Form, Meaning, and Comprehension/Production Asymmetries in Agreement

Erin Zaroukian

Children's knowledge of subject-verb agreement was investigated through a comprehension task, which examined whether children 24 to 40 months display a form-to-meaning mapping between an auditory stimulus with singular or plural subject-verb agreement and a video representation of that stimulus showing one or two actors performing the indicated action. Results suggest that comprehension may emerge late, perhaps even as late as 5 years (Johnson, de Villiers, & Seymour, 2005). These and previously established results are evaluated in terms of children's underlying competence, and it is suggested that children's acquisition of the form of subject-verb agreement may be later than generally assumed.

Introduction

For the communicative system that is human language, there is a general intuition that comprehension and production go hand in hand. Both, it seems, must involve an integration of linguistic structure and meaning for communication to be successful. There are many cases, however, where comprehension and production abilities deviate from one another. Language acquisition research, as a case in point, has found many instances where children show differences between their comprehension and their production abilities in the form of a comprehension-before-production asymmetry. In the realm of morphology, for example, children have been shown to score higher on tests of comprehension than on production for "–er" as a marker of agent and instrumental nouns (Clark & Hecht, 1983) and of "–s" as a marker of noun plurality (Winitz, Sanders, & Kort, 1981). This asymmetry calls out for an explanation.

Given that linguistic performance does not necessarily reflect linguistic competence, this comprehension-before-production asymmetry may reflect children's difficulty, not with any deep underlying knowledge, but with production-specific performance mechanisms. For example, to produce an utterance speakers may first activate forms in their lexicon to construct an utterance which is executed via their articulatory system and may then be subject to a self-monitoring mechanism, all of which may be specific to production (Levelt, 1989). Production may require such additional skills or resources such that children's knowledge may first emerge in the form of comprehension, leading to a comprehension-before-production asymmetry. On the other hand, comprehension may involve its own independent set of skills, such as some sort of anticipatory mechanism (Phillips, 2009). For example, ERP studies show evidence of anticipation in comprehension, such that a passage like "They wanted to make the hotel look more like a tropical resort. So, along the driveway, they planted rows of __ " elicits a greater deviation in neural activity if the blank is filled with the less congruent "pines" than the more congruent "palms", suggesting that during comprehension listeners anticipate a word like "palm" in this context (Federmeier, 2007). An anticipatory mechanism such as this suggests that the opposite pattern, a production-before-comprehension asymmetry, could also arise if in certain tasks children have greater difficulty with a comprehension-specific mechanism. These comprehension- and production-specific mechanisms and their differing importance in different tasks, then, may account for asymmetries in comprehension and production seen in language acquisition studies.

While comprehension- and production-specific mechanisms suggest that these asymmetries are somewhat superficial, an explanation involving competence and deeper

asymmetries has been offered by Naigles (2002). She suggested that, overall, form is easy and meaning is hard. This means that, while children may have learned certain forms (i.e. made relevant formal generalizations), learning the meanings of those forms may be so difficult that it may mask their knowledge of form. With this in mind, several seemingly contradictory findings in language acquisition can be explained. For example, infants appear to have abstract knowledge about word order and frames, but toddlers' production of verb frames seems item-based. The important difference between these findings, Naigles claimed, is that infant studies tested abstraction over form while toddler studies tested abstraction over form *and* meaning. When learning verbs, meaning must be learned in a rather item-based way, and the less complete a child's knowledge of the meaning of a particular verb, the less likely he may be to use that verb productively (e.g. in frames other than those in which the verb has been frequently heard), since meaning is an important predictor of a verb's argument structure. This can lead a child's knowledge of verb frames to appear item-based, though this is simply a reflection of their meaning knowledge. Additionally, children converse to convey meaning, not to show off their knowledge of form, so there is little reason for children to be boldly productive. This may be exacerbated by children's limited computational resources, which is supported by toddlers' increased ability to show abstract frame-based knowledge when a task of discriminating frame-based meaning is made easier (Naigles et al., 2002). In summary, Naigles's claim is that form is easy and meaning is hard, and conflicting experimental results arise from inconsistency in testing form or form and meaning, along with other computational constraints.

