referential cues. *Journal of Cognition and Development, 12*(1), 12-31.

Zinbarg, R. E., Suzuki, S., Uliaszek, A. A., & Lewis, A. R. (2010). Biased parameter estimates and inflated Type I error rates in analysis of covariance (and analysis of partial variance) arising from unreliability: Alternatives and remedial strategies. *Journal of Abnormal Psychology, 119*(2), 307-319.

Table 1: Bivariate correlations (Spearman's rho) between implicature scores, executive control and background variables[1].

	Relevance	Metaphor	Manner	Scalars act-out	Scalars binary	Implicature total	Version	Gender	Age	IQ	SES	Language ability-1	WM
Metaphor	.12												
Manner	.20*	.13											
Scalars act-out	.37**	.23**	.09										
Scalars binary	.18*	.16	.14	.30**									
Implicature total	.30**	.60**	.51**	.64**	.69**								
Version	-.02	-.33**	-.10	.01	-.07	-.23**							
Gender	-.03	.05	.08	.10	-.01	.09	-.10						
Age	.48**	.25**	.24**	.48**	.29**	.48**	.011	.03					
IQ	.37**	.14	.26**	.23**	.22*	.34**	.07	.02	.58**				
SES	.01	.07	.02	-.08	.12	.08	-.18*	-.02	-.12	.00			
Language ability-1	.46**	.20*	.21*	.44**	.31**	.46**	-.05	-.03	.47**	.36**	-.02		
WM	.26**	.12	.23**	.38**	.23**	.40**	.01	.01	.64**	.57**	-.04	.39**	
Inhibition[2]	.36**	.02	.07	.11	.13	.14	.12	-.02	.29**	.28**	.14	.23**	.25**

* Correlation is significant at the 0.05 level (2-tailed), **. Correlation is significant at the 0.01 level (2-tailed).

[1] All implicature and language scores are from tests taken in Cypriot Greek by multilinguals and bilectals and in Standard Modern Greek by monolinguals.

[2] Measure was reversed scored so that a higher value indicates better performance.

Note: Relevance=scores in the sub-test on relevance implicatures, Metaphor=scores in the sub-test on metaphors, Manner=scores in the sub-test on manner implicatures, Scalars act-out=scores in the act-out sub-test on scalar implicatures, Scalars binary=scores in the binary judgment sub-test on scalar implicatures, Implicature total=implicature composite score, Version=task version of the implicatures test, Age=participants' age in years, IQ=scores in the WASI matrix reasoning test, SES=socioeconomic status composite score, General language ability-1=general language ability composite score.

Table 2: Summary results of exploratory factor analysis on the four implicature measures.

Task (Measure)	Factor Loadings Factor 1: Pragmatics
Metaphor (n of correct responses)	**.56**
Manner (n of correct responses)	**.44**
Scalars binary (n of correct responses)	**.69**
Scalars act-out (n of correct responses)	**.70**
Eigenvalues	1.47
% of variance	36.79

Note: n=number, Metaphor=scores in the implicature sub-test on metaphors, Manner=scores in the implicature sub-test on manner implicatures, Scalars binary=scores in the binary judgment task on scalar implicatures, Scalars act-out=scores in the act-out task on scalar implicatures.

Table 3: Descriptive statistics (means and standard deviations) on background measures (raw values) in comparison 1 by language group.

Group	n		Age	FAS	LoPE	LoME	IQ	LC-Greek	WFVT-Greek
Bilectals	44	Mean	7;6	6	3	3.2	13.8	0.08	32.4
		(SD)	(0;9)	(1.9)	(1)	(0.8)	(6.5)	(0.7)	(5.2)
Multilinguals	26	Mean	7;7	6.9	3.8	4	16.4	0.17	19.8
		(SD)	(0;9)	(1.8)	(0.6)	(0)	(7.5)	(0.5)	(10.8)
Monolinguals	25	Mean	7;4	5.6	3.6	3.7	11.8	0.17	37.8
		(SD)	(0;9)	(1.5)	(0.8)	(0.6)	(5)	(0.4)	(4.9)

Note: n=number, Age=participants' age in years, IQ=scores in the WASI matrix reasoning test, LC-Greek=scores in the language comprehension test-Greek (taken in Cypriot Greek by bilectals and multilinguals and in Standard Modern Greek by monolinguals), FAS=scores in the Family Affluence Scale, LoPE=Level of Paternal Education, LoME=Level of Maternal Education, WFVT-Greek=scores in the Word Finding expressive Vocabulary Test (taken in Greek).

