BUCLD 42 Proceedings
To be published in 2018 by Cascadilla Press
Rights forms signed by all authors

The Impact of Childhood Bilingualism and Bi-dialectalism on Pragmatic Interpretation and Processing

Kyriakos Antoniou, Alma Veenstra, Mikhail Kissine, and Napoleon Katsos*

1. Introduction

In recent years, a great bulk of research has explored the impact of bilingualism on children's cognitive development (e.g., Barac, Bialystok, Castro, & Sanchez, 2014; Hammer, Hoff, Uchikoshi, Gillanders, Castro, & Sandilos, 2014). This body of work has reported two main patterns of results. First, bilingual children tend to show delays in some facets of language acquisition (e.g., Nicoladis, Palnmer, & Marentette, 2007; Oller, Pearson, & Cobo-Lewis, 2007). For instance, studies that have compared bilingual and monolingual children's vocabulary development have typically found that bilingual children possess smaller vocabularies in each of their two languages as compared to monolingual children in their only language (e.g., Bialystok, Luk, Peets, & Yang, 2010). Second, some evidence suggests superior bilingual skill in aspects of non-verbal cognition and social-communicative functioning. Particularly, studies have reported bilingual advantages in executive control (e.g., de Abreu, Cruz-Santos, Tourinho, Martin, & Bialystok, 2012), Theory of Mind (e.g., Kovacs, 2008), in the use of pragmatic cues (such as the speaker's perspective or gaze direction) to understand the speaker's referential intention (Liberman, Woodward, Keysar, & Kinzler, 2017; Yow & Markman, 2015), and in repairing communication failures (Wermelinger, Gampe, & Daum, 2017).

In this context, our study set out to test the hypothesis of a bilingual advantage in pragmatic understanding by comparing groups of bilingual, bi-dialectal, and monolingual children in the comprehension of various types of pragmatically implied meanings. In the next sections, we describe the theoretical background of our study by introducing some key theoretical concepts from Grice's (1989) work on meaning and implicature. We then review past research on bilingual children's pragmatic skills and present our own experiment.

2. Theoretical Background

* Kyriakos Antoniou, University of Cambridge, Université libre de Bruxelles, Hellenic Open University, ka353@cam.ac.uk. Alma Veenstra, University of Cambridge, Université libre de Bruxelles. Napoleon Katsos, University of Cambridge. Mikhail Kissine, Université libre de Bruxelles. This research was supported by grants from the Fondation Wiener-Anspach, Brussels and the Issac Newton Trust, Cambridge, UK

In Grice's account of meaning (1989), speaker's meaning can be analysed in terms of the speaker having certain intentions. To mean something with an utterance, is for the speaker to intend to produce a result to an addressee via the addressee's recognition of the speaker's intention. Grice's ideas on meaning formed the foundation for inferential models of communication according to which communicating involves the expression and understanding of intentions (Wilson & Sperber, 2005).

Grice (1989) further provided a philosophical analysis for cases of language use where speaker's meaning diverges from what the speaker says. He proposed that conversations are cooperative enterprises during which speakers are expected to abide to certain conversational rules; particularly, to the maxims of quantity ("Make your contribution as informative as is required" and "Do not make your contribution more informative than is required"), quality ("Do not say what you believe to be false" and "Do not say that for which you lack adequate evidence"), relation ("Be relevant") and manner ("Be perspicuous") (Grice, 1975: 45-47).

During communication, these maxims are often violated at the literal level. This invites listeners to infer non-literal meanings that preserve the maxims at the implicit level. These implicated meanings are what Grice called *conversational implicatures*. For example, when a speaker says *George's father was an erupted volcano*, s/he violates the maxim of quality at the level of what is said. This prompts the addressee to infer an implicated interpretation, such as for example that "George's father was very angry" (or a related figurative meaning).

Of course, Grice (1989) never linked his ideas to cognitive-psychological models of utterance processing and language development. Post-Gricean pragmatic theorists (e.g., Sperber & Wilson, 1986/1995; Levinson, 2000), however, adopted a cognitive conception of pragmatics and share a strong interest in issues related to the representation, processing and development of pragmatics.

3. Bilingualism and children's pragmatic skills

Recently, researchers have started to investigate bilingual children's pragmatic skills more systematically, with some studies reporting superior bilingual performance and others showing no differences with monolingual development.

