



Scalar implicatures in second language acquisition

Roumyana Slabakova*

Department of Linguistics, University of Iowa, 557 English Philosophy Building, Iowa City, IA 52242, USA

Received 4 November 2007; received in revised form 30 May 2009; accepted 25 June 2009

Abstract

This study investigates the second language acquisition (L2A) of scalar implicatures (Grice, 1989; Horn, 1972), implicatures based on a range of quantifiers ordered in terms of informational strength: *some* . . . *most* . . . *all*. We compared acceptance rates on scalar implicatures with *some* compared to universally true and false sentences with *all* by Korean natives in Korean, English natives in English, and two groups of advanced and intermediate Korean learners of English. Experiment 1 presented scalar sentences without context. Pragmatically infelicitous but logically correct sentences with *some* were the crucial test items. Results indicate that L2 learners derive scalar implicatures significantly more often than the native English and Korean speakers. Experiment 2 used statements in the rich context of stories presented with pictures and text. The pragmatically felicitous responses of the learners increased to over 90%. It is concluded that scalar implicatures present no problem to L2 learners, and that linguistic pragmatic principles are universal. These results are interpreted in light of two influential accounts of scalar implicature calculation and situate them among recent L2A theories.

© 2009 Elsevier B.V. All rights reserved.

Keywords: Second language acquisition; Scalar implicatures; Conversational implicature; Acquisition of pragmatics

1. Introduction

Human communication involves much more than straight encoding and decoding of messages. It is striking that speakers often intend to convey far more than the words they utter, and hearers manage to invoke intended interpretations far beyond the literal meaning of what they hear. Consider the following example of a well-known pragmatic inference:

(1) Some professors are smart.

Most people would agree that, in hearing the utterance in (1), they understand that the speaker has conveyed the assumption in (2).

(2) Not all professors are smart.

Notice that (2) is not encoded by the speaker's utterance, nor is it part of what the speaker has *said*. Rather, (2) is an assumption inferentially derived by the hearer on the basis of what the speaker has said. Logically speaking, *some*

* Tel.: +1 319 335 0208; fax: +1 319 335 3971.

E-mail address: Roumyana-Slabakova@uiowa.edu.

means *some and possibly all*. But if the speaker of (1) had meant *all professors are smart*, she would have uttered (3), being maximally informative. Since she didn't, then we can safely assume she means (2).¹

(3) All professors are smart.

The first systematic attempt to explain how the inference in (2) is derived is due to the philosopher of language Paul Grice. In a series of lectures presented at Harvard in 1967, published later as Grice (1989), he offered a comprehensive framework of the mechanisms of inferential communication. More specifically, he suggested that communication is essentially governed by certain rational expectations about how a conversational exchange should be conducted, which he called “maxims”. According to Grice’s maxims, interlocutors are normally expected to offer contributions that are truthful, informative, relevant to the goals of the conversation and appropriately phrased. These expectations about rational conversational conduct constrain the range of interpretations hearers are entitled to entertain in interpreting utterances. Furthermore, these expectations can be violated (or exploited) to create a variety of effects. According to Grice’s maxims, in producing (1) and meaning (2), the speaker has used part of the following maxim:

(4) Quantity Maxim

- i. make your contribution as informative as is required.
- ii. do not make your contribution more informative than is required.

The speaker has chosen a relatively weak term among a range of words ordered in terms of informational strength: *some . . . most . . . all*. Assuming that the speaker is trying to be cooperative and will say as much as she truthfully can, the fact that she chose the weaker term (*some*) gives the listener reason to think that she is not prepared to make the stronger statement in (3). This leads to the inference that the stronger statement does not hold, that is, to (2). The assumption in (2) is called a conversational implicature, and more specifically, a scalar implicature, since the propositions which *some . . . most . . . all* give rise to are ordered on a scale (Horn, 1972; Gazdar, 1979).

Examples such as these have been studied extensively in post-Gricean philosophical and linguistic work on verbal communication (Horn, 1984; Levinson, 2000; Carston, 1998; Chierchia, 2004; Wilson and Sperber, 2004). Much of this work has tried to tease apart in a principled way the contributions of semantics (what is linguistically encoded) versus pragmatics (what is inferred on the basis of the linguistically encoded meaning) and how exactly this computation happens. Being at the interface of semantics and pragmatics, scalar implicatures and their calculation have often been used to support one or another theory of pragmatics. They are also amenable to experimental study, that is, their linguistic representation and psychological reality in children and adult native speakers is testable using various psycholinguistic procedures (see e.g., chapters in the edited volume Noveck and Sperber, 2004, as well as Storto and Tanenhaus, 2005).

While there is a large literature on how children and adults compute implicatures (see section 2 below), scalar implicatures have not been directly tested in second language acquisition (but see section 3). In recent years, some consensus has been achieved in the field of generative second language acquisition that the challenging properties to acquire lie at the linguistic interfaces: the areas where syntax and morphology, syntax and semantics, syntax and pragmatics meet and interact. In other words, after they acquire the words and the syntactic computation in a second language, do learners know what specific word orders actually mean (the syntax–semantics interface)? And do they know what these word orders actually imply (the syntax–pragmatics interface)? It has recently been proposed that persistent challenges for L2 learners lie at the syntax–discourse pragmatics interface (Sorace, 2003). For example, Belletti et al. (2007) studied the comprehension and usage of null subjects by near-native speakers of Italian whose native language is English. They found that although their subjects showed complete mastery of null subject *pro* licensing in syntax, they continued to diverge on the discourse–pragmatic constraints of the null subject parameter, producing significantly more overt subjects than native speakers. Similar findings are reported by Gürel (2006) and Lozano (2006). The conclusion based on such findings is that pragmatics does present a challenge to second language learners, and divergence persists even in near-native speakers.

¹ The full-blown inference “it’s not the case that all professors are smart” depends on the speaker being epistemically adept about the situation at hand. Thanks go to an anonymous reviewer for making this point.

However, it is not clear what the actual source of this divergence is. Current syntactic theory argues that language variation is encoded in the formal features of the functional lexicon (Borer, 1984; Chomsky, 1995). It follows that linguistic procedures applying at the interfaces should be computed using a universal meaning-computation procedure. Therefore, the issue of transfer from the native language plays out in an interesting way in the area of linguistic pragmatics. The mechanisms of scalar implicature computation, whatever they are, are purported to be universal, that is, used by all human languages and all normally developed individuals. Hence, they are readily available from UG or transferable from the native language of the learner.

Does scalar implicature computation development depend on the maturation of some cognitive capacity, or of processing abilities? If scalar implicature calculation depends on the maturation of some cognitive capacity in children, we expect adult learners to be much better at it than children learning their mother tongue. Not only are they cognitively mature individuals but their native language is in a position to assist them in inference calculation. If, on the other hand, scalar implicature computation depends on processing capacity because it involves choice of an optimal competitor within a narrowly constructed set of options (Reinhart's (2006) reference set computation), we could expect adult learners to have more difficulty than adult and young native speakers.² This is one of the research questions that the present experimental study addresses. It could turn out to be the case that not the meaning computation mechanism itself, but its complex usage and the attending processing costs present persisting challenges to second language learners. The more properties at the syntax–semantics and the semantics–pragmatics interface are studied, the more detailed and informed our insight grows of language acquisition at the interfaces.

This article begins with a discussion of how children (section 2) and adult second language learners (section 3) acquire the computation of implied meaning. Next, I present two influential accounts of scalar implicatures in the current literature, what I dub the Integrative approach (Levinson, Chierchia and other neo-Griceans) and the Relevance Theory account (Sperber, Wilson, Carston, Noveck). In section 5 I present the Korean quantifiers roughly equivalent to the English *some*. Section 6 describes the two experimental studies while section 7 discusses their implications for second language acquisition theory and for the two accounts of scalar implicatures.

2. How do children acquire pragmatic competence?

How and when do children become aware of implicated meanings? Recent experimental investigations into children's interpretations of scalar terms have concluded that preschool children and even children 8–10 years old are often insensitive to these implicational meanings (Chierchia et al., 2001; Musolino and Lidz, 2002; Noveck, 2001). They treat the weaker, underinformative term in the scale *logically*, without being aware of its pragmatic potential and message.

I shall illustrate with a recent study on Greek children and adult pragmatic competence. Papafragou and Musolino (2003) acted out with toys the following (type of) situation in front of their subjects: a number of horses are playing in a meadow, then one by one they jump over a white fence, at the end of the happening all the horses are on the other side of the fence, having jumped over it. The reasonable expectation is to describe the situation with a sentence like (5):

(5) All the horses jumped over the fence.

However, the researchers had a puppet describe the situation with (6):

(6) Some of the horses jumped over the fence.

