Do adults show an effect of delayed first language acquisition when calculating scalar implicatures?

Kathryn Davidson
Yale University

Rachel I. Mayberry
University of California, San Diego

Contact Information:
Kathryn Davidson
Yale University
Program in Cognitive Science and Department of Linguistics
370 Temple St.
New Haven, CT 06520-8366
Tel: 203-432-2450
Email: kathryn.davidson@yale.edu

Abstract

Language acquisition involves learning not only grammatical rules and a lexicon, but also what someone is intending to convey with their utterance: the semantic/pragmatic component of language. In this paper we separate the contributions of linguistic development and cognitive maturity to the acquisition of the semantic/pragmatic component of language by comparing deaf adults who had either early or late first exposure to their first language (ASL). We focus on the particular type of meaning at the semantic/pragmatic interface called *scalar implicature*, for which preschool-age children typically differ from adults. Children's behavior has been attributed to either their not knowing appropriate linguistic alternatives to consider or to cognitive developmental differences between children and adults. Unlike children, deaf adults with late language exposure are cognitively mature, although they never fully acquire some complex linguistic structures, and thus serve as a test for the role of language in such interpretations. Our results indicate an overall high performance by late learners, especially when implicatures are not based on conventionalized items. However, compared to early language learners, late language learners compute fewer implicatures when conventionalized linguistic alternatives are involved (e.g. <all, some>). We conclude that (i) in general, Gricean pragmatic reasoning does not seem to be impacted by delayed first language acquisition and can account for multiple quantity implicatures, but (ii) the creation of a scale based on lexical items can lead to ease in alternative creation that may be advantageously learned early in life, and that this may be one of several factors contributing to differences between adults and children on scalar implicature tasks.

Key Words: Scalar Implicatures, American Sign Language, Age of Acquisition, Experimental Pragmatics, Late Language Learning

To become fluent in a language, a learner must acquire different levels of linguistic structure: the arbitrary mapping of form to meanings that make up the lexicon, the phonological and morphological processes at work shaping the form of these lexical items, the syntactic rules that combine words into clauses, and the semantic and pragmatic rules for interpreting what someone intends to be convey with their utterance. Decades of research suggest that the learning patterns for each of these levels of structure vary depending on degree of maturation and previous linguistic experience. So, for example, children learning a first language (L1), and older children and adults learning a second language later in life (L2) must both build arbitrary lexical mappings in the target language to learn the meanings of words. On the other hand, phonological and syntactic rules in the target language appear to be more easily learned by young L1 learners compared to older L2 learners, who may be beyond a sensitive period for language learning and have already acquired such rules in their L1 (Johnson and Newport 1991, Flege, Yeni-Komshian, & Liu 1999, Mayberry, Lock, & Kazmi, 2002. but see Birdsong & Molis, 2001). At the level of semantic/pragmatic acquisition the reverse pattern is sometimes observed: adult L2 learners succeed where childhood L1 learners have more difficulty (Newport 2002, Siegal and Surian 2004, Slabakova 2009). In this case, the older L2 advantage may arise from having a previously built conceptual framework in the L1 that can be transferred to meanings in the L2. The question we investigate here is whether this advantage also occurs for individuals who were born deaf and have had incomplete access to language during early childhood. This unique population can help dissociate the effects of cognitive maturity from those of language development when it comes to acquiring semantic/pragmatic knowledge.

Because children are developing both cognitively and linguistically, it can be difficult to determine for many semantic/pragmatic phenomena whether conceptual development and world knowledge aid linguistic development, or vice versa. For example, when children misinterpret a sentence with a quantifier like *some*, it could be the case that they lack language or world knowledge, such as the abstract and complex meaning of *some*, required to determine the intended meaning of utterances. It could also be the case that they have acquired the linguistic knowledge but are still developing other non-linguistic cognitive skills that are required for processing or reporting language meaning. Unlike children, L2 learners have adult-like cognitive skills and are still learning the L2, but they present a confound in teasing apart the conceptual and linguistic contributions to semantic and pragmatic acquisition because they may be able to transfer to their L2 pre-existing knowledge from their L1 about the meaning of numbers, quantifiers, and other structures that are similar cross-linguistically between their languages. Thus, understanding how learners initially create the conceptual and linguistic structures that provide language meanings remains a difficult problem.

In the present study, we investigate the role of cognitive and linguistic contributions to semantic/pragmatic acquisition by studying individuals whose linguistic and life experiences are distinct: delayed L1 learners. Our focus is on deaf adults who were not exposed to a sign language from their hearing parents, but consider American Sign Language (ASL) their first language when they did learn it later in life. This group of later L1 language learners seems to pattern like childhood L1 learners in some areas of language acquisition, while in other areas they behave quite differently from child L1 learners. For example, early lexical items appear to be acquired in a similar fashion to first language learning even in severely delayed first language learners (Ferjan Ramirez et al. 2013). Although little is known about the syntactic development of late L1 leaners, they appear to process grammatical rules less well compared to native signers after many years of experience, including those who have had language exposure beginning at age 5 (Mayberry & Lock, 2003, Mayberry et al. 2002, Boudreault & Mayberry 2006). To our knowledge, there is little research on the effects of delayed L1 acquisition on computation at the interface of compositional semantic and pragmatics. Not only can such an investigation provide a more complete picture of delayed first language acquisition, but it also has the potential to clarify the role of language input in the acquisition of meaning in more typical language development.

To investigate this question, we focus on a single, well-studied phenomenon at the semantic/pragmatics interface: *scalar implicature*, the name given to the phenomenon in which most adults will take (1b) to be true if they hear (1a).

(1) a. **Some** teas contain caffeine.

b. **Not all** teas contain caffeine.

Although people often assume (1b) when someone says (1a), (1b) is not *entailed* by (1a): “some” doesn't necessarily exclude “all” in the same way that “some” excludes “none”. This point is illustrated in (2a) and (2b), where it is possible to add further clarification to a sentence with “some” to indicate that “all” is true (2a). Attempting to do the same for other quantifiers, like “none” in (2b), leads to an absurd/infelicitous statement (notated with #). This ability to cancel the inference in (1) (i.e. that "some" implies *not all*) in the face of further information is one trademark of the type of pragmatic inference known as *implicature*, in contrast to a non-cancelable *entailment*.

(2) a. Some teas contain caffeine- in fact, all do!

b. #Some teas contain caffeine- in fact, none do!

Grice (1989) proposed that the basic meaning of a statement like (1a) is something like *At least some* teas contain caffeine; that is, (1a) is logically true in any situation where at least some tea contains caffeine. This is a rather weak existential statement: the only possibility that it rules out is that no teas contain caffeine. Under this view, the stronger meaning (that *some, but not all*, teas contain caffeine) comes about through pragmatic reasoning: the listener considers that if *All teas contain caffeine* were true, then the speaker should have said that, because it would have been just as easy to say as (1a) but it would have been a better description of the facts. Since the speaker didn't choose to say that, and participants in a conversation generally give the most informative description that they can easily provide (Grice’s “Maxim of Quantity”), then it must be false. Hence, the listener concludes (1b). Further work by Gazdar (1979) and Horn (1989) on this type of inference pattern highlighted the role of a conventionalized "scale" like <all, some> in the generation of the implicature: some alternative statements appear to have a privileged status, such that the listener very frequently considers substitution of the statement with a stronger value on the same scale as an alternative. For example, one considers *all* as an alternative to *some*, but does not consider *some and not all,* the negation of which would create exactly the wrong implicature (a sentence with *some* would erroneously implicate *all*). These generalizable scalar alternatives (<all, some>) should be ordered (in this case, by entailment), share the same semantic class, be no more difficult to say than the base term (Matsuomo 1995), and be no more syntactically complex (Katzir 2008)(ruling out *some but not all* as an alternative to *some*). In our view, the question of how one actually knows and *learns* the scales has been overshadowed by a larger debate concerning how much of the generation of the implicature is due to the grammar, and how much is part of extra-linguistic reasoning (Levinson 2000, van Rooij & Schulz 2004, Russell 2006, Katsos 2007, Geurts 2010, Chierchia et al. 2011, Sauerland 2012, among many others), although these vary in the extent to which a scale is grammaticalized or generalized.

The psycholinguistic literature discussing the processing of scalar implicatures has also primarily focused on the question of how much of the implicature is conventionalized and how much may happen as part of a grammatical calculation. There is evidence that the computation of implicature is largely an automatic inference but with a processing cost (Noveck & Posada 2003, Bott and Noveck 2004), although as we saw they are also cancelable (as in (2a)) and vary in their strength depending on the context (Breheny et al. 2006). For the purposes of the current paper, it is important to bear in mind that adults typically generate implicature interpretations more often than not, but that this is dependent on many factors and is rarely uniform: few studies find 100% rates of implicature calculation even among native speaking adults.