Table 4: Language characteristics of the multilingual group based on the Language Background Questionnaire in comparison 1.

Group	n		AoO CG	AoO En	CG Home	En Home	SMG	DoBE
Multilinguals	**26**	Mean	**1.9**	**9.9**	**2.6**	**2.5**	**2**	**.6**
		(SD)	**(4.6)**	**(20.6)**	**(0.9)**	**(0.9)**	**(0.9)**	**(.2)**

Note: n= number, SD=Standard Deviation, AoO CG=Age (in months) of onset of exposure to Cypriot Greek, AoO En=Age (in months) of onset of exposure to English, CG Home=amount of exposure to/use of Cypriot Greek at home (maximum score: 4), En Home=amount of exposure to/use of English at home (maximum score: 4), SMG=total amount of exposure to/use of Standard Modern Greek (maximum score: 4), DoBB=Degree of balanced exposure to/use of CG and English (range from 0 to 1, with 1 indicating more balanced exposure to/use of two languages).

Table 5: Descriptive statistics (means and standard deviations) on measures from the implicatures test (raw values) in comparison 1 by language group[1]. Results of between-group ANCOVAs (multilinguals versus bilectals versus monolinguals) on these measures are also reported.

Task	Bilectals (n=44)		Multilinguals (n=26)		Monolinguals (n=25)		F[5]	Effect size partial η^2
	Mean	(SD)	Mean	SD	Mean	SD		
All implicature measures combined	2.2	(0.5)	2.3	(0.6)	2.3	(0.6)	0.81[2]	.02
Relevance	2.9	(0.3)	2.9	(0.4)	3	(0)	n.a.	n.a.
Metaphor	2.1	(0.8)	2.1	(1)	2.4	(0.8)	1.73[2]	.04
Manner	2.1	(0.8)	2.1	(0.8)	1.9	(0.8)	0.56[3]	.01
Scalars act-out	2.4	(1.1)	2.6	(0.9)	2.6	(1)	1.92[3]	.04
Scalars binary	2.2	(1.2)	2.4	(1)	2.2	(1.1)	0.61[3]	.02
Fillers picture-selection	6.8	(0.4)	6.8	(0.5)	6.8	(0.4)	0.36[4]	.01
Fillers act-out	23.3	(2.4)	23.6	(1.4)	23.6	(1.3)	1.51[4]	.03
Fillers binary	16.2	(1.9)	16.6	(1.3)	16.6	(2)	0.46[4]	.01

Note: n.a.=not applicable, n=number, SD=Standard Deviation, Relevance=scores in the sub-test on relevance implicatures, Metaphor=scores in the sub-test on metaphors, Manner=scores in the sub-test on manner implicatures, Scalars act-out=scores in the act-out task on scalar impicatures, Scalars binary=scores in the binary judgment task on scalar implicatures, Fillers picture-selection=scores in the filler items of the picture-selection part of the implicatures test, Fillers act-out=scores in the filler items of the act-out task on scalar implicatures, Fillers binary=scores in the filler items of the binary judgment task on scalar implicatures.

[1]All implicature scores are from the implicatures test received in Cypriot Greek by multiliguals and bilectals and in Standard Modern Greek by monolinguals.

[2]F ratio resulting from an ANCOVA with Age, IQ, SES, and General language ability as covariates and Task Version and Group as between-subjects factors.

[3]F ratio resulting from an ANCOVA with Age, IQ, SES, and General language ability as covariates and Group as a between-subjects factor.

[4]F ratio resulting from an ANCOVA with Age, IQ, and SES as covariates and Group as a between-subjects factor.

[5]Degrees of freedom for the error term range from 79 to 83 depending on the analysis.

Table 6: Descriptive statistics (means and standard deviations) on background measures (raw values) in comparison 2 by language group.