In a pioneering set of studies, Siegal and colleagues (Siegal, Iozzi, & Surian, 2009; Siegal, Matsuo, Pond, & Otsu, 2007; Siegal, Surian, Matsuo, Geracci, Iozzi, et al., 2010) provided strong evidence for precocious pragmatic skills in bilingual preschool-aged children. First, they showed that bilingual children were better than monolinguals in interpreting scalar implicatures (i.e., to make the inference that *some* implicates "not all")[1]. Moreover, in their two most recent studies, they

[1]Such inferences are known as *scalar implicatures* because according to Horn (1972) they are generated based on linguistic scales (e.g., <some, all>) in which certain terms are ordered with respect to informativeness. The use of an informationally weaker term in the

also found that different groups of bilingual children outperformed monolinguals in their ability to detect pragmatically infelicitous utterances that violated Grice's maxims of conversation (e.g., sentences such as *I know your name*, which, as a reply to the question *What game do you like?*, violates the maxim of relation). Siegal and colleagues (2009) suggest that these pragmatic advantages possibly stem either from bilinguals' enhanced executive control skills or from a compensation mechanism that enhances pragmatic development to balance for bilingual children's often-reported delays in language development.

Three subsequent studies, however, by Antoniou and Katsos (2017) and Syrett and collaborators (Syrett, Austin, Sánchez, Germak, Lingwall, et al., 2016; Syrett, Lingwall, Perez-Cortes, Austin, Sánchez, et al., 2017) did not find evidence for superior pragmatic skills in bilingual children. Antoniou and Katsos (2017) tested school-aged (6-9 years of age) multilingual, bi-dialectal and monolingual Greek-speaking children in the comprehension of relevance, scalar, manner implicatures, and novel metaphors. They reported no group differences, even though multilinguals and bi-dialectals had lower language proficiency. Syrett et al. (2016) and Syrett et al. (2017) compared bilingual preschool-aged children (speakers of Spanish and English) and monolingual English children in scalar implicature comprehension. In line with Antoniou and Katsos (2017), they found no group differences in pragmatic understanding.

To sum, the evidence from the literature on bilingual children's pragmatic skills is inconclusive. There are some indications for a bilingual pragmatic advantage, but recent studies have reported largely no differences between bilingual and monolingual children.

4. The Present Study

Against this background, our study aimed to conduct another test of the hypothesis of a bilingual pragmatic advantage. Specifically, our goal was to examine this hypothesis by testing pragmatic meanings that have not been previously examined in bilinguals (contrastive implicatures, irony), by investigating pragmatic interpretation at the processing level (in terms of speed of comprehension) and by using a new sample of children who spoke different languages or dialects than previously examined (French-Dutch bilingual, Dutch-West Flemish bi-dialectal, and Dutch monolingual children). Besides bilingual children, we also tested a group of bi-dialectal children. The linguistic profile of bi-dialectals as speakers of two very similar dialects of the same language offers an opportunity to examine the effect of typological distance on the (possible) cognitive outcomes of bilingualism.

4.1. Method

4.1.1 Participants

scale (e.g., *some*) implicates the negation of the informationally stronger term in the same scale (i.e., "not all").

Three different groups of participants were tested: 46 bi-dialectal children who spoke Dutch and West Flemish (23 boys; 121–155 months old, mean age 136.2, *SD* 8.9 months); 48 French-Dutch bilingual children (20 boys; age range 121–144 months, mean age 132.5, *SD* 6.7 months); and 44 Dutch-speaking monolinguals (19 boys; ages 121–145, mean age 132.5 months, *SD* 6.8 months). Bi-dialectal and bilingual children were recruited from Belgium and monolinguals from the Netherlands. Dutch was the language of instruction at school for all children. Bi-dialectals spoke West Flemish at home. All bilinguals had French as their dominant language and used predominantly French at home.

4.1.2. Materials and Procedure

Testing was conducted in three sessions which lasted approximately 45 minutes each. All participants were tested in two vocabulary measures and an extensive pragmatics test examining the comprehension of various types of non-literal meanings (relevance, scalar, manner, contrastive implicatures, novel metaphors, and irony). Language of testing for these tests was Dutch. With regards to vocabulary, children were given the Peabody Picture Vocabulary Test III-NL (PPVT; Dunn, Dunn, & Schlichting, 2005) for receptive vocabulary and the Word Definitions task from the CELF 4-NL (Semel, Wiig, Secord, & Kort, 2008) for expressive vocabulary. The pragmatics test is described in detail in the following sections.