When it comes to typical first language acquisition, children famously fail to exhibit scalar implicature interpretations to the same extent as adults. This is shown in behavioral tasks, where children accept sentences like *Some animals got a snack* in situations where all got snacks. By contrast, adults and older children will typically reject that description as false or underinformative (Noveck 2001; Papafragou & Musolino 2003; Chierchia et al. 2001; Gualmini, et al. 2001; Huang & Snedeker 2009, among many others). Moreover, there is variation in children’s behavior among different proposed scales: at three years old, children reject "two" as a description for cases where "three" would also be true, while implicatures based on scales like <all, some> and <and, or> are acquired in the late preschool years (Papafragou & Musolino 2003), and implicatures based on modals <must, may> may not be acquired until elementary school (Noveck 2001). There is also some variation reported cross-linguistically in the timing of scalar implicature acquisition, especially when different languages do not have scales that map to the same underlying meanings (Siegal et al. 2007). Despite this variability, the general pattern observed throughout different scales and different languages finds young children behaving as if they had more "logical" interpretations than adults (where, for example, *some* has the simple existential meaning found in first order logic). Only later do children successfully exhibit adult-like behavior consistent with the appropriate scalar implicature (e.g. where "some" means *some and not all*).

Hypotheses vary for why children are more accepting than adults of underinformative descriptions (and thus not exhibiting scalar implicatures). At least three types of hypotheses have been suggested. Gualmini et al. (2001) and Chierchia et al. (2001) show that overtly presenting alternative utterances to children (i.e. "Which is a better way to say it, X or Y?") increases the chance that children reject the underinformative description. They conclude that the scalar implicature presents demands on overall cognitive processing and that, with changes to make the task easier, children can calculate implicatures at adult-like levels, results which are further supported by Foppolo et al. (2012). In contrast, Katsos & Smith (2010) and Katsos & Bishop (2010) argue that children have a different sensitivity to underinformativity, i.e. a higher "pragmatic tolerance" for underinformative descriptions, and so will be less likely to reject what they consider to be an underinformative utterance. They show that when children are given gradient answer options instead of binary "accept/yes" or "reject/no", children will pick a middle value in many cases of scalar implicature. They interpret this to mean that children simply have a lower threshold for acceptance than adults. In yet another view, Barner et al. (2010) argue that the difference between adults and children is that adults have had more experience with the language, and so know what words are scalar alternatives of each other, while young children have not yet acquired the words and their orderings particular to each scale. They show that children have a higher pragmatically appropriate rejection rate of underinformative sentences using *ad hoc* scales like (3), which share all of the required cognitive and behavioral requirements of a scalar implicature but are not based on a context generalizable lexical scale (see also Papafragou & Tantalou 2004) and for which the alternatives are not equally complex (e.g. <(tea, hot chocolate, and Coke), (tea, hot chocolate, Coke, and coffee)>.)

(3) a. Mary drank **tea, hot chocolate, and Coke**.
b. Mary did **not** drink **coffee**.
(where tea, hot chocolate, Coke, and coffee are the relevant drinks)

Clearly, such inferences naturally involve a combination of domain-specific linguistic knowledge (what negation and conjunction mean in (3), for example, and what quantifiers mean in (2)) and domain-general cognitive reasoning (that interlocutors usually try to be as informative as possible and what it means to be informative in a given context). However, because children are both socially and cognitively less mature than adults, and also have less language experience, studies in first language development remain divided on the issue. We propose that by studying scalar implicatures in deaf signers with delayed language exposure, we can begin to separate these factors and thereby illuminate the development of implicatures in typical children.

The second language (L2) acquisition literature could potentially provide this kind of information. Unfortunately, the literature on L2 scalar implicatures is more sparse than the first language (L1) literature, but existing studies suggest that adult second language learners have no difficulty with scalar implicature. This might suggest that cognitive maturity, and not linguistic development, is the key to success. However, there is a significant problem with taking this at face value: most scales used in scalar implicature studies have similar conceptual structure across languages. Common examples of scales are numbers (<four, three, two>), coordinators <and, or>, modals <must, may>, and quantifiers <all, some>. Slabakova (2010) reports that given an appropriate context, Korean second language learners of English are able to appropriately reject underinformative sentences to the same extent as native speakers. However, because of the similarities in scales across languages, it is possible that they are able to transfer knowledge from their L1 to succeed on the task, and so may not actually be computing the implicature based on the L2 (English) scale. There are some cases when there is not a complete overlap between languages in semantic scales, such as Japanese, which lacks a scale equivalent to English <the, a> because it does not make a distinction between definite and indefinite determiners. In the case of L2 Japanese learners of English, L2 learners have difficulty in the early stages of acquisition, and at first incorrectly transfer knowledge about scales in their L1 (Japanese) to their L2 (English), but advanced learners do not differ in proficiency from native speakers (Gruter et al. 2008). Thus, although second language learners generally do well on tasks that require scalar implicatures in their L2, and have less experience with their L2 relative to adults in their L1, they do not make the best test case for determining the contributions of general cognition vs. linguistic-specific knowledge to the acquisition of scalar implicatures because these inferences may involve a large degree of cross-linguistic transfer.

We propose that the variation in age of L1 acquisition among deaf adults provides a unique population for testing the relative contribution of mature reasoning vs. linguistic development in successful scalar implicature calculation. Davidson (2011) assessed scalar implicature calculation in American Sign Language by testing Deaf adults who grew up with Deaf, signing parents, and thus were native users of ASL. When presented with a prototypical scale like <all, some> in ASL, native signers calculated scalar implicatures to the same extent as did adult native speakers on the same scale in English (77% vs. 84%, in English and ASL respectively, which were not significantly different). By using the ASL signing population, we can thus compare the baseline performance of adult, Deaf native signers to Deaf adults who learned ASL as their first language later in life because they did not have signing family members and were unable to acquire English in a native-like fashion. This provides a means to hold cognitive maturity constant and vary linguistic experience. Within the population of deaf adults in the United States, in fact a large proportion of ASL users learn ASL as a first language later in life because the majority of deaf children are born to hearing parents who do not sign (exact numbers are difficult to determine, see Mitchell & Karchmer 2004). Because these individuals' family members did not use a sign language, and English was not fully accessible to them, these individuals were unable to acquire the language of their caregivers from birth, and so lacked a typical native language acquisition (Mayberry et al 2002). In the present study, we included signers who were exposed to ASL at various ages from birth (native signer) to 18 years old (late learner). Because the dominant culture in the USA uses English, all were exposed to a mixture of English and ASL throughout their education, so for all of the deaf participants in this study there was an ever-present semi-bilingualism, with English as a read language and ASL as a primary language. However, in the present study all participants considered themselves to be more dominant in ASL than in English and all participants considered *ASL to be their first language* (as reported in a background questionnaire). Therefore, in the case of deaf participants with hearing parents who learn ASL at a late age, there was not a first language to transfer to the language we tested (ASL *was* the L1), leading to a decreased possibility of transfer compared to L2 learners.

Because neither English nor ASL is learned in a typical, native-like manner, many non-native signers have been shown to have deficits as adults compared to native signers in processing complex linguistic structures in ASL, even after many years of using ASL (Boudreault & Mayberry, 2006; Mayberry et al. 2002). Unlike typical L1 acquisition, late L1 learners do not always proceed through all of the final stages of typical development, but are often permanently lacking the full extent of native language competency. As noted by Siegal & Surian (2004), this means that deaf, late L1 learners provide a potential test case for the role of language-specific knowledge in semantic/pragmatic phenomena because they have a decreased possibility for transfer from a native language due to incompletely developed language skills, without the extreme social neglect that accompanies cases of delayed L1 development in typically hearing children. In the present study, we extend our understanding of the effects of delayed L1 exposure to an unexplored level of linguistic structure: the semantic/pragmatic domain.

As described above, typical L1 learners exhibit pragmatically appropriate scalar implicature behavior at different ages depending on the scale involved, so in the present study we tested three different potential scales. The first is based on the quantifiers <all, some>, a prototypical scale in both English and ASL, described below in Section 1.1. To confirm that signers are not just translating from their knowledge of English throughout this ASL task, we also include an *ad hoc* scale that involves spatial language specific to language in the manual/visual modality, described in more detail in Section 1.2. A second scale that is prototypical in English is the coordination scale <and, or>, described in Section 1.3. Davidson (2013) has shown that conjunction ("and") and disjunction ("or") in ASL need not be expressed by two separate lexical items, unlike in English, and consequently sentences with disjunction trigger far fewer implicatures in ASL than does the corresponding scale in English. Nonnative signers' behavior on this scale, relative to the quantifier scale, can help determine exactly what aspects of scalar implicatures, if any, are difficult for late learners. In particular, the coordination scale can illuminate the role of lexical contrast. If nonnative signers struggle with the quantification scale but not the coordination scale, then part of the reason may be incomplete acquisition of the lexical scale structure (which is present in the quantifier scale, but not in the coordination scale). In addition, the coordination scale provides a means to test the extent of signers' translation into English, in which far more implicatures are computed for the coordination scale. Finally, we also include a language task to test signers' proficiency on difficult but unrelated aspects of ASL using a fourth sentence type, relative clauses, discussed in Section 1.4.