Group	n		Age	FAS	LoPE	LoME	IQ	LC-Greek	WFVT-Greek	PPVT-Greek
Bilectals	17	Mean	7;6	5.9	3	3.3	14.1	0.1	30.6	95
		(SD)	(0;9)	(2)	(1.1)	(1)	(7.3)	(0.4)	(5.6)	(16.4)
Monolinguals	25	Mean	7;4	5.6	3.6	3.7	11.8	0.17	37.8	106.4
		(SD)	(0;9)	(1.5)	(0.8)	(0.6)	(5)	(0.4)	4.9	(16.5)

Note: n=number, SD=Standard Deviation, Age=Age in years, FAS=scores in the Family Affluence Scale, LoPE=Level of Paternal Education, LoME=Level of Maternal Education, IQ=scores in the WASI matrix reasoning test, LC-Greek=scores in the language comprehension test-Greek (taken in Cypriot Greek by bilectals and multilinguals and in Standard Modern Greek by monolinguals), WFVT=score in the Word Finding expressive Vocabulary Test (taken in Greek), PPVT-Greek=score in Peabody Picture receptive Vocabulary Test (taken in Cypriot Greek by bilectals and in Standard Modern Greek by monolinguals).

Table 7: Descriptive statistics (means and standard deviations) on measures from the implicatures test (raw values) in comparison 2 by language group[1]. Results of between-group ANCOVAs (bilectals versus monolinguals) on these measures are also reported.

Task	Bilectals (n=17)		Monolinguals (n=25)		F[6]	Effect size partial η^2
	Mean	(SD)	Mean	(SD)		
All implicature measures combined	**2.2**	**(0.5)**	**2.3**	**(0.6)**	**0.52**[2]	**.01**
Relevance	**2.9**	**(0.2)**	**3**	**(0)**	**n.a.**	**n.a.**
Metaphor	**2.1**	**(0.6)**	**2.4**	**(0.8)**	**6.03**[3*]	**.14**
Manner	**2.1**	**(0.8)**	**1.9**	**(0.8)**	**1.7**[4]	**.05**
Scalars act-out	**2.4**	**(1.1)**	**2.6**	**(1)**	**0.30**[4]	**.01**
Scalars binary	**2.4**	**(1.2)**	**2.2**	**(1.1)**	**2.27**[4]	**.05**
Fillers picture-selection	**6.8**	**(0.4)**	**6.8**	**(0.4)**	**0.09**[5]	**.00**
Fillers act-out	**23.9**	**(0.3)**	**23.6**	**(1.3)**	**0.53**[5]	**.01**
Fillers binary	**15.9**	**(1.8)**	**16.6**	**(2)**	**1.36**[5]	**.03**

Note: n.a.=not applicable, n=number, SD=Standard Deviation, Relevance=scores in the sub-test on relevance implicatures, Metaphor=scores in the sub-test on metaphors, Manner=scores in the sub-test on manner implicatures, Scalars act-out=scores in the act-out task on scalar impicatures, Scalars binary=scores in the binary judgment task on scalar implicatures, Fillers picture selection=scores in the filler items of the picture-selection part of the implicatures test, Fillers act-out=scores in the filler items of the act-out task on scalar implicatures, Fillers binary=scores in the filler items of the binary judgment task on scalar implicatures.

[1] All implicature scores are from the implicatures test received in Cypriot Greek by bilectals and in Standard Modern Greek by monolinguals.

[2] F ratio resulting from an ANCOVA with Age, IQ, and General language ability-2 as covariates and Task Version and Group as between-subjects factors.

[3] F ratio resulting from an ANCOVA with Age and General language ability-2 as covariates and Task Version and Group as between-subjects factors.

[4] F ratio resulting from an ANCOVA with Age, IQ, and General language ability-2 as covariates and Group as a between-subjects factor.

[5] F ratio resulting from an ANCOVA with Age and IQ as covariates and Group as a between-subjects factor.

[6] Degrees of freedom for the error term range from 34 to 38 depending on the analysis.

*p<.05

Table 8: Descriptive statistics (means and standard deviations) on measures from the implicatures test (raw values) in comparison 2 by language group[1]. Results of between-group ANCOVAs (bilectals versus monolinguals) on these measures are also reported.