The parents of all children were also asked to complete a questionnaire asking for information regarding the children's language use and the family's socioeconomic status. Socioeconomic status (SES) was measured through the Family Affluence Scale (FAS; Boyce, Torsheim, Currie, & Zambon, 2006) and the parents' educational levels. Various other tasks were used but are not reported here (see Veenstra, Antoniou, Katsos, & Kissine, accepted).

Pragmatics test. The pragmatics test was designed on E-Prime (Psychology Software Tools, 2012) so that both reaction times and accuracy could be recorded. It included six sub-tests, each examining a different type of pragmatic meaning based on one of Grice's maxims of conversation. Specifically, the task included one sub-test on relevance implicatures (e.g., the utterance *It's raining* as a reply to the question *What kind of item do you want?* is irrelevant at the literal level and thus implicates that "The speaker wants the umbrella"); one on novel metaphors (e.g., the metaphorical utterance *He was a sinking ship* is false at the literal level and thus implicates the meaning "He was feeling sad"); one on irony (e.g., the statement *Yes, you know how much I like red clothes!* uttered in an ironic intonation violates the maxim of quality for listeners who know that the speaker does not like red clothes, and thus implicates the interpretation "No, you know that I definitely don't like red clothes"); one on manner implicatures (e.g., the statement *In this picture there is a shape with dots* is more prolix than the phrase *Square* and thus the listener can infer that with the former statement the speaker refers to an atypical shape with dots rather than to a square with dots); a sub-test

on scalar implicatures (e.g., in the sentence *There are stars on some of the cards*, the speaker used the less informative term *some* and thus implicates that the more informative term *all* does not hold); and a sub-test on contrastive implicatures (e.g., the phrase *open window* is over-informative if there is only one window in context, and thus implicates that there is another window in context).

For each sub-test, there was a critical condition with two items each (where the generation of a pragmatic meaning was required for accurate responding) and two literal conditions (Literal-1 and Literal-2) with two control or filler items each (where accurate responding required the children to understand only the explicit meaning of the target utterance).

There were three versions of the pragmatics test and each version was administered to an approximately equal number of participants overall and across groups. The children were asked to respond as fast and accurate as possible using a response box. Sample critical trials from a picture-selection sub-test (novel metaphors) and from the scalar implicature binary judgment task are shown in Tables 1 and 2, respectively.

Table 1. Sample critical trial from the sub-test on novel metaphors

Sequence	Pictorial stimulus	Auditory stimulus
1	Man looking at an expensive car	Niels's father had bought an expensive car. A few days ago, however, Niels accidentally kicked the ball on the car and broke one of its windows.
2	Man looking at an expensive car	Niels, how was your father feeling when he found out?
3	Man looking at an expensive car [Press the SPACE BAR to hear Niels's reply…][1]	
4	Man looking at an expensive car [Press the SPACE BAR to hear Niels's reply…][1]	He was a thunderous storm cloud.
5	Sad man, thunder falling from a cloud and man standing in the rain and holding an umbrella, angry man	

Note. Stimuli in brackets were presented in written form.
[1]Stimulus was presented at the bottom centre of the screen, below the picture.

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

Sequence	Pictorial stimulus	Auditory stimulus
1	[Press the SPACE BAR to continue…]	
2	Five cards face down	There are hearts on some of the cards.

3	Five cards 'turned over', each depicting a heart	

Note. Stimuli in brackets were presented in written form.

Relevance implicatures. Children heard short conversations between a character, named Thomas, and a female speaker, who asked Thomas questions about or requested from him to describe which of three items/pictures he wanted. Participants had to give Thomas one of the three items/pictures based on his reply.

Each trial was composed of various slides. For critical trials, Thomas was initially prompted to put on a blindfold (to motivate the indirectness of Thomas's replies in the critical trials of this sub-test). In the second slide, the target question was asked (e.g., *What kind of item do you want?*). Next, participants heard the target statement (e.g., *It's raining*) and, then, saw the three options (e.g., a winter hat, an umbrella, and a pair of shorts).

Control trials (Literal-1 condition; e.g., *It's raining*) differed from critical trials in that the speaker introduced three pictures (e.g., *Thomas, here are three pictures*) before asking the target question (e.g., *How is the weather in the picture that you want?*). For filler items (Literal-2 condition; e.g., *It's shining*), the speaker presented three items in separate slides (e.g., *Thomas, here is a cup of tea, here is an ice-cream, and here is a ring*) and, then, in the final slide, she requested from Thomas to describe to her what he wanted.