In Section 1 below we present details about each of these sentence types. Then, in Section 2 we describe the participants and the experimental procedure, which is a modified version of a Truth Value Judgment paradigm for scalar implicature (Noveck 2001, Papafragou and Musolino 2003), with the one difference that we use video stimuli. Section 3 presents our results, which, to foreshadow, suggest that native and non-native signers may differ on their calculation of a prototypical, lexically-based scale, but not in general on other quantity implicatures. Finally, in Section 4 we discuss our findings.

1. Sentence Types

1.1 Quantifiers

Davidson (2011) showed that a scale based on quantifiers (<ALL, SOME[1]>) in ASL generates implicatures for native signers that are not significantly different from scalar implicatures based on the <all, some> scale in English. Therefore, this quantifier scale was used as an example of a prototypical scale for the current experiment, as used in sentences of the form CANS, ALL RED `All of the cans are red' or CANS, SOME RED `Some of the cans are red.' The version of the quantifiers SOME and ALL that were used are shown in Figure 1. ASL has another sign for ALL that involves two "B" handshapes which does not seem to behave differently from the sign used here. Both quantifiers appear in sentences with topicalized restrictors (see Appendix for transcriptions), an especially natural structure in ASL, where word order is frequently influenced by information structure.

a. SOME b. ALL

Figure 1. SOME and ALL used in the Quantifier sentence type.

The interpretation of these quantifier scale sentences by late-learning ASL signers forms the crucial test for the role of language in scalar implicature acquisition: if a lack of cognitive maturity is the main reason that children fail at scalar implicatures, then we would expect all adult signers to perform similarly on computing scalar implicatures based on this scale. If, however, children's behavior is related to their incomplete acquisition of the linguistic relationship between words like "some" and "all" in English, and if this specific type of relationship (a lexical scale) is beneficially learned early in life,

[1] Signs in ASL are transcribed in capital letters with their closest English translation

then adult signers may vary in their calculation of these prototypical scalar implicatures in ASL based on whether they were exposed to ASL early in childhood or later.

1.2 Spatial Ad hoc

To contrast with the prototypical quantifier scale, we also included a sentence type based on an ad hoc scale, similar to example (3) above, except that all items in the context were of the same type (e.g. all bears instead of tea and hot chocolate, etc.). We included this *spatial ad hoc* scale in the current experiment because (a) it includes a spatial component that cannot easily be translated from English, and (b) it does not require knowledge of a conventionalized scalar relationship between two specific words like "some" and "all", but does require the remaining steps in the implicature calculation process (Barner et al. 2010, Stiller et al. 2011). These non-scale-based steps include the ability to evaluate the given description relative to other possible descriptions and to decide whether to accept or reject an underinformative description. If both early and late language-learning deaf adults perform well on this task, this would mean that domain general cognitive mechanisms are not an impediment to calculating scalar implicatures for late signers.

In addition to not relying on a generalized scale, the ad hoc scale in ASL used here includes spatial information via the locational classifier construction system. Davidson (2011) found that native signers more frequently rejected underinformative ad hoc scales in ASL than did native speakers in English, who were more flexible in their ad hoc interpretations of the analogous sentences in English. The difference between the trial types in the two languages in that experiment was that the scale used in the ASL sentences makes use of the visual/manual modality to include information about the spatial layout of the scene. For example, to describe the configuration of two forks on a table, an English speaker can combine gesture and words to say "There is a fork here [gesture to location X] and here [gesture to location Y]". In ASL, this locational information is incorporated directly into the grammar, as in the sentence shown in (4).

____br

(4) BEAR CL:5(claw)$_X$, CL:5(claw)$_Y$, CL:5(claw)$_Z$.

"There is a (toy) bear here (at location X) and here (at Y) and here (at Z)."

In (4), the signer begins with a topicalized noun, the sign for the noun "bear", which is accompanied by eye-brow raising that marks topicalization (a type of *nonmanual marking*, which is roughly analogous to spoken language intonation and includes movements of the face and body), notated by the "br" for "brow raising" with the line above the sign indicating the duration of brow raising (see the first frames of Fig. 2a-b). She then uses the classifier in ASL appropriate to the physical size and shape of the toy bears, which is a clawed 5 handshape. The classifier is always notated as "CL", and what follows the colon is the shape of the classifier (here, "5(claw)"), following the conventions used in Sandler and Lillo-Martin (2006). Finally, the classifier is established in an area of the signing space, of which there are three in this sentence, marked "X", "Y", and "Z", notated as a subscript following the handshape. Frame-by-frame signs are shown in Figure 2(a) for the sentence in (4), and in Figure 2(b) for the alternative sentence that only mentions two of the toy bears.

____br

a. BEAR CL:5(claw)$_X$, CL:5(claw)$_Y$, CL:5(claw)$_Z$.

`There is a (toy) bear here (at location X) and here (at Y) and here (at Z).'

____br

b. BEAR CL:5(claw)$_Y$, CL:5(claw)$_Z$.

`There is a (toy) bear here (at location X) and here (at Y).'

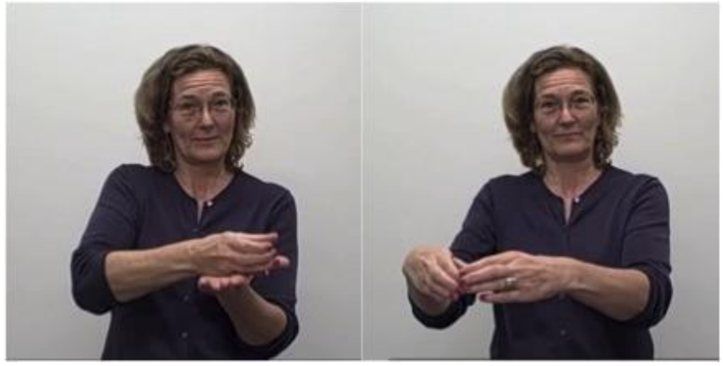
Figure 2: 3 item and 2 item trials used in the Spatial sentence type.

Through these spatial ad hoc scales we can test not only signers' interpretations of ad hoc scales to see if they have the correct tools for implicature calculation outside of a lexical contrast, but we can also be more certain that participants are not directly transferring linguistic information from English, because the spatial information included in these sentences using classifiers is not found in English.

1.3 Coordination

We also investigated signers' interpretations of a third set of sentences that were based on coordination, specifically logical conjunction (which is typically conveyed by "and" in English) and disjunction (which is typically conveyed by "or" in English). In ASL, a frequently utilized strategy for conveying these concepts is general use coordination (notated COORD) which can be interpreted either conjunctively or disjunctively depending on the context and other linguistic cues like nonmanual marking (see Davidson 2013 for more on general use coordination in ASL and its relationship to other forms of coordination in the language). For example, MARY DRINK TEA COORD COFFEE can either be interpreted as 'Mary drinks tea and coffee' or 'Mary drinks tea or coffee.' We included these coordinators as another sentence type because they do not involve lexical contrast in ASL (similar to both English and ASL ad hoc scales), but unlike ad hoc scales, the meanings of conjunction and disjunction have a stable, context-independent logical relationship.

There are two types of general use coordinators in ASL, but we chose to focus on the coordinator we notate as "COORD-shift" because it was reported to be the most natural for connecting the simple sentences in the current experiment. Though notated as if it is a manual sign, "COORD-shift" simply marks the timing of a shift in location from one area in space to another, so that each of the coordinates is placed in a different location in space. It is more than simple juxtaposition, because without the change in location the coordination in ungrammatical (for more information, see Davidson 2013). Nonmanual markings can be one method of distinguishing the disjunctive and conjunctive interpretations of COORD-shift. Conjunctive interpretations (*COORD-shift(and)*) can be signaled with nonmanual marking that includes head nodding at each coordinate, while disjunctive interpretations can be indicated with a squint and a bit lip throughout the coordinate clause. In English, disjunction is well known to generate scalar implicatures, so we might expect that when COORD-shift is interpreted as disjunction in ASL (*COORD-shift(or)*) with the squint and bit lip nonmanual marking, it will do the same. Figure 3 shows nonmanual marking distinguishing the conjunctive (a.) and disjunctive (b.) interpretations of COORD-shift. Davidson (2013) shows that while these nonmanual markings do distinguish the two readings, disjunction via nonmanual marking is much less likely to be interpreted along with a scalar implicature than is the lexical item "or" in English. This difference is taken to be due to the lack of lexical contrast in the conjunctive and disjunctive forms of COORD-shift. Thus, importantly, the coordination scale serves as yet another case in which native-like target responses by late L1 learners cannot be due to transfer from another language like English, because the same scale in English involves a high rate of implicature calculation, unlike ASL.



a. CUP COORD-shift(and) BOWL

`cup and bowl'

b. CUP COORD-shift(or) BOWL

'cup or bowl'

Figure 3: COORD-shift(and) and COORD-shift(or) used in the Coordination sentence type.

If implicatures based on COORD-shift(or) by late learners of ASL pattern with their calculation of implicatures based on quantifiers, and both are below the performance of early signers, this would suggest that late learning signers may not have completely acquired the generalizable, context-independent semantic relationships between scalar items. If, however, later learning-ASL signers pattern similarly on COORD-shift and the spatial ad hoc scale, and only show a difference from early signers on the quantifier scale, this would indicate that the contrast in *lexical* forms is a more crucial distinction.