Then they asked their child and adult participants whether the puppet had “answered well.” Note that the utterance in (6) is *logically* true; if all the horses jumped over the fence, then some of them jumped over the fence, too. However, the utterance in (6) is not *maximally informative*, and thus violates Grice's Quantity Maxim. In other words, the description in (6) is true but pragmatically infelicitous. Greek adults rejected this infelicitous description with 92.5% accuracy. However, Greek 5-year-olds rejected the description in (6) only 12.5% of the time. In a subsequent

² It is true that we do not have independent evidence that reference set computation as described in Reinhart (2006) is harder for second language learners than for adult native speakers. However, there is ample evidence that learners are slower and less accurate in L2 processing than they are in processing their native language (Juffs, 2001; Clahsen and Felser, 2006). Therefore the research question posited here is warranted.

experiment, the researchers trained the child participants in pragmatic anomaly detection. After changing the experimental design so that the children could be more aware of the goals of the task (judging felicity versus truth), 5.7-year-old children rejected the puppet's infelicitous description 52.5% of the time.³ The authors concluded that children have not developed the ability to calculate pragmatic felicity on-line, and discussed explanations in terms of children's awareness of the goals and of the general nature of the task. One explanation offered by Papafragou and Musolino (2003) is that children may have a different assessment of the communicative expectations raised by the experimental set-up, that is, they do not perceive the task to be about the maximal felicity but about the truth value of the statements. Another explanation involves the necessity to invoke two alternatives, a weaker and a stronger one, and to choose between them, which may be too effort-demanding for the children.

Guasti et al. (2005) argue that pragmatic enrichments should be as common among 7-year-olds as they are among adults. They set out to test the Pragmatic Delay hypothesis, according to which children cannot calculate implicatures almost until puberty, and the Pragmatic Limitation hypothesis, according to which children compute fewer implicatures in experimental situations of higher demand on their processing resources. To start with, their Experiment 1 used Noveck's (2001) underinformative factual universal statements (e.g., *Some giraffes have long necks*) and established a baseline for the subsequent experiments. 7-year-olds accepted these statements more than adults did (87% versus 50%). Experiment 2 manipulated experimental demands on the children and showed that training them to give the most specific description of a situation can indeed produce significant gains in pragmatic enrichment. However, Experiment 3 demonstrated that these gains were not long lasting, and had almost disappeared a week after the training session. Experiment 4 used a video-taped Truth Value Judgment Task. The stimuli were not universal factual statements but tightly controlled experimental situations, in which the children saw exactly how many objects were manipulated. The results indicated that when the pragmatic enrichment of a situation was highly relevant to the context, 7-year-old children computed them in an adult-like manner.⁴ Guasti et al. (2005) concluded that the Pragmatic Limitation hypothesis, but not the Pragmatic Delay hypothesis, was supported.

In discussing why 7-year-olds judge universal factual statements so differently from context-enriched experimental situations, Guasti et al. (2005) point out that the former are much more cognitively challenging than the latter.⁵ Furthermore, some adults also accept underinformative sentences. What can be going on? In order to reject an underinformative *some* statement, a speaker must evaluate the "evidence" for two statements: one containing *all* and one containing *some*. For example, in judging the statement *some giraffes have long necks*, speakers may come up with a subset of giraffes, say baby giraffes, which do not have long necks. Alternatively, in evaluating *some bikes have a handle-bar*, speakers may think of broken bikes which could be lacking a handle-bar. Since universal factual statements depend on world knowledge (which may vary widely among individuals), the acceptance of underinformative *some* sentences, at least for adult speakers, may be due to logical calculations of alternative sets to the ones described in the test sentence. Thus, considering and evaluating two alternatives, as well as possibly coming up with evidence for these alternatives, makes scalar implicature computation much more demanding on processing

³ The authors tested four separate utterances in each of three different scales in two experiments, introducing control items as well, to make sure that the children can say No in experimental situations. In their second experiment, they introduced enriched situations in which children were shown, for example, that Mickey put all of the hoops around a pole and they heard the sentence *Mickey put some of the hoops around the pole*. However, children were also told that Mickey claimed to be especially good at throwing hoops and that another character challenged him to succeed with all three hoops. With this type of context enrichment, 5-year-olds were more likely to produce pragmatically enriched judgments. However, they were still significantly different from the adults.

⁴ The results of Pouscoulous et al. (2007) largely confirm the findings of Guasti et al.'s (2005) Experiment 4. These authors manipulated the number of distracters, the type of response expected (action versus statement), and the linguistic complexity of the scalar term (*certain*s versus *quelques*). They generally found that when the task and the linguistic stimuli were simplified, children as young as 4 and 5 produced more pragmatic responses, although still not as many as adults.

⁵ An anonymous reviewer points out that presenting children and adults with sentences out of context, as in Noveck (2001), Guasti et al. (2005), Experiment 1 of this study, and almost any other study of scalar implicatures, does not tell us anything about hearers' real interpretations. The reviewer invokes Crain and Steedman (1985) experimental results (which showed that test sentences without context give us the wrong picture of children's interpretations) and calls for researchers to abandon this testing paradigm. While I am completely sympathetic to the reviewer's objections and would agree with them in most cases, this testing paradigm is only the baseline here and in most other publications on the topic. Furthermore, I will later offer the explanations of one particularly articulate English subject of mine, which are some, although anecdotal, indication of all subjects' thought process in judging underinformative sentences (see section 7). The ratiocinations of this participant underscore the fact that in interpreting sentences without context, he was indeed invoking alternative explanations and comparing them, that is, he was creating his own context.

resources. One might reasonably expect to see reflexes of this high processing demand in the behavior of children as well as second language learners.

In summary, there is a relative consensus in the literature on L1 acquisition of scalar terms that children do not lack the cognitive capacity to make scalar implicatures spontaneously. At the same time, it has been demonstrated that having sufficient cognitive resources is crucial in achieving the pragmatic enrichment of a statement.

3. Second language acquisition of scalar implicatures?

Currently, very few studies have examined whether second language learners have access to implicit contexts: input that is not explicitly stated but implied (Bouton, 1992, 1994; Carrell, 1981; Kasper, 1984; Takahashi and Roitblat, 1994). For example, Bouton analyzed whether L2 learners of English could comprehend conversational implicatures. He found that, compared with native speakers, learners' performance in comprehending the implicatures was significantly poorer, although it improved over time. Takahashi and Roitblat (1994) also investigated L2 comprehension of indirect requests. Their findings suggest that, although learners were able to arrive at the correct interpretations, they took longer to comprehend indirect requests than native speakers.

More recently, Dekydtspotter and Hathorn (2005) and Dekydtspotter et al. (2005) have explored L2 knowledge of implicatures as part of calculating meanings in complex constructions. Dekydtspotter and Hathorn (2005) tested knowledge of implied interpretation in the case of continuous (as in (7)) and discontinuous quantifiers (as in (8)) in the French interlanguage of English native speakers.

(7) *Quelque chose de remarquable a été observé par chacun des chercheurs.*
something of remarkable was observed by each of-the researchers
'Something remarkable was observed by each of the researchers.'

(8) *Quelque chose a été observé de remarquable par chacun des chercheurs.*
something was observed of remarkable by each of-the researchers
'Something remarkable was observed by each of the researchers.'

Both (7) and (8) may describe a situation where each researcher observed a different remarkable object. However, when French people want to express that each researcher observed the same remarkable object, they produce (7), but not (8). When uttering (8), native French speakers imply that the same remarkable object was not observed by all researchers. The correct interpretation of (7) and (8) depend on scalar implicature. The two sentences give rise to the following semantic configurations:

(9) $\forall y [\text{researcher}(y) \rightarrow \exists x [\text{thing}(x) \wedge \text{remarkable}(x) \wedge \text{observe}(x)(y)]]$ ($\forall\exists$)
'For every researcher there is a remarkable thing such that he/she observed it.'

(10) $\exists x [\text{thing}(x) \wedge \text{remarkable}(x) \wedge \forall y [\text{researcher}(y) \rightarrow \text{observe}(x)(y)]]$ ($\exists\forall$)
'There is a remarkable thing such that every researcher observed it.'

However, following de Swart's work (1992) on constraints over discontinuous constituent constructions, the researchers assume that the discontinuous constituent as in (8) only gives rise to an $\forall\exists$ interpretation (9). Semantically, the representations in (9) and (10) stand in an asymmetrical entailment relation: whenever (10) is true, (9) is true, but when (9) is true, (10) need not be true. In other words, (9) and (10) are in a superset–subset relationship. This asymmetric entailment creates an informational scale in which the entailing $\exists\forall$ representation in (10) is informationally stronger than the entailed $\forall\exists$ representation in (9). Using the Maxim of Quantity, the implication arises that the stronger (10) does not obtain. Dekydtspotter and Hathorn (2005) investigate whether second language learners of French come to know that the discontinuous sentence (8) excludes situations where the same object was observed by all researchers.