Novel metaphors. This sub-test was created drawing on previous work by Waggoner and Palermo (1989). It included only novel metaphors. These metaphors were selected from a pool of 12 metaphors rated from 32 adult Dutch speakers for novelty and aptness on a five-point scale (with a score of *5* indicating high novelty and aptness). All metaphors used in this sub-test had a mean novelty rating equal to or higher than 3.09/5 and a mean aptness rating equal to or higher than 1.97/5.

In this sub-test, children heard short dialogues between a character, named Niels, and a female speaker who asked Niels questions about how his father felt or what had happened at the end of a story. Children were shown three pictures and they had to select the picture that indicated how Niels's father felt or what had happened at the end of each story. For the critical items, one of the pictures depicted a sad man, a second depicted an angry man, and a third depicted a situation that was compatible with a literal-like interpretation of the target statement.

In the critical condition, the children heard stories about Niels's father. After the end of each story, the speaker asked the target question. In the next slide, participants heard the target statement, and after that, saw a visual display of three pictures. The same process was roughly followed in the other two conditions. In the Literal-1 condition, however, the target statement was an explicit description of how Niels's father felt (e.g., *He was an angry man*). Moreover, in the Literal-2 condition (e.g., *It was a thunderous storm cloud*), the children were instructed that they had to select the picture which indicated what had happened at the end of each story.

Manner implicatures. The general design of this sub-test was a sentence-to-picture matching task. Participants were initially introduced to the main character of this sub-test, a boy named Sander. They were informed that a speaker would ask Sander to describe a picture, they would then hear Sander's description, and that they had to select the picture described by Sander.

All conditions followed the same process. The children initially heard the female speaker asking Sander to describe one of three items (e.g., a picture of a motorcycle, a picture of an atypical vehicle with two wheels, and a picture of a car). In the final slide, participants heard the target description while viewing the three pictures.

Two of the control items included the simpler, alternative description of the critical statements within the same task version (e.g., *In this picture, there is a motorcycle* for the critical sentence *Vehicle with two wheels*) (Literal-1 condition) and the other two fillers described an item from the visual display of a critical item in another task version (e.g., *In this picture there is a triangle*) (Literal-2 condition).

In the critical condition, one of the pictures was compatible with a pragmatic interpretation of the target statement (e.g., it depicted an atypical vehicle with two wheels), the second picture corresponded to a literal but pragmatically infelicitous interpretation (e.g., picture of a motorcycle; competitor) and the third picture depicted an incorrect referent of the target description (e.g., a car). For critical trials, the corresponding simpler noun phrase was always presented in a previous trial. This was important to make explicit to participants that the speaker knew the word for the competitor and thus he did not just use the periphrastic description because he did not have a label for the competitor.

Contrastive Implicatures. This sub-test was designed drawing on the work of Kronmüller, Morisseau, and Noveck (2014). Children were initially presented with a character, named Martijn, who held two cards. The front side of the cards was visible only to the character. The children were instructed that Martijn would describe one of his cards and that their task was to figure what his second card depicted. For each trial, the children initially heard Mrtijn's description. In the next slide, the children read a question about Martijn's second card (*In your opinion, which one is Martijn's second picture?*). In the final slide, participants saw a visual display of three pictures.

In the critical condition, the correct picture could be pragmatically inferred (e.g., *In this picture there is an open window*). In the Literal-2 condition, Martijn explicitly provided a hint about what his second card depicted (e.g., *In the first picture, there is a sandwich. In the second picture, there is something black*). In the Literal-1 condition, Martijn's description did not provide any hints (either pragmatically or explicitly) about his second card (e.g., *In this picture there is lion*) and participants had to select between the other two cards of the visual display at chance.

Irony. The sub-test on irony comprehension was designed drawing on previous research by Kowatch, Whalen, and Pexman (2013). Children were introduced to a character, named Wouter. Once more, a female speaker asked

Wouter questions about which of three items he wanted. Participants were shown three items and they had to select the item they believed Wouter wanted.