This idea that form is easy and meaning is hard has consequences for how comprehension/production asymmetries are viewed. Production may be possible with neither form nor meaning if a child relies on memorized or item-based knowledge. To be productive and apply knowledge to novel forms, however, form is necessary. Meaning, though, does not appear to be necessary for production (see description of agreement below). For comprehension, on the other hand, meaning is vital. If meaning is more difficult than form, this suggests that production-before-comprehension asymmetries should be possible in language acquisition. Comprehension-before-production asymmetries may still exist, however, due to constraints on different comprehension- and production-specific mechanisms and to various task-specific issues. Underlyingly children may still instantiate Naigles's 'form is easy, meaning is hard' hypothesis.

Both comprehension-before-production asymmetries and production-before-comprehension asymmetries have been posited for children's acquisition of subject-verb agreement. The remainder of this section will discuss these claims and how they might be explained given the ideas presented above. This will involve how children comprehend and produce language, how they acquire form and meaning, and whether these asymmetries reflect lack of underlying knowledge or merely a masking of it.

In discussions of agreement below, a roughly Chomksyan view will be taken in which agreement is the realization of one category's interpretable feature as an uninterpretable feature on another category. Interpretability corresponds to the ability to carry meaning, such that interpretable features carry meaning and uninterpretable features do not. The form of agreement then will involve generalizations in terms of these features. Agreement here will mostly be discussed in terms of the feature NUMBER, which is interpretable on nouns, but not on verbs. For example, in (1a) the subject is singular, and this feature is realized on the verb as the suffix "–s". In (1b) the subject is plural, and this feature is realized on the verb as the null suffix "–ø".

(1)   a. The boy dances.
        b. The boys dance.

As this example demonstrates, lexical subject-verb agreement in English is only marked overtly for third person singular, as in (1a). Unless otherwise specified, this English lexical subject-verb agreement is the type of agreement that will be discussed.

Note that knowledge of the form of agreement does not require knowledge of the associated meaning. Children, then, could theoretically master production of agreement before comprehending it. It is also possible, however, that children's production may not reflect their underlying knowledge. For example, they may conserve computational resources by dropping inflectional markers on verbs, which are not important for conveying meaning. If children do not have the motivation and resources to preserve inflection on their verbs until after they have associated meaning with agreement, they may show a comprehension-before-production asymmetry. Alternatively, children could potentially appear to have good production without having learned the form of agreement by relying on item-based knowledge. The demands any given task places on a child may then have a drastic effect on the pattern of comprehension and production found.

Literature concerning the acquisition of subject-verb agreement is discussed below. Special attention is given to how results on production and comprehension bear on Naigles's 'form is easy, meaning is hard' hypothesis, as well as how contradictory asymmetries in production and comprehension may be explained and what they reveal about the nature of language learners.

Production
Mastery of agreement production is generally considered to occur between 26 and 46 months. This is based on a corpus study by Brown (1973), who examined the natural production of several children and found that they began using agreement in at least 90% of obligatory contexts between 26 and 46 months. This relatively young mastery, however, stands in contrast to poor production in elicitation tasks through the age of 7. This evidence comes from Berko (2004), who attempted to elicit agreement on nonce verbs, using prompts such as "This is a man who knows how to loodge. He is loodging. He does it every day. Every day he ____." Children 4-5 and 5;6-7 were tested and performed similarly, correctly inflecting the verb at best 56% of the time. These studies suggest that some knowledge of form may emerge early, but it appears to be limited, as children have difficulty producing agreement on novel verbs for several years after they have begun producing it on known verbs. It may be the case that poor performance on the elicitation task reflects a masking of form knowledge due to the additional computational demands associated with producing elicited responses and working with novel verbs, but the sharp contrast in performance and the vast age difference across Brown's study and Berko's study suggests that this may not be the whole explanation. Instead, these children could be relying on item-based knowledge, such that they are successful with known verbs, for which they have item-based knowledge, but unsuccessful with novel verb, for which they have no item-based knowledge.