The study used a Truth Value Judgment Task in quadruples, where a story representing explicit context (either about one remarkable object or about many remarkable objects being observed by different researchers, presented in English) was followed by a test sentence in French, either as in (7) or as in (8). Participants were asked to say whether

the test sentence was adequate to the story, and they could say Yes, No, or I don't know. Dekydtspotter and Hathorn (2005) found that native French speakers accept both one-object and different-object answers equally in continuous constructions, while they significantly preferred the different-object answer in discontinuous constructions. No effect for word order was found in the performance of the low intermediate proficiency learners of French ($n = 40$), while high intermediate learners ($n = 20$) showed the same pattern as the natives. The asymmetries in that response pattern cannot logically be accounted for unless implicatures are calculated. The authors concluded that scalar implicatures appear to be observed in a native-like fashion, therefore supporting the claim that L2 acquisition is constrained largely by a universal meaning computation processor.

Dekydtspotter and Hathorn's (2005) study investigates L2 knowledge of implicature arising from grammatical (LF) representations. Successful performance on their test depends on correctly interpreting discontinuous constituents, a complex syntactic construction. The study I report on here is the first, to my knowledge, that investigates scalar implicatures related to a lexical scale and is thus directly comparable to the interesting findings from child language acquisition.

4. Two theoretical accounts of SI in children and adult native speakers

Grice's original idea of implicatures involved a less informative term being used in an utterance in a way that does not satisfy the first Maxim of Quantity. If a speaker is seen to use the weaker of two alternative utterances, she is taken to implicate that the proposition that would have been expressed by the use of the stronger term in the scale is false. This analysis has significant implications for language processing. According to Grice's original account, the inference from the utterance to its scalar implicature goes through an evaluation not just of what the speaker has said and the context but also of what the speaker might have said but did not, which is an arduous process. The Gricean account of implicature derivation seems to be implausible from a cognitive and developmental point of view (Noveck and Sperber, 2007). Two influential accounts of scalar implicature computation have been proposed in the literature, which envisage different sequences of processes and implicate different architecture of the linguistic modules: the (neo-Gricean) Integrative account (Levinson, 2000; Chierchia, 2004, 2006) and the Relevance Theory account (Sperber and Wilson, 1986/1995; Carston, 1998). These accounts also make different predictions with respect to the relative speed and ease of implicature computation.

Levinson (2000) elaborates on Grice's original and somewhat vague description of implicatures. For Levinson, generalized conversational implicatures are *default inferences*, that is, inferences that are automatically generated and that may be cancelled if context appears to call for it. Levinson treats scalar implicatures as exemplary cases of generalized conversational implicatures (GCI) (whereas Grice's own examples of GCIs don't include scalar implicatures). Levinson's theory of scalar implicatures as default GCIs argues that they exploit pre-existing scales such as *<some, all>*, *<or, and>*, *<possible, necessary>*, *<start, finish>*, etc. The most important claim from a psycholinguistic perspective is that these pragmatically enriched meanings arise by default, as part of the lexical meanings of the scale items irrespective of the context, and can be cancelled when the context demands. The fact that these inferences are made by default adds to the speed and efficiency of communication. In the same vein, Chierchia (2004, 2006) posits that pragmatic enrichment using a scale does not happen at the end of an utterance, when semantic calculation is finished and pragmatic calculation kicks in. Implicature calculation is relatively easy because it is a one-step process that does not really depend on context and does not involve the arduous calculation of what was said and what might have been said. Implicatures are factored in throughout the computation of meaning, with semantic and pragmatic calculation operating in a recursive, compositional manner.

Given this view of the language architecture, observed differences between adults and children are more likely to arise because the undoing of the scalar meanings adds to the complexity of the language processing by consuming processing resources. Recall from the previous section that Guasti et al. (2005) argues that logical interpretations may be given when implicatures are undone by conjuring up alternative contexts. For example, in evaluating the sentence *Some giraffes have long necks*, both adult and child participants can conjure an alternative set of giraffes, say, baby giraffes, that do not have long neck, and they do grow long necks later on. Since coming up with alternative sets of referents is more work, logical answers to infelicitous *some* statements are predicted to be more effortful.

In the Relevance Theory framework (Sperber and Wilson, 1986/1995; Carston, 1998; Wilson and Sperber, 2004), an implicature is defined as an inference that the speaker intends and expects the hearer to draw in order to arrive at an interpretation of the utterance that is relevant enough. In particular, a scalar implicature is derived when a relatively

weak statement fails to meet the hearer's expectation of relevance. Linguistic expressions serve not to *encode* the speaker's meaning but to *indicate* it. The speaker's meaning is inferred from the linguistic meaning of the words and expressions, but only taken together with the context. Furthermore, inferring the speaker's explicit and implicit meaning is not done sequentially but in parallel. Although context can have local effects, it can only be factored in completely when the whole utterance meaning is calculated. The final overall interpretation of an utterance results from a mutual adjustment of implicatures and explicatures guided by expectations of relevance (Noveck and Sperber, 2007). From a relevance-theoretic perspective, so called 'scalar implicatures' are not scalar, nor even necessarily implicatures. The speaker's meaning is always inferred, even when it consists in a literal interpretation of the linguistic expressions used. The inferences involved, however, differ in the time and effort they require. Given that a scalar implicature is an inference that goes further than the semantic origins of words like *some* and that it is carried out to achieve relevance and has to consider context, Relevance Theory considers implicature an effortful, non-necessary inference.

How do the two accounts square with the child language data on pragmatic enrichment? If pragmatic inference is a relatively easy one-step process while logical undoing of implicatures necessitates context and evaluating of alternative propositions as the Integrative account suggests, who should be better at implicature, children or adults? The latter account is apparently contradicted by data indicating that children accept more logical than pragmatically felicitous interpretations, compared to adults. Many recent studies, some of which I summarized in section 2 above, point to such a conclusion. Furthermore, if training and context helps children be more pragmatically felicitous (e.g., Papafragou and Musolino, 2003; Guasti et al., 2005), then the Relevance Theory account seems to be supported? I believe that children and adults may achieve pragmatic enrichment using the same mechanism but may accept logical meanings of underinformative sentences for different reasons. The mechanism Guasti et al. (2005) suggest (undoing implicature by conjuring up a subset of described objects for which the statement may not hold, e.g., baby giraffes may not have long necks) is probably more available to adults than to 5–7-year-olds. Since no child language experimental study so far has addressed this issue, we turn to adult second language learners to tease apart the two accounts.

What are the opposing predictions of the two accounts with respect to the psychological reality and the development of scalar implicatures in a second language? It is well known that L2 learners have less computational, processing resources at their disposal than they do in their native language. Psycholinguistic studies of L2 processing show that even if they demonstrate the same pattern of processing as native speakers, for example as measured by word-by-word reaction times in the moving window technique, L2 learners are often much slower and sometimes less accurate compared to native speakers (see White and Juffs, 1998; Juffs, 2001 and Clahsen and Felser, 2006 for reviews). At the same time, they are cognitively mature individuals and have access to the universal implicature computation mechanism, whatever it may be.

The learning task for L2 acquisition involves transferring the purportedly universal scalar implicature computation mechanism from the L1. Therefore, we expect L2 learners to be accurate in the interpretation of the scalar terms, but that processing resources may have an impact on accuracy (and speed). As Table 1, largely based on Table 1 from Noveck and Sperber (2007) illustrates, the Integrative account predicts that, assuming less processing power in L2 than in L1, L2 learners will give a smaller percentage of logical answers to scalar implicatures without context, compared with adult native speakers. The Relevance Theory account, on the other hand, predicts that L2 learners will produce a smaller percentage of pragmatic answers (hence more logical answers) compared with native adults.

Table 1
Predictions of the two accounts for interpreting "some" as *some and possibly all* (the logical choice) (adapted from Noveck and Sperber, 2007).

Interpretation of the scalar term	The Integrative account	The Relevance Theory account
Literal	default enrichment + context-sensitive cancellation, <i>more difficult</i> hence slower	no enrichment, <i>easier</i> hence faster
Enriched	default enrichment, <i>easier</i> hence faster	context-sensitive enrichment, <i>more difficult</i> hence slower

5. Korean quantifiers

Before discussing the interpretation of English underinformative sentences by Korean native speakers, we must consider the array of lexical items that would map English quantifier *some*. Korean has two quantifiers of similar meaning, *ilbu* and *etten*. It seems that they are close equivalents, as (11) illustrates:

- (11) Ilbu/etten haksayng-tul-un yanghwasalul silh-e-ha-n-ta
Some/some student-pl-Top quantifier-Acc hate-Link-do-Pres-Decl
'Some students hate quantifiers.'

The quantifier *ilbu* is Sino-Korean and its literal meaning is 'one part'. In some sentences (and for some speakers), its usage is somewhat different from *etten*.