Again, each trial was composed of various slides. For critical trials, the speaker initially described Wouter's preferences (e.g., *Wouter, I know that blue is your favourite colour for clothes and that you definitely don't like red clothes. But a red jacket would be nice to wear*). Then, she introduced three items, each in a different slide (e.g., *Here is a blue jacket, here is a red jacket, and here is a green jacket*) and the target question was heard (e.g., *Would you like to wear the red jacket, now?*). In the next slide, participants heard the target statement (e.g., *Yes, you know how much I like red clothes!* with an ironic intonation) and, after that, the three items were shown on the screen.

The Literal-1 and Literal-2 conditions included control literal-yes items (e.g., *Yes, you know how much I like red clothes!*) and control literal-no items (e.g., *No, you know how much I hate red clothes!*), respectively.

Scalar Implicatures. In this sub-test (modified from Antoniou, Cummins, & Katsos, 2016), participants were instructed that they will hear a character, named Bram, describing various visual displays. They had to indicate whether Bram's utterances were correct or incorrect descriptions of the respective pictures.

In each trial, participants initially saw the back side of five cards. The target statement was then heard (using the quantifiers *some*, *all*, and *none*) and, after that, the front side of the cards (each depicting one item) was revealed.

There were two critical under-informative cases using the quantifier *some* (e.g., *There are stars on some of the cards* for a visual display where all cards depicted a star). True and informative *some* and false *some* statements (Literal-1 and Literal-2 conditions, respectively) served as control trials.

4.2. Results

4.2.1 Preliminary analyses

There was sufficient variability in most pragmatic sub-tests (accuracies ranging from 44% in the metaphor and irony sub-tests to 82% for manner), besides for relevance where ceiling performance was observed (93%). For this reason, performance in the relevance sub-test was excluded from the subsequent analysis on accuracy (but not for reaction times). In addition, the three groups performed at ceiling in all literal conditions (lowest accuracy was 90%).

We also created composite scores from variables that were statistically and conceptually related to increase reliability of measurement for these variables (Rushton, Brainerd, & Pressley, 1983). A vocabulary proficiency composite score was computed based on children's scores in the PPVT and the Word Definitions task and a SES composite score was calculated based on maternal level of education, paternal level of education, and FAS score. The composites were formed by transforming the individual measures into *z* scores and then averaging the relevant measures.

4.2.2. Main analyses

Background measures. There were significant differences between the three groups in age in months ($F(2, 135)=3.625$, $p<.05$), SES ($F(2, 135)=80.56$, $p<.05$), and Vocabulary proficiency ($F(2, 135)=9.316$, $p<.05$).

In terms of age, bi-dialectal children tended to be older than both bilingual ($p=.08$, Bonferroni correction applied) and monolingual children ($p=.06$, Bonferroni correction applied). With regards to SES, bilingual children were of a higher SES than both bi-dialectal and monolingual children; and monolingual children had a higher SES than bi-dialectal children (all $ps<.05$, Bonferroni correction applied). Finally, with respect to vocabulary, monolingual children had a significantly higher score than both bi-dialectal and bilingual children (all $ps<.05$, Bonferroni correction applied). To account for these background differences, Age, Vocabulary, and SES were included as covariates in the between-group analyses on pragmatic performance.

Pragmatics accuracy. Descriptive statistics for accuracy in critical conditions of the pragmatics test by Type of pragmatic meaning (Metaphor, Irony, Manner, Contrastive, Scalars, Relevance) and Group (Bilinguals, Bi-dialectals, Monolinguals) are shown in Table 3.

An Analysis of Covariance (ANCOVA) was conducted with Type of pragmatic meaning as a within-subjects factor, Group (Bilinguals, Bi-dalectals, Monolinguals) and Version (One, Two, Three) as between-subjects factors, and Age, Vocabulary, and SES as covariates. Results indicated no significant differences between the three groups (for Group: $F(2, 119)=1.67$, $p>.05$; for the Version by Group interaction: $F(4, 119)=0.70$, $p>.05$; for the Type by Group interaction: $F(7.55, 448.98)=0.75$, $p>.05$; for the Type by Group by Version interaction: $F(15.09, 448.98)=1.35$, $p>.05$). Similar results were found when Vocabulary was not covaried in the analysis suggesting that bilingual and bi-dalectal children exhibit equivalent to monolinguals pragmatic performance despite their lower vocabulary (as measured by formal language tests).

