THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Are children with Specific Language Impairment competent with the pragmatics and logic of quantification?

Citation for published version:
Katsos, N, Roqueta, CA, Clemente Estevan, RA & Cummins, C 2011, 'Are children with Specific Language Impairment competent with the pragmatics and logic of quantification?' Cognition, vol 119, no. 1, pp. 43-57. DOI: 10.1016/j.cognition.2010.12.004

Digital Object Identifier (DOI):
10.1016/j.cognition.2010.12.004

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Cognition

Cognition xxx (2011) xxx–xxx

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Are children with Specific Language Impairment competent with the pragmatics and logic of quantification?

Napoleon Katsos [a,*], Clara Andrés Roqueta [a,b], Rosa Ana Clemente Estevan [b], Chris Cummins [a]

[a] Research Centre for English and Applied Linguistics, University of Cambridge, United Kingdom
[b] Department of Developmental and Educational Psychology, Universitat Jaume I de Castelló, Spain

ARTICLE INFO

Article history:
Received 22 October 2009
Revised 22 November 2010
Accepted 5 December 2010
Available online xxxx

Keywords:
Specific Language Impairment
Pragmatics
Informativeness
Scalar implicature
Logic
Quantification
Quantifiers

ABSTRACT

Specific Language Impairment (SLI) is understood to be a disorder that predominantly affects phonology, morphosyntax and/or lexical semantics. There is little conclusive evidence on whether children with SLI are challenged with regard to Gricean pragmatic maxims and on whether children with SLI are competent with the logical meaning of quantifying expressions. We use the comprehension of statements quantified with 'all', 'none', 'some', 'some…not', 'most' and 'not all' as a paradigm to study whether Spanish-speaking children with SLI are competent with the pragmatic maxim of informativeness, as well as with the logical meaning of these expressions.

Children with SLI performed more poorly than a group of age-matched typically-developing peers, and both groups performed more poorly with pragmatics than with logical meaning. Moreover, children with SLI were disproportionately challenged by pragmatic meaning compared to their age-matched peers. However, the performance of children with SLI was comparable to that of a group of younger language-matched typically-developing children. The findings document that children with SLI do face difficulties with employing the maxim of informativeness, as well as with understanding the logical meaning of quantifiers, but also that these difficulties are in keeping with their overall language difficulties rather than exceeding them. The implications of these findings for SLI, linguistic theory, and clinical practice are discussed.

1. Introduction

Specific Language Impairment (SLI) is a disorder or delay in the development of expressive and/or receptive language in the absence of other cognitive, motor or auditory impairments (American Psychiatric American Psychiatric Association, 1994; see also Bishop, 1997; Bishop, 2000; Leonard, 1998). Hallmark features of the disorder include a pronounced delay in the production and/or comprehension of morphosyntax, phonology and vocabulary. In the current understanding of SLI no significant impairments are expected to manifest themselves in non-core domains of language such as pragmatics. Instead, pragmatics, non-verbal communication and social cognition are considered to be areas of relative strength (Bishop, 2002; Leonard, 1998; among others). As such, children with SLI should not face particular difficulties with employing Gricean maxims (1975/1989), which are conversational rules that enjoin interlocutors to be cooperative, and specifically informative, true, concise and relevant. That is, in the event that they do face difficulties, these should not exceed the level of difficulty that could be expected given the children's overall language delay.

In this paper we review previous research on the competence of children with SLI with regard to Gricean pragmatic maxims and we highlight methodological and conceptual issues that warrant further investigation. We then use the comprehension of quantified statements as

* Corresponding author.
E-mail address: nk248@cam.ac.uk (N. Katsos).

0010-0277/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2010.12.004

a paradigm to study whether children with SLI are competent with the Gricean maxim of informativeness.

Besides investigating pragmatic meaning, the present experiment also assesses children's competence with the logical meaning of quantifiers. This aspect of meaning is underpinned by set-theoretical relations rather than pragmatic principles and has not been studied extensively in children with SLI. Because pragmatic and logical competence with quantifying expressions is a major aspect of everyday communication and reasoning (Moxey & Sanford, 2000; Newton, Donlan, & Roberts, 2010; Sanford & Moxey, 2003; among others), documenting whether children with SLI face difficulties in these areas is critical from a clinical point of view. Moreover, documenting the extent and nature of any difficulties would enhance our understanding of whether SLI is a predominantly grammatical impairment. This contributes to the broad debate on whether the communicative difficulties that are reported in SLI (e.g. in understanding and constructing narratives) are proportionate to or in excess of what would be expected given the children's general language difficulties (Bishop, 1997; Leonard, 1998). Research along these lines also has implications for linguistic theory and the role of grammar in determining how pragmatic and logical inferences are generated.

In the following sections we review previous work on children with SLI and their competence with Gricean maxims and logical reasoning. We then introduce some basic concepts in the semantics and pragmatics of quantification and we present the paradigm used in this study.

1.1. Previous investigations on the pragmatic skills of children with SLI and other atypically-developing populations

Surian, Baron-Cohen, and van der Lely (1996) investigated the mastery of Gricean maxims of quantity, quality, and relation, as well as politeness, by English-speaking children with SLI, children with High Functioning Autism, and a language-matched typically-developing control group. Participants were presented with two puppets who answered a question, one of them in a pragmatically appropriate way and one in a way that violated expectations of informativeness, truthfulness, relevance, or politeness. They were then asked to point to the puppet that said something silly. For example, with regard to the first maxim of quantity, which enjoins interlocutors not to be under-informative, the puppets were asked "How would you like your tea?" One of the puppets would respond "With milk", while the other would respond "In a cup", which is presumably true but nevertheless under-informative. Surian et al. reported that children with SLI pointed to the pragmatically-inappropriate puppet at rates significantly above chance overall, and at the same levels as the control group. Thus, they concluded that children with SLI do not face exceptional challenges with Gricean maxims.

However, there are two issues that should be explored further before endorsing this conclusion. First, the typically-developing language-matched group's language level was almost a year below their chronological age (verbal age on the TROG, Bishop, 1983, was 5.8, compared to a chronological age of 6.7). It is possible in principle that a comparison with a truly average typically developing group could have yielded statistically significant differences where this study did not, and shown the SLI group to be pragmatically impaired by comparison.

Second, Surian et al. (1996) report that both children with SLI and the control group were above chance in the task overall, but that there were significant differences when looking at each maxim separately. Children with SLI were at ceiling for statements that violated expectations of truthfulness, relation or politeness (100%, 100% and 98% respectively). However, they performed at rates of only 63% and 65% with the first and second maxims of quantity (roughly paraphrased as 'do not be under-informative' and 'do not be over-informative' respectively). The control group also performed at ceiling rates for truthfulness, relation and politeness but at 58% and 78% with under- and over-informativeness respectively. In fact, neither the control group nor children with SLI performed at levels above chance with the first maxim of quantity, while children with SLI did not perform above chance with the second maxim either.

These differences across conditions could be explained by the fact that successful performance in this task did not necessarily rely solely on mastery of the Gricean maxims themselves. Violations of truthfulness and relation were blatantly incongruous statements: for example, given the question "Where do you live?", the statement that violated truthfulness was "On the moon". Participants could reject these statements based on the implausibility of the statement at the factual level, which relies on encyclopedic knowledge about the world, without having to consider whether or not the speaker knowingly violated a conversational maxim.

Moreover, violations of politeness involve competence with culture-specific social norms, rather than with the culture-independent considerations of communicative efficiency that underpin the maxims of quantity, relation and manner (von Fintel & Matthewson, 2008). It is not necessarily surprising that children with SLI performed well with this maxim, especially bearing in mind that social cognition is considered an area of relative strength (Bishop, 2002).

Furthermore, the cases where successful performance does depend solely on mastery of the maxims themselves – that is, the violations of the first and second maxim of quantity – are also the only ones where chance performance is obtained. In any case, the ceiling or chance performance reported for each of the maxims that were tested cannot reveal whether the pragmatic skills of children with SLI are at comparable levels with their language-matched counterparts. It merely documents that children found the task sometimes too difficult and sometimes too easy.

In a more recent study on children with autistic spectrum conditions and/or language impairment, Norbury (2005) documents that children with SLI (as well as children with autistic spectrum conditions and low language skills) have problems with understanding metaphors such as "Mum left the bread out overnight. This morning it was a brick". Comprehension of metaphorical expressions relies on the Gricean maxim which enjoins interlocutors to be truthful, and on the recognition that the comprehender should look for non-literal interpretations of the metaphoric expression in order to preserve truthfulness. Comprehension further relies on the maxim of relation which mandates that the non-literal interpretations to be considered ought to be relevant to the context. Norbury (2005) reports that performance in the pragmatic task is primarily predicted by advanced world-knowledge and richness of lexical representations for the critical words, which includes the encyclopedic knowledge that allows the metaphoric interpretations. This result demonstrates empirically what we argued for at the conceptual level, namely that performance in a task that investigates Gricean pragmatics can be mediated by non-Gricean pragmatic competence such as world-knowledge. In the next section, we discuss the linguistic concepts underlying a task that we designed to investigate competence with the first maxim of quantity, in which we attempt to minimise the extent to which pragmatic task performance depends upon other non-pragmatic skills.