- (12) Etten salam-tul-un khi-ka khu-ta.
some salam-Pl-Top heigh-Nom tall-Decl
(i) 'Some people are tall.'
(ii) 'Some among the people are tall.' (with reference to a specific subset of all people)
- (13) Ilbu salam-tul-un khi-ka khu-ta.
one part salam-Pl-Top heigh-Nom tall-Decl
(i) # 'Some people are tall.'
(ii) 'Some among the people are tall.' (with reference to a specific subset of all people)

According to three native Korean informants,⁶ *ilbu* always refers to plural entities, while *etten* can refer either to singular or plural entities. In (12) above, *etten* has both a general existential meaning as in (i) but also a subset of a specific set meaning as in (ii). *Ilbu* only has the second meaning. Thus it was felt by my informants that *etten* is more likely to be a counterpart of English *some*.

6. The present study

6.1. Experiment 1

6.1.1. Materials and procedure

Experiment 1 replicated Noveck's (2001) study on French *certaines* (see also Guasti et al.'s 2005 Experiment 1 on Italian children; Pouscoulous et al., 2007; Experiment 1 from Feeney et al., 2004; among others). This particular experimental paradigm has become the standard baseline in child language studies of implicature and the child language results across studies have been largely consistent. Four groups participated in Experiment 1: 30 Korean natives, who took the test translated in Korean, 23 native speakers of English, 30 advanced and 20 intermediate Korean learners of English, who took the English test. The Korean native group was added to establish the baseline for the perception of implicatures by Korean adults. Scalar implicatures are argued to be universal, but of course such a claim is stronger if it enjoys empirical support.

All the Korean participants were either graduate or undergraduate students at a Midwestern University in the United States. They had all started to learn English in the classroom at 12 or 13, but they were not exposed to English for communication purposes before they began their studies in the US. Their proficiency level in English was established based on their TOEFL scores presented at the university in which they were enrolled. The advanced learners had attained the proficiency allowing them to study as undergraduates or graduate students at a US university. Their TOEFL scores averaged 598 points. Most of the intermediate students were admitted to ESL classes at the same institution, and their available TOEFL scores at the time of testing had a mean of 486. As it turned out, the performance of the two groups did not vary according to proficiency.

⁶ I am grateful to Jae Shim, Kum Young Lee, and Danny Kang for judgments and discussion.

Table 2
Percentage of logical responses and standard deviations across participants groups in Experiment 1.

Groups	True <i>all</i>	False <i>all</i>	Felicitous <i>some</i>	Infelicitous <i>some</i>
English adults ($n = 23$)	75.5 (29.8)	98.9 (3.4)	96.7 (5.6)	55.4 (46.1)
Korean adults ($n = 30$)	88 (18.5)	98.5 (3.8)	99 (3.8)	61.2 (43.7)
L2 advanced ($n = 30$)	82 (22.4)	98 (5.4)	98 (5.4)	39.2 (38.9)
L2 intermediate ($n = 20$)	78 (26.8)	97 (6.8)	90 (17)	41.8 (40.6)
Kruskal–Wallis	$H_{(3)} = 2.5$	$H_{(3)} = 1.57$	$H_{(3)} = 5.71$	$H_{(3)} = 7.54$
H statistic	$p = .461$	$p = .665$	$p = .12$	$p = .043$

The test items, factually universal and factually existential statements, were presented to the learners in a written form. Following the established procedure, they had to circle whether they *Agree* or *Disagree* with the statement. True or False answers were not used because underinformative *some* statements are logically true but pragmatically infelicitous. Scalar implicature researchers, who have used this test in a number of languages, do not want to bias participants in judging test sentences, either along logical lines or along pragmatic lines, so a test answer along the lines of *Agree/Disagree* has been felt to be neutral in this respect. The test items included 8 universally true sentences as in (14), 8 sentences infelicitous with *some* as in (15), 8 sentences felicitous with *some* as in (16), 8 sentences false with *all* as in (17), and 8 absurd fillers as in (18).

- (14) All elephants have trunks.
 (15) Some elephants have trunks.
 (16) Some books have color pictures.
 (17) All books have color pictures.
 (18) All/some garages sing.

6.1.2. Group results

The responses were coded for logical correctness, that is, answering “I agree” to the underinformative sentences as in (15). Table 2 presents the percentage of logical responses. Since there is little variance in the data in some conditions, non-parametric tests were used.⁷ Results of the Kruskal–Wallis test together with means and standard deviations are given in Table 2. We are most interested in the acceptance (pragmatically felicitous) or rejection (pragmatically infelicitous) of the underinformative *some* statements (*Some elephants have trunks*). This is the only condition in which we find significant differences between groups. The differences are due to the fact that the two learner groups demonstrate considerably less logical responses than the English and Korean adults judging the same sentences in their native languages. There are no differences between the performance of the two native speaker groups ($p = .91$ by post hoc Tukey HSD test), which suggests that the scalar implicature computation mechanism is similar in English and Korean. These results are also roughly in agreement with Feeney et al.’s (2004) results (English adults rejected underinformative sentences 65% of the time) and Guasti et al.’s (2005) results (Italian adults rejected them 50% of the time). The performance of the two learner groups is also similar, which suggests that there is no development with respect to this property based on increased proficiency. Once the lexical items are learned, computation of implicatures is supported by the native language and somewhat elevated due to processing resources (see section 7).

To reiterate the main finding in these results, Korean learners of English accept less logical answers, hence more implicatures (~60%) than they do in their native language, and more than English and Korean native adults.

Looking at group results, there is still the question of whether native speakers and learners were responding at chance (50%), particularly on Infelicitous *some* items. This issue will be addressed in two ways. First, a reliability

⁷ I am grateful to an anonymous reviewer for suggesting this procedure.

R. Slabakova/Lingua xxx (2009) xxx–xxx

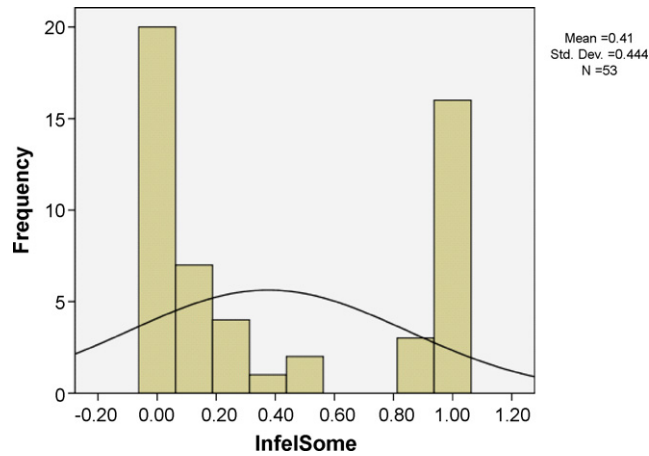


Fig. 1. A histogram on native speaker frequency of logical and pragmatic choices in the infelicitous *some* condition.

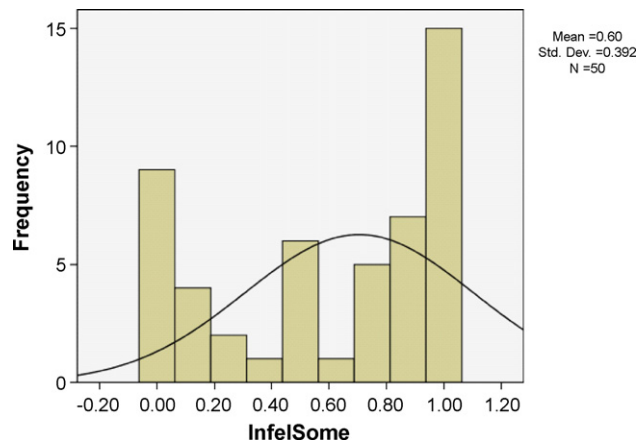


Fig. 2. A histogram on non-native speaker frequency of logical and pragmatic choices in the infelicitous *some* condition.

alpha analysis was performed, and the Cronbach’s alpha statistics is .856, which is considered significant (Field, 2005). More importantly, none of the item values is bigger than the actual alpha, suggesting that the removal of this item would not improve overall reliability. Secondly, and much more visually convincing, histograms were plotted of the frequencies of the different scores for native speakers and non-native speakers separately. Figs. 1 and 2 give these histograms with the normal curve superimposed. As is evident from the histograms, the frequencies have a clearly bimodal distribution, with the biggest majority of speakers choosing clearly “logical” or clearly “pragmatic” answers, and not too many people choosing one or the other 50 percent of the time. See also Table 3 below for another presentation of individual results.

Table 3
Number and percentage of “Pragmatic” and “Logical” individuals per participant group in Experiment 1.

