Bilingualism: Language and Cognition

cambridge.org/bil

How does childhood bilingualism and bi-dialectalism affect the interpretation and processing of pragmatic meanings?

Kyriakos Antoniou[1,2], Alma Veenstra[3,4], Mikhail Kissine[5] and Napoleon Katsos[6]

[1]University of Cyprus, Center for Applied Neuroscience; [2]Hellenic Open University, School of Humanities; [3]Université libre de Bruxelles, Center for Research in Linguistics LaDisco; [4]University of Cambridge, Department of Theoretical and Applied Linguistics; [5]Université libre de Bruxelles, Center for Research in Linguistics LaDisco and [6]University of Cambridge, Department of Theoretical and Applied Linguistics

Research Article

Cite this article: Antoniou K, Veenstra A, Kissine M, Katsos N (2019). How does childhood bilingualism and bi-dialectalism affect the interpretation and processing of pragmatic meanings? *Bilingualism: Language and Cognition* 1–18. https://doi.org/10.1017/S1366728918001189

Received: 27 January 2018
Revised: 3 December 2018
Accepted: 9 December 2018

Key words:
pragmatic processing; pragmatic interpretation; implicature; bi-dialectalism; bilingualism

Address for correspondence:
Kyriakos Antoniou,
E-mail: antoniou.kyriakos@ucy.ac.cy

Abstract

Recent research has reported superior socio-communicative skills in bilingual children. We examined the hypothesis of a bilingual pragmatic advantage by testing bilingual, bi-dialectal and monolingual children on the comprehension and processing of various pragmatic meanings: relevance, scalar, contrastive, manner implicatures, novel metaphors and irony. Pragmatic responses were slower than literal responses to control items. Furthermore, children were least accurate with metaphors and irony. Metaphors and irony were also the most difficult to process; for these meanings, pragmatic responses were slower than literal responses to the same critical items. Finally, pragmatic performance positively correlated with working memory. Despite this variation, we found no bilingual or bi-dialectal advantage over monolinguals in pragmatic responses or speed of pragmatic processing. This was also true despite bilinguals' and bi-dialectals' lower vocabularies as measured by formal tests. We conclude that bilingual children exhibit monolingual-like pragmatic interpretation, despite their often-reported weaker language knowledge in the target language.

1. Introduction

A growing body of research has recently focused on the cognitive development of bilingual children (e.g., Barac, Bialystok, Castro & Sanchez, 2014; Hammer, Hoff, Uchikoshi, Gillanders, Castro & Sandilos, 2014; Paap, Johnson & Sawi, 2015). Two main trends emerge from this body of work. On the one hand, bilingual children are often reported to lag behind monolinguals in aspects of language development, such as vocabulary proficiency, when each of their languages is considered separately (e.g., Bialystok, Luk, Peets & Yang, 2010). This is, however, only a normal reflection of the different quantity and quality of language input that bilinguals receive in each of their languages as compared to monolinguals (e.g., because their language input is divided in two languages). When bilinguals' total or conceptual vocabulary (across their two languages) is taken into account, for instance, vocabulary differences with monolinguals often disappear (e.g., Hammer et al., 2014). On the other hand, some evidence suggests a positive effect of bilingualism on executive functions (Bialystok, 2017; Costa & Sebastián-Gallés, 2014; though see Paap et al., 2015 for counter-evidence and criticism), a set of cognitive processes that include working memory, inhibition, and switching (Miyake, Friedman, Emerson, Witzki, Howerter & Wager, 2000).

In the past years, language and executive functions have dominated the literature on the cognitive effects of bilingualism. However, other facets of bilingual children's cognitive development, such as pragmatic-communicative skills, have received relatively little research attention. Likewise, very few studies have examined the cognitive development of bi-dialectal children – that is, children who grow up with two structurally, lexically similar, genetically related, and (to some degree) mutually intelligible linguistic varieties, known as *dialects* (though see, e.g., Antoniou, Grohmann, Kambanaros & Katsos, 2016; Blom, Boerma, Bosma, Cornips & Everaert, 2017; Garraffa, Beveridge & Sorace, 2015; Grohmann & Kambanaros, 2016; Kirk, Fiala, Scott-Brown & Kempe, 2014; Ross & Melinger, 2017). In this study, we examined how bilingualism and bi-dialectalism affect children's understanding and processing of pragmatic meanings.

1.1 Theoretical background

According to Grice (1989), speaker's meaning can be rationally reconstructed in terms of intention attribution. For a speaker to mean something, with an utterance, is for this speaker to intend with that utterance to provoke a response (e.g., to induce a belief) in an addressee, such that the reason for the addressee's response is the recognition of this very intention.

CAMBRIDGE UNIVERSITY PRESS

Grice (1989) also accounted for cases of conversation where the speaker's meaning diverges from the conventional meaning of the words she/he used. In this account, communicative exchanges are presumed to abide by the Principle of Cooperation, which is subdivided in four conversational Maxims. According to these maxims, speakers are expected to be relevant (maxim of relevance), give no less and no more information than is needed for the purpose of the talk exchange (maxim of quantity), be brief, orderly and avoid ambiguity and obscurity (maxim of manner), and provide information that is true and for which they have adequate evidence (maxim of quality).

Conversational maxims are often not fulfilled at the literal level. This invites the listener to infer an interpretation that goes beyond what the speaker explicitly said. These implied meanings are what Grice called CONVERSATIONAL IMPLICATURES. The utterance *He was a sinking ship*, for instance, when taken literally, violates the maxim of quality. This invites the listener to infer an interpretation that makes the utterance truthful – i.e., that "He is feeling very sad" or a related figurative interpretation.

Grice himself did not intend for his ideas to be associated to cognitive-psycholinguistic models of language processing and development. Nevertheless, pragmatic theorists who have followed up on Grice (e.g., Levinson, 2000; Sperber & Wilson, 2002) share a cognitive-psychological perspective, and aim to provide an explanatory account of the mechanisms that deliver different aspects of meaning and of how these mechanisms develop in children.

The cognitive factors that underpin the development and interpretation of non-literal meanings are a matter of on-going debate. One proposal suggests that the relevant factor affecting pragmatic skills is language proficiency and that level of pragmatic development is in keeping with levels of structural language (see Andrés-Roqueta & Katsos, 2017). An alternative view is that understanding pragmatic language also depends on executive control (e.g., Huang & Snedeker, 2009b; Siegal & Surian, 2007) and/or Theory of Mind (e.g., Sperber & Wilson, 2002). Executive control skills, for instance, might be required for inhibiting the literal meaning of an utterance or for coordinating different pieces of information – literal meaning, encyclopedic knowledge, pragmatic maxims, speaker's knowledge and intentions – that are necessary for accessing a pragmatic interpretation. There is also some experimental evidence indicating that pragmatic understanding draws on working memory (WM). For instance, studies on adults' understanding of scalar implicatures (specifically, on the derivation of the inference that *some* implicates "not all")[1] have shown that burdening participants' WM resources with a secondary task resulted in fewer scalar implicature interpretations (De Neys & Schaeken, 2007; Marty & Chemla, 2013; Marty, Chemla & Spector, 2013; see also Antoniou, Cummins & Katsos, 2016; in addition, Chiappe & Chiappe, 2007 and Mashal, 2013 for WM effects on metaphor interpretation).

On the other hand, theoretical accounts inspired by Grice (e.g., Sperber & Wilson, 1986/1995; 2002) propose that pragmatic understanding involves reasoning about and considering the speaker's beliefs, knowledge and intentions (but see Breheny, 2006; Kissine, 2013, 2016; Levinson, 2000; Recanati, 2002). In this respect, it has been suggested that pragmatic intrerpretation depends on a Theory of Mind (ToM), the cognitive component that is responsible for ascribing mental states (e.g., beliefs, intentions) to oneself and/or to others and for interpreting the behaviour of others based on these mental states. Again, experimental evidence has shown that ToM-related processes (e.g., taking into account speaker's reliability or knowledge) are involved during online pragmatic interpretation (see e.g., Breheny, Ferguson & Katsos, 2013; Grodner & Sedivy, 2011; Spotorno & Noveck, 2014) and that successful pragmatic understanding in children may depend on the development of a Theory of Mind (Filippova, 2014).

1.2 Bilingualism and socio-communicative skills

Three studies conducted by Siegal and colleagues (Siegal, Iozzi & Surian, 2009; Siegal, Matsuo, Pond & Otsu, 2007; Siegal, Surian, Matsuo, Geraci, Iozzi, Okumura & Itakura, 2010) suggest enhanced pragmatic skills in bilingual preschool-aged children. Siegal et al. (2007) report that bilingual children were more likely than monolinguals to understand the scalar implicature associated with *some*. Similarly, in subsequent research, Siegal et al. (2009; 2010) found that bilingual children were better than monolinguals at indicating which of two statements was a pragmatically inappropriate reply to a question (i.e., violated a conversational maxim).[2] Siegal et al. (2009) propose two possible explanations for the bilingual pragmatic advantage in their study. First, they suggest that bilingual children exhibited enhanced pragmatic abilities because of their superior executive control (EC) skills. The second explanation is that bilingual children develop better pragmatic skills as a compensation for the initial lag they often exhibit in aspects of language acquisition.

Bilingual advantages have been also reported for other facets of socio-communicative development. Yow and Markman (2015), for instance, found that bilingual toddlers were more adept at using multiple cues to understand a speaker's communicative intent. More recently, Liberman, Woodward, Keysar and Kinzler (2017) reported that bilingual infants were more likely than monolinguals to interpret an experimenter's request as referring to a mutually visible object rather than to an identical object in their privileged perspective. Finally, various studies reported superior bilingual performance in ToM tasks for preschool-aged children (Barac et al., 2014).

Other research, however, has failed to find evidence for better pragmatic-communicative skills in bilinguals. Antoniou and Katsos (2017) tested Greek-speaking multilingual, bi-dialectal and monolingual children (aged six to nine years) in various pragmatic meanings (relevance, scalar, manner implicatures and novel metaphors). They hypothesised that multilingual children would exhibit better pragmatic understanding skills than monolinguals for two reasons. First, because the previous study by Siegal et al. (2007) found that bilingual children outperformed monolinguals in scalar implicature (SI) understanding; and second, because previous evidence indicated that bilingual children excel in almost all factors that have been suggested to underlie pragmatic interpretation (ToM, EC, sensitivity to Gricean maxims). However, Antoniou and Katsos (2017) found no group differences in pragmatic comprehension. Null differences were also reported by Syrett, Austin, Sánchez, Germak, Lingwall, Perez-Cortes, Arias-Amaya and Baker (2016) and Syrett, Lingwall, Perez-Cortes, Austin, Sánchez,

[1]Such inferences are known as scalar implicatures because they are generated based on scales which order lexical terms (e.g., *some*, *all*) with respect to the strength of the information that they convey (Horn, 1972). Since the speaker used the less informative term *some*, then, by Grice's maxim of quantity, this implicates that the more informative term *all* does not hold.

[2]A meta-analysis of the studies by Siegal and colleagues indicates a medium-to-large bilingual effect on pragmatics (r=.39; Antoniou & Katsos, 2017).

. https://doi.org/10.1017/S1366728918001189

Baker, Germak and Arias-Amaya (2017), who compared Spanish-speaking bilingual and monolingual children (aged three to six years) in SI interpretation.[3]

To sum, bilingual children have been found to exhibit advantages in various pragmatic-communicative and social skills, though not necessarily in pragmatic comprehension. Moreover, bilingual benefits have been most consistently reported for preschool-aged children. Thus, type of social or pragmatic skills and age seem to be two factors that can possibly explain why the bilingual benefit is found in some studies but not in others.