1.2. Quantification and informativeness

In natural language, quantifying expressions like 'all', 'most' and 'some' are used to express relations between sets. For example, the statement 'all the apples are in the boxes' is true if all the apples under discussion are inside the boxes, and it is false in every other circumstance. The statement 'most of the apples are in the boxes' is true if the number of apples inside the boxes is greater than the number of apples outside and false otherwise. Similarly, the statement 'some of the apples are inside the boxes' (where the noun phrase and the copula verb are inflected for plural) is true if two or more apples are in the boxes, and false otherwise (Gamut, 1991; Horn, 1972; among others). In addition to being evaluated for their truth or falsity, the latter two statements can also be evaluated for whether they are sufficiently informative or not. Take for example the situation in (1a) and compare its descriptions in (1b) and (1c):

(1) a. Situation: All the apples are in the boxes
b. Description 1: Some of the apples are in the boxes
c. Description 2: All the apples are in the boxes

When we employ the standard definition for 'some', all that (1b) says is that two or more of the apples are in the boxes, which is strictly speaking true of the situation in (1a). The robust intuition that (1b) is not an appropriate description for (1a) (even though it is true of that situation) is captured by Grice's Cooperative Principle and the first maxim of quantity, also called the maxim of informativeness (1975/1989; see also Horn, 1984; Levinson, 1983; among others). Since it is known that in the present situation all the apples are inside the boxes, it is more informative to say so explicitly, i.e. to say (1c). This is preferred over using a statement like (1b), which is logically compatible with all the apples being in the boxes, but is less informative (to confirm this, note that (1c) entails (1b) but not the other way round). A pragmatically competent speaker who used (1b) would typically be implicitly communicating that not all of the apples are inside the boxes. This implicitly communicated aspect of meaning is known in the literature as an *implicature*. Since this implicature relies on logical scales of informativeness (in this case the scale formed by 'some' and 'all') it is known as a *scalar implicature*.

Thus, describing situation (1a) with a statement like (1b) ought to be rejected by pragmatically competent interlocutors for at least one of the two following reasons: because they have detected that their interlocutor is under-informative and is thus violating a fundamental pragmatic principle; or because they have understood (1b) with the scalar implicature that 'not all the apples are in the boxes', a proposition which is obviously not true in this situation.

Similarly, a statement such as 'most of the apples are inside the boxes' is logically true but pragmatically under-informative in a situation like (1a). Also, statements such as 'some of the apples are not in the boxes' and 'not all the apples are in the boxes' are logically true in a situation where all of the apples are outside the boxes. However, they are under-informative, since a pragmatically competent speaker should have described this situation by stating that 'none of the apples are in the boxes'. Therefore, a pragmatically competent speaker using 'some…not' or 'not all' would typically be implicating that it is not the case that no apples are in the boxes, i.e. that some of the apples are in the boxes.

Several investigations in various languages have studied children's competence with informativeness and their ability to generate scalar implicatures (English: Feeney, Scrafton, Duckworth, & Handley, 2004; Huang & Snedeker, 2009; Hurewitz, Papafragou, Gleitman, & Gelman, 2006; Katsos, 2009; French: Noveck, 2001; Pouscoulous, Noveck, Politzer, & Bastide, 2007; Greek: Papafragou & Musolino, 2003; Papafragou & Tantalou, 2004; Italian: Guasti et al., 2005; among others). These studies have mainly investigated typically-developing children's acceptance or rejection of statements with 'some' in contexts where the corresponding statement with 'all' would also be true (and hence more informative). The degree to which children's performance is adult-like has been shown to depend upon factors such as the age of children studied, the provision of a visual context against which utterances are evaluated rather than reliance upon world-knowledge (Guasti et al., 2005), directing children's attention to detecting under-informativeness (Papafragou & Musolino, 2003) and the choice of specific lexical items (Pouscoulous et al., 2007).

In all cases, however, it has been documented that children under 7 years old do not reject under-informative statements with 'some' at the same rates as adults. Moreover, they do not reject under-informative statements at the same level as they reject logically false ones. That is, while children are perfectly able to reject (1b) in a situation where none of the apples are in the boxes, they predominantly accept it in a situation where all the apples are in the boxes, i.e. where the statement is logically true but under-informative. Notice that successful performance in the former condition relies solely on mastery of the

Please cite this article in press as: Katsos, N., et al. Are children with Specific Language Impairment competent with the pragmatics and logic of quantification?. *Cognition* (2011), doi:10.1016/j.cognition.2010.12.004

logical meaning of the quantifier, while in the latter condition, relying on the logical meaning leads to an interpretation that is strictly speaking true. This should lead to acceptance, unless the child has mastered the maxim of informativeness, which mandates that the statement ought to be rejected in this context. Hence, in the task that we describe below, any difference between performance in the logically false condition on the one hand, and in the logically true but pragmatically under-informative condition on the other, can be straightforwardly attributed to mastery (or lack thereof) of the maxim of informativeness.

1.3. Quantification and logical meaning

Having drawn a distinction between informativeness and truth with respect to principles of pragmatics and logic, we should also note that quantifiers like 'all' and 'none', 'some' and 'some…not', 'most' and 'not all' exhibit several differences and similarities with regard to their logical meaning (for an overview, see Chierchia & McConnell-Ginet, 2000; Gamut, 1991; among others).

First, 'all' and 'some' contrast with 'none' and 'some-not' in terms of polarity. The former two are positive quantifiers, in the sense that they are used to state properties that the sets under discussion do have, while the latter two are negative, used to state properties that the sets under discussion do not have. Since negation is linguistically and psycholinguistically complex compared to affirmation (see Horn, 1989; Just & Carpenter, 1971; among others), it is predicted that positive quantifiers are easier to master than negative ones.[1]

Second, quantifiers like 'some' and 'most' can be compared in terms of the complexity of their logical meaning. While 'some of the apples are in the boxes' requires that an unspecified quantity of apples are in the boxes, 'most of the apples are in the boxes' further specifies that this quantity must be more than half of the apples under discussion. This difference in logical complexity has been argued to give rise to cognitive, psychologically relevant, complexity (van Benthem, 1986; van Rooij, 2008, i.a.). Indeed, children acquire simple quantifiers such as 'some' and 'all' earlier than complex ones like 'most' (see Barner, Chow, & Yang, 2009; Hanlon, 1987; Hanlon, 1988 and references therein) and the link between logical complexity, adult reaction times and children's order of acquisition has been documented for a wide range of quantifying expressions (see Cummins & Katsos, 2010; Geurts, Katsos, Cummins, Moons, & Noordman, 2010; Musolino, 2004). Thus, with regard to child performance we predict that participants will perform better with the logical meaning of 'some' than with 'most'.

However, while the arithmetical and mathematical challenges that children with SLI face are well documented (Fazio, 1994; Fazio, 1996; Fazio, 1999; see also Cowan, Donlan, Newton, & Lloyd, 2005; Donlan, Cowan, Newton, & Lloyd, 2007), much less is known about SLI and logical language. Recently, Newton et al. (2010) reported that children with SLI have difficulties in deductive reasoning. Language level, non-verbal intelligence and working memory were implicated in these difficulties, depending on the specifics of the reasoning tasks used. In an investigation that studied logical language more closely, Newton (2007) reported that children with SLI performed poorly in a categorical syllogism task. For example, given (2a) and (2b), children were asked to choose from the four options in (2i–iv):

(2) a. All clowns hold blue balloons
b. All people standing on boxes are clowns
i. All people standing on boxes hold blue balloons
ii. Some people standing on boxes hold blue balloons
iii. No people standing on boxes hold blue balloons
iv. None of the above

Children with SLI children scored lower than age-matched typically-developing children. However, they scored comparably to a group of language-matched children, suggesting that their difficulties are proportionate with their overall language level. It is an open question whether the difficulties children with SLI face are due to problems in grasping the logical meaning of the quantifiers 'all', 'some' and 'none', or with the demands of the syllogistic process itself.

In the following section, we report a new study which was designed to investigate competence with informativeness and logical meaning in Spanish-speaking children with SLI. The task we used is the Spanish translation of the Quantifier Comprehension task by Katsos and Smith (2010). The Spanish expressions 'todos/todas', 'ninguno/ninguna (de los/las)', 'algunos/algunas', 'algunos/algunas…no', 'la mayoría (de los/las)', and 'no todos/todas' have the same logical properties as the English expressions 'all', 'none (of the)', 'some', 'some…not', 'most (of the)' and 'not all' respectively.

From the consideration above, we expect that the predictions in Table 1 will hold for typically-developing children. In the present study, this is a group of typically-developing children matched for age (henceforth *AM-TD*) to a group of children with SLI and a group of younger typically-developing children matched for language (henceforth *LM-TD*). Whether the predictions will be borne out for the group of children with SLI as well is an empirical issue.

In addition to studying within-group effects, between-group comparisons will also be performed to reveal any

Table 1
Within-group predictions with regard to pragmatic and logical competence. '>' stands for "yields higher performance than".