1.4 Relative Clauses

Signers were assessed for their interpretations of a fourth sentence type which was not a test of implicature calculation, but was a potential challenge for a different reason: signers must use nonmanual marking to disambiguate two strings of signs which are exactly the same but differ in meaning based on nonmanual marking. Consider Figure 4, where the string of signs BEAR HUG DOG HAVE EARPHONES can mean ‘The bear who was hugging the dog had headphones’ with the nonmanual marking seen in (a), which marks BEAR as the subject and HUG DOG as a subject relative clause. Alternatively, the same string can mean ‘The bear was hugging the dog who had the headphones’, with the nonmanual marking seen in (b), indicated an object relative clause. Because of the linguistic subtlety required to detect the difference in meaning, we included these sentences in the current experiment as a control to create an independent measure of the ASL proficiency of the participants, and especially their proficiency with nonmanual marking. It is also directly relevant to the coordination trials, because nonmanual marking is distinguishing the conjunctive and disjunctive interpretations of COORD-shift.

______________br

a. BEAR HUG DOG HAVE EARPHONES.

‘The bear who is hugging the dog has headphones.’

b. BEAR HUG DOG HAVE HEADPHONES.

'The bear is hugging the dog who has headphones.'

Figure 4: The Relative Clause sentence type

To summarize our research design, if scalar implicatures are based on linguistic knowledge of scales that, like other linguistic structures, are difficult to learn without being able to map onto a native L1, then we expect to see a difference between early learning and later learning deaf signers on the quantifier scale. If, on the other hand, scalar implicatures are based primarily on world knowledge and cognitive weighing of the meanings of alternatives, we should see no differences among signers on any of the scales, including quantifiers. Both the spatial and coordination sentences require knowledge that cannot be transferred from English to ASL, because it is unique to ASL, and require all of the steps of scalar implicature calculation *except* knowledge of a lexical scale, because ASL doesn't express the difference between these contrasting meanings through the lexicon. If ordered, lexical contrast is the crucial information to acquire early, then we expect to see no difference between participants on these spatial and coordination scales in ASL. Finally, signers' performance on the relative clause scale allows us an independent measure of their receptive nonmanual competency.

2. Methods

2.1 Participants

Participants were twenty-three adults from the San Diego area who use American Sign Language and self-identified as deaf. All were unable to hear speech, and all used ASL in their home, at work, or both. Participants were recruited directly through email requests from a laboratory database of interested participants or indirectly through recommendations by friends, and were reimbursed with cash or gift cards. Their chronological ages ranged from 19 to 59. Age of ASL acquisition ranged from birth (native signers) to 18. Visual inspection of chronological age plotted with age of ASL acquisition in Figure 5 shows two clearly separate groups: "Native" early signers who learned ASL from birth (Age of Acquisition, AoA, = 0, n=10) and "Late Nonnative" signers who learned ASL in their late teenage years (AoA = 16-19, n=7). The remaining six "Early Nonnative" signers had AoA ranging from 2 to 11. In what follows we will analyze the participants' data in two motivated groupings: Native versus Nonnative (early and late) signers, and Early (native and nonnative) versus Late signers.

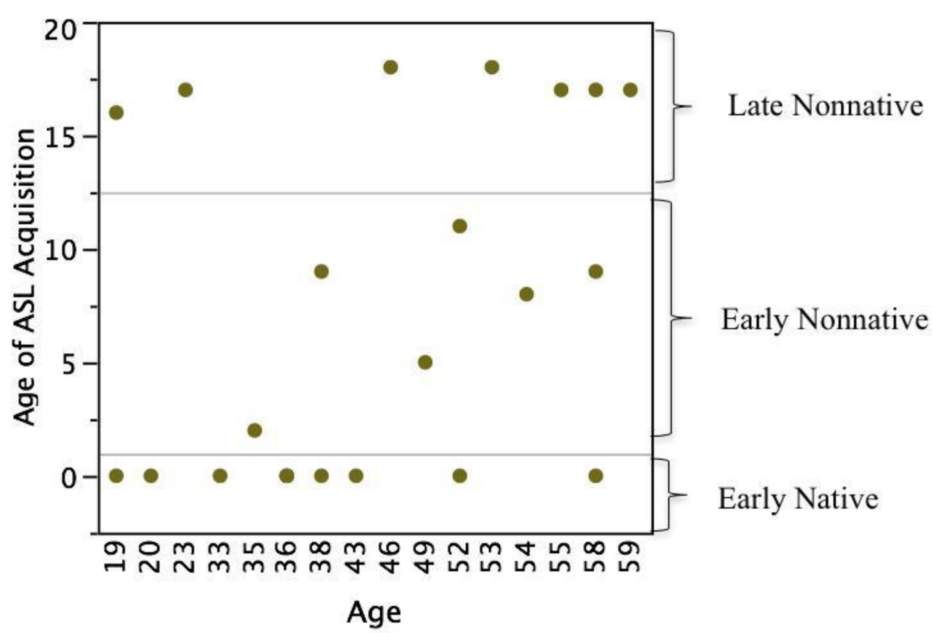


Figure 5: Participants' ages of ASL acquisition relative to (ordered) age at testing.

Despite having learned ASL at varying ages and having a wide range of chronological ages at time of testing, all participants had extensive experience with ASL. The participant with the least ASL experience was 19 years old and learned ASL at age 16, while the next least experienced signer was 23 and learned ASL at age 17. Means for age, age of acquisition, and years of ASL are shown in Table 1.

Table 1: Participant groups and mean (standard deviation) age in years, AoA in years, and years of ASL experience

	Female / n	**Age**	**AoA**	**Years of ASL**
Native Early ASL	7 / 10	37.1 (12.2)	0	37.1 (12.2)
Nonnative Early ASL	2 / 6	47.6 (9.2)	7.3 (3.3)	40.3 (7.8)
Nonnative Late ASL	4 / 7	44.7 (16.8)	17.1 (0.7)	27.6 (16.4)

2.2 Procedures and Stimuli

Each testing session lasted 30-35 minutes. Participants were tested either in a laboratory at UCSD or at various meeting places throughout San Diego county. Both the instructions and the task itself were presented in ASL via a laptop in video format by a native signer who was deaf. Participants were instructed in ASL that, for each trial of the experiment, a picture would appear on the screen and that after they had viewed the picture, they should press the Space Bar key and a video description would begin to play next to the picture. Participants were instructed to press the smile face (the "1" key covered with a smile face sticker, directly below the picture of a smile face on screen) if they were "satisfied" that the description "matches" the picture. If they were "not satisfied", and thought that the description "does not match" the picture, they were instructed to press the frown face (the "0" key covered with a frown, directly below the picture of the frowning face on screen) (Figure 6). It was not possible to replay a video after it had begun playing.

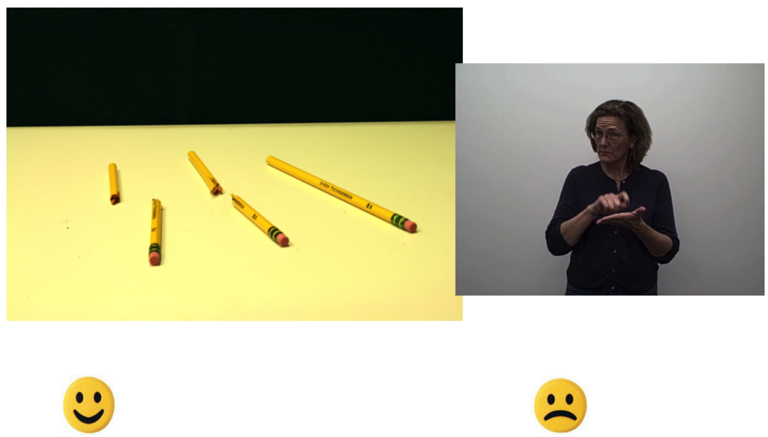


Figure 6: Screenshot during a Quantifier experimental trial. The picture appeared on the left, and the signed description on the right after the Spacebar is pressed.

Participants saw three practice trials to acquaint them with the task: (1) a picture of a red bowl, and a video description THAT BOWL, RED "the/that bowl is red"; (2) a picture of a white shoe, and a video description THAT SHOE, BLACK "the/that shoe is black"; (3) a picture of a wooden spoon, and a video description THAT SPOON, WOODEN "the/that spoon is wooden." Participants then had an opportunity to ask questions if anything about the task or playing the videos was unclear. Practice trials were followed by further instructions, and a confirmation that the task was understood. Finally, 48 trials were presented, consisting of 36 trials of "scale" sentence types: (a) quantifiers (12 total), (b) coordination (12 total), (c) spatial ad hoc (12 total), distributed equally among three Trial Types (Match, Mismatch, and Test), and 12 trials of (d) relative clause structures (6 Match and 6 Mismatch). Responses were recorded using Psyscope software on a 13-inch Macbook laptop. Trials were counterbalanced with a Latin Square design.