1.3 The present study

In this context, our study aimed to achieve four main goals. Firstly, we wanted to test the hypothesis of a bilingual advantage in pragmatic comprehension using a more sensitive experimental design than previous studies. Our study included a larger sample of participants ($n = 138$) and more critical items ($n = 12$) than past research on bilingual children's understanding of non-literal meanings (apart from Syrett et al., 2017 who tested 139 children in two experiments and Antoniou & Katsos, 2017, who used 15 critical items). In addition, we examined the comprehension of irony, which is generally thought to be the most demanding pragmatic meaning for children (Filippova, 2014). Thus, the items on irony are possibly more sensitive to group differences. Finally, our study examined two different measures of pragmatic comprehension (accuracy and reaction times) while previous studies focused only on accuracy. All the above in combination suggest that the current experiment is more likely to detect a bilingual advantage in pragmatic understanding, if such an effect exists.

Second, we aimed to examine how bilingualism/bi-dialectalism affects the interpretation of pragmatic meanings that have not been previously tested in bilinguals/bi-dialectals (specifically, contrastive implicatures and irony). We expected a bilingual advantage particularly in irony comprehension for various reasons. As already noted, irony is the most difficult and late-developing pragmatic meaning for children (Filippova, 2014). It is possible that a bilingual advantage in older children is found only when using very demanding tasks. This is because older children have possibly reached an advance level of cognitive development and, hence, any bilingual advantages are more difficult to detect. Moreover, a previous study by Yow and Markman (2011) reported that bilingual toddlers were more adept than monolinguals at using tone of voice to judge a speaker's emotion behind an utterance when semantic content and intonation conflicted. This situation resembles cases of irony where intonation indicates an interpretation different from the utterance's literal meaning.[4]

Third, we wanted to examine how bilingualism/bi-dialectalism affects pragmatic processing. We hypothesised a bilingual advantage in pragmatic processing (i.e., faster pragmatic interpretations) for various reasons. First, experimental evidence suggests that proficient bilinguals can achieve native-like semantic and grammatical processing, even in their second language (e.g., Clahsen & Felser, 2006b). On the other hand, there is some evidence that bilinguals are more sensitive to pragmatic cues during language processing than native speakers, especially in their second language (Roberts & Felser, 2011; Roberts, Gullberg & Indfrey, 2008; see also in Clahsen & Felser, 2006a; 2006b; Foucart, Garcia, Ayguasanosa, Thierry, Martin & Costa, 2015). Moreover, other research suggests that pragmatic processing draws on cognitive resources like WM and ToM, which are thought to be domains of bilingual strength. Findings with adults, for instance, indicate that the processing of pragmatic meanings is associated with a reaction-time cost as compared to literal meanings (e.g., Bott & Noveck, 2004; Breheny, Katsos & Williams, 2006; Deliens, Antoniou, Clin, Ostaschenko & Kissine, 2018; Huang & Snedeker, 2009a). This further suggests that pragmatic interpretation is a non-automatic process that draws on executive functions. Indeed, as already noted, studies with adults have shown that SI processing depends on WM resources (De Neys & Schaeken, 2007; Marty & Chemla, 2013). Finally, experimental evidence on irony has revealed that irony processing activates brain regions associated with ToM (Spotorno & Noveck, 2014).

Our final goal was to contribute to the debate regarding the cognitive factors that underpin pragmatic understanding. If bilingual and/or bi-dialectal children exhibit the expected pattern of lower language proficiency in the language of testing but better (or equal) pragmatic skills as compared to monolinguals, this will be evidence against an account which predicts that pragmatic understanding depends solely on structural language skills in the target language. We should note that, in this study, we were interested in the effect of language proficiency only in the language of testing because, based on previous research, we had reasons to believe that bilinguals and bi-dialectals will exhibit lower language performance when tested in only one of their languages. This naturally poses the question of whether having lower proficiency in a given language affects pragmatic interpretation in that language. We were also interested in possible effects of language knowledge per se (not some other aspect of general language use or cognitive skill that possibly correlates with language proficiency/knowledge). For instance, when interpreting implicatures in a given language, listeners need to recruit language knowledge from that language to process and understand (at least to some degree) the semantics of target utterances. There is no reason to suppose that language knowledge in the other language (or total language knowledge across languages) is necessary in this process.

As previously mentioned, bilingual and bi-dialectal children tend to exhibit enhanced EC skills as compared to monolinguals. If pragmatic comprehension draws on executive functions, then we would expect bilinguals and/or bi-dialectals to also excel in pragmatic understanding. Our study, however, offers the opportunity to examine this question irrespective of group differences in language proficiency or EC. By directly looking at the correlations between executive functions, language proficiency and pragmatic performance, we can provide direct evidence regarding the independent effects of these two factors (if any).

Finally, we were especially interested in whether bilingual/bi-dialectal children understand pragmatic language by relying

[3]The studies conducted by Syrett and colleagues (2016; 2017) are also interesting because the bilingual children spoke two languages (Spanish and English) which differ in the number of lexical items that correspond to the English term *some*. While English has only the lexical item *some*, Spanish has two such terms, *unos* and *algunos*. As Syrett et al. (2016; 2017) point out, *unos* is more tightly associated with a "some and possibly all" semantic interpretation, while *algunos* is more tightly connected to the scalar meaning "some but not all" (which still seems to be pragmatic in nature given that it is cancellable). Despite this difference between the languages spoken by bilinguals, however, their results revealed largely equivalent scalar implicature responses in bilingual and monolingual children.

[4]We do not think that experiential factors could possibly confound the group comparisons on irony. We have no reason to expect that the three groups in our study differed in their experience with irony, especially since the concept of irony is generally considered to be universal (e.g., Wilson & Sperber, 2003) and since our bilinguals spoke an additional language (French) or dialect (West Flemish) which are relatively closely related to Dutch, the language of testing (and the language of the monolinguals).

. https://doi.org/10.1017/S1366728918001189

differently (e.g., more heavily) as compared to monolinguals on their EC resources.

In this study, a bilingual, a bi-dialectal and a monolingual group of Dutch-speaking children were administered a novel task (in Dutch) on various pragmatic meanings: relevance, scalar, contrastive, manner implicatures, novel metaphors and irony. The task was designed on E-Prime (Psychology Software Tools, Pittsburgh, PA, 2012) and both accuracy and reaction times (RTs) were recorded. Children were further given two vocabulary tests only in Dutch. Various EC tasks were also administered as part of another study (Veenstra, Antoniou, Katsos & Kissine, 2018).

We tested children between 10–12 years of age for three reasons. First, given that bilingual and bi-dialectal children in our study were exposed to Dutch primarily through education, we wanted to ensure that they had sufficient experience and knowledge of that language. Second, we tested older children because we wanted to examine pragmatic meanings such as irony, which previous research has shown to develop until late in childhood (Filippova, 2014). Thirdly, we wanted to test children who would be old and competent enough to understand the pragmatic meanings examined and to perform unaided a computerised task recording RTs.

We also tested bi-dialectal children because we wanted to examine how bi-dialectals pattern in terms of their cognitive skills as compared to monolinguals and bilinguals. Moreover, we were interested in how the close language similarity between the varieties spoken by bi-dialectals possibly affects the cognitive outcomes of bilingualism.

2. Method

2.1 Participants

Approval for this study was obtained from the Ethical board of the Université libre de Bruxelles. All parents gave informed consent for their children's participation. Participants included 46 bi-dialectal children (in Dutch and West Flemish; 23 girls; aged 121–155 months, mean age 136.2, *SD* 8.9 months), 48 bilinguals (in Dutch and French; 28 girls; aged 121–144 months, mean age 132.5, *SD* 6.7 months), and 44 monolingual children (speakers of Dutch; 25 girls; aged 121–145 months, mean age 132.5 months, *SD* 6.8 months). A power calculation using G* Power (Faul, Erdfelder, Lang & Buchner, 2007) indicated that our study had a power of .99 to detect a group effect of $r = .39$ (see Antoniou & Katsos, 2017).[5] This level of power is well above Cohen's (1988) recommended level of .8.

Bi-dialectal children were recruited in West-Flanders (Belgium), bilinguals in Brussels (Belgium), and monolinguals in Eindhoven (the Netherlands). All children were educated in Dutch. Dutch instruction started in daycare and kindergarten, around the age of 2;6 years. Bilingual children were dominant in French and spoke exclusively French at home, while bi-dialectals used exclusively West Flemish at home.

2.2 Materials and procedure

Participants were tested in three sessions taking approximately 45 minutes each. For EC, we used the following measures (tasks are given in parentheses): interference effect (Attentional Networks Task; Rueda, Fan, McCandliss, Halparin, Gruber, Lercari & Posner, 2004), switch cost (Colour-Shape task; Ellefson, Shapiro & Chater, 2006), and number of correctly recalled trials (in the forward and backward Corsi Blocks task and Digit Recall task; Mueller & Piper, 2014 and Kort, Schittekatte & Compaan, 2008, respectively). These tasks and measures are described in Veenstra et al. (2018).

Pragmatics test

Each type of pragmatic meaning was based on one of Grice's (1989) maxims. In the relevance sub-test, an utterance such as *It's raining* as a reply to the question *What kind of item do you want?*, when taken literally, violates the maxim of relevance. This invites the listener to infer a relevant interpretation, i.e., that (between a winter hat, an umbrella, and a pair of shorts) the speaker wants the umbrella. For SIs, the use of the term *some* invites the listener to infer (by Grice's maxim of quantity) that *all* does not hold. For contrastive implicatures, the use of a modified noun phrase such as *Open window* to describe a single window would be over-informative. This implicates that there is another window in context. For manner implicatures, a description such as *Vehicle with two wheels* (in the context of a motorcycle and an atypical vehicle with two wheels) implicates that "The speaker refers to the atypical vehicle". This is because the speaker used a more prolix sentence instead of the simpler description *Motorcycle*. Finally, for novel metaphors and irony, the violation of the maxim of quality at the literal level (e.g., when a speaker utters *Yes, you know how much I like red clothes!*, but interlocutors know that she/he definitely hates red clothes) invites the listener to infer a truthful interpretation (e.g., "No, you know that I definitely hate red clothes!"). Each sub-test is detailed below.

All sub-tests, apart from SIs, had a picture-selection format. The participant heard a target statement, and was then presented with three pictures from which she/he had to select one. For SIs, children had to judge whether a statement was a correct or incorrect description of a visual display. Items were recorded in Dutch by a native Dutch speaker for monolinguals and by a native Belgian Dutch speaker for the other groups. All verbal stimuli, apart from ironic sentences, were pronounced with a neutral intonation without any contrastive focus accent or non-default prosody. Ironic statements were produced with a distinctive, exaggerated, ironic prosody.

For critical items of picture-selection sub-tests, one picture corresponded to the pragmatic meaning. For novel metaphors, irony, and manner implicatures, a second picture was compatible with a literal interpretation. The other pictures were incorrect matches to both the semantic and pragmatic meaning of target statements. For the SI sub-test, rejection of the critical statements as incorrect (e.g., *There are stars on some of the cards* as a description of a display where all cards depicted stars) indicated a SI interpretation whereas acceptance of the same items indicated a literal interpretation.[6]

In sum, there were six sub-tests with two critical and four control or filler items each. We used only two critical items for each pragmatic meaning because we wanted to avoid effects of fatigue or the children losing interest in the task (the task took approximately 25 minutes). In addition, item-level power is higher when considering the analyses on overall pragmatic performance which average responses to all critical items.