Prediction 1: Polarity	Positive ('all', 'some') > negative ('none', 'some...not')
Prediction 2: Logical Complexity	'Some' > 'most'
Prediction 3: Type of meaning	Logically false > pragmatically under-informative (for 'some', 'some…not', 'most', 'not all')

[1] In principle "some Xs are not Y" exhibits a scope ambiguity, but the so-called isomorphic reading, which can be paraphrased as "there exist some Xs such that they are not Y", is strongly preferred over the non-isomorphic (paraphrased as "it is not the case that there exist some Xs such that they are Y" (Musolino & Lidz, 2006, i.a.).

particular strengths or weaknesses of the SLI group relative to the typically-developing children. Thus, if pragmatics and/or logical meaning is an area of particular strength for children with SLI, they are expected to perform similarly to the AM-TD group or at least better than the LM-TD; while if any one of these skills is an area of exceptional SLI difficulty, they are expected to perform even more poorly than the LM-TD group. Performance at the same level as the LM-TD group would indicate that the competence of children with SLI is commensurate with their overall language level – that is, neither an area of particular strength nor of particular weakness.

2. The experiment

2.1. Method

2.1.1. Participants

Twenty-nine Spanish-speaking children with SLI (mean age 6.5, range 4.0–9.1) were recruited from language centres attached to state primary schools in the north-east of Spain. These language centres are attended by children with various developmental or language delays or disorders. The children in this study were recommended for participation by their attending speech and language therapist and/or educational psychologist based on the following criteria: they had a history of language delay in their clinical record; language problems were mentioned as their primary reason for attending the language centre; they did not have sensorineural hearing loss, or any other known mental or physical handicap; no medical condition likely to affect language was reported in their clinical record, such as a diagnosis of Autistic Spectrum Conditions; all were native speakers of Spanish; and finally, they were on the caseloads of the speech and language therapist and/or educational psychologist of the centre at the time of their participation in the study.

In addition to these criteria, inclusion in the SLI group was conditional on scoring 1SD below age level on either of two standardised language measures administered by the researchers. The first was a receptive language measure, the 'Comprensión de Estructuras Gramaticales' (*CEG*; Mendoza, Carballo, Muñoz, & Fresneda, 2005), which is similar in aims and structure to the Test for Reception of Grammar for English-2 (*TROG-2*; Bishop, 2003), and measures the comprehension of grammatical constructions with sentences of varying length and complexity. The second was an expressive language measure, the Sentence Recall subtest of the 'Evaluación del Lenguaje Infantil' battery (*ELI*; Saborit & Julián, 2005), which measures expressive language ability and short term auditory memory, which are measures that are highly sensitive to language impairment (e.g. Conti-Ramsden, Botting, & Faragher, 2001). Inclusion in the SLI group was also conditional upon having non-verbal IQ abilities within 1SD of the mean on the Raven Coloured Progressive Matrices test (Raven, Raven, & Court, 1998).

As mentioned above, two typically-developing control groups were also selected: an age-matched group (AM-TD; mean age 6.4, range 3.8–9.3); and a language-matched group (LM-TD; mean age 4.7, range 3.2–7.7). Both groups were matched one-to-one to the children with SLI on the basis of gender and school that they attended (the latter was done to minimise the effects of educational and social variation). The typically-developing controls were also native speakers of Spanish, and had no sensorineural hearing loss, or any mental or physical handicap or medical condition likely to affect language. They did not have a highly-gifted student recommendation on their school record and their performance on the Raven Coloured Progressive Matrices test was within 1SD of the age-appropriate mean. The AM-TD children were matched to the children with SLI within ±3 months of age. Because the quantifier task reported in the following section assesses comprehension rather than production, the LM-TD children were matched to the children with SLI on their raw scores on the receptive language test.

A series of pairwise comparisons between the SLI, AM-TD and LM-TD groups on age, expressive and receptive language and vocabulary as well as non-verbal IQ raw test scores and percentiles by means is reported in Table 2. Besides the measures used as selection criteria, the SLI group was matched to the LM-TD group on raw score on expressive and receptive vocabulary (measured by the subtests of the ELI battery). However, the AM-TD group outperformed both the SLI and the LM-TD group on raw scores on these measures.

2.1.2. Procedure

2.1.2.1. Task administration. All children were assessed in school during school-time in three or four sessions. The first two sessions of roughly 30 min each were used to administer the selection and matching measures as well as other tasks not reported in this paper. The quantifier task, which took about 35 min, was administered in one or two subsequent sessions, depending on how quickly the participating child's completed the task and how tired the child appeared to be.

2.2. Quantifier task: experimental procedure

Participants were seated in front of a laptop computer and they were shown visual displays on the screen while listening to sentences. Before the experiment started, the experimenter introduced a fictional character, the Cavegirl, and explained that they were going to help the Cavegirl to learn better Spanish so that she could speak as well as they do. In this particular game, the Cavegirl will say how many toys are in the boxes. If what she says is right, the participant is instructed to tell her "that was right". If what she says is wrong, the participant should tell her "that was wrong" and, in order to help her learn, they should tell her why it was wrong. In the first two visual displays, the experimenter familiarised the participant with the boxes and the objects that might appear in the boxes.

Overall, this experimental paradigm is designed to test quantified utterances in the context of a game where being true and being informative are intended to be understood as important. Based on previous research on the factors that affect child performance with quantified statements, these are two of the main conditions that elicit high

Table 2
Mean, standard deviation ('SD') and median ('Me') of descriptive, selection and matching measures for the SLI, AM-TD and LM-TD groups.

	(a) SLI ($n = 29$)		(b) AM-TD ($n = 29$)		(c) LM-TD ($n = 29$)		*Pairwise*
	Mean (SD)	Me	Mean (SD)	Me	Mean (SD)	Me	
Descriptive measures							
Age (months)	78.20 (18.65)	83	77.51 (17.51)	83	57.00 (14.83)	51	$a \approx b$ $c < a^{**}$ $c < b^{**}$
Gender (M/F)	20/9		20/9		20/9		–
Selection measures							
Non-verbal IQ percentile score	71.37 (20.04)	75	75.86 (16.98)	75	78.68 (16.53)	85	$a \approx b$ $a \approx c$ $b \approx c$
Non-verbal IQ raw score	22.41 (5.98)	24	23.65 (4.77)	23	18.03 (5.35)	18	$a \approx b$ $c < a^{*}$ $c < b^{**}$
CEG percentile score	17.13 (13.47)	15	67.20 (21.60)	70	54.86 (20.51)	50	$a < b^{**}$ $a < c^{**}$ $c < b^{*}$
CEG raw score	49.17 (11.28)	53	63.58 (9.49)	64	52.00 (10.18)	53	$a < b^{**}$ $a \approx c$ $c < b^{**}$
Sentence Recall percentile score	11.79 (13.6)	5	72.31 (29.84)	90	60.72 (28.70)	60	$a < b^{**}$ $a < c^{**}$ $c < b^{\blacklozenge}$
Sentence Recall raw score	5.55 (1.91)	6	7.89 (1.54)	9	6.41 (1.74)	7	$a < b^{**}$ $a < c^{\blacklozenge}$ $c < b^{*}$
Matching measures							
Vocab-receptive percentile score	61.38 (30.76)	70	86.24 (13.28)	95	73.03 (24.83)	80	$a < b^{**}$ $a \approx c$ $c < b^{*}$
Vocab-receptive raw score	20.24 (6.32)	22	23.93 (4.41)	26	18.48 (5.06)	18	$a < b^{*}$ $a \approx c$ $c < b^{**}$
Vocab-expressive percentile score	41.76 (25.89)	30	72.90 (23.77)	80	69.66 (24.05)	75	$a < b^{**}$ $a < c^{**}$ $b \approx c$
Vocab-expressive raw score	18.10 (7.02)	19	22.00 (6.40)	24	16.13 (6.25)	16	$a < b^{*}$ $a \approx c$ $c < b^{*}$

Note: By two-tailed Mann-Whitney *U*-test. CEG is 'Comprensión de Estructuras Gramaticales' language measure (Mendoza et al., 2005).
* $P < 0.05$; ** $P < 0.001$; ♦ $P < 0.1$.

performance from children (see Guasti et al., 2005; Papafragou & Musolino, 2003; among others).

2.3. Items

The six quantifying expressions mentioned above were tested. These quantifiers were combined with an object to form sentences of the type '[quantifier] (of the) [objects] are (not) in the boxes'. Each sentence was combined with a visual display in which a certain number of boxes contained an instance of the object that is mentioned. Sentences with 'all' were combined with displays where five out of five or two out of five boxes had an object inside. Sentences with 'none' were combined with displays where no boxes or two out of five boxes had an object inside. Sentences with 'some', 'some…not' and 'not all' were combined with displays where zero, two, or five out of five boxes had an object inside. Sentences with 'most' were combined with displays where two, four, or five out of five boxes had an object inside. These combinations created the following conditions. For 'some', 'some…not', 'not all' and 'most' there were three conditions, one in which the statement was logically true, one in which it was logically false, and one in which it was logically true but pragmatically under-informative. For 'all' and 'none' there were two conditions: a condition where the quantifier was logically true and a condition where it was logically false (there cannot be an under-informative but true condition for 'all' and 'none'). In total, there were six items for each quantifier in which the correct response was mandated by the logical meaning of the quantifier (three logically true and three logically false). There were another six items for the under-informative condition for 'some', 'some…not', 'not all' and 'most' in which the correct response is pragmatically determined. The appropriate response for each quantifier given each arrangement is shown in Table 3. A list of all items used is presented in Appendix A and one sample display for each condition for 'all', 'some' and 'most' is presented in Appendix B.