In *quantifier trials*, each picture consisted of a set of three objects of which either some of the objects or all of them fulfilled a characterization about that object (e.g. red cans, lit candles, full glasses, etc.). The entire list is given in the Appendix. Under the **Match** condition, the characterization applied to all of the objects (e.g. three cans, all red), and the description was accurate (e.g. CANS, ALL RED 'All of the cans are red.'). Under the **Mismatch** condition, the characterization applied to only two of the objects (e.g. three cans, only two are red), and the description was not accurate (e.g. CANS, ALL RED 'All of the cans are red.'). Finally, under the **Test** condition, the characterization applied to all of the objects (e.g. three cans, all red), and the description was not maximally informative (e.g. CANS, SOME RED 'Some of the cans are red.'). Trials for quantifiers, spatial, and coordination sentence types, described below, were counterbalanced so that each sentence frame (e.g. red cans) appeared in only one trial type (Match, Test, Mismatch) for each participant, and each third of participants saw the sentence frame in a different trial type.

In each of the *spatial ad hoc trials*, the picture showed either three or two objects. In the **Match** condition, there were three objects laid out on the table, and the description was accurate both in number and in location of the items (e.g. BEAR CL:5(claw)$_X$, CL:5(claw)$_Y$, CL:5(claw)$_Z$. 'There are three bears.'). In the **Mismatch** condition, the picture showed two objects, but the description was inaccurate and said there were three items (e.g. BEAR CL:5(claw)$_X$, CL:5(claw)$_Y$, CL:5(claw)$_Z$. 'There are three bears.'). Finally, under the **Test** condition, the picture showed three objects, except that the description was not maximally informative, mentioning only two of the objects (e.g. BEAR CL:5(claw)$_X$, CL:5(claw) $_Z$. 'There are two bears.'). In each description, the locations of all items were accurate.

In the *coordination trials*, the picture consisted of two different objects (e.g. a mug and a bowl), and then either one or two of the same type of object (e.g. spoons) in some visible relationship (e.g. containment, support, etc.) to the first objects. In the **Match** condition in these trials, each of the two different objects were related to one of the similar objects, and the description was accurate (e.g. HAVE SPOON IN CUP COORD-shift(and) SPOON IN BOWL. 'A spoon is in the mug and a spoon is in the bowl.’). Under the **Mismatch** condition, only one of the two different objects were related to one of the similar objects, but the description said that they both were equally related (e.g. HAVE SPOON IN CUP COORD-shift(and) SPOON IN BOWL. ‘A spoon is in the mug and a spoon is in the bowl.’). Under the **Test** condition, each of the two different objects were related to one of the similar objects, but the description was not maximally informative due to the disjunctive nonmanual marking on the general use coordinator (e.g. HAVE SPOON IN CUP COORD-shift(or) SPOON IN BOWL. ‘A spoon is in the mug or a spoon is in the bowl.’).

In the *relative clause* sentence types, each picture consisted of two characters (e.g. a dog and a bear), a relationship between the characters (e.g. hugging) and one object (e.g. headphones). In the **Match** condition, the signed description involves a subject relative clause (e.g. [BEAR HUG DOG] HAVE EARPHONES), which is always a true description of the situation shown in the picture (e.g. the bear is hugging the dog and has the headphones). In the **Mismatch** condition, the signed description is an object relative clause (e.g. BEAR HUG [DOG HAVE HEADPHONES]) but the picture is the same as in the Match condition, so the description is false. Pictures were created in which the Mismatch sentence was true (e.g. the bear is hugging the dog and the dog has the headphones) and these were used to elicit the signed sentences for the Mismatch condition, but these pictures were never shown to participants. Participants each saw 6 Mismatch trials and 6 Match trials of the Relative Clause sentence type. Trials were counterbalanced so that each set of characters/relationship/object appeared in either the Match or Mismatch condition for each participant, and each half of participants saw them in opposite conditions.

3. Results

In Section 1.4, we described the relative clause sentence trials, which were intended as a separate measure of language skills in late learning signers, especially their sensitivity to meaningful non-manual distinctions. The mean score on this task for early signers was 0.59 (SD= 0.10), suggesting that this was a difficult task. However, the mean score for the late learners was 0.58 (SD= 0.14), which was not significantly different than early signers ($t(8.7)=0.09$, $p=0.92$), and the performance of all signers as a group was significantly above chance ($t(22)=3.48$ $p < 0.01$). We take this to indicate that our late learning signers were not differentially sensitive to non-manual prosody in an experimental setting, which in particular is required for the coordination sentence types described in Section 1.3, although the task was more difficult than originally intended.

Scalar implicature calculation was analyzed using mixed logit models, which are well-suited for modeling data like the binary felicity judgments that the participants made in this study. Such models allow us to generalize beyond the specific conditions of the current experiment by including subjects and items as random factors in our model (Baayen 2007, Jaeger 2008), and unlike ANOVAs or t-tests, these models do not require normally distributed data. They also provide information about effect size for each factor in the magnitude of the coefficients β, eliminating the need for additional post-hoc tests.

Each participant's responses for each sentence type described above in Sections 1.1-1.3 (quantifiers, spatial ad hoc, and coordination) and trial type (Match, Mismatch, Test) were analyzed, with language background (Early or Late ASL exposure) fixed for each participant. Rejection (i.e., calculating the scalar implicature and rejecting the infelicitous description) was taken to be the "accurate" response for the Test trial types. Of course, acceptance was the correct Match response, while rejection was the accurate Mismatch response. First, to situate our test trials, we determine overall accuracy in control trials (entailment) vs. test trials (implicatures) among the three scales, which we present first. Then, to test for scalar implicature calculation, acceptances of the logically true but pragmatically infelicitous Test trials and the logically true Match trials were compared across sentence types. Finally, we directly address our question of how age of acquisition affects interpretation of various sentence types, building a model of all Test trials based on language background and sentence type to determine if there is an interaction between scales and language background.

Both types of controls (Match and Mismatch trials) relied on judgments of logical truth: the correct response required the participant simply to compute the truth or falsity of the utterance. By contrast, the Test trials relied on an additional level of pragmatic infelicity: correctly rejecting the description was based on under-informativity, despite being logically true. A mixed logit model of accuracy on all trial types was created with trial type (Match, Mismatch, Test) as a fixed factor and subjects and items as random factors. This model found no significant difference between overall accuracy on Match and Mismatch trials ($\beta = 0.44$, $z = 1.21$, $p = 0.23$), but there was a significant difference between overall accuracy on Test and Match trials, with Test trials significantly lower ($\beta = -1.67$, $z = -5.66$, $p < 0.001$), indicating more variability in responses to stimuli that were pragmatically inappropriate, but true. When a fixed effect of Sentence Type was added to the same model, there was still a main effect of Trial type, with Test trials being less accurate ($\beta = -2.49$, $z = -6.54$, $p < 0.001$), as well as a main effect of Sentence Type, with the spatial ad hoc sentence type being significantly more accurate than other sentence types ($\beta = 2.85$, $z = 2.42$, $p < 0.05$). There was also a significant interaction: quantifiers were less accurate than other scales on the Test trial type ($\beta = 2.29$, $z = 3.67$, $p < 0.001$), a point we return to below. In sum, we found lower accuracy on the Test trials, indicating more flexibility with implicatures compared to entailment.

Next, we investigated the extent of scalar (and ad hoc) implicature interpretations. Recall that the ASL description of a Test trial is logically true (like the Match trial) but pragmatically infelicitous (unlike the Match trial). Therefore, a mixed logit model was created for the acceptances of the responses on Match and Test trial types with Trial Type and Sentence Type as fixed effects and with subjects and items as random effects. This model found a significant main effect of Trial type, with significantly lower acceptances (i.e., more rejections) on the Test trials overall compared to Match trials ($\beta = -1.36$, $z = -3.54$, $p < 0.001$), indicating an overall significant number of pragmatic interpretations. There was also some variation in scales: our model found a significant main effect of Sentence Type, with spatial ad hoc scales accepted more than quantifiers and coordination ($\beta = 2.82$, $z = 2.34$, $p < 0.05$), and a significant interaction, such that quantifier test cases were accepted significantly less often than coordination ($\beta = -3.20$, $z = -5.01$, $p < 0.001$) and spatial ad hoc ($\beta = -8.17$, $z = -4.99$, $p < 0.001$).

Table 2: Mean rejection (=accuracy for Test case) rates (with standard deviations) by Early and Late ASL Learning signers for each sentence type.

		Match	Mismatch	Test
Quantifiers	**Early**	0.08 (0.18)	0.95 (0.10)	0.94 (0.11)
	Late	0.14 (0.20)	1	0.75 (0.29)
Spatial	**Early**	0	1	0.98 (0.06)
	Late	0.04 (0.09)	1	1
Coordination	**Early**	0.16 (0.18)	0.88 (0.16)	0.39 (0.26)
	Late	0.14 (0.13)	0.79 (0.17)	0.32 (0.19)

Recall that our main question is how late learning signers compare to early learning signers when computing implicatures based on different scales (Table 2). To this end, a mixed logit model was created for the accuracy of all Test responses, with Sentence Type and Background Type as fixed effects and subjects and items as random effects. This model found a significant main effect of Sentence Type, with lower accuracy (and hence, fewer implicatures) on the coordination sentence type compared to quantifiers ($\beta = 3.80$, $z = 4.81$, $p < 0.001$) or spatial ad hoc ($\beta = 5.21$, $z = 3.89$, $p < 0.001$). There was no significant main effect of Early vs. Late learners ($\beta = -0.28$, $z = -0.53$, $p = 0.60$). However, there was a marginally significant interaction between Sentence Type and Background Type, such that Late learners were less accurate than Early learners on the Quantifier scale ($\beta = -1.56$, $z = -1.71$, $p = 0.09$), the only scale that required a learned lexical contrast. As shown in Figure 7 there were ceiling effects for the spatial scale.