For relevance, scalar, manner implicatures, novel metaphors and irony, (some) non-implicature items were control items;

[5]We first converted the effect size *r* to an effect size *f* (the effect used by G*Power). The conversion indicated that r=.39 equals f=.42.

[6]See Horn (1972) that *some* is semantically compatible with *all*.

. https://doi.org/10.1017/S1366728918001189

that is, they were comparable to the implicature items in semantic meaning and linguistic form (for SIs, irony and two items from novel metaphors and relevance implicatures), or in terms of communicated meaning (for two items from novel metaphors), or in terms of semantic meaning (for two items from manner implicatures). For contrastive implicatures, all non-implicature items were fillers.

In each sub-test, the children performed two practice trials. In all sub-tests, there were three conditions (Pragmatic, Literal-1, Literal-2). For metaphors, each condition began with different instructions and its own practice trial because initial pilot testing showed that participants found the general instructions confusing.

There were three versions of the pragmatics test. For each critical trial in one version, the second and/or third versions included the control trial(s) for that critical trial. An approximately equal number of children (within each group) was tested with each version.

Children were instructed to respond as fast and accurate as possible. Reaction times were recorded as soon as the final response slide appeared (after the target statement). The SI sub-test was administered first, and the other sub-tests were then randomly presented. The linguistic stimuli from the pragmatics test can be found in the online supplementary material (Supplementary Material). Tables 1 and 2 provide sample trials from the (picture-selection) sub-test on irony and the SI judgment task, respectively.

Relevance implicatures

Children were presented with a character, named Thomas and a female speaker, who asked Thomas to describe which of three items/pictures he wanted. They had to select one item/picture based on Thomas's reply. All trials were randomly presented.

For critical trials, the speaker initially urged Thomas to wear a blindfold. This was necessary to motivate why Thomas replied indirectly. The target question was then heard (e.g., *What kind of item do you want?*). Next, participants heard the target statement (e.g., *It's snowing*) and, immediately after that, saw the final slide depicting three items (e.g., a coat, a belt and a hat).

The same procedure was followed for Literal-1 control trials (e.g., *It's raining*) with the difference that the speaker did not urge Thomas to wear a blindfold; she introduced three pictures, and then asked the target question (e.g., *How is the weather in the picture that you want?*). For the Literal-2 filler trials (e.g., *It's shining*), the speaker introduced three items in separate slides and, then, asked Thomas to describe what he wanted.

Scalar implicatures

Participants heard a character, named Bram, describing various visual displays with five cards. They had to judge whether Bram's utterances were correct or incorrect descriptions of the visual displays.

In each trial, an auditory stimulus was played, *There are* < *X* > *on* < *Q* > *of the cards*, where X was the item (moons, squares, rings, hearts, suns, stars) and Q the quantifier (*all*, *some*, *none*). The cards were then immediately 'turned over' to reveal the items.

There were two critical items with *some*. True-and-informative *some* and false *some* statements (Literal-1 and Literal-2 conditions, respectively) served as controls.

Contrastive implicatures

This sub-test was based on a task previously used by Kronmüller, Morisseau and Noveck (2014). Children were instructed that a character named Martijn would describe one of two cards (visible only to himself) and they had to guess what his second card depicted. Trials were presented randomly.

Critical statements included modified or bare nouns such as *Open window* or *Window*. For each trial, the children heard Martijn describing one card. The next slide included a written question *In your opinion, which one is Martijn's second picture?* In the final slide, participants saw three pictures (e.g., a closed window, a dog and an open window).

In the Literal-2 condition, Martijn explicitly provided a hint about what his second card depicted. In the Literal-1 condition, he did not provide any hints about his second card and participants had to select a card at chance.

Manner implicatures

Participants were informed that for each trial they had to select the picture described by a character named Sander. Critical statements included modified noun phrases with generic nouns (e.g., *Vehicle with two wheels*). For each trial, the children heard a female speaker asking Sander to describe a picture. They heard the target description and, finally, viewed three pictures (e.g., a motorcycle, an atypical vehicle with two wheels, a car).

For critical trials, one picture corresponded to a pragmatic interpretation, the second was compatible with a literally true but pragmatically inappropriate interpretation (competitor), and the third depicted another incorrect referent. Trials were presented in a fixed order so that, for each critical statement, the corresponding simpler noun phrase (e.g., *Motorcycle*) was introduced in a previous trial. This was necessary to avoid the participants thinking that the speaker did not know the label for the competitor, which would make the critical statement ambiguous.

Novel metaphors

The sub-test was designed based on Waggoner and Palermo (1989). In this sub-test, we used six (from a list of 12) metaphors that were judged as novel (mean novelty rating of 3.09/5) but also ranked high for aptness (lowest mean rating was 1.97/5) by 32 adult Dutch native speakers.

Children were presented with a character, named Niels, and a female speaker who asked questions. They had to select one of three pictures depicting a sad man, an angry man or a situation compatible with a literal-like interpretation of a metaphor.

In the Metaphor condition, children heard stories about Niels's father. The speaker asked the target question (e.g., *Niels, how was your father feeling when he found out?*) and the target statement was heard (e.g., *He was a thunderous storm cloud*). Children had to select the picture showing how Niels's father felt.

The same procedure was followed in Literal-1 (e.g., *He was a sad man*) and Literal-2 (e.g., *It was a sinking ship*) control trials. In the Literal-2 condition, however, the children were instructed to select the picture that indicated what happened at the end of each story.

There were three blocks, one for each condition. The blocks and the trials within each block were randomly presented.

Irony

This sub-test was based on a task previously used by Kowatch, Whalen, and Pexman (2013). A female speaker asked a character, named Wouter, questions about which item he wanted. Participants had to give Wouter the item he wanted. Trials were presented in a fixed order, so that each critical trial appeared only after a Literal-1 (e.g., *Yes, you know how much I like*

. https://doi.org/10.1017/S1366728918001189

Table 1. Example critical trial from the irony sub-test

Sequence	Pictorial stimulus	Auditory stimulus
1		Wouter, I know that you like playing football and that you definitely don't like reading in your free time. But reading a story book could be interesting.
2	Toy car	Here is a toy car
3	Story book	Here is a story book
4	Football	And here is a football
5		Would you like to read the story book, now?
6	[Press the SPACE BAR to hear Wouter's reply…]	
7	[Press the SPACE BAR to hear Wouter's reply…]	Yes, you know how much I like reading in my free time!
8	Toy car, story book, football	

Note. Stimuli in brackets were presented in written form.

Table 2. Example critical trial from the scalar implicatures sub-test

Sequence	Pictorial stimulus	Auditory stimulus
1	[Press the SPACE BAR to continue…]	
2	Five cards face down	There are moons on some of the cards.
3	Five cards 'turned over', each depicting a moon	

Note. Stimuli in brackets were presented in written form.

vegetables for lunch!) and Literal-2 (e.g., *No, you know how much I hate red clothes!*) control trial.

For all trials, the speaker initially gave contextual information regarding Wouter's preferences. She then introduced three items and the target question was heard. Next, the target statement was presented and, then, three items appeared on the screen.

Socioeconomic status and language background questionnaire

This questionnaire (based on Paradis, 2011; Paradis, Emmerzael & Duncan, 2010) asked for information regarding the child's language use, among other topics. It also included three indicators of socioeconomic status (SES): the Family Affluence Scale (FAS; Boyce, Torsheim, Currie & Zambon, 2006) and parents' levels of education.

Language measures

Receptive vocaburaly in Dutch was tested using the Peabody Picture Vocabulary Test III-NL (PPVT; Dunn, Dunn & Schlichting, 2005) and Dutch expressive vocabulary with the Word Definitions task from the CELF 4-NL (Semel, Wiig, Secord & Kort, 2008).

3. Results

Accuracy in the pragmatics test was analysed with generalized linear mixed-effects models using the logit link function from the lme4 package in RStudio (Bates, Maechler, Bolker & Walker, 2015; RStudio Team, 2016). Reaction times were analysed with linear mixed-effects models, using the same package. The significance of the fixed effects was assessed with likelihood ratio tests where a model containing the fixed effect was compared to an identical model without it (Barr, Levy, Scheepers & Tily, 2013). For follow-up contrasts, we used the multcomp package in R (Hothorn, Bretz, Westfall, Heiberger, Schuetzenmeister, Scheibe & Hothorn, 2016) and the default single-step method to correct for multiple comparisons.

For the comparison models, we attempted to specify the maximal random effects structure (Barr et al., 2013). Moreover, we did not include random slopes for control variables (Barr et al., 2013). To deal with non-convergence, we followed the procedure described in Barr et al. (2013). Where singularity was observed, we dropped random effects associated with zero variance or a random correlation of +/-1 (Bolker, Brooks, Clark, Geange, Poulsen, Stevens & White, 2009).

All other between-group comparisons were conducted using Analyses of (Co)Variance (Bonferroni correction was applied for multiple constrasts). Finally, all between-group analyses included as control variables background measures for which significant group differences were found.

3.1 Preliminary analyses

Figure 1 shows accuracy by Condition (Pragmatic, Literal-1, Literal-2) and sub-test. Accuracy data from the relevance sub-test was excluded from subsequent analyses because a ceiling effect was observed (93%).

To reduce the number of variables entered into subsequent analyses and to increase reliability of measurement for these variables (Carlson, 2003; Rushton, Brainerd & Pressley, 1983), we created composite scores for variables that were conceptually and statistically related. These were calculated by transforming into z scores and averaging the individual measures (e.g., Carlson & Meltzoff, 2008). A vocabulary composite score was computed based on the PPVT and the Word Definitions test, and a SES composite measure was calculated based on maternal, paternal level of education and FAS. Moreover, Veenstra et al. (2018) created composite scores (as above) for three EC factors (based on a Principal Component Analysis): Non-verbal WM (from scores in the forward and backward Corsi Blocks task), Verbal WM (from scores in the forward and backward Digit Recall task) and Inhibition (from the switching cost and interference effect). Correlations between accuracy in the pragmatic sub-tests, Age, Vocabulary, SES, Task Version and EC are presented in Table 3.

. https://doi.org/10.1017/S1366728918001189

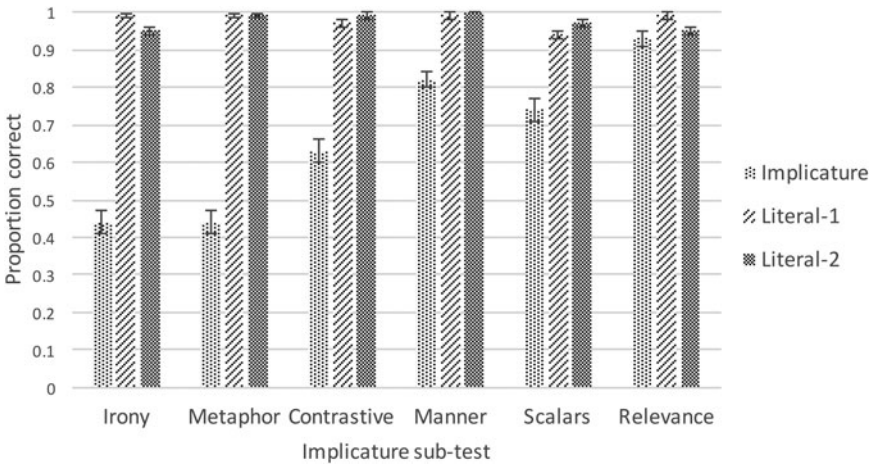

Fig. 1. Accuracy (in proportions) by Condition (Literal-1, Literal-2, Pragmatic) in each pragmatic sub-test (Irony, Metaphor, Contrastive, Manner, Relevance and Scalar implicatures).