In order to avoid any effects from previous mention of a quantifier and object combination, no quantifier was combined more than once with the same type of object. Moreover, to avoid participants rejecting a statement on grounds other than falsity or under-informativeness, the

Please cite this article in press as: Katsos, N., et al. Are children with Specific Language Impairment competent with the pragmatics and logic of quantification?. *Cognition* (2011), doi:10.1016/j.cognition.2010.12.004

Table 3
Proportion of correct responses for each quantifier for the three groups.

	Logically true correct response: Accept	Logically false correct response: Reject	Under-informative correct response: Reject
All	Arrangement 5/5	Arrangement 2/5	Not available for 'all'
SLI	0.99	0.87	
AM-TD	1.00	0.95	
LM-TD	1.00	0.86	
None	Arrangement 0/5	Arrangement 2/5	Not available for 'none'
SLI	0.84	0.82	
AM-TD	0.95	0.98	
LM-TD	0.64	0.75	
Some	Arrangement 2/5	Arrangement 0/5	Arrangement 5/5
SLI	0.75	0.85	0.34
AM-TD	0.93	0.99	0.87
LM-TD	0.83	0.95	0.36
Some. . .not	Arrangement 2/5	Arrangement 5/5	Arrangement 0/5
SLI	0.64	0.69	0.56
AM-TD	0.91	0.94	0.85
LM-TD	0.78	0.68	0.64
Most	Arrangement 4/5	Arrangement 2/5	Arrangement 5/5
SLI	0.60	0.56	0.17
AM-TD	0.67	0.60	0.53
LM-TD	0.57	0.52	0.13
Not all	Arrangement 2/5	Arrangement 5/5	Arrangement 0/5
SLI	0.74	0.59	0.57
AM-TD	0.92	0.93	0.85
LM-TD	0.72	0.59	0.64

statements were always correct with regard to the type of object that was mentioned and there were no objects of any other type in the display. Also, the objects were chosen from the domain of household items, food, musical instruments and toys, and teachers confirmed that these objects were known by the children. Furthermore, the children were shown an instance of each of the objects on the computer screen before the main experimental session where the experimenter pointed to each of them and named them. This was done in order to familiarise the participant with the exact term that would be used in the experiment to refer to this object. Finally, as mentioned above, if a child rejected a statement, they were asked why it was wrong. This question was asked to check whether the participant was rejecting a statement for reasons unrelated to falsity or informativeness.

Six blocks of items were created, with the order of items pseudo-randomised within each block to avoid mention of the same quantifier or the same object in adjacent items. The blocks were presented in one of three different orders.

After participants were familiarised with the task, and before the main experimental session, their competence with numbers was tested with five statements, one for each number from one to five, three of which were true and two false. This was done to check that participants could count up to five, which is the maximum number of objects that appeared in the boxes in the task.

2.4. Results

Every group performed at high rates with the number-sentences that were administered to test the comprehension of numbers from one to five (94%, 92% and 100% for the SLI, the LM-TD and AM-TD group respectively).

The results from the primary task are presented in Table 3. Because the outcome response in the task is a categorical judgment ('right' or 'wrong') rather than a continuous response ANOVAs (with or without transformation of the data) are not a suitable tool for statistical analysis. Instead, we used a series of mixed logistic regressions with the Laplace approximation as recommended by Baayen, Davidson, and Bates (2008), Jaeger (2008), and Manning (2003), chap. 5.7) among others. All analyses were performed using the statistics software package *R* (R Development Core Team, 2005). We favoured this approach over the more familiar non-parametric alternatives to ANOVAs (Friedman's ANOVAs followed by Wilcoxon signed-ranks tests for within-group comparisons, and Kruskal-Wallis and Mann-Whitney *U*-tests for between-group comparisons) because these allow us to probe more fully for interactions between conditions.

2.4.1. Omnibus analysis

We conducted an omnibus analysis of these data using a logistic regression in *R*. In this analysis, each of the 16 test conditions in Table 3 was coded as a distinct variable ('Some-true', 'Some-false', 'Some-under(informative)', etc.) Each variable was set to 1 for items belonging to that condition, and 0 elsewhere. Each of the three groups (SLI, AM-TD and LM-TD) was coded as a distinct variable, set to 1 for all items for participants belonging to that group, and 0 elsewhere. Taking the 'All-true' condition for AM-TD participants as a baseline, we implemented a model in which all the other variables were evaluated as potential

Please cite this article in press as: Katsos, N., et al. Are children with Specific Language Impairment competent with the pragmatics and logic of quantification?. *Cognition* (2011), doi:10.1016/j.cognition.2010.12.004

sources of fixed effects, and participant identity ('ID') was considered as a random effect.

Results for this model are presented in Table 4. In accordance with convention for logistic regressions, we report the odds ratio as a measure of effect size.

This analysis demonstrates highly significant effects for all conditions and groups under test. That is to say, performance on every other condition, and by every other group, is different from the performance of the AM-TD group in the 'All-true' condition. However, this does not necessarily indicate that there are significant pairwise differences between these other conditions and groups. In the following sections we further explore the precise nature of these differences in order to address the research hypotheses discussed earlier in this paper.

Hypothesis 1 (*Polarity*). To test the effect of positive vs. negative quantifiers, we implemented a further series of mixed logistic regressions. In these we considered the semantically true and false cases for 'all', 'some', 'none' and 'some…not'. We tested for fixed effects of polarity (positive vs. negative), as well as quantifier type (universal: 'all', 'none' vs. existential: 'some', 'some…not') and included participant ID as a random effect in the model. Anticipating the possibility of a response bias, whereby young children might default to accepting statements for which they were uncertain, we also included truth-value (true vs. false) as a fixed effect. We implemented a balanced coding scheme, as the conditions were balanced, in which 'all' and 'some' were coded as +0.5 and 'none' and 'some…not' as −0.5 for polarity, 'all' and 'none' were coded as +0.5 and 'some' and 'some…not' as −0.5 for quantifier type, and true cases were coded as +0.5 and false as −0.5 for truth-value. We implemented this model for all three groups of participants without distinguishing between them. The results are presented in Table 5.

These results indicate a highly significant advantage for positive over negative quantifiers, in addition to a smaller but still highly significant advantage for universal over existential quantifiers. We revised this model by removing the non-significant truth-value and introducing a cross-term to probe for interactions between polarity and universality. In this model, the interaction between polarity and universality was marginally non-significant (Coefficient = 0.535, Odds ratio = 1.71, SE of coefficient = 0.314, $Z = 1.71$, $P = 0.088$).

We further wished to test the effect of SLI on the comprehension of these semantically determined quantifiers. To do so, we compared the SLI and AM-TD group by performing further logistic regression analyses considering group as a source of fixed effects. The analysis reported in Table 6 uses polarity, universality and group as potential fixed effects, and participant ID as a random effect. Here the SLI group is coded +0.5 and the AM-TD group −0.5 for group.

The analysis indicates a highly significant effect of group, with performance for the SLI participants lower than that of their age-matched controls. To test whether this effect holds across conditions, we implemented a further logistic regression to test for interactions between group and the other two variables, polarity and universality. However, neither of the interactions reached significance in this model (Polarity × Group: Coefficient = 0.512, Odds ratio = 1.67, SE of coefficient = 0.49, $Z = 1.05$, $P = 0.29$. Universality × Group: Coefficient = 0.601, Odds ratio = 1.82, SE of coefficient = 0.49, $Z = 1.21$, $P = 0.22$). This suggests that the SLI group performed more poorly across all the conditions.

We also ran a parallel regression to compare the performance of SLI participants with their language-matched control group. SLI participants were coded as +0.5 and LM-TD participants as −0.5. The findings are reported in Table 7. There is no significant effect of group between the SLI and LM-TD groups, indicating that their performance cannot be differentiated across these conditions as a whole. Taken collectively, these two groups continue to exhibit the main effects of polarity and universality identified earlier.

Testing for first-level interactions as above, we find a significant interaction between group and universality

Table 4
Omnibus analysis.

Variable	Coefficient	Odds ratio	SE of coeff.	*Z*	*P*
(Intercept)	7.66	2130	1.05	7.27	<0.001
All-false	−3.50	0.0301	1.05	−3.35	<0.001
Some-true	−4.11	0.0164	1.04	−3.95	<0.001
Some-false	−3.02	0.0486	1.05	−2.87	<0.001
Some-under	−6.06	0.00233	1.03	−5.89	<0.001
Most-true	−5.57	0.00382	1.03	−5.38	<0.001
Most-false	−5.88	0.00281	1.03	−5.68	<0.001
Most-under	−7.68	0.000462	1.03	−7.44	<0.001
None-true	−4.30	0.0136	1.04	−4.14	<0.001
None-false	−4.01	0.0181	1.04	−3.86	<0.001
Some-not-true	−4.55	0.0106	1.04	−4.39	<0.001
Some-not-false	−4.58	0.0103	1.04	−4.41	<0.001
Some-not-under	−5.16	0.00572	1.03	−5.02	<0.001
Not-all-true	−4.44	0.0118	1.04	−4.28	<0.001
Not-all-false	−5.05	0.00642	1.04	−4.88	<0.001
Not-all-under	−5.12	0.00599	1.03	−4.97	<0.001
SLI	−2.00	0.136	0.344	−5.80	<0.001
LM-TD	−1.96	0.141	0.344	−5.69	<0.001

Please cite this article in press as: Katsos, N., et al. Are children with Specific Language Impairment competent with the pragmatics and logic of quantification?. *Cognition* (2011), doi:10.1016/j.cognition.2010.12.004

Table 5
Testing polarity, quantifier type and response bias across all three groups.