Task	Bilectals (n=17)		Monolinguals (n=25)		F^4	Effect size partial η^2
	Mean	(SD)	Mean	(SD)		
All implicature measures combined	2.4	(0.7)	2.3	(0.6)	2.58^2	.06
Relevance	3	(0)	3	(0)	n.a.	n.a.
Metaphor	2.3	(0.8)	2.4	(0.8)	0.18^2	.01
Manner	2.3	(0.6)	1.9	(0.8)	5.48^{2*}	.13
Scalars act-out	2.6	(1)	2.6	(1)	0.51^2	.01
Scalars binary	2.3	(1.1)	2.2	(1.1)	0.83^2	.02
Fillers picture-selection	6.8	(0.4)	6.8	(0.4)	0.01^3	.00
Fillers act-out	23.9	(0.3)	23.6	(1.3)	0.95^3	.02
Fillers binary	16.6	(1.8)	16.6	(2)	0.02^3	.00

Note: n.a.=not applicable, n=number, SD=Standard Deviation, Relevance=scores in the sub-test on relevance implicatures, Metaphor=scores in the sub-test on metaphors, Manner=scores in the sub-test on manner implicatures, Scalars act-out=scores in the act-out task on scalar impicatures, Scalars binary=scores in the binary judgment task on scalar implicatures, Fillers picture selection=scores in the filler items of the picture-selection part of the implicatures test, Fillers act-out=scores in the filler items of the act-out task on scalar implicatures, Fillers binary=scores in the filler items of the binary judgment task on scalar implicatures.
[1]All implicature scores are from the implicatures test received in Standard Modern Greek by both bilectals and monolinguals.
[2]F ratio resulting from an ANCOVA with General language ability-SMG as a covariate and Group as a between-subjects factor.
[3]F ratio resulting from an ANOVA with Group as a between-subjects factor.
[4]Degrees of freedom for the error term range from 38 to 40 depending on the analysis.
*p<.05

Table 9: Results of multiple regression analysis on implicature performance based on the whole sample of children.

	B (SE)	B
Constant	1.629 (0.306)	
Age	0.135** (0.040)	0.333
WM	0.083 (0.073)	0.110
IQ	0.002 (0.008)	0.026
Version	-0.334** (0.88)	-0.27
General language ability-1	0.196** (0.071)	0.227

Note 1: R^2= .38. F-Test (5, 127)=15.465 (p<.05). * p<0.05, **p<.01

Note 2: Age=age in years, WM=working memory composite score, IQ=scores in the WASI matrix reasoning test, General language ability-1=general language ability composite score, Version=task version of implicatures test.

Table 10: Summary of studies included in the basic meta-analysis and associated effect sizes.

Study	Experiment	Bilinguals (n)	Monolinguals (n)	Effect size (Pearson's r)[1]	Total n
Siegal et al., 2007	n.a.	English-Japanese (20)	English (21)	.26*	41
Siegal et al., 2007	n.a.	English-Japanese (20)	Japanese (23)	.6**	43
Siegal et al., 2009	1	Slovenian-Italian (22)	Slovenian (19)	.5**	41
Siegal et al., 2009	2	Slovenian-Italian (38)	Slovenian (41)	.29*[2]	79
Siegal et al., 2009	2	Slovenian-Italian (38)	Italian (43)	.28*[2]	81
Siegal et al., 2010	1	German-Italian (36)	Italian (41)	.64**	77
Siegal et al., 2010	2	English-Japanese (33)	Japanese (59)	.30**[3]	92
Siegal et al., 2010	2	English-Japanese (33)	Japanese (59)	.27**[4]	92

Note: n=number, n.a.=not applicable

[1] Effect size for the comparison between bilinguals and monolinguals on overall performance in the Conversational Violations Test. A positive effect indicates better bilingual performance.

[2] Effect size calculated by averaging the effect sizes for the between-group comparisons on each individual maxim.

[3] Effect size for the comparison between bilinguals and monolinguals when bilinguals took the test in English and monolinguals in Japanese (and verbal mental age in English for bilinguals and in Japanese for monolinguals was covaried in the analysis).

[4] Effect size for the comparison between bilinguals and monolinguals when both groups took the test in Japanese (and verbal mental age in Japanese was covaried in the analysis).

* p<.05
** p<.01