3.2 Main analyses

Background measures

There were significant group differences in age in months ($F(2, 135) = 3.625$, $p < .05$), SES ($F(2, 135) = 80.56$, $p < .05$) and Vocabulary ($F(2, 133) = 9.944$, $p < .05$). Regarding age, there was a trend for bi-dialectals being older than both bilingual ($p = .06$) and monolingual children ($p = .07$). For SES, bilinguals had a higher SES than bi-dialectal and monolingual children, and monolinguals had a higher SES than bi-dialectals ($ps < .05$). Moreover, monolinguals had a higher Dutch vocabulary than bi-dialectal and bilingual children (all $ps < .05$). Finally, as reported in Veenstra et al. (2018), there were no significant group differences in EC.

Pragmatics test: Accuracy

Descriptive statistics for accuracy by Type of Pragmatic Meaning (Relevance, Scalar, Contrastive, Manner, Metaphors, Irony), Condition (Pragmatic, Literal-1, Literal-2) and Group (Monolinguals, Bi-dialectals, Bilinguals) are reported in Table 4.

Effect of Type of Pragmatic Meaning and Task Version

Models with by-subject and by-item random intercepts revealed a significant effect of Type ($\chi^2(4) = 30.898$, $p < .05$), but a non-significant effect of Version ($\chi^2(2) = 1.21$, $p > .05$) or Type by Version interaction ($\chi^2(8) = 15.271$, $p = .054$). Pairwise contrasts for Type indicated that accuracy in the scalars sub-test was significantly higher than for metaphors (estimate=−1.41, $SE = 0.32$, z=−4.38, p < .05) and irony (estimate=−1.41, $SE = 0.32$, z=−4.38, p < .05); performance for contrastive implicatures was significantly lower than for manner (estimate = 1.06, $SE = 0.33$, $z = 3.16$, $p < .05$), but did not significantly differ from metaphors (estimate=−0.87, $SE = 0.32$, z=−2.73, $p = .07$) and irony (estimate=−0.87, $SE = 0.32$, z=−2.73, $p = .07$); and accuracy was significantly higher in manner than in the metaphor (estimate=−1.93, $SE = 0.34$, z=−5.73, $p < .05$) and irony sub-tests (estimate=−1.92, $SE = 0.34$, z=−5.7, $p < .05$).

Effect of Group

We specified a model with Group, Type, and their interaction, SES and Age as covariates, by-subject random slopes for Type and by-item random intercepts.

Results showed that neither Group ($\chi^2(2) = 5.04$, $p = .08$) nor the interaction were significant ($\chi^2(8) = 6$, $p > .05$). Given the group differences in vocabulary, we also included Vocabulary as an additional covariate in the models. These analyses indicated again a non-significant Group effect ($\chi^2(2) = 4.289$, $p > .05$) and a non-significant interaction ($\chi^2(8) = 8.42$, $p > .05$). These results are presented in Table 5.

The relation between pragmatic accuracy, SES, Age, Vocabulary and Executive Functions

Initial analyses on pragmatic accuracy with Type, Verbal WM, Non-verbal WM, Inhibition, Vocabulary, SES and Age as predictors showed no significant effects (besides for Type). However, given that Verbal and Non-Verbal WM significantly correlated with each other (see Table 3) and to benefit from the increased reliability resulting from combining four WM indicators, we repeated the above analyses using a single WM composite score. This analysis showed that only WM positively predicted pragmatic performance ($\chi^2(1) = 3.77$, $p = .052$; estimate = 0.19, $SE = 0.09$, $z = 1.98$, $p = .048$; see also Table 3). These results are presented in Table 6.

Moreover, we examined whether the WM effect was qualified by Group and/or Type. The three-way interaction between Group, Type and WM was significant ($\chi^2(8) = 18.88$, $p < .05$). The interaction is illustrated in Figure 2. It indicates positive relations between WM and accuracy in most sub-tests, particularly for bilinguals and monolinguals. However, in the manner sub-test, there were negative correlations between WM and accuracy for bilinguals and bi-dialectals and a positive relation for monolinguals.

Thus, WM positively predicted pragmatic performance, but there is no clear evidence that bilinguals/bi-dialectals achieve pragmatic interpretation by relying differently on these resources relative to monolinguals.

Pragmatics test: Reaction times

Table 7 reports descriptive statistics for RTs by Type (Relevance, Scalar, Contrastive, Manner, Metaphors, Irony), Condition (Pragmatic, Literal-1, Literal-2) and Group (Monolinguals, Bi-dialectals, Bilinguals).

Pragmatic versus literal interpretations

We first examined whether pragmatic interpretations were slower than literal responses. However, for the reader's convenience, we only present a summary of these results here. The full analyses are presented in the online supplementary material (Supplementary Material).

Table 3. Bivariate correlations (Spearman's rho) between pragmatic scores, task version, executive control, vocabulary, and background variables

	Relevance	Metaphor	Manner	Contrastive	Scalars	Irony	Pragmatics	Version	Age	SES	Vocabulary	Verbal WM	Non-verbal WM	Inhibition[1]
Metaphor	.12													
Manner	.13	−.12												
Contrastive	.09	.19*	.16											
Scalars	.07	.10	.04	.07										
Irony	−.01	−.03	−.10	.12	−.05									
Pragmatics	.17*	.51**	.29**	.57**	.45**	.49**								
Version	−.12	−.21*	.06	−.15	−.02	−.05	−.17*							
Age	.01	.02	−.003	.15	−.06	−.11	.06	−.06						
SES	.06	.07	.01	.19*	.07	.03	.11	.05	−.14					
Vocabulary	−.06	.12	−.18*	.09	.06	.06	.07	−.04	.20*	.10				
Verbal WM	−.06	.13	−.08	.19*	.03	.09	.18*	.08	.12	.29**	.20*			
Non-verbal WM	.10	.17	−.09	.14	.09	−.03	.14	−.06	.07	.25**	.14	.22*		
Inhibition[1]	−.05	.05	−.13	−.05	.04	.04	−.04	−.06	.10	−.08	.04	−.06	.01	
WM	.01	.20*	−10	.24**	.06	.06	.23*	.01	.12	.35*	.21*	.77**	.75**	−.03

*. Correlation is significant at the 0.05 level (two-tailed), **. Correlation is significant at the 0.01 level (two-tailed).
[1] Measure was reverse scored so that a higher value indicates better performance.
Note. Relevance = accuracy for the critical items of the sub-test on relevance implicatures, Metaphor = accuracy for the critical items of the sub-test on metaphors, Manner = accuracy for the critical items of the sub-test on manner implicatures, Contrastive = accuracy for the critical items of the sub-test on contrastive implicatures, Scalars = accuracy for the critical items of the sub-test on scalar implicatures, Irony = accuracy for the critical items of the sub-test on irony, Pragmatics = overall accuracy for the critical items of the pragmatics test (calculated by averaging accuracy in all pragmatic sub-tests, besides relevance), Version = task version of the pragmatics test, Age = participants' age in months, SES = socioeconomic status composite score, Vocabulary = vocabulary composite score, Verbal WM = verbal working memory composite score, Non-verbal WM = non-verbal working memory composite score, Inhibition = inhibition composite score.

Table 4. Descriptive statistics (proportion correct and standard deviations) from the pragmatics test (raw values) by type of pragmatic meaning, condition and language group

Type/ Condition	Bi-dialectals ($n = 46$)		Monolinguals ($n = 44$)		Bilinguals ($n = 48$)	
	Proportion	(*SD*)	Proportion	(*SD*)	Proportion	(*SD*)
Scalars						
Pragmatic	.67	(.47)	.80	(.41)	.76	(.43)
Literal-1	.90	(.3)	.97	(.18)	.96	(.2)
Literal-2	.97	(.18)	.98	(.15)	.98	(.14)
Contrastive						
Pragmatic	.57	(.5)	.72	(.45)	.61	(.49)
Literal-1	.95	(.22)	.97	(.18)	.99	(.1)
Literal-2	1	n.a.	1	n.a.	.96	(.20)
Manner						
Pragmatic	.85	(.36)	.78	(.41)	.81	(.39)
Literal-1	1	n.a.	1	n.a.	.96	(.2)
Literal-2	1	n.a.	1	n.a.	1	n.a.
Metaphor						
Pragmatic	.40	(.49)	.51	(.5)	.41	(.5)
Literal-1	.99	(.11)	.99	(.11)	.99	(.1)
Literal-2	.99	(.11)	.99	(.11)	1	n.a.
Irony						
Pragmatic	.45	(.5)	.45	(.5)	.41	(.5)
Literal-1	1	n.a.	1	n.a.	.98	(.14)
Literal-2	.91	(.29)	.97	(.18)	.98	(.14)
Relevance						
Pragmatic	.95	(.23)	.92	(.27)	.92	(.28)
Literal-1	.99	(.1)	1	n.a.	.99	(.1)
Literal-2	.95	(.23)	.97	(.18)	.95	(.22)

Note. n.a.=not applicable, *n* = number, *SD* = standard deviation, Proportion = proportion correct, Relevance = sub-test on relevance implicatures, Metaphor = sub-test on metaphors, Manner = sub-test on manner implicatures, Scalars = sub-test on scalar impicatures, Contrastive = sub-test on contrastive implicatures, Pragmatic = condition with critical implicature items, Literal-1 = first condition with literal items, Literal-2 = second condition with literal items.

Overall, there was strong evidence that ironic and figurative interpetations take longer to process than literal interpretations. This was manifested both in the comparisons between correct pragmatic responses to critical items and correct literal responses to control items, and in the comparisons between pragmatic (correct) and literal (incorrect) responses (i.e., responses where the literal picture was selected) to critical items. Similarly, there was some evidence that relevance, manner, and SIs incur processing delays. This was evident in the comparisons between correct pragmatic responses to critical items and correct literal responses to control items. The evidence for the latter three pragmatic meanings, however, is weaker as, for relevance implicatures, our items were not designed to allow comparisons between pragmatic and literal responses to the same critical items (there was no picture that corresponded to a literal interpretation); for SIs, this comparison showed no differences in processing time; and for manner, it showed that literal responses were slower than pragmatic interpretations.

Regarding manner, we believe that these results indicate that children calculated the implicature in all cases, but, for literal responses, they cancelled it in favour of the literal interpretation (since critical statements were true, though pragmatically inappropriate descriptions of the literal picture). Nevertheless, we still report the group results on manner RTs below.

Global analyses on reaction times

We first conducted a global analysis on RTs with Group as a between-subject factor, Type and Condition as within-subject factors, SES and Age as covariates and by-subject and by-item random intercepts. This analysis showed a significant three-way interaction between Group, Type and Condition ($\chi^2(20) = 71.01$, $p < .05$). As Figure 3 shows, the interaction was due to bilinguals performing slower than the other two groups in the critical condition of the irony sub-test. This difference is more prominent between bilinguals and bi-dialectals.