	Number who chose pragmatic answers over 75% of the time	Number who chose logical answers over 75% of the time	Unclear choices (between 62.5% and 37.5% logical answers)
English natives	9 (37%)	13 (54%)	1 (4%)
Korean natives	10 (33%)	19 (64%)	1 (3%)
ESL advanced	17 (57%)	8 (27%)	5 (16%)
ESL intermediate	10 (50%)	7 (35%)	3 (15%)

6.1.3. Individual results

Group results may hide important individual differences, especially in this area of meaning computation where individuals may take different approaches to felicity. That is why individual accuracy on the infelicitous *some* condition was calculated with a 75% cut-off point, or six out of eight items. If a participant chose pragmatic answers 75% and above, she was classified as a predominantly “pragmatic” individual; if a participant chose 25% and less pragmatic answers on the Infelicitous *some* condition, she was classified as a “logical” individual. The tally is given in Table 3.

The main conclusion we can draw from Table 3 and the histograms in Figs. 1 and 2 is that most English and Korean adults as well as L2 learners fall into one of two groups, Logical or Pragmatic individuals. The similarity in the patterns is striking. That is, most individuals do not demonstrate haphazard behavior, but are consistent in their choices of either being logically correct, or being pragmatically felicitous, in judging underinformative sentences. The same split was observed with French adults in Noveck’s (2001) study. The fact that roughly the same categorical distribution obtains for the two learner groups suggests that people are still either logical or pragmatic with respect to scalar implicature calculation, in their second language as well as in their native language. However, it is important to notice that the number and percentage of “logical” individuals diminishes when learners are judging underinformative sentences in their second language. The percentage of individuals who do not have very clear judgments also rises to about 16%. These individual results are entirely in keeping with the group findings, and confirm that the logical choices in judging underinformative sentences are much more common in the native language (63–67%) than they are in a second language (27–35%).

6.2. Experiment 2

6.2.1. Participants and procedure

A group of 20 native English speakers and another one of 35 Korean natives participated in Experiment 2, each tested in their own language. The learner groups consisted of 36 advanced and 20 intermediate Korean learners of English. The same characteristics of the learners obtain as in Experiment 1. Some of the same individuals participated in the second experiment, but testing happened more than six months after the first experiment.

The purpose of Experiment 2⁸ was to provide participants with the necessary and sufficient context for judging underinformative sentences. They saw a series of pictures mounted on black cardboard, with written sentences below each picture. All stories were about Charlotte, who interacted with two out of three, or with three out of three objects. For example, in a *some* infelicitous story, Charlotte was seen eating all three candies that she found on the kitchen table. Her mother asked: “Charlotte, what have you been doing with the candies?” Charlotte told her mother that she ate *some* of the candies. The intention was to show that Charlotte was trying to hide the extent of her naughtiness from her mother, at least some of the time. The participants were alone in a quiet room with a research assistant, who showed them the storyboards one by one and read aloud the sentences describing the events in each picture. The question learners had to answer was again *Do you agree?* The research assistant recorded the oral answers. There were four tokens each of true *all*, false *all*, felicitous *some* and infelicitous *some* story–answer combinations (Fig. 3).

6.2.2. Group results

Responses were again coded for logical correctness, that is, answering “I agree” to *I have eaten some of the candies* when Charlotte had eaten all three. The reliability statistic Cronbach’s alpha is .702, which is not unusual for questionnaires in the social sciences with a small(er) number of questions (Field, 2005:667–668). Only one item (T-all-3) had a value that is bigger than the actual alpha, suggesting that the removal of this item would improve overall reliability. Table 4 presents the percentage of logical responses. For the same reasons as in Experiment 1, non-parametric Kruskal–Wallis tests were run and the results are in the last row of the table.

As Table 4 shows, Korean learners of English are highly accurate on the True *all*, False *all*, and Felicitous *some* sentences. They give only about 10% logical answers, that is, about 90% pragmatic answers, on the crucial Infelicitous *some* condition. In this respect, they are significantly more “pragmatic” than the English native speakers with 62.5%

⁸ This format of the test was a replication of Experiment 2 from Feeney et al (2004). I am indebted to Susan Crafton, who made the original pictures and test sentences available to me.



Charlotte finds three pieces of candy on the kitchen table. Charlotte likes candy.



Charlotte eats the first candy.



Charlotte eats the second candy.



Charlotte eats the third candy.



Charlotte's mum says, "Charlotte what have you been doing with the candies?"
Charlotte says, "I've eaten some of/all the candies."

Fig. 3. An example of a storyboard used in Experiment 2.

pragmatic answers ($F(2, 74) = 8.73, p < .0001$) and the Korean native speakers with 75% pragmatic answers ($F(2, 90) = 6.47, p < .001$).⁹ These results point to the fact that even with the explicit context of this experiment, the means in the two native groups still contain some logical answers, although to a lesser extent than the means of underinformative sentences in Experiment 1. In other words, a number of judgments gave Charlotte, the fictional character in the experimental items, the benefit of the doubt and accepted that she was telling the (logical) truth even

⁹ A parametric one-way ANOVA is justified in this condition.

Table 4
Percentage and standard deviations of Logical Responses in Experiment 2.

Groups	True <i>all</i>	False <i>all</i>	Felicitous <i>some</i>	Infelicitous <i>some</i>
English natives ($n = 20$)	96 (12)	99 (5.6)	98 (7.7)	37.5 (42.45)
Korean natives ($n = 35$)	94 (17.2)	91 (17.1)	91(18.1)	25 (33.4)
Advanced ($n = 36$)	95 (10)	94 (10.9)	98 (14)	9 (24.7)
Intermediate ($n = 20$)	86 (15.9)	93 (14.3)	95 (10.2)	12.5 (26.2)
Kruskal–Wallis	$H_{(3)} = 1.58$	$H_{(3)} = 3.72$	$H_{(3)} = 2.72$	$H_{(3)} = 10.99$
H statistic	$p = .678$	$p = .294$	$p = .437$	$p = .012$

Table 5
Number and percentage of “Pragmatic” and “Logical” individuals per participant group in Experiment 2.

	Number who chose pragmatic answers over 75% of the time	Number who chose logical answers over 75% of the time	Unclear choices (50% logical answers)
English natives	12 (60%)	7 (35%)	1 (5%)
Korean natives	27 (77%)	6 (17%)	3 (8.5%)
ESL advanced	33 (92%)	2 (5%)	1 (2.7%)
ESL intermediate	17 (85%)	1 (5%)	2 (10%)

she when had interfered with three items but told her mother that she had handled only two of them. The Korean learners of English were not so “generous” to Charlotte and suggested she was trying to cheat significantly more often.

6.2.3. Individual results

The results of Experiment 2 from the point of the choices of individual learners are presented in Table 5. The same split exists in individuals, in the sense that most natives and learners are consistent in their choices. When comparing the individual results on infelicitous *some* sentences in the two experiments, see Tables 3 and 5, we notice that some individuals who judged underinformative sentences logically were convinced by the presented context and judged the sentences pragmatically. But a certain number of individuals (7 English natives, 6 Korean natives and 3 out of 56 learners) retained their logical treatment of the weak quantifier *some*. In this respect, the individual results confirm the group results in Experiment 2: individual participants behave exactly as suggested by the group means.

7. Discussion

Let me summarize the main findings of both experiments in this study. The experimental materials asked speakers to choose between two freely available interpretations of underinformative sentences with the scalar terms *some* and *all*. They could interpret *some* to mean *some and possibly all*, the logical interpretation, or *some but not all*, the informationally stronger pragmatic interpretation. The most important finding is that, when given a free choice, Korean learners attribute more pragmatic interpretations to scalar implicatures without context (~60%) than they do in their native Korean (38%), and significantly more than English adults (45%, see Experiment 1). When asked to judge sentences with *some* in context which makes that use infelicitous, they offer pragmatic judgments in around 90% of the times, and again significantly more than they would in their native language (75%) and more than English native speakers (62.5%). In other words, learners behave more pragmatically in their second language than in the native language, no matter if the judgments are supported by context or not. These conclusions are strongly confirmed by the individual results as well. These findings may address the two accounts of scalar implicatures introduced in section 4. To reiterate, the Integrative account (Chierchia, 2004; Levinson, 2000) proposes that scalar implicatures are computed automatically and integrated into the semantics. Thus, logical interpretations are given only when implicatures are *undone* by conjuring up alternative contexts, hence logical answers to infelicitous *some* statements are more effortful. Contra the Integrative account, the Relevance Theory account (Sperber and Wilson, 1986/1995; Carston, 1998) holds that scalar implicature computation comes after the logical form of the sentence is complete and only if the context is relevant. Thus, implicatures are more effortful. Table 1 summarized these predictions and is based on Noveck and Sperber (2007).