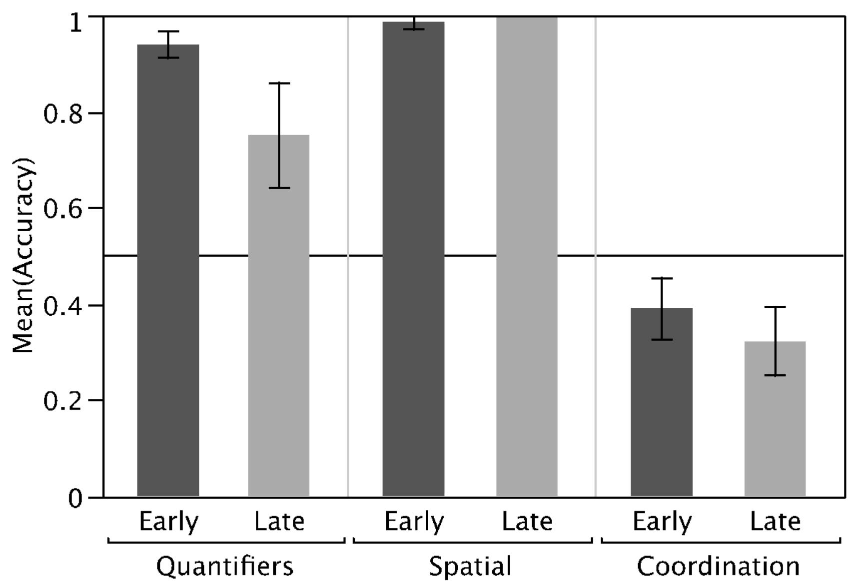


Figure 7: Rejection of underinformative sentences, indicating scalar implicature calculation as a function of scale type and AoA.

To probe the nature of this interaction in more detail, we ran the same model, but used Acquisition Type (Native or Nonnative) instead of Background Type (Early or Native). This analysis grouped early nonnative signers with late learners (all "nonnative signers") instead of with native signers (see Figure 5). This modeled accuracy of Test responses, with Sentence Type and Acquisition Type as fixed effects and subjects and items as random effects, and found the same main effects as the previous model. Quantifiers ($\beta = 3.44$, $z = 4.23$, $p < 0.001$) and spatial ad hoc ($\beta = 4.85$, $z = 3.56$, $p < 0.001$) scales were significantly more accurate than coordination. There was also no main effect of acquisition status ($\beta = 0.06$, $z = 0.12$, $p = 0.91$), and unlike the previous model, this model found no interactions with acquisition status, including with quantifiers ($\beta = -0.57$, $z = -0.66$, $p = 0.51$). Thus the two models yielded contrasting results for acquisition status: in the previous model the signers were categorized as having learned ASL in early childhood but in the second model the signers were categorized as having learned ASL in infancy from deaf parents. The results suggest that AoA effects on quantifier scale calculation may be more related to the age onset of L1 acquisition in childhood than language input from deaf parents.

Finally, in addition to comparing the performance of native versus nonnative signers and early learning versus late learning ASL signers, we also analyzed participants' mean accuracy on scalar implicature calculation in a linear regression model with age of ASL acquisition, age, and relative clause task score as main factors in a full factorial analysis. Although age itself is not a typical predictor of language-related behavior, it was included as a factor in our model because our older signers tended to have later first ages of ASL acquisition and we asked whether this could interact with their performance on the linguistic measures. Results of the analysis indicated that together age of ASL acquisition, age, and relative clause task score did not explain a significant proportion of variance in accuracy scores for either the spatial sentence type, $R^2 = .10$, $F(7, 15) = 0.25$, $p > .1$, or the coordination sentence type, $R^2 = .23$, $F(7, 15) = 0.64$, $p > .1$, and no individual factor significantly predicted accuracy scores for either sentence type. By contrast, and consistent with the results of the first logit model above, the results of the linear regression for the quantifier sentence type did explain a significant proportion of variance for the accuracy scores, $R^2 = .76$, $F(7, 15) = 6.78$, $p < 0.001$. There was a significant main effect of both age of acquisition (slope = -0.012, $t(22) = -3.45$, $p < 0.01$) and age (slope = 0.008, $t(22)=3.79$, $p < 0.01$) and a significant interaction of age and age of acquisition (slope = 0.001, $t(22)=2.69$, $p < 0.05$), indicating that while older signers calculated more implicatures than younger signers, this was mainly true of early learning signers, and overall earlier learning signers interpreted descriptions with more scalar implicatures than later learning signers.

4. Discussion

The goal of this study was to investigate language acquisition factors involved in scalar implicature calculation by looking closely at the pragmatic skills of a set of adults: deaf signers of American Sign Language who were highly experienced with the language but first exposed to a sign language at varying ages. Late exposed signers are cognitively mature adults with adult-like world knowledge, but previous research indicates that, as a group, they have reduced language proficiency due to acquiring their first language(s) later in life. Thus, they can help disentangle the contributions of linguistic and general cognitive/world knowledge factors in the failure of typically developing children to provide adult-like pragmatic judgments for scalar implicature. In the course of this investigation, we also studied a new aspect of the language of late L1 learners: the semantic/pragmatic interface. Our results suggest overall high performance by all non-native signers, indistinguishable from native signers on two sentences types and the ASL relative clause task. Nonnative signers are not a homogenous group: half of our participants learned ASL quite late (either in high school or college), while the other half learned ASL as young children. A re-classification of participants into early and late childhood ASL acquisition groups, as well as a linear regression model, revealed a possible age of ASL acquisition effect for one sentence type, one based on the prototypical lexically-based scale <all, some>.

When we consider the overall performance of delayed L1 learners on this semantic/pragmatic study, there were high rates of success and few differences among deaf signers of varying ages of ASL acquisition. This is surprising in light of other research showing late learners' decreased performance on *syntactic* grammaticality judgments (Boudreault & Mayberry 2006), and suggests that many of the pragmatic reasoning steps involved in quantity implicature tasks can be accounted for by general cognitive principles and/or world knowledge. Delayed L1 learner's high performance is consistent with an overall Gricean (1989) account of pragmatic abilities, which are not posited to be language-specific, and should be independent from the learning of any language. In particular, the adults in this study had a large amount of world knowledge (far more than 3-year-olds typically tested in acquisition studies) and so their success with the quantity implicatures in this study is consistent with a large (if not complete) role of world and contextual knowledge leading to scalar implicature interpretations. On the other hand, our trial sentences were for the most part grammatically simple sentences, so more research is needed on complex sentence types to determine whether pragmatic abilities among signers of various backgrounds remains parallel when task difficulty increases. For example, “embedded scalar implicatures” (Chierchia et al. 2011) may involve significantly more grammatical complexity and semantic computation than the simple unembedded implicatures studied here, and could potentially show further differences between early and late learners (and, for that matter, between children and adults).

The second conclusion we draw from these results is that despite widespread success with pragmatics, one particular aspect of scalar implicature generation may be affected by the age-onset of language experience during child development: the generalizable lexical scale. Recall that the only place in which there was even a marginal difference between early and late signers was on the quantifier scale. From the point of view of what is required to appropriately reject a sentence that makes use of the quantifier scale, all of the same components are necessary as in the spatial ad hoc scale-based sentence, except for the concept of a context-general "scale" based on two or more regularly contrasting lexical items. For quantifiers, participants can possibly activate stored knowledge, in the form of a pragmatic connection between <all, some> in addition to their logical meanings, to know that there is a better description of the picture (involving *all* instead of *some*), and then reject the underinformative description. For the spatial sentence type, there can be no automatic reference to a stronger pragmatic alternative that holds in a variety of context; rather, rejection involves reasoning about entire alternative propositions specific to the context of utterance. In this respect, the spatial scale in ASL is like the ad hoc scale where adults infer from *Mary drank tea, hot chocolate, and Coke* that *Mary didn't drink coffee*. Children have no difficulty with these inferences (Papafragou & Tantalou 2004, Stiller et al. 2011), and we find here that late learning ASL signers also show no problems with the ad hoc sentence type. As previously described, these late learners have significant, albeit atypical, experience with English, but crucially, the spatial sentence types are unique to ASL (involving “classifier constructions”), and cannot have been computed via transfer from English.