. https://doi.org/10.1017/S1366728918001189

Table 5. Summary of the generalized linear mixed-effects regression model for the effect of Group on pragmatic accuracy

Effect	Coefficient	*SE*	*z*-value	Pr(>|z|)	Random Slope(s)
Intercept	1.23	1.41	0.87	>.05	no
SES	0.2	0.11	1.87	=.06	n.a.
Age	−0.26	1.35	−0.20	>.05	n.a.
Vocabulary	−0.01	0.08	−0.18	>.05	n.a.
Group 2	0.47	0.39	1.20	>.05	no
Group 3	0.11	0.40	0.28	>.05	no
Type 2	−0.49	0.43	−1.14	>.05	no
Type 3	1.12	0.47	2.36	<.05	no
Type 4	−1.24	0.44	−2.83	<.001	no
Type 5	−0.97	0.44	−2.24	<.05	no
Type 2 : Group 2	0.06	0.48	0.13	>.05	no
Type 3 : Group 2	−1.11	0.54	−2.08	<.05	no
Type 4 : Group 2	−0.17	0.48	−0.35	>.05	no
Type 5 : Group 2	−0.68	0.48	−1.42	>.05	no
Type 2 : Group 3	−0.23	0.46	−0.51	>.05	no
Type 3 : Group 3	−0.71	0.52	−1.35	>.05	no
Type 4 : Group 3	−0.42	0.47	−0.90	>.05	no
Type 5 : Group 3	−0.66	0.46	−0.43	>.05	no

Note. n.a.=not applicable, SES = socioeconomic status composite score, Age = participants' age in months, Vocabulary = vocabulary composite score, Type 1 = accuracy in the scalar implicatures sub-test, Type 2 = accuracy in the contrastive implicatures sub-test, Type 3 = accuracy in the manner implicatures sub-test, Type 4 = accuracy in the metaphors sub-test, Type 5 = accuracy in the irony sub-test, Group 1 = bi-dialectals, Group 2 = monolinguals, Group 3 = bilinguals.

We also conducted an analysis on RTs with the same variables as above but collapsing across Type to benefit from the increased item-level power resulting from averaging RTs from 12 items. This analysis showed a significant Group by Condition interaction ($\chi^2(4) = 20.27$, $p < .05$). As Figure 4 shows, the interaction is driven by differences between bi-dialectals and the other two groups (in that bi-dialectals were faster) in the Literal-1 and Pragmatic conditions. Again, these differences were more prominent between bi-dialectal and bilingual children. Subsequent between-group analyses for each condition showed a significant Group effect only for the Literal-1 condition ($\chi^2(2) = 6.74$, $p < .05$) in that bi-dialectals were faster than bilinguals (estimate = 626, $SE = 254$, $z = 2.46$, $p < .05$). The above analyses were repeated with vocabulary as an additional covariate and revealed largely the same results.

Finally, to further explore the significant three-way interaction between Group, Type and Condition reported above, we performed analyses on RTs in each pragmatic sub-test separately. These analyses showed the following significant results. For relevance, there was a significant Group effect in the analysis with Vocabulary covaried ($\chi^2(2) = 7.60$, $p < .05$) in that bi-dialectals were overall faster (i.e., across conditions, not only in the pragmatic one) than monolinguals (estimate = 524.7, $SE = 195$, $z = 2.68$, $p < .05$). Similarly, for SIs, there was a significant Group effect ($\chi^2(2) = 6.77$, $p < .05$, for the analysis without vocabulary covaried) due to bi-dialectals responding overall faster (i.e., across conditions, not only in the pragmatic one) than monolinguals (estimate = 296.5, $SE = 116$, $z = 2.547$, $p < .05$). Finally, the analysis on irony revealed a significant interaction beween Group and Condition ($\chi^2(4) = 27.46$, $p < .05$, without vocabulary covaried). Subsequent analyses indicated that there was a significant Group effect only in the ironic condition ($\chi^2(2) = 6.11$, $p < .05$) in that bi-dialectals responded faster than bilinguals (estimate = 7818, $SE = 3291$, $z = 2.376$, $p < .05$). For scalars and irony, the results above were the same when vocabulary was covaried in the analyses.

Correlations between pragmatic performance and aspects of bilinguals'/bi-dialectals' language experience

We initially looked at the bivariate correlations between bilinguals'/bi-dialectals' pragmatic performance and amount of exposure to Dutch (language of testing), second language or dialect and degree of balanced bilingualism or bi-dialectalism (see Supplementary Material, for how these were quantified). There was only a significant negative correlation between RTs for scalar implicature responses and exposure to second dialect (Spearman's rho(36)= −.33, p(two-tailed) < .05), suggesting faster performance for bi-dialectals with more exposure to a second dialect.

Moreover, we further explored whether those bilingual/bi-dialectal children who had more exposure to Dutch (language of testing) might exhibit better pragmatic performance. To do this, we divided the bilingual and bi-dialectal groups into subgroups based on their exposure to Dutch (West Flemish-dominant bi-dialectals, Dutch-dominant bi-dialectals, unbalanced bilinguals, balanced bilinguals; see Appendix SB in the online Supplementary Material for how this was achieved).

Preliminary analyses showed that the five groups did not differ in age ($F(4, 131) = 2.36$, $p = .057$). However, there were significant group differences in SES ($F(4, 131) = 43.62$, $p < .05$) in that West-Flemish dominant and Dutch-dominant bi-dialectals had

Table 6. Summary of the generalized linear mixed-effects regression model for the effect of Type, Working Memory, Inhibition, Vocabulary, Age and Socioeconomic Status on pragmatic accuracy

Effect	Coefficient	*SE*	*z*-value	Pr(>|z|)	Random Slope(s)
Intercept	2.21	1.39	1.60	>.05	no
Type 2	−0.57	0.35	−1.63	>.05	n.a.
Type 3	0.52	0.36	1.43	>.05	n.a.
Type 4	−1.46	0.35	−4.16	<.0001	n.a
Type 5	−1.47	0.35	−4.21	<.0001	n.a.
SES	0.05	0.08	0.69	>.05	no
Vocabulary	0.002	0.09	0.03	>.05	yes, by-item
WM	0.19	0.09	1.98	<.05	yes, by-item
Inhibition	−0.02	0.08	−0.30	>.05	yes, by-item
Age	−1.05	1.36	−0.77	>.05	no

Note. n.a.=not applicable, Type 2 = accuracy for critical items in the contrastive implicatures sub-test, Type 3 = accuracy for critical items in the manner implicatures sub-test, Type 4 = accuracy for the critical items in the metaphors sub-test, Type 5 = accuracy for critical items in the irony sub-test, SES = socioeconomic status composite score, Vocabulary = vocabulary composite score, WM = working memory composite score, Inhibition = Inhibition composite score, Age = participants' age in months.

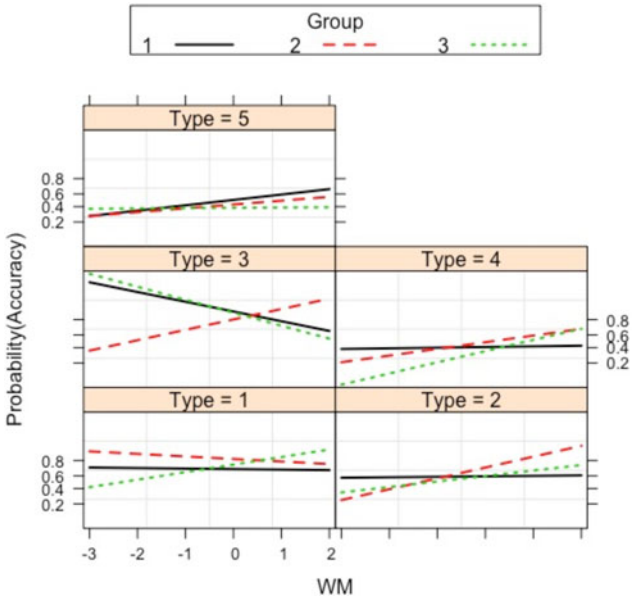

Fig. 2. Three-way interaction between Group (1 = bi-dialectals, 2 = monolinguals, 3 = bilinguals), Type (Type 1 = scalar implicatures, Type 2 = contrastive implicatures, Type 3 = manner implicatures, Type 4 = metaphors, Type 5 = irony) and Working Memory (WM) from the generalized linear mixed-effects regression model on pragmatic accuracy.

a lower SES than the other groups ($ps < .05$) and did not differ from each other ($p > .05$); and that unbalanced bilinguals had a higher SES than monolinguals ($p < .05$). Moreover, there was a significant Group effect on vocabulary ($F(4, 129) = 5.18$, $p < .05$) in that only the bilingual groups with the least exposure to Dutch (West Flemish-dominant bi-dialectals and unbalanced bilinguals) had lower vocabularies than monolinguals ($ps < .05$). The results on vocabulary are in line with accounts which suggest that bilinguals' vocabulary gap in a given language might close with sufficient exposure to that language (Thordardottir, 2011).

The following significant results were observed in the analyses on pragmatic performance (see online Supplementary Material, Table S1, for descriptive statistics). For accuracy, there was a significant Group by Type interaction ($\chi^2(16) = 27.4$, $p < .05$, with vocabulary included as a covariate or not). Subsequent analyses on each type of pragmatic meaning showed a significant Group effect only for metaphors (with vocabulary included: $\chi^2(4) = 9.93$, $p < .05$; without vocabulary included: $\chi^2(4) = 10.18$, $p < .05$). Pairwise contrasts, however, revealed no significant differences between the five groups (all $ps > .05$).

For RTs, the Group by Type by Condition interaction in the global analysis on RTs was significant (without vocabulary covaried or not: $\chi^2(40) = 85.33$, $p < .05$). To further explore the interaction, we conducted analyses for each sub-test separately. There was a significant Group by Condition interaction in the analysis for irony (with vocabulary included ($\chi^2(8) = 33.62$, $p < .05$) or not ($\chi^2(8) = 33.54$, $p < .05$)). Subsequent analyses, however, for each condition separately (Pragmatic, Literal-1, Literal-2) indicated no significant group effects (all $ps > .05$). Moreover, there was a significant Group effect in the analysis for SIs (whether vocabulary was covaried ($\chi^2(4) = 9.86$, $p < .05$) or not ($\chi^2(4) = 13.6$, $p < .05$)), even though, again, post-hoc contrasts showed no significant group differences (all $ps > .05$).

Finally, the Group by Condition interaction in the global analysis on RTs without the Type factor, was also significant ($\chi^2(8) = 30.22$, $p < .05$ whether vocabulary was covaried or not). Subsequent group comparisons, however, for each condition separately showed no significant results (with vocabulary covaried or not).

4. Discussion

In this study, we tested the hypothesis of a bilingual and/or bi-dialectal advantage in children's understanding and processing of pragmatic meanings.