Variable	Coefficient	Odds ratio	SE of coeff.	*Z*	*P*
(Intercept)	2.76	15.8	0.211	13.1	<0.001
Truth-value	−0.074	0.929	0.144	−0.511	0.61
Positive	1.23	3.42	0.153	8.04	<0.001
Universal	0.642	1.90	0.147	4.38	<0.001

Table 6
Testing polarity and quantifier type in SLI and AM-TD groups.

Variable	Coefficient	Odds ratio	SE of coeff.	*Z*	*P*
(Intercept)	3.37	29.0	0.284	11.8	<0.001
Positive	0.954	2.60	0.207	4.61	<0.001
Universal	1.19	3.30	0.211	5.65	<0.001
Group	−2.24	0.107	0.558	−4.01	<0.001

(Universality × Group: Coefficient = 1.25, Odds ratio = 3.48, SE of coefficient = 0.324, $Z = 1.54$, $P < 0.001$). The other interactions are non-significant. A comparison between the means in Table 2 suggests that this interaction arises from a difference in performance between the SLI and LM-TD groups, with the SLI group performing better with universal quantifiers and the LM-TD group performing better with existential quantifiers.

Hypothesis 2 (*Quantifier complexity*). We predicted that because 'most' is logically more complex than 'some' it will be harder to acquire. We performed logistic regression analyses restricting ourselves to the semantically true or false conditions (that is, setting aside the under-informative conditions whose meaning does not rely exclusively on logical principles). We implemented a mixed model in which quantifier complexity ('some' vs. 'most') and truth-value are tested as potential sources of fixed effects, and participant ID is a random effect. In this model, 'most' is coded as +0.5 and 'some' as −0.5 for quantifier complexity, with true conditions coded as +0.5 and false conditions as −0.5 for truth-value. Considering all three groups of participants together, the results are presented in Table 8. The findings indicate significantly better performance for 'some' than for 'most', and no significant response bias.

We further compared the SLI group with the age-matched controls by adding group as a predictor variable to the above model (SLI coded as +0.5, AM-TD as −0.5), and removing the non-significant truth-value as a variable. The results presented in Table 9 indicate a main effect of group, with SLI participants underperforming their AM-TD counterparts, in addition to the same effect of quantifier complexity reported before.

Table 7
Testing polarity and quantifier type in SLI and LM-TD.

Variable	Coefficient	Odds ratio	SE of coeff.	*Z*	*P*
(Intercept)	2.06	7.85	0.207	9.96	<0.001
Positive	1.32	3.76	0.164	8.08	<0.001
Universal	0.63	1.88	0.156	4.03	<0.001
Group	0.131	1.14	0.408	0.321	0.748

A further model tested for an interaction between quantifier complexity and group, which reached significance (Coefficient = 1.63, Odds ratio = 5.10, SE of coefficient = 0.501, $Z = 3.25$, $P < 0.01$). With reference to the mean scores in Table 2, it appears that this interaction arises from the SLI participants' performance being disproportionally worse in the 'some' case compared to the 'most' case.

Comparing the SLI participants with their language-matched control group (coding SLI as +0.5 and LM-TD as −0.5 for group), a parallel analysis (Table 10) yields no significant difference between the groups, across the quantifiers 'most' and 'some'. Taken collectively, these groups exhibit significantly lower performance for 'most' than for 'some'. The interaction between group and quantifier complexity was again significant (Coefficient = 0.883, Odds ratio = 2.42, SE of coefficient = 0.382, $Z = 2.31$, $P < 0.05$). Examination of the means in Table 2 indicates that this effect is driven by SLI participants performing relatively better with 'most'.

Hypothesis 3 (*Pragmatic vs. logical meaning*). Further logistic regressions were implemented to test the hypothesis that it is particularly difficult to acquire pragmatic competence (in this case, to reject under-informative utterances) compared to acquiring competence with logical meaning (specifically, rejecting logically false utterances). Under-informative cases were tested for the quantifiers 'some', 'some…not', 'not all' and 'most'. The following analyses investigate the participants' performance on under-informative vs. semantically false items across these quantifiers.

First, we test SLI participants against their age-matched counterparts (see Table 11). The predictor variable 'group' is set to +0.5 for SLI participants and to −0.5 for AM-TD participants. The predictor variable 'meaning' is set to +0.5 for pragmatic (under-informative) cases and −0.5 for logical (logically false) cases. These are mixed models, with a random effect of participant ID in addition to the fixed effects.

Augmenting this model with a meaning by group interaction term shows that the interaction is significant (Coefficient = −0.546, Odds ratio = 0.579, SE of coefficient = 0.255, $Z = -2.14$, $P < 0.05$). Hence, comparing SLI and AM-TD participants, we observe a significant effect of meaning (with lower performance in the pragmatic than the logically false condition) and group (lower performance for the SLI than the AM-TD group). We also observe a significant interaction, apparently due to SLI participants exhibiting lower performance in the under-informative conditions. Further analyses in which the quantifiers were treated separately showed that this pattern was robust

Table 8
Testing quantifier complexity ('some' vs. 'most') and response bias across all three groups.

Variable	Coefficient	Odds ratio	SE of coeff.	*Z*	*P*
(Intercept)	1.26	3.54	0.105	12.1	<0.001
Truth-value	−0.130	0.878	0.152	−0.856	0.392
Quantifier complexity	−1.77	0.170	0.167	−10.6	<0.001

Table 9
Testing quantifier complexity ('some' vs 'most') in SLI and AM-TD.

Variable	Coefficient	Odds ratio	SE of coeff.	*Z*	*P*
(Intercept)	1.32	3.73	0.130	10.2	<0.001
Quantifier complexity	−1.67	0.188	0.206	8.15	<0.001
Group	−0.666	0.514	0.246	−2.71	<0.01

Table 10
Testing quantifier complexity ('some' vs. 'most') in SLI and LM-TD.

Variable	Coefficient	Odds ratio	SE of coeff.	*Z*	*P*
(Intercept)	1.01	2.76	0.109	9.28	<0.001
Quantifier complexity	−1.49	0.224	0.186	−8.03	<0.001
Group	−0.156	0.856	0.211	−0.738	0.46

Table 11
Testing pragmatic vs. logical meaning in SLI and AM-TD groups.

Variable	Coefficient	Odds ratio	SE of coeff.	*Z*	*P*
(Intercept)	1.16	3.19	0.175	6.60	<0.001
Meaning	−1.16	0.313	0.125	−9.28	<0.001
Group	−1.88	0.153	0.346	−5.45	<0.001

across quantifiers, except that the main effect of group did not reach significance in the "not all" case, and the interaction only reached significance with "most".

Table 12 presents the parallel comparison between the SLI group and the language-matched group.

In this case, when a group by meaning interaction term is included in the analysis, it does not reach significance (Coefficient = −0.147, Odds ratio = 0.863, SE of coefficient = 0.216, $Z = -0.681$, $P = 0.496$). Hence, for SLI and LM-TD participants, performance was significantly lower for items requiring a pragmatic inference on the basis of under-informativeness, but group was not a significant factor. SLI participants appear to behave quantitatively like their age-matched counterparts on these items. Again this pattern is robust across quantifiers, with the exception that there is no significant main effect of meaning for 'not all'.

We can also consider participant distribution for under-informative statements. Bott and Noveck (2004) and Guasti et al. (2005) among others report that individual participants tend either to reject every instance of under-informative statements with 'some', or to accept all of them, rather than switching responses between trials. We labelled our participants as consistently informative if they rejected five or six out of the six under-informative statements with 'some', consistently under-informative if they rejected at most one under-informative statement, or inconsistent otherwise. In the AM-TD group, most of the children were consistently informative (26 out of 29), with the remaining three being inconsistent. In the LM-TD group, most of the children were consistent (23 out of 29), with 16 being consistently under-informative and seven consistently informative. A parallel analysis for the SLI group reveals that most of the children were consistent (25 out of 29), with 16 being consistently under-informative and 9 consistently informative. Thus, the SLI children appear to conform to broadly the same pattern of participant distribution as their typically-developing peers.

We performed a similar analysis for the logical conditions for 'most', which was an exceptionally difficult quantifier. Participants were classified as consistently correct if they were correct in five or six out of six responses in the logical conditions, consistently incorrect if they gave zero or one correct response correct, and inconsistent otherwise. In every group, the overwhelming majority of participants were inconsistent (20, 22 and 20 out of 29 in the SLI, the LM-TD and the AM-TD group respectively) with a very small number providing systematically wrong responses (one in

Table 12
Testing pragmatic vs. logical meaning in SLI and LM-TD groups.