I suggest that this study's Experiments 1 and 2 results are more consistent with the Integrative account than they are with the Relevance Theory account. Clearly, L2 learners derive more implicatures than native speakers, with or without the support of context (although the percentages differ in the two experiments). Why would that be? Implicature derivation in Experiment 1 depends on world knowledge. To explain why adult natives (and possibly some children) give logical responses, Guasti et al. (2005) argue that these speakers conjure up alternative contexts, which make the underinformative *some* statements plausible (e.g., *Some elephants have trunks* may be considered felicitous if baby elephants do not have trunks). I suggest that the overall depressed accuracy on True *all* sentences (for all participant groups) is in support of Guasti et al.'s stipulation. I want to repeat here the justification one English participant (#12) gave to some of his answers. In judging sentences like *All elephants have trunks* and *All cats have ears*, this participant did not agree with the statements, and added: "I assume that there are some elephants that have had their trunks removed, for whatever reason." The same individual was 100% consistently logical on underinformative sentences like *Some elephants have trunks*, accepting them for the same logic. This person made similar remarks in writing on three separate sentences, which points to the fact that he was really aware of and considering alternative explanations to the universal and existential statements without context. Similarly, when faced with explicit context suggesting that Charlotte, the fictional character of Experiment 2, is trying to misrepresent facts to her mother, the biggest majority of learners (53 out of 56) indicate that this misrepresentation is pragmatically infelicitous. In contrast, a sizable minority of Korean and English native speakers (17% and 35%, respectively) still maintain their logical answers, even in defiance of the contextual factual situation and in a way finding "logical" excuses for Charlotte's economy with the truth.

To explain the native speaker–learner discrepancy, I would like to propose that L2 learners may not be able to come up with alternative contexts as often as they do in their native language, hence they derive L2 implicatures significantly more often than native adults. Relevance Theory cannot easily explain these findings. If implicatures were more effortful in terms of processing resources, the fact that L2 learners make them more often than native speakers would go against everything we know about processing in a second language (see more on this below). It is hardly plausible to imagine that speakers are more sensitive to context while speaking their second language, as compared to speaking their native language, as the Relevance Theory account would have it. According to this account, pragmatic readings arise when hearers and readers are not satisfied with the relevance of the logical readings, which appear by default. A much more credible explanation is that L2 learners lack the processing resources to undo automatic pragmatic interpretations. Thus, default calculation of implicature appears to be supported. These results from second language acquisition may be in a position to contribute to the theoretical debate on how exactly scalar implicature computation proceeds, by indicating that the Integrative approach offers a better explanation for the findings.¹⁰

These findings also point to the fact that the higher accuracy in computing the pragmatically felicitous meaning of underinformative sentences by L2 learners is not due to some (ill-defined) cognitive maturation. If it were a cognitive maturity problem, one would expect all the adult participants in this experimental study to do equally well, since they are all cognitively mature individuals.¹¹ Instead, we see a marked difference in the behavior of native speakers as

¹⁰ Of course, undoing default implicatures by conjuring up alternative contexts is only one possible explanation of my findings. An anonymous reviewer suggests that second language learners may generate more implicatures because they allow a different level of felicity in their L2 as compared to their native language. The reviewer conjectures that "native groups who are presented with underinformative utterances do note that the utterance is only partially true but they do not consider this infelicity to be grave enough to warrant the rejection of the utterance. On the contrary, second-language groups, who may be very much aware that they are performing a task in a test language, might be more sensitive to any deviation from the accepted norm..." The issue of whether speakers are more sensitive to linguistic infelicity in their native or their second language is addressed by many studies investigating knowledge of semantics and pragmatics in L2A. For a review of such studies see Slabakova (2008), especially studies discussed in chapter 6. I would like to give just one example of a learning situation at the syntax–semantics interface (there are not enough L2A studies on the semantics–discourse interface, see section 3 above). White et al. (1997) studied L2 knowledge of binding of reflexives in French and Japanese learners of English. In both English and French, binding of reflexives is local (where the Governing Category is the same tensed clause) and binding to subject and object is possible (*Jane_i told Mary_j the truth about herself_{ij}*). It is also noted that native speakers prefer subject binding to object binding; thus binding by object is a grammatically available but dispreferred option (when there is a pragmatically valid choice). White et al. found that English native speakers allow object binding in 87.5% of the cases, while French-native learners of English only allow it 75% of the time, although French works in exactly the same way as English as far as binding of reflexives is concerned. When we compare the rejection of long-distance antecedents for reflexives, the English natives were accurate 95% of the time while the French ESL speakers only 67.5%. This is only one of many experimental works which document the fact that L2 speakers are less sensitive, not more sensitive, than native speakers, to felicity and ungrammaticality in language even on properties directly transferred from their native language.

¹¹ It is of course true that younger children behave differently from adults, as demonstrated by many other experiments, reviewed in section 2.

opposed to L2 learners, the latter being consistently more “pragmatic” in their judgments. Individual analysis in Tables 3 and 5 revealed that the vast majority of participants fall categorically into either “logical” or “pragmatic” respondents.¹² That is, some people readily conjure up alternative contexts and evaluate multi-member reference sets (Reinhart, 2006), while others simply stay with the default enrichment provided by the scale. However, lesser percentages of learners are “logical” across experiments as opposed to natives. The source of this division of native speakers and of learners may well be availability of processing resources. It takes processing resources to calculate a meaning, compare it to a situation, and then discard this meaning in favor of another. The more processing resources an individual possesses, the easier it is to keep in short term memory and compare the alternative meanings for the same sentence. The performance of the Korean learners of English is in support of this explanation. First of all, note that there are no proficiency level effects in the data, suggesting that learners are not experiencing any lexical or grammatical difficulties with the structure of the test sentences. Secondly, it is well established in the psycholinguistic literature on L2 acquisition that when learners interpret complex sentences, they demonstrate largely the same parsing decisions and patterns as native comprehenders, but they are significantly slower and less accurate (Juffs, 2004, 2005, 2006; Clahsen and Felser, 2006). Researchers argue that the basic mechanisms of grammar remain intact in L2 learners, in spite of inferior performance on judgments and speed. Thus, the inferior performance is attributed to processing difficulties, not competence differences with native speakers (White and Juffs, 1998), which could very well be due to more limited processing capacity. We see another reflection of this limited processing capacity, but no difficulty with grammatical structure, in this experimental study.

We must consider another possible explanation of the experimental results: the effects of bilingualism on executive functioning (involving attention, inhibition and focusing) in children and adults. Bialystok (2001), Bialystok and Senman (2004), Bialystok and Martin (2004) and others have shown that bilinguals often exhibit significantly superior executive functioning and attentional abilities that are associated with better responses on metacognitive and metalinguistic tasks. Thus it is possible that our adult learners exhibit superior pragmatic competence compared to the native participants. Siegal et al. (2009) is particularly relevant to the present discussion, because it examines pragmatic competence in bilingual and monolingual children. Children participating in this experiment were bilingual in Italian and Slovenian, or monolingual in either language. The researchers tested 3–6-year-old children on a conversational violations test to find out whether they would obey Gricean maxims. On a laptop, children were shown short conversational exchanges involving three dolls. One doll asked a question, the other two provided alternative responses: one violating a maxim and the other pragmatically felicitous. Children were asked to point to the doll which said something “silly” or “rude”.

Results of both experiments in Siegal et al. (2009) (which follow the same procedure but have different participants) show that there is a definite advantage of the bilingual children over the monolingual ones on four Gricean maxims: Quantity II, Quality, Relation and Politeness. Bilingual children were more accurate in choosing non-redundant answers, true answers over false ones, answers that were relevant to the questions, and polite answers over rude ones. The only maxim on which all the children performed equally well and hovered at around 60% pragmatic responses was the Maxim of Quantity I. Here is an explanation of those experimental items: “The target utterances were designed to fall short of providing an informative enough answer, as in the following:

Question: “What did you get for your birthday?”

Answer: “A present.” (Alternative appropriate answer: “A bicycle.”) (Siegal et al., 2009:116)

As the reader can ascertain, these test items are very close in meaning to the scalar implicature contexts tested in the present experiment. Results of 60% pragmatic answers for children before the age of six are largely in line with other studies in the literature on scalar implicature computation in children (see section 2). More importantly, however,

¹² It is interesting to note that 7–8-year-old children show a different pattern of response from L2 learners (see section 2). Children’s, native adults’, and adult L2 learners’ choices fall on a continuum, with children being the most logical and adult L2 learners being the most pragmatic. The difference could be due to processing resources. For example, Reinhart (2006:289) argues that children do invoke the two interpretations in the necessary reference set, but get stuck there. Then they use an arbitrary default strategy, either “minimal commitment” or “maximal commitment” to the stronger of the two alternatives. Thus children’s behavior differs substantially from native adults’ behavior in that adults depend on context in their choice of alternative while children use an arbitrary default. Since the focus of this article is not child but adult L2 learner behavior, I will not consider these options further.

Siegal et al. (2009) do not establish an advantage for bilingual children comprehending underinformative sentences. It is possible that comprehending underinformative sentences involves different semantic-pragmatic calculations than detecting relevance and rudeness. I suggest that, even if we disregard the differences between the child bilinguals tested by Siegal et al. (2009) and adult L2 learners tested in the present study, the superior attentional and executive functioning advantages of bilingualism cannot explain our results.