4.1 Bilingualism, bi-dialectalism, pragmatic interpretation and processing

Children were tested on a wide range of pragmatic meanings and exhibited variability in pragmatic performance at various levels. Pragmatic responses to critical items in all sub-tests (besides

Table 7. Descriptive statistics (mean reaction times and standard deviations for correct responses) from the pragmatics test (raw values) by type of pragmatic meaning, condition and language group

Type	Bi-dialectals		Monolinguals		Bilinguals	
	(*n* = 46)		(*n* = 44)		(*n* = 48)	
	Mean	(*SD*)	Mean	(*SD*)	Mean	(*SD*)
Scalars						
Pragmatic	1611	(758)	1893	(1359)	1744	(1037)
Literal-1	1213	(450)	1279	(523)	1324	(957)
Literal-2	1345	(458)	1512	(782)	1384	(645)
Contrastive						
Pragmatic	3105	(2206)	3083	(2012)	3401	(2816)
Literal-1	3787	(2205)	4395	(2747)	5329	(5439)
Literal-2	1384	(778)	1383	(667)	1523	(864)
Manner						
Pragmatic	2271	(2594)	3056	(5651)	3260	(6202)
Literal-1	1275	(1652)	1399	(1775)	1392	(1448)
Literal-2	991	(1174)	856	(785)	814	(658)
Metaphor						
Pragmatic	5003	(4442)	5202	(3372)	5661	(3507)
Literal-1	2019	(830)	2072	(835)	2335	(1006)
Literal-2	2055	(947)	1749	(868)	2058	(996)
Irony						
Pragmatic	3273	(3953)	4507	(5613)	8356	(13517)
Literal-1	1127	(498)	1223	(1016)	1338	(1762)
Literal-2	1935	(1911)	1993	(1891)	1932	(2624)
Relevance						
Pragmatic	2089	(1242)	2766	(3289)	2451	(1435)
Literal-1	1456	(751)	1347	(959)	1527	(943)
Literal-2	1282	(991)	1379	(1517)	1338	(1025)

Note. *n* = number, *SD* = standard deviation, Relevance = sub-test on relevance implicatures, Metaphor = sub-test on metaphors, Manner = sub-test on manner implicatures, Scalars = sub-test on scalar impicatures, Contrastive = sub-test on contrastive implicatures, Pragmatic = condition with critical implicature items, Literal-1 = first condition with literal items, Literal-2 = second condition with literal items.

contrastive implicatures) were significantly slower than literal responses to control items. In addition, overall pragmatic performance positively correlated with working memory. Crucially, these two findings are also consistent with past research on pragmatic interpretation in adults suggesting that our task was successful in tapping into pragmatic processing.[7] Moreover, there was significant variability in the difficulty level of the pragmatic meanings tested. With regards to accuracy, children exhibited ceiling performance with relevance implicatures, very high performance for manner, moderate performance with scalar and constrastive implicatures, and were least accurate with novel metaphors and irony (see Figure 1). Novel metaphors and irony were also the most demanding at the processing level, in that, for these pragmatic meanings, pragmatic responses to critical items were significantly slower than both correct literal responses to control trials and incorrect literal responses to critical trials.

However, despite this variation in pragmatic performance, there was no evidence for a bilingual or bi-dialectal pragmatic advantage over monolinguals. Thus, we consider these results as strong evidence indicating that bilingual and bi-dialectal children do not differ from monolinguals in pragmatic understanding. This holds for various types of pragmatic meanings (including late-developing types, like irony) and is true at both the interpretation and processing level. It is also true despite bilinguals' and

[7]Both of these findings are also predicted by pragmatic theory and particularly Relevance Theory (Sperber & Wilson, 1986). Sperber and Wilson (1986), for instance, suggest that pragmatic interpretation is an effortful process. Even though they are not very clear about the exact nature of this cognitive effort, several researchers have interpreted it either in terms of pragmatic meanings requiring extra processing time or in terms of involving additional cognitive resources (like executive functions) relative to literal interpretations (e.g., Bott & Noveck, 2004; De Neys & Schaeken, 2007).

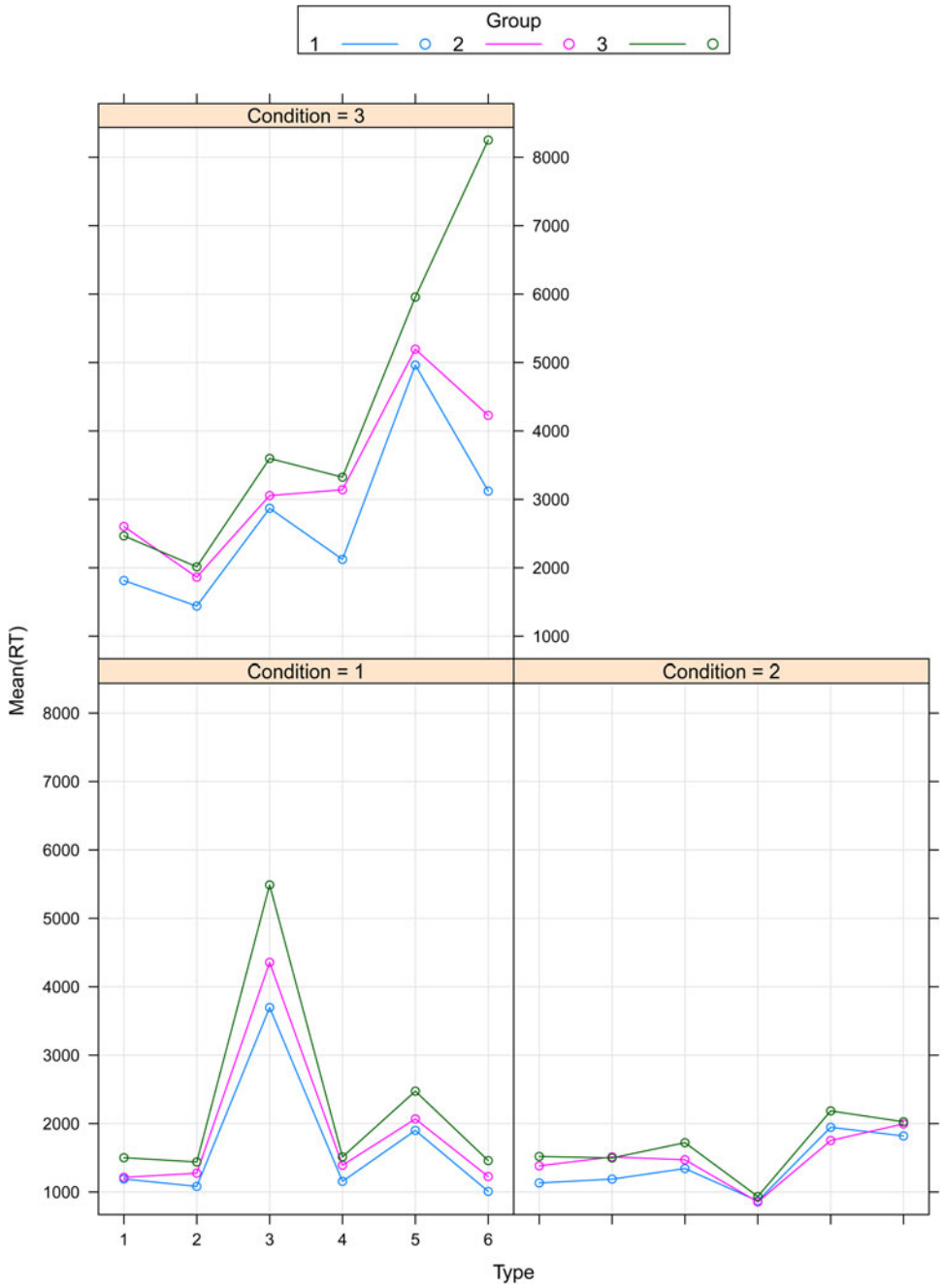

Fig. 3. Three-way interaction between Group (1 = bi-dialectals, 2 = monolinguals, 3 = bilinguals), Type (Type 1 = relevance implicatures, Type 2 = scalar implicatures, Type 3 = contrastive implicatures, Type 4 = manner implicatures, Type 5 = metaphors, Type 6 = irony) and Condition (Condition 1 = Literal-1, Condition 2 = Literal-2, Condition 3 = Pragmatic) from the linear mixed-effects regression model on reaction times.

bi-dialectals' lower vocabulary as measured by formal language tests. Moreover, this holds irrespective of factors such as typological proximity between the language pairs spoken by bilinguals (bilingual and bi-dialectal children performed largely comparably to monolinguals) and language dominance or balance (Dutch-dominant bi-dialectals, Flemish-dominant bi-dialectals, balanced and unbalanced bilinguals also performed largely comparably to monolinguals).

At face value, our results contradict past research that did report superior social-pragmatic skills in bilinguals. How is it possible to explain these contradictory findings? A close examination of the studies that reported superior social-pragmatic skills in bilinguals suggests that the strict majority of them have been conducted with preschool-aged children. As Antoniou and Katsos (2017) suggest, it is possible that a bilingual pragmatic advantage is found only in the first years of life, a period during which there is possibly more room for plasticity or acceleration of pragmatic development because of bilingualism. Nevertheless, even though bilingual advantages in the preschool years have been found for certain socio-pragmatic skills, such specific benefits are not very likely to extend to pragmatic comprehension. This is because the studies by Syrett et al. (2016; 2017) investigated SI comprehension in preschool-aged

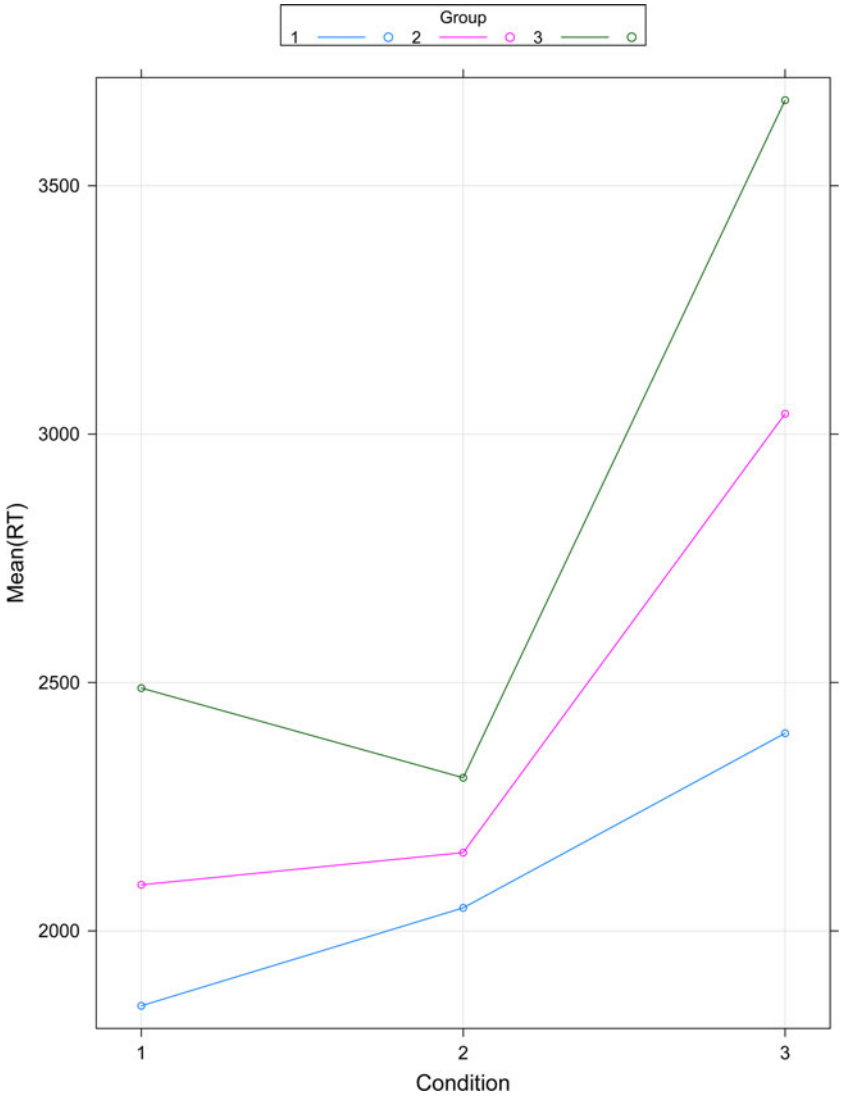

Fig. 4. Two-way interaction between Group (1 = bi-dialectals, 2 = monolinguals, 3 = bilinguals), and Condition (Condition 1 = Literal-1, Condition 2 = Literal-2, Condition 3 = Pragmatic) from the linear mixed-effects regression model on reaction times (with the factor Type excluded).

children and did not find differences between bilinguals and monolinguals (but see Siegal et al., 2007).