Table 3. *Descriptive statistics (proportion correct, mean reaction times for correct responses, and standard deviations) from the pragmatics test (raw values) by type of pragmatic meaning and language group*

Type	Bi-dialectals (n=46)		Monolinguals (n=44)		Bilinguals (n=48)	
	Mean	(*SD*)	Mean	(*SD*)	Mean	(*SD*)
			Accuracy			
Scalars	.67	(.47)	.80	(.41)	.76	(.43)
Contrastive	.57	(.5)	.72	(.45)	.61	(.49)
Manner	.85	(.36)	.78	(.41)	.81	(.39)
Metaphor	.40	(.49)	.51	(.5)	.41	(.5)
Irony	.45	(.5)	.45	(.5)	.41	(.5)
Relevance	.95	(.23)	.92	(.27)	.92	(.28)
			Reaction Times			

Scalars	1611	(758)	1893	(1359)	1744	(1037)
Literal-1	1213	(450)	1279	(523)	1324	(957)
Literal-2	1345	(458)	1512	(782)	1384	(645)
Contrastive	3105	(2206)	3083	(2012)	3401	(2816)
Literal-1	3787	(2205)	4395	(2747)	5329	(5439)
Literal-2	1384	(778)	1383	(667)	1523	(864)
Manner	2271	(2594)	3056	(5651)	3260	(6202)
Literal-1	1275	(1652)	1399	(1775)	1392	(1448)
Literal-2	991	(1174)	856	(785)	814	(658)
Metaphor	5003	(4442)	5202	(3372)	5661	(3507)
Literal-1	2019	(830)	2072	(835)	2335	(1006)
Literal-2	2055	(947)	1749	(868)	2058	(996)
Irony	3273	(3953)	4507	(5613)	8356	(13517)
Literal-1	1127	(498)	1223	(1016)	1338	(1762)
Literal-2	1935	(1911)	1993	(1891)	1932	(2624)
Relevance	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, Literal-1: first condition with literal items, Literal-2: second condition with literal items.

Pragmatics test: reaction times. Average reaction times by Type of pragmatic meaning (Metaphor, Irony, Manner, Contrastive, Scalars, Relevance), Condition (Pragmatic, Literal-1, Literal-2) and Group (Bilinguals, Bi-dialectals, Monolinguals) are shown in Table 3.

Irony. An ANCOVA was conducted on reaction times with Group (Bilinguals, Bi-dialectals, Monolinguals) as a between-subjects factor, Condition (Pragmatic, Literal-1, Literal-2) as a within-subjects factor, and Age, Vocabulary, and SES as covariates. Results indicated no significant group differences (for the Group effect: $F(2, 61)=2.62$, $p=.08$; for the Condition by Group interaction: $F(2.11, 64.41)=2.74$, $p=.07$). The same results were largely found when Vocabulary was not covaried in the analyses. Moreover, the correlation between accuracy and reaction times for critical items was not significant (Spearman's rho=-.04, p(two-tailed)>.05), excluding the possibility of speed-accuracy trade-offs.

Metaphors. Similar analyses as above were conducted on reaction times in the Metaphor sub-test. Once more, results indicated a non-significant effect of Group ($F(2, 68)=1.063$, $p>.05$) and a non-significant Group by Condition interaction ($F(2.23, 75.75)=1.175$, $p>.05$). Results were the same when Vocabulary was not covaried into the analysis. The correlation between accuracy and reaction times for critical items was significant (Spearman's rho=.27, p(two-tailed)<.05), which suggests that some participants were responding slower in order to perform more accurately. To control for possible speed-accuracy trade-offs in the critical condition, we calculated efficiency scores by dividing mean reaction times by the percentage of accurate responses in critical trials. Similar ANCOVAs as above on the efficiency scores (with Vocabulary covaried but without the Condition factor) indicated no significant group differences ($F(2, 68)=0.77$, $p>.05$; without Vocabulary covaried: $F(2, 69)=0.59$, $p>.05$).

Scalars. An ANCOVA was conducted on reaction times with Group as a between-subjects factor, Condition as a within-subjects factor and Age, Vocabulary, and SES as covariates. Results indicated a significant Group effect ($F(2, 105)=5.21$, $p<.05$) but a non-significant Group by Condition interaction ($F(2.98, 156.41)=2.54$, $p=.059$). Post-hoc pairwise comparisons with Bonferroni correction applied showed only a significant difference between bi-dialectals and monolinguals in that bi-dialectals were overall faster ($p<.05$; for the comparison between bi-dialectals and bilinguals, $p=.051$). The analysis without Vocabulary covaried showed no significant results. Finally, the correlation between accuracy and reaction times in critical trials was not significant (Spearman's rho=.16, p(two-tailed)>.05), excluding the possibility of speed-accuracy trade-offs.