Variable	Coefficient	Odds ratio	SE of coeff.	*Z*	*P*
(Intercept)	0.297	1.35	0.149	2.00	<0.05
Meaning	−1.26	0.284	0.108	−11.7	<0.001
Group	−0.132	0.876	0.294	−0.447	0.655

Please cite this article in press as: Katsos, N., et al. Are children with Specific Language Impairment competent with the pragmatics and logic of quantification?. *Cognition* (2011), doi:10.1016/j.cognition.2010.12.004

each group) and the remaining being consistently correct (eight, six and eight in the SLI, the LM-TD and the AM-TD group respectively). This reveals a majority of participants who have not mastered the correct logical meaning of 'most', but also a substantial sub-group of competence of roughly the same size in all three groups. As for under-informative utterances, the distribution of children with SLI conforms to the pattern expected given their overall language level.

3. General discussion

In this study we set out to investigate whether Spanish-speaking children with SLI are competent with the pragmatic maxim of informativeness and with the logical meaning of the expressions 'all', 'none', 'some', 'some-not', 'most' and 'not all'. Children with SLI and their age-matched peers both performed better with the logical than the pragmatic meaning of quantified statements. They also performed better with positive than negative quantifiers, and better with the logically simpler 'some' than with the more complex 'most'. Moreover, it was documented that both groups performed better with universal ('all', 'none') compared to existential ('some', 'some…not') quantifiers. This effect was not predicted on the basis of the logical analysis of these quantifiers, as universality and existentiality are equally fundamental and simple logical properties. One possibility is that, because 'some' and 'some…not' apply to a wide range of possible proportions (e.g. 'some' may apply from 'almost none of' to even 'all of', while 'all' and 'none' are true of only one proportion each), the full meaning of 'some' is comparatively difficult to acquire. This hypothesis is compatible with the proposal by Hanlon (1987; also 1988: 67ff) according to which quantifiers that select the entirety of the set under discussion are computationally simpler and hence easier to acquire than quantifiers that select a subset.

Critically, in all of the primary analyses, we found that children with SLI perform lower than age-matched peers. Nevertheless, in every case where children with SLI performed more poorly than their age-matched peers, they performed approximately at the same level as a younger language-matched group. The emerging picture is one where competence with informativeness and the logical meaning of quantified statements is indeed an area of significant difficulty for children with SLI, but this difficulty is proportionate to their overall language difficulties.

Moreover, an analysis of the justifications for rejection given by all three groups revealed no qualitative differences in the type of justification given by the groups, as all rejections were with reference to the number of objects in the boxes. For instance, the most common justification for rejecting logically true 'some', both for the SLI and LM-TD children, was that "two [toys] are in", suggesting that two is not a numerosity that children associated with 'some'. This error has been reported for English- and Greek-speaking typically-developing 5-year-old children (Katsos & Bishop, accepted for publication; Katsos et al., 2009). Hence, even in an analysis of specific error patterns, the SLI children seem to be unexceptional compared to the language-matched group.

Nevertheless, this should not conceal more subtle differences between the children with SLI and the language-matched group: even though there was no main effect of group across all comparisons tested, there was an interaction of group with universality (the language-matched group found existential quantifiers easier), and with quantifier complexity (the language-matched group found logically simple ones easier). Given the lack of obvious differences between the SLI and LM-TD groups, what could account for these interactions? The data in Table 2 appear to suggest that the interactions arise from the SLI group performing better on 'none' and, to a lesser extent, 'most', while the LM-TD group perform better on 'some' and 'some…not'. A possible explanation of this is that performance with 'none' and 'most' is dependent on mathematical competence, which was not specifically investigated in this study. On this account, mathematical competence should facilitate the acquisition of the concept of 'zero', which might be a pre-requisite for mastery of the quantifier 'none'. Similarly, as 'most' requires the precise evaluation of cardinal quantities, its mastery might be connected to mathematical abilities. We should note that, if this hypothesis were to be borne out by further study, it might transpire that the performance of SLI children with quantifiers is in fact worse than that of language-matched controls, and an advantage in mathematical skills (due to chronological age) has obscured that pattern in our experiment.

Overall, the present investigation contributes to the view in which children with SLI may face difficulties in several aspects of communication, but especially striking and disproportionate impairments are most likely to be found in the areas of morphosyntax, phonology and general vocabulary (Bishop, 2002; Leonard, 1998; i.a.) rather than pragmatic or logical meaning. This should not conceal the fact that competence with pragmatics and the logical meaning of quantifiers is significantly impaired in children with SLI when compared to age-matched controls and that in certain cases children with SLI perform exceptionally poorly. As competence with expressions of quantity is a major part of everyday communication skills, clinicians might want to target aspects of pragmatics and logical meaning alongside the traditional areas where children with SLI are known to be impaired (a point also made by Norbury, Nash, Baird, and Bishop (2004)).

Finally, let us consider the implications of these findings for the architecture of the linguistic system, and the relation between grammatical and pragmatic competence. With regard to pragmatics, we reported that children with SLI performed at the same level as a group of younger typically-developing children with whom they were matched on a receptive grammar test (which measured several morphosyntactic and semantic constructions). The SLI children did not differ from the language-matched control group on comparisons between the mean scores of the two groups for under-informative statements, nor on a qualitative analysis of the justifications given for their answers, nor in their distribution.

A similar picture of the relation between overall language level and pragmatic competence is reported for participants with Autism Spectrum Disorders (age-range 19–40 years) by Pijnacker, Hagoort, Buitelaar, Teunisse,

Please cite this article in press as: Katsos, N., et al. Are children with Specific Language Impairment competent with the pragmatics and logic of quantification?. *Cognition* (2011), doi:10.1016/j.cognition.2010.12.004

and Geurts (2009). In that study, following the design of Bott and Noveck (2004), participants were administered an under-informative utterances paradigm where the correct response relies on mastery of the logical and pragmatic meaning of quantifiers as well as category membership knowledge (e.g. a critical item for under-informativeness was "Some sparrows are birds"). In this paradigm, where participants need to invoke world-knowledge from long term memory, it is consistently reported that even typical adults do not perform at ceiling (Bott & Noveck, 2004; Guasti et al., 2005; Noveck, 2001; Noveck & Posada, 2003; among others). Indeed, the control group accepted under-informative utterances around 30% of the time, while the group of High Functioning Autistic participants, which had a lower overall language score, accepted these utterances at rates of around 40%. However, both groups were outperformed by the group with Asperger's syndrome, whose members had a higher language score, and accepted less than 15% of under-informative utterances. Therefore, this study provides evidence that groups with lower levels of linguistic ability exhibit lower pragmatic competence, even in adulthood.

What are the implications of this finding for the grammar–pragmatics interface? Post-Gricean pragmaticians such as Levinson (2000) and Sperber and Wilson (1986)/1995, 2002) argue for a modular view of pragmatics and grammar (comprising of phonology, syntax and semantics), whereby each system operates with distinct types of rules that have been evolved to serve different purposes. In this view it is possible that even though children with SLI are impaired on the comprehension of the logical meaning of quantifiers (which is part of linguistic semantics, and hence the grammar), they would nevertheless perform at the same level as their age-matched peers with pragmatics. The findings of this study do not document this pattern. On the other hand, Chierchia (2004), Chierchia (2006) and Spector (2007) (see also Chierchia, Fox, & Spector, in press) have advocated a radically different approach whereby some version of the maxim of informativeness is ultimately embedded within the semantic system, and hence the grammar. The straightforward prediction of this view would be that children with SLI would be as challenged with the logical meaning of quantifiers as with the pragmatic, which is indeed the picture that emerges from the present study. However, it should be highlighted that the modular view of the grammar–pragmatics interface could also predict the same finding: since the input to the pragmatic system (the representation of the quantifier's logical meaning) is impaired in children with SLI, it is reasonable to expect that the output of this system will also be impaired.

As such, the findings of this study are ultimately compatible with both views. Ongoing investigations on the pragmatic competence of children with Grammatical-SLI who exhibit particular impairments with grammar over and beyond those found in children with typical SLI may be a further promising avenue for disentangling the empirical predictions made by the modular and the grammatical approach (Skordi, Katsos, Marshall, & van der Lely, submitted for publication).

Acknowledgments

This research was supported by three travel grants for scientific collaboration between the first two co-authors awarded by the European Union's COST Action A33 'Cross-linguistically Robust Stages of Children's Linguistic Performance' http://www.zas.gwz-berlin.de/cost/. The first author would like to acknowledge support provided by the British Academy (SG-47135). The second author was also supported by a grant for short-term visits to research centres by Fundació Caixa-Castelló Bancaixa (Action 2.1. Jaume I University of Castellón, Spain). Moreover, the second and third co-authors would also like to acknowledge support by Grant P1-1B2008-56 by the Fundació Caixa-Castelló, and Grant SEJ2006-03593 by the Ministerio de Educación y Ciencia of Spain. The fourth author was supported by a University of Cambridge Domestic Research Studentship.

We would also like to thank Courtenay Norbury, and two anonymous reviewers of *Cognition*, for valuable comments and suggestions.

Appendix A

Item list: The first column presents the quantified statement, the second column presents the arrangement ('ARR': how many of the five objects were inside the five boxes), and the third column codes which response is correct.