A second explanation might be the degree to which the pragmatic skill examined depends on language. Pragmatic competence involves a wide range of skills, some of which are non-verbal (e.g., following a pointing gesture) and some of which are verbal to a lesser or greater extent (e.g., implicatures). Given that bilingual children often exhibit lower linguistic performance in each of their languages separately, it is possible that the degree to which a pragmatic skill involves language modulates the emergence of the bilingual pragmatic advantage.

All in all, then, our results, in conjunction with the findings of past research, support the following conclusion. Bilingual preschool-aged children seem to enjoy benefits in various social and pragmatic skills. These advantages, however, do not extend to children's comprehension of verbal-pragmatic inferences and this seems to hold irrespective of age.

As a final note, we should clarify that we do not believe that testing pragmatic comprehension in older children was a wrong methodological decision for our study. Different experiences with two languages, at different developmental stages, and in diverging environmental situations (e.g., sociolinguistic contexts, SES layers) lead to unique bilingual experiences that possibly affect cognitive functioning in different ways. We maintain that it is important to examine the possibility of a cognitive advantage across the range of bilingual experiences and to uncover the conditions under which or the specific skills for which the bilingual benefit might or might not be evident.

4.2 The cognitive foundations of pragmatic development

In this study, we also investigated the effects of language proficiency and executive functions on pragmatic performance. The group results provided evidence against a view of pragmatic comprehension as depending only on structural language skills in the target language: bilingual and bi-dialectal children exhibited lower vocabularies than monolinguals but comparable pragmatic performance. The factorial analyses, however, could not directly answer the question of whether pragmatic understanding depends on executive functions, given that, in this study, we did not find better executive control skills in bilingual and/or bi-dialectal children. The lack of EC differences in our sample suggests either a complete absence of a bilingual EC advantage (e.g., Paap et al., 2015) or that the cognitive benefit in bilinguals is found only under certain conditions. Factors that have been suggested to modulate the emergence of the bilingual cognitive advantage, for instance, include age,

patterns of everyday language use, the sociolinguistic context of language use, language proficiency, and type of tasks used (see Dong & Li, 2015; Green & Abutalebi, 2013).

The findings of the correlational analysis, however, revealed a positive effect of WM. The positive WM effect possibly indicates that pragmatic comprehension involves a process whereby listeners combine different pieces of information (e.g., the speaker's knowledge, literal features, encyclopedic knowledge, pragmatic maxims) in working memory. A related possibility is that the WM effect is linked to Theory of Mind. As already noted, some theoretical accounts suggest that pragmatic comprehension is a mind-reading process that involves understanding the intentions of a speaker behind an utterance, and, hence, requires the use of a ToM (e.g., Sperber & Wilson, 1986). In addition, there is some experimental evidence that shows that using ToM to understand language is a slow and costly process that draws on executive functions (including WM; e.g., Apperly, Riggs, Simpson, Chiavarino & Samson, 2006; Lin, Keysar & Epley, 2010; Schneider, Lam, Bayliss & Dux, 2012). Of course, it is also possible that both WM and ToM have independent effects on pragmatic comprehension. Future studies should more closely examine which exact aspect of pragmatic comprehension makes it dependent on WM and whether pragmatic understanding is independently affected by ToM.

Moreover, we did not find any strong indication that the relation between pragmatics and WM differed in the three groups. Thus, our findings largely suggest that pragmatic interpretation in bilinguals does not draw on executive functions in a different way relative to monolinguals.

Finally, our results showed no significant effect of inhibition on pragmatic performance. Inhibition would be expected to play a role, for instance, within theoretical models that suggest that listeners (always or sometimes) consider a literal (or at least some features associated with a literal) representation of a word or utterance before or in parallel to accessing a pragmatic interpretation (e.g., Giora, Givoni & Fein, 2015; Grice, 1989; Rubio-Fernandez, 2007). Inhibition skills would then be required to suppress these irrelevant literal features. The finding that inhibition does not affect pragmatic interpretation suggests that these literal features are either not considered at all or at least that they are not actively inhibited during pragmatic comprehension.

4.3 A model of bilingual pragmatic development, representation and processing

What do the above findings, then, mean regarding the development, processing and representation of pragmatics in bilinguals? First, our results indicate that there is a point in pragmatic development during which language proficiency in the target language no longer affects pragmatic interpretation in the same language (even though there is still variability in children's pragmatic performance). This finding provides some (perhaps partial support) to the view that pragmatic competence (at least the pragmatic knowledge that is needed for the comprehension of pragmatic meanings) is to some extent independent of structural language skills (see also Katsos, Roqueta, Estevan & Cummins, 2011; Levinson, 2016).

Second, pragmatic principles such as Gricean maxims possibly reflect universal properties of communication, whether verbal or not (Grice, 1989; Prince, 1982; Sperber & Wilson, 1986; Stivers, Enfield, Brown, Englert, Hayashi, Heinemann, Hoymann, Rossano, de Ruiter, Yoon & Levinson, 2009; see also Antoniou, in press; Katsos, Cummins, Ezeizabarrena, Gavarró, Kraljević, Hrzica, Grohmann, Skordi, de López, Sundahl, van Hout, Hollebrandse, Overweg, Faber, van Koert, Smith, Vija, Zupping, Kunnari, Morisseau, Rusieshvili, Yatsushiro, Fengler, Varlokosta, Konstantzou, Farby, Guasti, Vernice, Okabe, Isobe, Crosthwaite, Hong, Balčiūnienė, Nizar, Grech, Gatt, Cheong, Asbjørnsen, von Koss Torkildsen, Haman, Miękisz, Gagarina, Puzanova, Anđelković, Savić, Jošić, Slančová, Kapalková, Barberán & Özge, 2016; Slabakova, 2010). This further suggests that bilingual children possibly have equivalent to monolinguals exposure to pragmatic principles across their languages and (verbal or non-verbal) communicative experiences in general.

The independence of pragmatics from language is further supported by additional considerations. Given the claim that pragmatic maxims are universal, it follows that pragmatic principles are possibly represented independently of the bilinguals' languages, since it would make little sense (at least from a perspective of cognitive efficiency) to represent the same knowledge twice. Moreover, independence of pragmatics from language is predicted by some pragmatic and language acquisition theories. Sperber and Wilson (2002), for instance, propose a pragmatic module (an autonomous system) that is responsible for interpreting communicative stimuli based on the pragmatic principle of relevance. Some theoretical accounts of language acquisition, on the other hand, suggest that language is acquired by using preexisting pragmatic knowledge (e.g., Grice's cooperative principle) to make inferences about the speaker's intentions (e.g., Tomasello, 2008). Finally, the claim that pragmatics can be separated from language(s) is further supported by the fact that pragmatic knowledge is also employed during non-verbal communication (Sperber & Wilson, 1986).

Third, our findings also revealed that pragmatic interpretation draws on working memory but that pragmatic understanding in bilinguals does not seem to depend differently on these non-verbal cognitive resources relative to monolinguals.

It is therefore possible that bilingual children have a single, language-independent pragmatic system that develops and functions in a similar way to monolinguals (see also Antoniou, in press; Kecskes, 2015; Slabakova, 2010). To the extent that bilinguals have sufficient proficiency in the relevant language to understand and process the explicit meaning of non-literal utterances, this account would predict bilingual skill in pragmatic interpretation that is comparable to monolinguals.

4.4. The role of confounding factors

In our study, there were significant group differences in age, SES and vocabulary. Is it possible that these background differences confounded our results? We find this highly unlikely. First, the three variables were included as covariates in all between-group analyses. Second, none of these variables significantly affected pragmatic performance. Finally, balanced bilinguals in our study did not differ from monolinguals in any of the background variables, had extensive exposure to Dutch (language of testing) and were balanced in their two languages to a higher degree (see online Supplementary Material, Appendix SB). Yet, again, no significant differences were observed between this group and monolinguals.

5. Conclusion

Using a comprehensive test of different types of pragmatic meanings, our study revealed no bilingual or bi-dialectal advantage over

. https://doi.org/10.1017/S1366728918001189

monolinguals in pragmatic comprehension or processing. Null differences between bilinguals, bi-dialectals and monolinguals are important findings; especially, when these are reported for pragmatic-communicative skills that draw on language and are evident in the presence of lower language proficiency (as measured by formal tests) for bilinguals and bi-dialectals. These findings suggest that bilingual and bi-dialectal children maintain equivalent to monolinguals' verbal pragmatic-communicative functioning, despite their often-reported weaker language knowledge in the target language.

Supplementary Material. Supplementary material can be found online at http://doi.org/10.1017/S1366728918001189

Acknowledgements. Research presented in this paper has been made possible by a Fondation Wiener-Anspach (Brussels, Belgium) and Issac Newton Trust (Cambridge, UK) grant "The impact of bilingualism and bi-dialectalism on linguistic and cognitive development" to N. Katsos and M. Kissine. The writing of this article has been also supported by a Marie Skłodowska-Curie fellowship (grant agreement No. 800305) to K. Antoniou. We also thank all our participants and their parents, and the teachers and principals of the following primary schools: Floralaan Eindhoven, 't Wilgenhof Vlamertinge, Capucienen Ieper, Groenheuvel Bellegem, De Baai Kortrijk, Sint-Amandscollege Noord Kortrijk, Mater Dei Brussel, and Mooi-Bos Brussel. Special thanks also go to Marlein Rusch and Dalila De Waele, for helping with the data collection.

References

Andrés-Roqueta C and Katsos N (2017). The Contribution of Grammar, Vocabulary and Theory of Mind in Pragmatic Language Competence in Children with Autistic Spectrum Disorders. *Frontiers in psychology* **8**, 996.

Antoniou K (in press). Multilingual pragmatics: Implicature comprehension in adult L2 learners and multilingual children. In N Taguchi (ed.), *The Routledge Handbook of Pragmatics and Second Language Acquisition*. Abingdon: Routledge.

Antoniou K, Cummins C and Katsos N (2016) Why only some adults reject under-informative utterances. *Journal of Pragmatics* **99**, 78–95.

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

Antoniou K and Katsos N (2017) The effect of childhood multilingualism and bilectalism on implicature understanding. *Applied Psycholinguistics* **38**, 787–833.

Apperly IA, Riggs KJ, Simpson A, Chiavarino C and Samson D (2006) Is belief reasoning automatic? *Psychological Science* **17**, 841–844.

Barac R, Bialystok E, Castro DC and 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

Barr DJ, Levy R, Scheepers C and Tily HJ (2013) Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* **68**, 255–278.

Bates D, Maechler M, Bolker B and Walker S (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* **67**, 1–48.

Bialystok E (2017) The bilingual adaptation: How minds accommodate experience. *Psychological bulletin* **143**, 233.

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

Blom E, Boerma T, Bosma E, Cornips L and Everaert E (2017) Cognitive advantages of bilingual children in different sociolinguistic contexts. *Frontiers in Psychology* **8**, 552.

Boyce W, Torsheim T, Currie C and Zambon A (2006) The family affluence scale as a measure of national wealth: validation of an adolescent self-report measure. *Social indicators research* **78**, 473–487.

Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Stevens MHH and White JSS (2009) Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in ecology & evolution* **24**, 127–135.

Bott L and Noveck IA (2004) Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of memory and language* **51**, 437–457.

Breheny R (2006) Communication and folk psychology. *Mind & Language* **21** (1), 74–107.

Breheny R, Ferguson HJ and Katsos N (2013) Taking the epistemic step: Toward a model of on-line access to conversational implicatures. *Cognition* **126**, 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**, 434–463.

Carlson SM (2003) Executive function in context: Development, measurement, theory, and experience. *Monographs of the Society for Research in Child Development* **68**, 138–151.