One alternative possibility to consider is that the later learners could have differed from early learners on the quantifier sentences due to the topicalized structure in the quantifier sentence trials rather than to the scale itself. Difficulty on topicalization is consistent with previous studies finding complex syntactic structures difficult for late learners. However, it is important to note that sentences in the spatial trials also began with topicalized nouns, and here there was no difference between groups. We take this to indicate that topicalization is not the reason why late learners performed less well, but instead that some aspect of the lexical scale may be advantageously learned early in life (although clearly is not completely inaccessible to late learners, either). More broadly, our results suggest that the linguistic effects associated with the age-onset of language learning are not entirely restricted to morpho-syntactic and phonological phenomena, as has been previously proposed, but extend to the semantic/pragmatic information contained in the lexicon as well.

Extending our findings to typical childhood L1 acquisition, we suggest that children's decreased performance on prototypical scalar implicatures compared to adults may be affected by their incomplete acquisition of the language, not just generally immature pragmatic skills. For prototypical scales like quantifiers, we suggest that specific linguistic learning must occur, namely the lexical terms and their relation on the scale itself. Our results suggest that early language experience is an important factor in the learning of these types of scalar implicatures and their semantic/pragmatic underpinnings. This does not rule out other contributing factors (e.g. pragmatic tolerance), but does suggest that in addition to these, we should consider the importance of a lexical scale. However, we only tested a single lexical scale (<all, some>). Clearly more research is needed to determine whether our results extend to other scales, since as we discussed there is variation among different scales for when children exhibit adult-like behavior.

In contrast to the other sentence types, all groups of signers calculated far fewer implicatures in the Coordination sentence type, accepting about as many descriptions as they rejected for the test case. Like the spatial sentence type, the Coordination sentence type is also not based on a lexical contrast, since the coordinators COORD-shift(and) and COORD-shift(or) differ only in non-manual marking, despite conveying traditionally “scalar” propositional content. Because signers showed no difficulty in rejecting false descriptions on the Mismatch sentence type, they appear to be aware of non-manual marking and, crucially, are able to make an appropriate distinction between conjunctive and disjunctive readings. We conclude then that when there is no explicit lexical contrast, as in the coordination scale in ASL, then there is a less strong bias towards negating a stronger contrasting lexical item, and hence drawing the implicature. This finding lends particularly strong support in favor of some stored information in the lexicon (about the pragmatic relationship between scalar lexical items) as contributing to high levels of scalar implicature interpretations. Given this, we are inclined to take the marginal difference between early and late signers on the quantifier scale more seriously, because the behavior of late signers on quantifiers and the behavior of all signers on the coordination scale point toward a role for stored lexical knowledge in scalar implicature calculation, potentially in addition to many other factors. Together, these results also suggest that privileged access to this information may show effects of the age-onset of language: because there is no lexical contrast to be learned in the Coordination sentences, there was no advantage for these types of scales in having learned the language at an early age.

Based on the native-like performance by late learners on spatial and coordination sentences but not on quantifiers, it is important to consider precisely what aspects of a scale are especially well-suited to learning early in life. On the one hand, early language learning may facilitate the automatic processing of the implicature given the scale; on the other hand, it could be that early language learning facilitates the creation of the scale itself through learning the meaning of individual lexical items and, at the same time, how they relate to one another. We speculate that the answer may lie in linking these two explanations. In particular, we suggest that learning a scale involves learning the lexical items and, crucially, their place in the entire scale, such that all of the alternatives on a scale are, or may easily be, activated and considered when one lexical item is used. Under such a view, at least some aspect of meaning of the lexical item contains information about how it is related to the entire scale, i.e. stored generalizable pragmatic information. One task of the language learner, then, is to learn the relationship of the lexical items to each other. If this structured relationship is learned at an early age as part of the meaning of scalar items, this can lead to more automatic triggering of the scalar implicature in adulthood. In particular, we suggest that while the mapping of forms to meaning in a lexicon may be successfully learned later in life (in an L2, Newport 2002), another part of the lexicon studied here, namely, the regular and generalizable relationship of words to each other in semantic scales, may be less easily learned for the first time later in life. This structure is similar, though not identical, across many languages (contributing to success for adult L2 learners on scalar implicatures that are based on scales that can be transferred from their L1). In this sense, deaf adults with delayed L1 are a unique group that allow us to directly probe the acquisition pattern we propose.

Many aspects of the process of scalar implicature calculation are still under debate, such as whether they are the result of a process of counterfactual reasoning about others' thoughts (Grice 1989; Russell 2006), of a conventionalized version of this reasoning (Levinson 2000), or of a grammaticalized process (Chierchia 2006; Chierchia et al. 2011). This question is clearly complex, but we believe that our results add one important piece of data to this debate: adults with delayed first language exposure pattern with young children in showing adult-like performance on ad hoc scales, but, unlike adults who acquired an L1 in early life, show occasionally degraded performance on lexically based scales. This lends empirical support to the important role of a generalized, context-independent scale in implicature calculation, and not toward proposals in which implicatures require a simple comparison of alternative propositional content. More research with deaf children who experience less linguistic input during childhood than the educated participants tested here (Ferjan Ramirez et al., 2013) would solidify these findings.

To summarize, we investigated scalar implicature calculation in adults who had decreased overall language experience in childhood relative to both their life experience and cognitive maturity. Our test group was deaf adults who were not exposed to a sign language early in life, consider American Sign Language (ASL) their first and dominant language, and consequently learned their first language much later than is typical in language acquisition. These adults showed reduced performance on the lexically based scale, but not on the other scales tested that did not involve lexical items, namely the spatial ad hoc and coordination scales. We conclude that the extent to which domain-general skills account for processing at the semantic/pragmatic interface may depend upon the type of scale, with lexically determined scales being more sensitive to language acquisition and ultimate language proficiency than ad hoc scales.

Acknowlegdments

Much gratitude to Marla Hatrak, Ivano Caponigro, Cami Miner, Corinne Brion, Carol Padden, Matt Hall, Susan Fischer, Brandon Scates, Peggy Lott, audiences at ESSLLI 2009 in Bordeaux, Gallaudet Brown Bag lunch, Polinsky Language Sciences lab at Harvard, the Center for Research in Language at UCSD, members of the Mayberry Multimodal Language Development lab at UCSD, San Diego Sign Language Reading Group, and UCSD Semantics Babble. This research was supported by a pre-doctoral fellowship supported by NIH predoctoral training grant #T32DC000041 and a UCSD Linguistic Department Research grant to K. Davidson and NIH grant R01DC012797 to R. Mayberry. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

Barner, D., Brooks, N., & Bale, A. (2011). Accessing the unsaid: the role of scalar alternatives in children's pragmatic inference. *Cognition*, *118*(1), 84–93.

Bott, L., & Noveck, I. (2004). "Some utterances are underinformative: The onset and time course of scalar inferences." *Journal of Memory and Language*, 51(3), 437-457

Breheny, R., Katsos, N., & Williams., J. (2006). "Are generalised scalar implicatures generated by default? An on-nline investigation into the role of context in generating pragmatic inferences" *Cognition, 100*(3), 434-363

Birdsong D, & Molis M. (2001). On the evidence for maturational constraints in second language acquisition. *Journal of Memory & Language, 44*, 235-49.

Boudreault, P., & Mayberry, R. I. (2006). Grammatical processing in American Sign Language: Age of first-language acquisition effects in relation to syntactic structure. *Language and Cognitive Processes*, *21*, 608–635.

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 A. H.-J. Do, L. Domínguez, & A. Johansen (Eds.), *BUCLD 25 Proceedings* (pp. 157–168). Somerville, MA: Cascadilla Press.

Chierchia, G., Fox, D., & Spector, B. (2011). The Grammatical View of Scalar Implicatures and the Relationship between Semantics and Pragmatics. In K. Von Heusinger, C. Maienborn, & P. Portner (Eds.), *Semantics: An International Handbook of Language Meaning* (Vol. 44, pp. 1–43). Mouton de Gruyter.

Davidson, K. (2013). And or Or: General Use Coordination in ASL. *Semantics & Pragmatics*, 6(4), 1-44.

Davidson, K. (2011). *The Nature of the Semantic Scale : Evidence from Sign Language Research*. University of California, San Diego.

Ferjan Ramirez, N., Lieberman, A. M., & Mayberry, R. I. (2013). The initial stages of first-language acquisition begun in adolescence: When late looks early. *Journal of Child Language*, *40*, 391–414.

Flege J.E., Yeni-Komshian G.H., Liu S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language. 41*, 78-104.

Foppolo, F., Guasti, M. T., & Chierchia, G. (2012). Scalar implicatures in child language: Give children a chance. *Language Learning and Development*, *8*(4), 364–394.

Gazdar, G. (1979). *Pragmatics: Implicature, presupposition, and logical form.* Academic Press, New York.

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

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

Grüter, T., Lieberman, M., & Gualmini, A. (2008). The acquisition of the scope properties of disjunction by Japanese learners of English: a test case for L1 versus UG as the L2 initial state. In R. Slabakova, J. Rothman, P. Kempchinsky, & E. Gavruseva (Eds.), *Proceedings of the 9th Generative Approaches to Second Language Acquisition Conference (GASLA 2007)* (pp. 47–56). Cascadilla Press.

Gualmini, A., Crain, S., Meroni, L., Chierchia, G., & Guasti, M. T. (2001). At the semantics/pragmatics interface in child language. *Proceedings of Semantics and Linguistic Theory XI*. Ithaca, NY: CLC Publications, Cornell University.