Relevance implicatures. An ANCOVA on reaction times with Group, Condition, age, Vocabulary, and SES included in the analysis revealed a significant Group effect ($F(2, 131)=3.65$, $p<.05$) but a non-significant Group by Condition interaction ($F(2.99, 195.67)=1.931$, $p>.05$). Post-hoc pairwise comparisons with Bonferroni correction applied showed that bi-dialectal children were significantly faster than both bilingual and monolingual children ($p<.05$). Moreover, the correlation between accuracy and reaction times for critical items was not significant (Spearman's rho=.13, p(two-tailed)>.05), excluding the possibility of speed-accuracy trade-offs.

Contrastive implicatures. An ANCOVA was conducted on reaction times with Group (Bilinguals, Bi-dialectals, Monolinguals) as a between-subjects factor. Condition (Pragmatic, Literal-1, Literal-2) as a within-subjects factor, and age, Vocabulary, and SES as covariates. Results indicated a non-significant effect of Group ($F(2, 123)=1.400$, $p>.05$) and a non-significant Group by Condition interaction ($F(3.55, 218.04)=0.908$, $p>.05$). Results were the same when Vocabulary was not covaried into the analysis.

Finally, the correlation between accuracy and reaction times for critical items was significant (Spearman's rho=.29, p(two-tailed)<.05) suggesting the possibility of speed-accuracy trade-offs in participants' performance. A between-group analysis on efficiency scores, however, showed, once more, no group differences ($F(2, 125)=1.137$, $p>.05$, for the analysis with vocabulary covaried).

Manner implicatures. An ANCOVA was conducted on reaction times with Group (Bilinguals, Bi-dialectals, Monolinguals) as a between-subjects factor, Condition (Pragmatic, Literal-1, Literal-2) as a within-subjects factor, and Age, Vocabulary, and SES as covariates. Results indicated a non-significant effect of Group ($F(2, 123)=0.116$, $p>.05$) and a non-significant Group by Condition interaction ($F(2.22, 136.45)=0.190$, $p>.05$). Results were the same when Vocabulary was not covaried into the analysis. Finally, the correlation between accuracy and reaction times for critical items was not significant (Spearman's rho=-.012, p(two-tailed)>.05), which suggests that no speed-accuracy trade-offs were present in this sub-test.

5. Discussion

To date, research on bilingual children's pragmatic skills has provided mixed results, with some research reporting enhanced pragmatic skills in bilingual children and other studies finding no differences in pragmatic development between bilingual and monolingual children. In the present study, we sought to investigate the hypothesis of a bilingual pragmatic advantage in a new sample of Dutch-speaking bilingual, bi-dialectal, and monolingual children. We used a comprehensive pragmatics test that examined the interpretation of a wide range of pragmatic meanings at both the accuracy and processing level.

Our results indicated largely no differences in pragmatic comprehension between bilingual, bi-dialectal, and monolingual children. Bi-dialectals were faster than monolinguals in the scalar and relevance implicatures sub-tests, but the difference was found in overall performance (i.e., across the three conditions) and, hence, cannot be attributed to more efficient pragmatic processing *per se*. Thus, there was no clear evidence in our data for a bilingual or bi-dialectal pragmatic advantage over monolinguals. Furthermore, bilingual and bi-dialectal children exhibited monolingual-like pragmatic comprehension skills despite their lower vocabularies as measured by standardised language tests. This further suggests that bilingual children's often-reported weaker language knowledge does not have any further negative implications in terms of ecologically valid aspects of communicative competence, such as pragmatic interpretation.

The results of our study are in line with past research conducted by Antoniou and Katsos (2017) and Syrett and colleagues (2016; 2017). Thus, the bulk of research to date suggests that the finding of null differences in pragmatic comprehension between bilingual and monolingual children is robust; it holds at both the accuracy and the processing level and it's true across types of pragmatic meanings, bilingual or bi-dialectal samples, and age groups.

References

Antoniou, Kyriakos., Cummins, Chris., & Katsos, Napoleon. (2016). Why only some adults reject under-informative utterances. *Journal of Pragmatics, 99*, 78-95.

Antoniou, Kyriakos, & Katsos, Napoleon. (2017). The effect of childhood multilingualism and bilectalism on implicature understanding. *Applied Psycholinguistics, 38*, 787-833.