What are the implications of the current findings for second language acquisition theory? Current theoretical proposals about second language development have highlighted the importance of studying linguistic phenomena at the interfaces between syntax and other cognitive domains (Sorace, 2003). In a recent study of anaphora resolution in near-native speakers of Italian, Sorace and Filiaci (2006) put forward the Interface Hypothesis (see also Belletti et al., 2007). According to this hypothesis, narrow syntactic properties in L2 acquisition are completely acquirable but linguistic knowledge that involves an interface may not be acquirable and can lead to lasting indeterminacy in L2 grammars, felt even at the near-native level. In agreement with White and Juffs (1998), Sorace and Filiaci also argue that L2 grammars may have insufficient processing resources at their disposal to integrate the multiple types of information involved at the interfaces between syntax and discourse. However, the picture that emerges from studies of null subjects and their appropriate usage is rather dark and pessimistic. It seems that L1 effects on L2 performance may be persistent and lead to permanent divergence from native speaker performance. Furthermore, the claims about acquisition at the interfaces (e.g., the Interface Hypothesis) are made largely based on data from null subject usage. It is appropriate, therefore, to consider L2 performance on another linguistic property, this time at the semantics–pragmatics interface.

I argued above that non-native participants in this experimental study did not have any difficulty with the lexical items and the grammatical constructions of sentences in Experiments 1 and 2 (as indicated by the lack of proficiency level effects on accuracy, and by their superior accuracy on True *all* and Felicitous *some* sentences as well as fillers). They also calculated scalar implicatures in an appropriately pragmatic manner with more than 90% accuracy, when these were supported by adequate context (Experiment 2). They did diverge from native speakers' rates of performance, but only from a subset of native speakers, the "logical" individuals. Therefore, I submit that the L2 performance on scalar implicature computation is a qualified success. It is clearly possible to understand implied meaning in a second language.

The crucial difference in the findings of this study compared to the results on near-native acquisition of null subject usage lies in the effect of the native language. I pointed out above that the scalar implicature calculation mechanism is purportedly universal, so that the either Universal Grammar or the native language aids the learners in their acquisition of this property. This claim was tested and supported by the judgments of the Korean native group, which were in striking accord with the English native group. In comparing null subjects with scalar implicatures at the interface of syntax/semantics and pragmatics/discourse, the following claims have been made by researchers. In the use of null subjects, persistent *L1 influence* on processing strategies may be due to *processing limitations*. In scalar implicature computation, there is *no adverse L1 influence* since the property is universal, but still some effects of *processing limitations* are detectable. The crucial difference between the two learning situations at the syntax–discourse interface (null subjects and scalar implicatures) is the effect of the native grammar, or the availability of L1 transfer in the first learning situation but not in the second. Thus it is clear that processing limitations are manifested *even in the absence* of L1 transfer. In addition, processing limitations are felt more when there is no sufficient context, and under heavier processing demands. In sum, one could argue that processing capacity is the common source of divergence at the syntax–pragmatics interface, which is exacerbated when the native language does not facilitate but impedes the computation.

What exactly are these processing limitations in children and in second language learners, and what are they due to? They may or may not have the same underlying sources in children's and adults' performance. We need more reaction time (RT) and ERP studies of individuals (both L2 learners and natives) to be able to tell exactly how individuals process potential implicatures. One recent study, Feeney et al. (2004) in their Experiment 3 tested 50 English native speakers on items similar to our Experiment 1 items and measured their RTs. They observed that logical responses to infelicitous *some* took significantly longer (1844 ms with a SD of 643) than logical responses to felicitous *some* sentences (1452 ms with a SD of 305). "This finding supports the idea that the logical response to infelicitous *some* is accompanied by cognitive processing over and above that needed for a logical response to felicitous *some*." (Feeney et al., 2004:129). They also found that the tendency to respond logically to infelicitous *some* was positively associated with a measure of cognitive capacity, a counting span task. The correlation between counting span and logical answers

to infelicitous *some* sentences was significant, while the correlations between counting span and the logical answers in the other three conditions were not significant. In combination, these findings suggest that the inhibition of a pragmatic response in favor of a logical one is likely to require superior processing resources.¹³

If this is the case with native speakers' processing of implicatures, what can we conjecture about individual differences in L2 processing? The research on this topic is in its infancy. Preliminary results on individual L2 processing (Juffs, 2004, 2005, 2006) suggest that working memory as measured by the word span and the reading span tests can only account for a fraction of the variation present in experimental data. Undoubtedly, investigating these L2 processing limitations (individual as well as group limitations) and formulating clear proposals about linguistic computation at the interfaces will be among the most exciting topics for research in the next decade.

Future research should address the default-based versus context-based accounts of scalar implicature derivation through reaction times (RTs) within individual speakers or within groups of speakers with similar behavior. As my experimental data show, native speakers as well as L2 speakers divide roughly in two groups: those that interpret underinformative sentences predominantly pragmatically, and those that interpret them predominantly logically. To date, RT studies of implicature derivation group all their participants together in calculating means and executing the statistical analyses. However, it is quite plausible that the time course of implicature calculation is different in the two groups of speakers. One can only claim that either pragmatic readings or logical readings of underinformative utterances take longer than the other if one investigates the processing of the same individual, or of more tightly-controlled groups of individuals who pattern the same way with respect to implicature calculation.

In conclusion, I have argued that, judging by L2 knowledge of scalar implicatures and interpretation of underinformative sentences, linguistic pragmatic competence in this respect does appear to be universal, but may also be dependent on processing resources. I have qualified the claims of the Interface Hypothesis by suggesting that the prime source of difficulty at the interfaces is limited processing capacity. Furthermore, I have argued that L2 acquisition data may constitute an important source of evidence for linguistic theory, capable of informing the theoretical debate on the pathway of scalar implicature computation.

Acknowledgements

The research presented in this article was conducted with the financial and logistical support of the Obermann Center for Interdisciplinary Studies at the University of Iowa. I am grateful to all the participants in the experiment, as well as the following research assistants: Kum Young Lee, Min-A Park and Jae-Eun Kim. Aidan Feeney, Susan Scafton, and Simon Handley generously made their test items and raw data available to me, thank you! Different aspects of this study were presented at the GASLA 6 conference, University of Calgary, Banff, 2006; the International Symposium on Bilingualism in Hamburg, 2007; and the McGill Linguistics Colloquium, 2007. I want to thank the various audiences at these meetings for their invaluable questions and feedback.

¹³ There are a number of recent psycholinguistic studies which trace the time line of implicature calculation on-line, measuring reaction times (RTs). These studies explicitly address the debate between the neo-Gricean (integrational) account and the context-based Relevance Theory account. In many such studies, researchers ask participants to adopt a particular interpretation of underinformative sentences without context and then measure their RTs in reading these sentences. I find this research design problematic because oral instruction as to interpretation does not ensure that individual participants do use this interpretation in their linguistic processing. One such study is Bott and Noveck (2004), particularly their Experiments 1 and 2. Bott and Noveck's third and fourth experiment present their participants with underinformative sentences without giving them an interpretation but ask them to respond to the sentences with True or False. However, underinformative sentences are *always* True, although they may be infelicitous when interpreted logically. Logical thinking and pragmatic felicity are confounded. Thus the fact that the researchers see higher RTs for pragmatic answers may be due to the extra work that their participants have to do to consider logical truth and then felicity. Another recent study is Breheny et al. (2006). Their second experiment manipulates the syntactic position of the *some* phrase, subject versus object, on the assumption that in Greek word order correlates with information structure (theme–rheme; old–new information). That may be true, but there is no independent measure of how much time it takes speakers to process themes and rhemes which are not underinformative but just regular NPs. Thus information structure and implicatures are again confounded, in my opinion. The study I cite here, Feeney et al. (2004), manages to evade such confounds. Thus I feel that the RT studies that offer support to the context-based account of scalar implicature derivation are less than conclusive at this time.

Appendix A. Appendix of test items in Experiment 1

Infelicitous some/True all

All/Some elephants have trunks.
All/Some books have pages.
All/Some hammers have a handle.
All/Some robins have wings.
All/Some cats have ears.
All/Some airplanes have wings.
All/Some giraffes have long necks.
All/Some televisions have screens.

Felicitous some/False all

All/Some books have color pictures.
All/Some clothes have zips.
All/Some dresses have pockets.
All/Some children are blonde.
All/Some flowers are yellow.
All/Some birds live in cages.
All/Some dogs have spots.
All/Some drinks are made of chocolate.

Absurd all/Some

All garages sing.
All crayons have noses.
All chairs tell time.
All birds have telephones.
Some fish are made of leaves.
Some fruits have computers.
Some books are good to eat.
Some children are made of feathers.