Numerals		
Two apples are in the boxes	ARR: two	RIGHT
Three apples are in the boxes	ARR: four	WRONG
One apple is in the boxes	ARR: one	RIGHT
Four apples are in the boxes	ARR: three	WRONG
Five apples are in the boxes	ARR: five	RIGHT
(Hay [dos/tres/una/cuatro/cinco] manzana(s) dentro de las cajas)		
'All'		
All the pens are in the boxes	ARR: all	RIGHT
All the balls are in the boxes	ARR: all	RIGHT
All the dinosaurs are in the boxes	ARR: all	RIGHT
All the t-shirts are in the boxes	ARR: two	WRONG
All the sandwiches are in the boxes	ARR: two	WRONG
All the shoes are in the boxes	ARR: two	WRONG
(Todos los/Todas las [bolis, pelotas, dinosaurios, camisetas, sándwiches, zapatos] están dentro de las cajas)		
'None'		
None of the clocks are in the boxes	ARR: none	RIGHT

Appendix A (*continued*)

None of the bananas are in the boxes	ARR: none	RIGHT
None of the balls are in the boxes	ARR: none	RIGHT
None of the cars are in the boxes	ARR: two	WRONG
None of the balloons are in the boxes	ARR: two	WRONG
None of the skirts are in the boxes	ARR: two	WRONG
(Ninguno de los/Ninguna de las [relojes, plátanos, pelotas, coches, globos, faldas] están dentro de las cajas)		
'Some'		
Some of the clocks are in the boxes	ARR: all	WRONG
Some of the bananas are in the boxes	ARR: all	WRONG
Some of the cars are in the boxes	ARR: all	WRONG
Some of the telephones are in the boxes	ARR: all	WRONG
Some of the dolls are in the boxes	ARR: all	WRONG
Some of the flowers are in the boxes	ARR: all	WRONG
Some of the oranges are in the boxes	ARR: two	RIGHT
Some of the pens are in the boxes	ARR: two	RIGHT
Some of the pears are in the boxes	ARR: two	RIGHT
Some of the skirts are in the boxes	ARR: none	WRONG
Some of the apples are in the boxes	ARR: none	WRONG
Some of the strawberries are in the boxes	ARR: none	WRONG
(Algunos de los/Algunas de las [relojes, plátanos, coches, teléfonos, muñecas, flores, naranjas, bolis, peras, faldas, manzanas, fresas] están dentro de las cajas)		
'Some not'		
Some of the apples are not in the boxes	ARR: none	WRONG
Some of the guitars are not in the boxes	ARR: none	WRONG
Some of the teddy-bears are not in the boxes	ARR: none	WRONG
Some of the telephones are not in the boxes	ARR: none	WRONG
Some of the bicycles are not in the boxes	ARR: none	WRONG
Some of the dinosaurs are not in the boxes	ARR: none	WRONG
Some of the vases are not in the boxes	ARR: two	RIGHT

Appendix A (*continued*)

Some of the shoes are not in the boxes	ARR: two	RIGHT
Some of the trains are not in the boxes	ARR: two	RIGHT
Some of the TVs are not in the boxes	ARR: all	WRONG
Some of the books are not in the boxes	ARR: all	WRONG
Some of the sandwiches are not in the boxes	ARR: all	WRONG
(Algunos de los/Algunas de las [manzanas, guitarras, osos de peluche, teléfonos, bicicletas, dinosaurios, jarrones, zapatos, trenes, teles, libros, sándwiches] no están dentro de las cajas)		
'Most'		
Most of the teddy-bears are in the boxes	ARR: all	WRONG
Most of the apples are in the boxes	ARR: all	WRONG
Most of the guitars are in the boxes	ARR: all	WRONG
Most of the trains are in the boxes	ARR: all	WRONG
Most of the oranges are in the boxes	ARR: all	WRONG
Most of the strawberries are in the boxes	ARR: all	WRONG
Most of the shoes are in the boxes	ARR: four	RIGHT
Most of the vases are in the boxes	ARR: four	RIGHT
Most of the dolls are in the boxes	ARR: four	RIGHT
Most of the cars are in the boxes	ARR: two	WRONG
Most of the sandwiches are in the boxes	ARR: two	WRONG
Most of the bicycles are in the boxes	ARR: two	WRONG
(La mayoría de los/La mayoria de las [osos de peluche, manzanas, guitarras, trenes, naranjas, fresas, zapatos, jarrones, muñecas, coches, sándwiches, bicicletas] están dentro de las cajas)		
'Not all'		
Not all the vases are in the boxes	ARR: none	WRONG
Not all the shoes are in the boxes	ARR: none	WRONG
Not all the TVs are in the boxes	ARR: none	WRONG
Not all the pears are in the boxes	ARR: none	WRONG
Not all the strawberries are in the boxes	ARR: none	WRONG

(*continued on next page*)

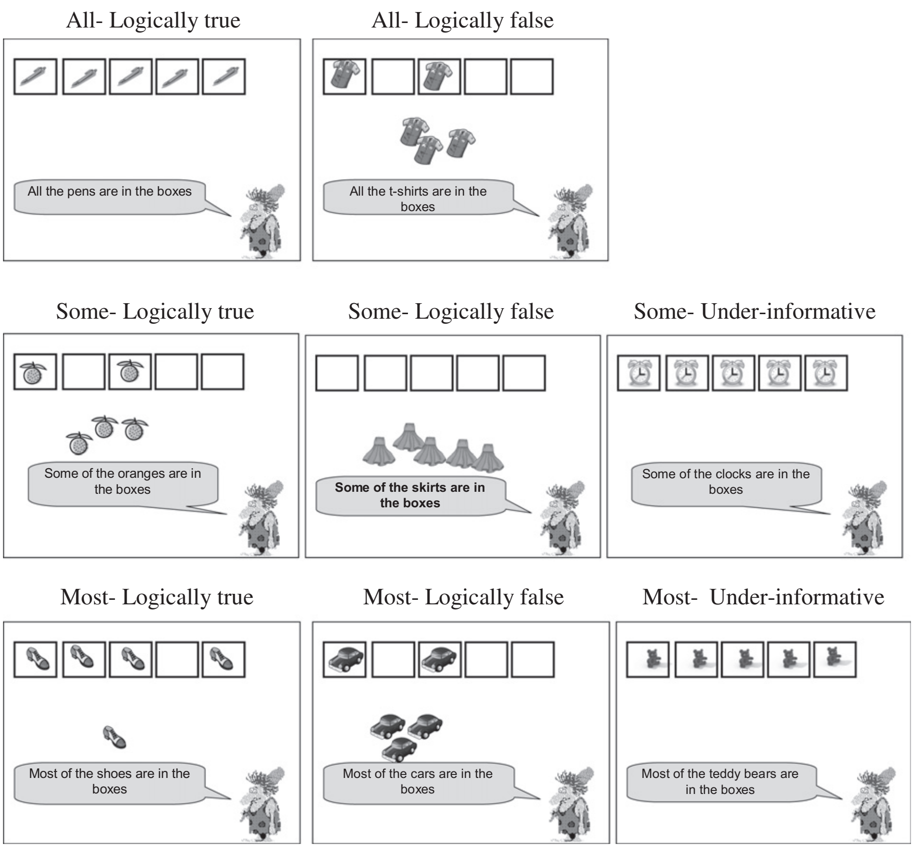
Please cite this article in press as: Katsos, N., et al. Are children with Specific Language Impairment competent with the pragmatics and logic of quantification?. *Cognition* (2011), doi:10.1016/j.cognition.2010.12.004

Appendix A (*continued*)

Not all the books are in the boxes	ARR: none	WRONG
Not all the clocks are in the boxes	ARR: two	RIGHT
Not all the bananas are in the boxes	ARR: two	RIGHT
Not all the telephones are in the boxes	ARR: two	RIGHT
Not all the skirts are in the boxes	ARR: all	WRONG
Not all the flowers are in the boxes	ARR: all	WRONG
Not all the t-shirts are in the boxes	ARR: all	WRONG
(No todos los/No todas las [jarrones, zapatos, teles, peras, fresas, libros, relojes, plátanos, teléfonos, faldas, flores, camisetas] están dentro de las cajas)		

Appendix B

One sample display for each arrangement for 'all', 'some' and 'most'. The arrangements for 'none', 'some-not' and 'not all' were designed in a similar fashion, as described in the Items section.



References

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders (DSMIV)*. Washington, DC: American Psychiatric Association.

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

Barner, D., Chow, K., & Yang, S-J. (2009). Finding one's meaning: A test of the relation between quantifiers and integers in language development. *Cognitive Psychology, 58*, 195–219.

Bishop, D. V. M. (1983). *The test for reception of grammar (TROG)*. Age and Cognitive Performance Research Centre, University of Manchester.

Bishop, D. V. M. (1997). *Uncommon understanding: Development and disorders of language comprehension in children*. Hove: Psychology Press.

Bishop, D. V. M. (2000). Pragmatic language impairment: A correlate of SLI, a distinct subgroup, or part of the autistic continuum? In D. V. M. Bishop & L. B. Leonard (Eds.), *Speech and language impairments in children: Causes, characteristics, intervention and outcome* (pp. 99–114). Hove, England: Psychology Press.

Bishop, D. V. M. (2002). Speech and language difficulties. In M. Rutter & E. Taylor (Eds.), *Child and adolescent psychiatry: Modern approaches* (pp. 664–681). Oxford: Blackwell Science.