Barac, Raluca, Bialystok, Ellen, Castro, Dina C., & Sanchez, Marta. (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

Bialystok, Ellen, Luk, Gigi, Peets, Kathleen F., & Yang, Sujin. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition, 13*, 525-531.

Boyce, William, Torsheim, Torbjorn, Currie, Candace, & Zambon, Alessio. (2006). The family affluence scale as a measure of national wealth: validation of an adolescent self-report measure. *Social indicators research, 78*, 473-487.

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

Dunn, Lloyd M., Dunn, Leota M., & Schlichting, Johanna E. P. T. (2005). *Peabody picture vocabulary test-III-NL*. Harcourt Test Publishers.

Grice, Paul H. (1975). Logic and conversation. In P. Cole and J.L. Morgan (Eds.), *Syntax and semantics*, (Vol. 3, pp. 225-242). New York: Seminar Press.

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

Hammer, Carol S., Hoff, Erika, Uchikoshi, Yuuko, Gillanders, Cristina, Castro, Dina C., & Sandilos, Lia E. (2014). The language and literacy development of young dual language learners: A critical review. *Early Childhood Research Quarterly, 29*, 715-733.

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

Kovács, Ágnes M. (2008). Early bilingualism enhances mechanisms of false-belief reasoning. *Developmental Science, 12*(1), 48-54.

Kowatch, Kristy, Whalen, Juanita M., & Pexman, Penny M. (2013). Irony comprehension in action: A new test of processing for verbal irony. *Discourse Processes, 50*, 301-315.

Kronmüller, Edmundo, Morisseau, Tiffany., & Noveck, Ira A. (2014). Show me the pragmatic contribution: a developmental investigation of contrastive inference. *Journal of child language, 41*, 985-1014.

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

Liberman, Zoe, Woodward, Amanda L., Keysar, Boaz, & Kinzler, Katherine D. (2017). Exposure to multiple languages enhances communication skills in infancy. *Developmental science, 20*.

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

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

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

Rushton, Philippe J.., Brainerd, Charles J., & Pressley, Michael. (1983). Behavioral development and construct validity: The principle of aggregation. *Psychological bulletin, 94*, 18.

Semel, Eleanor. M., Wiig, Elisabeth H., Secord, Wayne A., & Kort, Willem. (2008). *CELF® 4 NL: Clinical evaluation of language fundamentals: Nederlandse versie*. Pearson.

Siegal, Michael, Iozzi, Laura, & Surian, Luca. (2009). Bilingualism and conversational understanding in young children. *Cognition, 110*, 115-122.

Siegal, Michael, Matsuo, Ayumi, & Pond, Clair. (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 Michael, Surian Luca, Matsuo Ayumi, Geraci Alessandra, Iozzi Laura, et al. (2010). Bilingualism Accentuates Children's Conversational Understanding. *PLoS ONE 5*(2): e9004. doi:10.1371/journal.pone.0009004

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

Syrett, Kristen, Austin, Jennifer, Sanchez, Liliana, Germak, Christina, Lingwall, Anne, et al. (2016). The influence of conversational context and the developing lexicon on the calculation of scalar implicatures. *Linguistic Approaches to Bilingualism, 7*, 230-264.

Syrett, Kristen, Lingwall, Anne., Perez-Cortes, Silvia, Austin, Jennifer, Sánchez, Liliana, et al. (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*.

Veenstra, Alma, Antoniou, Kyriakos, Katsos, Napoleon, & Kissine, Mikhail. (accepted). 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, John E., & Palermo, David S. (1989). Betty is a bouncing bubble: Children's comprehension of emotion-descriptive metaphors. *Developmental psychology, 25*, 152.

Wermelinger, Stephanie, Gampe, Anja, & Daum, Moritz M. (2017). Bilingual toddlers have advanced abilities to repair communication failure. *Journal of experimental child psychology, 155*, 84-94.

Wilson, Deirdre, & Sperber, Dan. (2005). Pragmatics. In F. Jackson & M. Smith (Eds.), *Oxford Handbook of Contemporary Philosophy* (pp. 468-501). New York: Oxford University Press

Yow, Quin W, & Markman, Ellen M. (2015). A bilingual advantage in how children integrate multiple cues to understand a speaker's referential intent. *Bilingualism: Language and Cognition, 18*, 391-399.