Horn, L. (1989). *A Natural History of Negation*. Chicago: Chicago University Press.

Huang, Y. T., & Snedeker, J. (2009). Semantic meaning and pragmatic interpretation in 5-year-olds: Evidence from real-time spoken language comprehension. *Developmental Psychology*, *45*(6), 1723–1739.

Johnson, J., & Newport, E. L. (1991). Critical period effects on universal properties of language: The status of subjacency in the acquisition of a second language. *Cognition*, *39*(3), 215–258.

Katsos, N. (2007). The semantics/pragmatics interface from an experimental perspective: the case of scalar implicature. *Synthese*, *165*(3), 385–401.

Katsos, N., & Bishop, D. V. M. (2011). Pragmatic tolerance: implications for the acquisition of informativeness and implicature. *Cognition*, *120*(1), 67–81.

Katsos, N., & Smith, N. (2010). Pragmatic tolerance and speaker-comprehender asymmetries. In K. Franich, K. M. Iserman, & L. L. Keil (Eds.), *BUCLD 34 Proceedings*. Cascadilla Press.

Levinson, S. (2000). *Presumptive Meanings: the theory of generalized conversational implicatre*. Cambridge, MA: MIT Press.

Mayberry, R. I., & Lock, E. (2003). Age constraints on first versus second language acquisition: Evidence for linguistic plasticity and epigenesis. *Brain and Language*, *87*, 369–383.

Mayberry, R. I., Lock, E., & Kazmi, H. (2002). Linguistic ability and early language exposure. *Nature*, *417*, 38.

Mitchell, R., & Karchmer, M. A. (2004). Chasing the mythical ten percent: Parental hearing status of deaf and hard of hearing students in the United States. *Sign Language Studies*, *4*(2), 138–163.

Newport, E. (2002). Critical periods in language development. *Encyclopedia of Cognitive Science*. 737-740.

Noveck, I. A. (2001). When children are more logical than adults: experimental investigations of scalar implicature. *Cognition*, *78*(2), 165–188.

Noveck, I. A., & Posada, A. (2003). Characterizing the time course of an implicature: An evoked potentials study. *Brain and Language*, *85*(2), 203–210.

Papafragou, A., & Musolino, J. (2003). Scalar implicatures: experiments at the semantics-pragmatics interface. *Cognition*, *86*(3), 253–82.

Papafragou, A., & Tantalou, N. (2004). Children ' s Computation of Implicatures. *Psychology*, (776117726).

Russell, B. (2006). Against Grammatical Computation of Scalar Implicatures. *Journal of Semantics*, *23*, 361–382.

Sandler, W., & Lillo-Martin, D. (2006). *Sign Language and Linguistic Universals*. Cambridge: Cambridge University Press.

Sauerland, U. (2012). The Computation of Scalar Implicatures: Pragmatic, Lexical or Grammatical? *Language and Linguistics Compass*, *6*(1), 36–49.

Siegal, M., Matsuo, A., & Pond, C. (2007). Bilingualism and Cognitive Development: Evidence from Scalar Implicatures. In Y. Otsu (Ed.), *Proceedings of the Eighth Tokyo Conference on Psycholinguistics* (pp. 265–280). Tokyo: Hituzi Syobo.

Siegal, M., & Surian, L. (2004). Conceptual development and conversational understanding. *Trends in cognitive sciences*, *8*(12), 534–8.

Slabakova, R. (2010). Scalar implicatures in second language acquisition. *Lingua*, *120*(10), 2444–2462.

Stiller, A., Goodman, N. D., & Frank, M. C. (2011). Ad-hoc scalar implicature in adults and children. *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*.

Van Rooij, R., & Schulz, K. (2004). Exhaustive interpretation of complex sentences. *Journal of Logic, Language, and Information*, *13*, 491–519.

Appendix: Experimental Stimuli

Quantifiers ***(****N.B. the first noun is topicalized with brow raising nonmanual marking)*

BALL, ALL/SOME YELLOW.

BOOKS, ALL/SOME OPEN.

CANDLE, ALL/SOME LIT.

CAN, ALL/SOME RED.

PENCIL, ALL/SOME BROKE.

BOWL, ALL/SOME YELLOW.

(Playing)CARD, ALL/SOME BLACK.

CUP, ALL/SOME FULL.

HANGER, ALL/SOME BROWN.

SHOE, ALL/SOME BLACK.

SOCK, ALL/SOME BROWN.

SPOON, ALL/SOME WOODEN.

Spatial *(N.B. Many clauses involve spatial classifiers, noted with (X,Y,Z) to indicate spatial placements, and the first noun is frequently (maybe always) topicalized)*

BALL CL:5(claw)$_X$ CL:5(claw)$_Y$ CL:5(claw)$_Z$.

BEAR CL:5(claw)$_X$ CL:5(claw)$_Y$ CL:5(claw)$_Z$.

BOWL CL:C$_X$ CL:C$_Y$ CL:C$_Z$.

CANDLE CL:C$_X$ CL:C$_Y$ CL:C$_Z$.

CAN CL:C$_X$ CL:C$_Y$ CL:C$_Z$.

CUP CL:C_X CL:C_Y CL:C_Z.

D-V-D CL:B_X CL:B_Y CL:B_Z.

COLOR PEN CL:G_X CL:G_Y CL:G_Z.

PENCIL CL:G_X CL:G_Y CL:G_Z.

SHOE CL:B_X CL:B_Y CL:B_Z.

SPOON CL:G_X CL:G_Y CL:G_Z.

Coordination

HAVE BALL CL:5(claw)$_X$ GLOBE CL:5(claw)$_{Y(\text{behind X})}$ COORD-shift(and/or) SANDAL CL:B_W CL:5(claw)$_X$.

HAVE BEAR CL:V(curved)$_{X(\text{inside cupped nondominant hand})}$ WINE CL:C_X COORD-shift(and/or) BOWL$_Y$ CL:V(curved)$_{Y(\text{inside cupped nondominant hand})}$.

HAVE TOOTHBRUSH CL:$G_{X(\text{inside cupped nondominant hand})}$ MUG COORD-shift(and/or) BOTTLE CL:$G_{X(\text{inside cupped nondominant hand})}$.

HAVE PLATE CL:C(half)$_X$ CL:C(half)$_Y$ [WHITE CANDLE CL:$C_{X(\text{nondominant hand flat})}$ COORD-shift(and/or) BROWN CANDLE CL:$C_{Y(\text{nondominant hand flat})}$]

HAVE COLOR MARKER [WALLET CL:B_X CL:G_X COORD-shift(and/or) GLOVE CL:B_Y CL:G_Y]

HAVE PENCIL CL:G_X CEREAL B-O-X CL:B(curved)$_X$ COORD-shift(and/or) PENCIL CL:G_Y J-U-I-C-E CL:C_Y

HAVE BOX$_X$ PURSE$_Y$ CEREAL B-O-X CL:$B_{X(\text{inside C shaped nondominant hand})}$ COORD-shift(and/or) CEREAL B-O-X CL:$B_{Y(\text{inside C shaped nondominant hand})}$

HAVE MEASURE [CL:C(reduced)`long thin object'] $_X$ [CL:U$_{X(\text{inside C shaped nondominant hand})}$ BLACK CUP$_X$ COORD-shift(and/or) GLASS [CL:C(reduced)`tall round object']$_Y$ CL:U$_{Y(\text{inside C shaped nondominant hand})}$]

HAVE SCISSORS CL:3$_{X(\text{on B shaped nondominant hand})}$ BOOK COORD-shift(and/or) SHOE CL:3$_{Y(\text{on B shaped nondominant hand})}$

HAVE SOCK IN BLUE SHOE COORD-shift(and/or) SOCK IN BROWN SHOE.

HAVE SPOON IN CUP COORD-shift(and/or) SPOON IN BOWL.

HAVE #TOWEL PINEAPPLE CL:5(claw)$_X$ CL:B$_{X(\text{on top})}$ COORD-shift(and/or) BOOK CL:B$_Y$ CL:B$_{Y(\text{on top})}$

Relative Clause *(N.B. the first three signs are topicalized using brow raise nonmanual marking as shown in one condition, and are not topicalized in the other)*

(____________br)
BEAR HUG DOG HAVE HEADPHONES.

(___________________br)
BEAR STARE-AT TIGER HAVE CANDY.

(__________________br)
BEAR RIDING RABBIT HAVE COIN.

(__________________br)
WOMAN CARRY BEAR HAVE CANDY.

(_____________________br)
MAN CATCHING WOMAN HAVE PURSE.

(___________________br)
FATHER TOUCHES SON HAVE RING.

(_________________br)
BEAR HOLD WOMAN HAVE BOWTIE.

(______________br)
BEAR PUSH DOG HAVE BANDAGE.

(_____________br)
FISH BITE BEAR HAVE PURSE.

(________________br)
BOY RIDING BEAR HAVE KEY

(____________________br)
RABBIT CHASING DOG HAVE BALL

(__________________br)
WOMAN SIT-ON DOG HAVE CANDY.