References

- Belletti, A., Bennati, E., Sorace, A., 2007. Theoretical and developmental issues in the syntax of subjects: evidence from near-native Italian. *Natural Language and Linguistic Theory* 4, 657–689.
- Bialystok, E., 2001. *Bilingualism in Development: Language, Literacy, and Cognition*. Cambridge University Press, New York.
- Bialystok, E., Martin, M., 2004. Attention and inhibition in bilingual children: evidence from the dimensional change card sort task. *Developmental Science* 7, 325–339.
- Bialystok, E., Senman, L., 2004. Executive processes in appearance reality tasks: the role of inhibition of attention and symbolic representation. *Child Development* 75, 562–579.
- Borer, H., 1984. *Parametric Syntax*. Foris, Dordrecht.
- 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.
- Bouton, L., 1992. The interpretation of implicature in English by NNS: Does it come automatically without being explicitly taught? In: Bouton, Lawrence (Ed.), *Pragmatics and Language Learning Monograph Series*, vol. 3. DEIL, University of Illinois at Urbana-Champaign, pp. 64–77.
- Bouton, L., 1994. Can NNS skill in interpreting implicature in American English be improved through explicit instruction? A pilot study. In: Bouton, Lawrence (Ed.), *Pragmatics and Language Learning Monograph Series*, vol. 5. DEIL, University of Illinois at Urbana-Champaign, pp. 88–109.
- Breheny, R., Katsos, N., Williams, J., 2006. Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition* 100 (3), 434–463.
- Carrell, P., 1981. Relative difficulty of request forms in L1/L2 comprehension. In: Hines, Mary, Rutherford, William (Eds.), *On TESOL'81*. TESOL, Washington, DC, pp. 141–152.
- Carston, R., 1998. Informativeness, relevance and scalar implicature. In: Carston, R., Uchida, S. (Eds.), *Relevance Theory: Applications and Implications*. John Benjamins, Amsterdam, pp. 179–236.

- Chierchia, G., 2004. Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In: Belletti, A. (Ed.), *Structures and Beyond*. Oxford University Press.
- Chierchia, G., 2006. Broaden your views: implicatures of domain-widening and the “logicality” of language. *Linguistic Inquiry* 37 (4), 535–590.
- Chierchia, G., Crain, S., Guasti, M.T., Gualmini, A., Meroni, L., 2001. The acquisition of disjunction: evidence for a grammatical view of scalar implicatures. In: Do, A.H.-J., Dominguez, L., Johansen, A. (Eds.), *Proceedings of the 25th Boston University Conference on Language Development*. Cascadilla Press, Somerville, MA.
- Chomsky, N., 1995. *The Minimalist Program*. MIT Press, Cambridge, MA.
- Clahsen, H., Felser, C., 2006. Grammatical processing in language learners. *Applied Psycho-linguistics* 27, 3–42.
- Crain, S., Steedman, M., 1985. On not being led up the garden path: The use of context by the psychological parser. In: David Dowty, Lauri Karttunen, Arnold Zwicky, (Eds.), *Natural Language Parsing: Psychological, Computational and Theoretical Perspectives*. Cambridge University Press, Cambridge, pp. 94–128.
- Dekydtspotter, L., Hathorn, J., 2005. *Quelque chose . . . de remarquable* in English–French acquisition: mandatory, informationally encapsulated computations in second language interpretation. *Second Language Research* 21 (4), 291–323.
- Dekydtspotter, L., Sprouse, R., Meyer, T., 2005. Was für N interrogatives and quantifier scope In English-German interpretation. In: Dekydtspotter, L., et al. (Eds.), *Proceedings of the 7th Generative Approaches to Second Language Acquisition Conference (GASLA 2004)*. pp. 86–95.
- Feeney, A., Scafton, S., Duckworth, A., Handley, S., 2004. The story of *some*: everyday pragmatic inference by children and adults. *Canadian Journal of Experimental Psychology* 58 (2), 121–132.
- Field, A., 2005. *Discovering Statistics Using SPSS*, second edition. Sage Publications.
- Gazdar, G.A., 1979. *Pragmatics: Implicatures, Presuppositions and Logical Form*. Academic Press, New York, NY.
- Grice, P., 1989. *Studies in the Way of Words*. Harvard University Press, Cambridge, MA.
- 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.
- Gürel, A., 2006. L2 acquisition of pragmatic and syntactic constraints in the use of overt and null subject pronouns. In: Slabakova, R., Montrul, S., Prévost, P. (Eds.), *Inquiries in Linguistic Development*. John Benjamins, Amsterdam, pp. 259–282.
- Horn, L., 1972. *On the semantic properties of the logical operators in English*. Doctoral Dissertation. UCLA.
- Horn, L., 1984. Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In: Schiffrin, D. (Ed.), *Meaning, Form and Use in Context: Linguistic applications (GURT '84)*. Georgetown University Press, Washington, DC, pp. 11–42.
- Juffs, A., 2001. Psycholinguistically-oriented second language research. *Annual Review of Applied Linguistics* 21, 207–223.
- Juffs, A., 2004. Representation, processing and working memory in a second language. *Transactions of the Philological Society* 102, 199–226.
- Juffs, A., 2005. The influence of first language on the processing of wh-movement in English as a second language. *Second Language Research* 21, 121–151.
- Juffs, A., 2006. Processing reduced relatives versus main verb ambiguity in English as a second language: a replication study with working memory. In: Slabakova, R., Montrul, S., Prévost, P. (Eds.), *Inquiries in Linguistic Development*. John Benjamins, Amsterdam, pp. 213–232.
- Kasper, G., 1984. Pragmatic comprehension in learner-native speaker discourse. *Language Learning* 34 (1), 1–20.
- Levinson, S., 2000. *Presumptive Meanings*. MIT Press, Cambridge, MA.
- Lozano, C., 2006. Focus and split intransitivity: the acquisition of word order alternations in non-native Spanish. *Second Language Research* 22, 145–187.
- Musolino, J., Lidz, J., 2002. Preschool logic: truth and felicity in the acquisition of quantification. In: Skarabela, B., Fish, S., Do, A.H.-J. (Eds.), *Proceedings of the 26th Boston University Conference on Language Development*. Cascadilla Press, Somerville, MA, pp. 406–416.
- Noveck, I., 2001. When children are more logical than adults: experimental investigations of scalar implicature. *Cognition* 78, 165–188.
- Noveck, I., Sperber, D., 2004. *Experimental Pragmatics*. Palgrave, Basingstoke.
- Noveck, I., Sperber, D., 2007. The why and how of experimental pragmatics: the case of ‘scalar inferences’. In: Burton-Roberts, Noel (Ed.), *Advances in Pragmatics*. Palgrave, Basingstoke, pp. 184–212.
- Papafragou, A., Musolino, J., 2003. Scalar implicatures: experiments at the syntax semantics interface. *Cognition* 86, 253–282.
- Pouscoulous, N., Noveck, I., Politzer, G., Bastide, A., 2007. A developmental investigation of processing costs in implicature production. *Language Acquisition* 14 (4), 347–375.
- Reinhart, T., 2006. *Interface Strategies: Optimal and Costly Computations*. MIT Press, Cambridge, MA.
- Siegal, M., Iozzi, L., Surian, L., 2009. Bilingualism and conversational understanding in young children. *Cognition* 110, 115–122.
- Slabakova, R., 2008. *Meaning in the Second Language*. Mouton de Gruyter, Berlin.
- Sorace, A., 2003. Near-nativeness. In: Long, M., Doughty, C. (Eds.), *Handbook of Second Language Acquisition*. Blackwell, Oxford, pp. 130–151.
- Sorace, A., Filiaci, F., 2006. Anaphora resolution in near-native speakers of Italian. *Second Language Research* 22, 339–368.
- Sperber, D., Wilson, D., 1986/1995. *Relevance: Communication and Cognition*. Oxford Blackwell.
- Storto, G., Tanenhaus, M., 2005. Are scalar implicatures computed online? In: Maier, E., Bary, C., Huitink, J. (Eds.), *Proceedings of Sinn und Bedeutung*, 9, Nijmegen Centre for Semantics, Nijmegen, pp. 431–445.
- Takahashi, S., Roitblat, H., 1994. Comprehension process of second language indirect requests. *Applied Psycholinguistics* 15 (4), 475–506.
- White, Juffs, 1998. Constraints on wh-movement in two different contexts of nonnative language acquisition: Competence and processing. In: Suzanne, Flynn, Gita, Martohardjono, Wayne, O’Neil (Eds.), *The Generative Study of Second Language Acquisition*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 111–129.
- White, L., Bruh-Garavito, J., Kawasaki, T., Pater, J., Prevost, P., 1997. The researcher gave the subject a test about himself; problems of ambiguity and preference in the investigation of reflexive binding. *Language Learning* 47, 145–172.
- Wilson, D., Sperber, D., 2004. *Relevance Theory*. In: Horn, L., Ward, G. (Eds.), *Handbook of Pragmatics*. Blackwell, Oxford, pp. 607–632.