Bishop, D. V. M. (2003). *Test for the reception of grammar-2 (TROG-2)*. London: Pearson Assessment.

Bott, L., & Noveck, I. A. (2004). Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language, 51*, 437–457.

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

Chierchia, G. (2006). Broaden your views: Implicatures of domain widening and the 'logicality' of language. *Linguistic Inquiry, 37*, 535–590.

Chierchia, G., Fox, D., & Spector, B. (in press). The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In K. von Heusinger, C. Maienborn, & P. Portner (Eds.), *Handbook of semantics*. Mouton de Gruyter.

Please cite this article in press as: Katsos, N., et al. Are children with Specific Language Impairment competent with the pragmatics and logic of quantification?. *Cognition* (2011), doi:10.1016/j.cognition.2010.12.004

Chierchia, G., & McConnell-Ginet, S. (2000). *Meaning and grammar: An introduction to semantics* (2nd ed.). Cambridge, MA: MIT Press.

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

Cowan, R., Donlan, C., Newton, E. J., & Lloyd, D. (2005). Number skills and knowledge in children with specific language impairment. *Journal of Educational Psychology, 97*, 732–744.

Cummins, C., & Katsos, N. (2010). Comparative and superlative quantifiers: Pragmatic effects of comparison type. *Journal of Semantics*. doi:10.1093/jos/ffq006.

Donlan, C., Cowan, R., Newton, E. J., & Lloyd, D. (2007). The role of language in mathematical development: Evidence from children with specific language impairments. *Cognition, 103*, 23–33.

Fazio, B. B. (1994). The counting abilities of children with specific language impairment – A comparison of oral and gestural tasks. *Journal of Speech and Hearing Research, 37*, 358–368.

Fazio, B. B. (1996). Mathematical abilities of children with specific language impairment: A 2-year follow-up. *Journal of Speech and Hearing Research, 39*, 839–849.

Fazio, B. B. (1999). Arithmetic calculation, short-term memory, and language performance in children with specific language impairment: A 5-yr follow-up. *Journal of Speech, Language, and Hearing Research, 42*, 420–431.

Feeney, A., Scrafton, S., Duckworth, A., & Handley, S. J. (2004). The story of some: Everyday pragmatic inference by children and adults. *Canadian Journal of Experimental Psychology, 58*(2), 121–132.

Gamut, L. T. F. (1991). *Logic, language, and meaning* (Vol. 1). Chicago: University of Chicago Press.

Geurts, B., Katsos, N., Cummins, C., Moons, J., & Noordman, L. (2010). Scalar quantifiers: Logic, acquisition, and processing. *Language and Cognitive Processes, 25*(1), 130–148.

Grice, P. (1975). *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Guasti, M. T., Chierchia, G., Crain, S., Foppolo, F., Gualmini, A., & Meroni, L. (2005). Why children and adults sometimes (but not always) compute implicatures. *Language and Cognitive Processes, 20*(5), 667–696.

Hanlon, C. C. (1987). Acquisition of set-relational quantifiers in early childhood. *Genetic, Social and General Psychology Monographs, 113*(2), 215–264.

Hanlon, C. (1988). The emergence of set-relational quantifiers in early childhood. In F. S. Kessel (Ed.), *The development of language and language researchers: Essays in honor of Roger Brown* (pp. 65–78). Hillsdale, NJ: Erlbaum.

Horn, L. (1972). *On the semantic properties of logical operators in English*. UCLA dissertation, distributed by Indiana University Linguistics Club, 1976.

Horn, L. (1984). Toward a new taxonomy for pragmatic inference. In D. Schiffrin (Ed.), *Meaning, form and use in context: Linguistic applications, Proceedings of GURT '84* (pp. 11–42). Washington, DC: Georgetown University Press.

Horn, L. (1989). *A natural history of negation*. Chicago: University of Chicago Press.

Huang, Y., & Snedeker, J. (2009). Semantic meaning and pragmatic interpretation in five-year olds: Evidence from real time spoken language comprehension. *Developmental Psychology, 45*(6), 1723–1739.

Hurewitz, F., Papafragou, A., Gleitman, L., & Gelman, R. (2006). Asymmetries in the acquisition of numbers and quantifiers. *Language Learning and Development, 2*, 77–96.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*, 434–446.

Just, M. A., & Carpenter, P. A. (1971). Comprehension of negation with quantification. *Journal of Verbal Learning and Verbal Behavior, 10*, 244–253.

Katsos, N. (2009). Evaluating under-informative statements with context-dependent and context-independent scales: Experimental and theoretical implications. In U. Sauerland & K. Yatsushiro (Eds.), *Semantics and pragmatics: From experiment to theory*. Palgrave Studies in Pragmatics, Language and Cognition.

Katsos, N., & Bishop, D. V. M. (accepted for publication). Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition*.

Katsos, N., & Smith, N. (2010). *The quantifier comprehension task*. Cambridge, UK: Clinical Linguistic Applications.

Katsos, N., Smith, N., Haman, E., Miękisz, A., Jensen-de Lopez, K., Varlokosta, S., et al. (2009). Crosslinguistic investigations in the acquisition of quantification. Paper presented in the 50th conference of the Linguistics Association of Great Britain, Edinburgh.

Leonard, L. (1998). *Children with specific language impairment*. Cambridge, MA: MIT Press.

Levinson, S. (1983). *Pragmatics*. Cambridge, England: Cambridge University Press.

Levinson, S. (2000). *Presumptive meanings*. Cambridge, MA: MIT Press.

Manning, C. D. (2003). Probabilistic syntax. In R. Bod, J. Hay, & S. Jannedy (Eds.), *Probabilistic linguistics* (pp. 289–341). Cambridge, MA: MIT Press.

Mendoza, E., Carballo, G., Muñoz, J., & Fresneda, M. D. (2005). *(CEG) Test de Comprensión de Estructuras Gramaticales*. Madrid: TEA.

Moxey, L. M., & Sanford, A. J. (2000). Communicating quantities: A review of psycholinguistic evidence of how expressions determine perspectives. *Applied Cognitive Psychology, 14*(3), 237–255.

Musolino, J. (2004). The semantics and acquisition of number words: Integrating linguistic and developmental perspectives. *Cognition, 93*, 1–41.

Musolino, J., & Lidz, J. (2006). Why children aren't universally successful with quantification. *Linguistics, 44*(4), 817–852.

Newton, E. J. (2007). Reasoning abilities and SLI: The relationship between language and thought. In *Presented in the 4th AFASIC international symposium*, Warwick, UK.

Newton, E. J., Donlan, C., & Roberts, M. J. (2010). Deductive reasoning in children with specific language impairment. *British Journal of Developmental Psychology, 28*(1), 71–87.

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

Norbury, C. F., Nash, M., Baird, G., & Bishop, D. V. M. (2004). Using a parental checklist to identify diagnostic groups in children with communication impairment: A validation of the Children's Communication Checklist-2. *International Journal of Lan guage and Communication Impairments, 39*, 345–364.

Noveck, I. A. (2001). When children are more logical than adults. *Cognition, 86*, 253–282.

Noveck, I. A., & Posada, A. (2003). Characterizing the time course of an implicature: An evoked potentials study. *Brain and Language, 85*, 203–210.

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

Papafragou, A., & Tantalou, N. (2004). Children's computation of implicatures. *Language Acquisition, 12*(1), 71–82.

Pijnacker, J., Hagoort, P., Buitelaar, J., Teunisse, J.-P., & Geurts, B. (2009). Pragmatic inferences in high-functioning adults with autism and Asperger syndrome. *Journal of Autism and Developmental Disorders, 39*, 607–618.

Pouscoulous, N., Noveck, I., Politzer, G., & Bastide, A. (2007). A developmental investigation of processing costs in implicature production. *Language Acquisition, 14*, 347–376.

R development core team (2005). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. <http://www.R-project.org>.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Raven manual: Section 3. Standard progressive matrices*. Oxford: Oxford Psychologists Press.

Saborit, C., & Julián, J. P. (2005). *Evaluación del lenguaje Infantil – ELI*. Collecció, Educació: Universitat Jaume I de Castelló (UJI).

Sanford, A. J., & Moxey, L. M. (2003). New perspectives on the expression of quantity. *Current Directions in Psychological Sciences, 12*, 240–243.

Skordi, A., Katsos, N., Marshall, C., & van der Lely, H. (submitted for publication). Investigating scalar implicature in children with grammatical specific language impairment. *Language Learning and Development*.

Spector, B. (2007). Scalar implicatures: Exhaustivity and Gricean reasoning. In M. Aloni, P. De kker, & A. Butler (Eds.), *Questions in dynamic semantics: Vol. 17 of Current research in the semantics/pragmatics interface*. Elsevier.

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

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

Surian, L., Baron-Cohen, S., & van der Lely, H. K. J. (1996). Are children with autism deaf to Gricean maxims? *Cognitive Neuropsychiatry, 1*, 55–71.

van Benthem, J. (1986). *Essays in logical semantics*. Reidel.

van Rooij, I. (2008). The tractable cognition thesis. *Cognitive Science, 32*(6), 939–984.

von Fintel, K., & Matthewson, L. (2008). Universals in semantics. *The Linguistic Review, 25*(1-2), 139–201.