Carlson SM and Meltzoff AN (2008) Bilingual experience and executive functioning in young children. *Developmental science* **11**, 282–298.

Chiappe DL and Chiappe P (2007) The role of working memory in metaphor production and comprehension. *Journal of Memory and Language* **56**, 172–188.

Clahsen H and Felser C (2006a) Continuity and shallow structures in language processing. *Applied Psycholinguistics* **27**, 107–126.

Clahsen H and Felser C (2006b) How native-like is non-native language processing? *Trends in Cognitive Science* **10**, 564–570.

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

Costa A and Sebastián-Gallés N (2014) How does the bilingual experience sculpt the brain? Nature Reviews. *Neuroscience* **15**, 336.

De Neys W and Schaeken W (2007) When people are more logical under cognitive load: dual task impact on scalar implicature. *Experimental Psychology* **54**, 128–33.

Deliens G, Antoniou K, Clin E, Ostashchenko E and Kissine M (2018) Context, facial expression and prosody in irony processing. *Journal of Memory and Language* **99**, 35–48.

Dong Y and Li P (2015) The cognitive science of bilingualism. *Language and Linguistics Compass* **9**, 1–13.

Dunn LM, Dunn LM and Schlichting JEPT (2005) *Peabody picture vocabulary test-III-NL*. Harcourt Test Publishers.

Ellefson MR, Shapiro LR and Chater N (2006) Asymmetrical switch costs in children. *Cognitive Development* **21**, 108–130.

Faul F, Erdfelder E, Lang A-G and 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.

Filippova E (2014) Irony production and understanding. In D Matthews (ed.), *Pragmatic Development in First Language Acquisition*, pp. 261–278. Amsterdam/Philadelphia: John Benjamins.

Foucart A, Garcia X, Ayguasanosa M, Thierry G, Martin C and Costa A (2015) Does the speaker matter? Online processing of semantic and pragmatic information in L2 speech comprehension. *Neuropsychologia* **75**, 291–303.

Garraffa M, Beveridge M and Sorace A (2015) Linguistic and Cognitive Skills in Sardinian–Italian Bilingual Children. *Frontiers in Psychology* **6**, 1898.

Giora R, Givoni S and Fein O (2015) Defaultness reigns: The case of sarcasm. *Metaphor and Symbol* **30**, 290–313

Green DW and Abutalebi J (2013) Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology* **25**, 515–530.

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

Grodner D and Sedivy J (2011) The effects of speaker-specific information on pragmatic inferences. In N Pearlmutter & E. Gibson (Eds.), *The processing and acquisition of reference*. Cambridge, MA: MIT Press.

Grohmann KK and Kambanaros M (2016) The Gradience of Multilingualism in Typical and Impaired Language Development: Positioning Bilectalism within Comparative Bilingualism. *Frontiers in Psychology* **7**, 37.

Hammer CS, Hoff E, Uchikoshi Y, Gillanders C, Castro DC and Sandilos LE (2014) The language and literacy development of young dual language learners: A critical review. *Early Childhood Research Quarterly* **29**, 715–733.

Horn LR (1972) *On the Semantic Properties of Logical Operators in English*. Doctoral dissertation, University of California, LA

Hothorn T, Bretz F, Westfall P, Heiberger RM, Schuetzenmeister A, Scheibe S and Hothorn MT (2016) *Package 'multcomp'. Simultaneous inference in general parametric models*. Project for Statistical Computing, Vienna, Austria.

Huang Y and Snedeker J (2009a) On-line interpretation of scalar quantifiers: insight into the semantics-pragmatics interface. *Cognitive Psychology* **58**, 376–415.

Huang YT and Snedeker J (2009b) Semantic meaning and pragmatic interpretation in 5-year-olds: Evidence from real-time spoken language comprehension. *Developmental psychology* **45**, 1723.

Katsos N, Cummins C, Ezeizabarrena MJ, Gavarró A, Kraljević JK, Hrzica G, Grohmann KK, Skordi A, de López KJ, Sundahl L, van Hout A, Hollebrandse B, Overweg J, Faber M, van Koert M, Smith N, Vija M, Zupping S, Kunnari S, Morisseau T, Rusieshvili M, Yatsushiro K, Fengler A, Varlokosta S, Konstantzou K, Farby S, Guasti MT, Vernice M, Okabe R, Isobe M, Crosthwaite P, Hong Y, Balčiūnienė I, Nizar YMA, Grech H, Gatt D, Cheong WN, Asbjørnsen A, von Koss Torkildsen J, Haman E, Miękisz A, Gagarina N, Puzanova J, Anđelković D, Savić M, Jošić S, Slančová D, Kapalková S, Barberán T and Özge D (2016) Cross-linguistic patterns in the acquisition of quantifiers. *Proceedings of the National Academy of Sciences* **113**, 9244–9249.

Katsos N, Roqueta CA, Estevan RAC and Cummins C (2011) Are children with Specific Language Impairment competent with the pragmatics and logic of quantification?. *Cognition* **119**, 43–57.

Kecskes I (2015) How does pragmatic competence develop in bilinguals?. *International Journal of Multilingualism* **12**, 419–434.

Kort W, Schittekatte M and Compaan E (2008) *CELF-4-NL: Clinical Evaluation of Language Fundamentals-vierde-editie*. Amsterdam: Pearson Assessment and Information B.V.

Kowatch K, Whalen JM and Pexman PM (2013) Irony comprehension in action: A new test of processing for verbal irony. *Discourse Processes* **50**, 301–315.

Kirk NW, Fiala L, Scott-Brown KC and Kempe V (2014) No evidence for reduced Simon cost in elderly bilinguals and bidialectals. *Journal of Cognitive Psychology* **26**, 640–648.

Kissine M (2013) *From utterances to speech acts*. Cambridge University Press.

Kissine M (2016) Pragmatics as metacognitive control. *Frontiers in psychology* **6**, 2057.

Kronmüller E, Morisseau T and Noveck IA (2014) Show me the pragmatic contribution: a developmental investigation of contrastive inference. *Journal of child language* **41**, 985–1014.

Levinson SC (2016) Turn-taking in human communication–origins and implications for language processing. *Trends in cognitive sciences* **20**, 6–14.

Levinson SC (2000) *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT press.

Liberman Z, Woodward AL, Keysar B and Kinzler KD (2017) Exposure to multiple languages enhances communication skills in infancy. *Developmental science* **20**.

Lin S, Keysar B and Epley N (2010) Reflexively mindblind: using theory of mind to interpret behavior requires effortful attention. *Journal of Experimental Social Psychology* **46**, 551–556.

Marty PP and Chemla E (2013) Scalar implicatures: working memory and a comparison with only. *Frontiers in Psychology* **4**, 403. doi: 10.3389/fpsyg.2013.00403

Marty P, Chemla E and Spector B (2013) Interpreting numerals and scalar items under memory load. *Lingua* **133**, 152–163.

Mashal N (2013) The role of working memory in the comprehension of unfamiliar and familiar metaphors. *Language and Cognition* **5**, 409–436.

Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A and Wager TD (2000) The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive psychology* **41**, 49–100.

Mueller ST and Piper BJ (2014) The Psychology Experiment Building Language (PEBL) and PEBL test battery. *Journal of Neuroscience Methods* **222**, 250–259.

Paap KR, Johnson HA and Sawi O (2015) Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex* **69**, 265–278.

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

Paradis J, Emmerzael K and Duncan TS (2010) Assessment of English language learners: Using parent report on first language development. *Journal of communication disorders* **43**, 474–497.

Prince EF (1982) Grice and universality: A reappraisal. Ms., University of Pennsylvania.

Psychology Software Tools, Inc. [E-Prime 2.0]. (2012) Retrieved from https://www.pstnet.com.

Recanati F (2002) Does linguistic communication rest on inference?. *Mind & Language* **17**, 105–126.

Roberts L and Felser C (2011) Plausibility and recovery from garden paths in second language sentence processing. *Applied Psycholinguistics* **32**, 299–331.

Roberts L, Gullberg M and Indefrey P (2008) Online pronoun resolution in L2 dis- course. *Studies in Second Language Acquisition* **30**, 333–357.

Ross J and Melinger A (2017) Bilingual advantage, bidialectal advantage or neither? Comparing performance across three tests of executive function in middle childhood. *Developmental science* **20**.

RStudio Team (2016) *RStudio: Integrated Development for R*. RStudio, Inc., Boston, MA URL http://www.rstudio.com/.

Rubio-Fernandez P (2007) Suppression in metaphor interpretation: differences between meaning selection and meaning construction. *Journal of Semantics* **24**, 345–371.

Rueda MR, Fan J, McCandliss BD, Halparin JD, Gruber DB, Lercari LP and Posner MI (2004) Development of attentional networks in childhood. *Neuropsychologia* **42**, 1029–1040.

Rushton JP, Brainerd CJ and Pressley M (1983) Behavioral development and construct validity: The principle of aggregation. *Psychological bulletin* **94**, 18.

Schneider D, Lam R, Bayliss A.P. and Dux PE (2012) Cognitive load disrupts implicit theory-of-mind processing. *Psychological Science* **23**, 842–847.

Semel EM, Wiig EH, Secord WA and Kort W (2008) *CELF® 4 NL: Clinical evaluation of language fundamentals: Nederlandse versie*. Pearson.

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

Siegal M, Matsuo A, Pond C and 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 and 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, Okumura Y and Itakura S (2010) Bilingualism Accentuates Children's Conversational Understanding. *PLoS ONE* **5**(2): e9004. doi:10.1371/journal.pone.0009004

Slabakova R (2010) Scalar implicatures in second language acquisition. *Lingua* **120**, 2444–2462.

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

Sperber D and Wilson D (2002) Pragmatics, modularity and mind-reading. *Mind & Language* **17**, 3–23.

Spotorno N and Noveck IA (2014) When is irony effortful? *Journal of Experimental Psychology: General* **143**, 1649.

Stivers T, Enfield NJ, Brown P, Englert C, Hayashi M, Heinemann T, Hoymann G, Rossano F, de Ruiter JP, Yoon KE and Levinson SC (2009) Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences* **106**, 10587–10592.

Syrett K, Austin J, Sanchez L, Germak C, Lingwall A, Perez-Cortes S, Arias-Amaya A and Baker H (2016) The influence of conversational context and the developing lexicon on the calculation of scalar implicatures. *Linguistic Approaches to Bilingualism* **7**, 230–264.

Syrett K, Lingwall A, Perez-Cortes S, Austin J, Sánchez L, Baker H, Germak C and Arias-Amaya A (2017) Differences between Spanish monolingual and Spanish-English bilingual children in their calculation of entailment-based scalar implicatures. *Glossa: a journal of general linguistics* **2**.

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

Tomasello M (2008) *Origins of Human Communication*. MIT Press.

Veenstra A, Antoniou K, Katsos N and Kissine M (2018) Resisting attraction: Individual differences in executive control are associated with subject-verb agreement errors in production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Waggoner JE and Palermo DS (1989) Betty is a bouncing bubble: Children's comprehension of emotion-descriptive metaphors. *Developmental psychology* **25**, 152.

Wilson D and Sperber D (2003) Relevance theory. In: G. Ward & L Horn (Eds.). *Handbook of Pragmatics*, pp. 607–632. Oxford, UK: Blackwell Sci.

Yow WQ and Markman EM (2011) Bilingualism and children's use of paralinguistic cues to interpret emotion in speech. *Bilingualism: Language and Cognition* **14**, 562–569.

Yow WQ and Markman EM (2015) A bilingual advantage in how children integrate multiple cues to understand a speaker's referential intent. *Bilingualism: Language and Cognition* **18**, 391–399.