The correspondence between children's rate of agreement usage and their input, found by Theakston et al. (2003), provides evidence for early item-based knowledge. Theakston et al. trained children aged 2;6-3;0 on known and novel verbs in three conditions: marked (e.g. "This one tams!"), unmarked (e.g. "Will this one tam?") and mixed (e.g. both "This one tams!" and

"Will this one tam?"). When children were later asked to produce these verbs in their marked form (e.g. "tams"), agreement on known verbs was produced at the same rate across input conditions, but for novel verbs its rate of production was directly proportional to the rate of inflection in the input. Specifically, more inflection was produced on novel verbs when trained in the marked condition than in the mixed or unmarked conditions. This follows from an item-based account, which predicts that children will more or less reproduce their input. Since they have heard known nouns many times outside the experiment, the experimental input should have little effect on children's production, which was seen in this study. Novel verbs, on the other hand, were only heard in the experimental input, so children should learn the words as they are presented, which was seen in this study. Furthermore, children produced more agreement in the marked condition for novel than known verbs. This follows from an item-based account because the children had heard unmarked forms of the known verbs (e.g. in questions outside the experimental setting), but not of these novel verbs. Similarly, children produced more agreement in the unmarked condition for known than novel verbs. This follows because the children had heard marked forms of the known verbs, but not of these novel verbs. This overall input matching suggests that children 2;6-3;0 are not mapping these words onto abstract underlying structures, which should treat known and novel verbs the same, but rather are using item-based knowledge about each individual item. Brown's early success might then be due to children's item-based knowledge of the forms they produced, and Berko's lack of success may be due to children's incorrect or missing item-based knowledge for the novel verbs, since they heard uninflected but never inflected forms of the verbs they were asked to produce. The appearance of item-based knowledge, however, does not entail a lack of knowledge of the form of agreement.

Children in Theakston et al.'s study may have known the form of agreement, but this knowledge may have been masked due to demands placed on their computational resources. Consider the child's task of learning the inflectional paradigm of their language. It seems reasonable to suppose that children begin by memorizing whole word forms, as if every word form were irregular (cf. Pinker, 1991). Eventually children may learn the form of agreement, extracting rules from these stored word forms (e.g. add "–s"). Now consider what could happen when children are placed in demanding experimental situations. It does not seem unreasonable to propose that in such situations when their computational resources are pushed to their limits, children are unable to use their recently extracted form and instead rely on memory. If children are forced to rely on memory, input-matching would be expected. Thus, input matching in such experimental settings need not indicate a lack of form knowledge. A similar story can be used to account for Brown's and Berko's results. Brown's spontaneous speech placed relatively low demands on the children, so their young success may be due to use of form knowledge instead of simply good item-based knowledge. Berko's elicitations placed higher demands on children, and her novel verbs were only ever heard in their uninflected and progressive forms. Children may have been unable to supply an appropriate third-person singular form due to computational limitations and instead used a similar uninflected form from memory. Results showing item-based knowledge may simply show a masking of form knowledge.

Unfortunately, it is difficult to tell whether item-based results show a masking of form knowledge or whether they show a genuine lack of form knowledge. There are areas, however, in which masking and item-based accounts may make different predictions. In a task with lower demands on a child's computational resources, a masking explanation predicts better performance, since it is more likely that children will have the resources necessary to show their knowledge of form. An item-based explanation predicts that performance, which could improve

in some respects due to more computational resources available, will continue to match the input. Theakston et al. reported both elicited and spontaneous production data, and it would be interesting to know whether there was less input matching in the spontaneously-produced forms, which would favor a masking account. Unfortunately the two types of production were not reported separately. While it is difficult to determine whether results that appear item-based mask knowledge of form, there is reason to suspect that children have knowledge of the form of agreement by 26 to 46 months.

Comprehension

Abstract knowledge of the meaning of nominal number markers, which is likely a prerequisite to comprehension of agreement, appears to emerge by 36 months. This result comes from an Intermodal Preferential Looking Paradigm (IPLP) study by Kouider et al. (2006), where the displays featured one nonce object versus eight nonce objects, and audio stimuli were of the form "Look at the blickets" or "Look at the blicket". The 36-month-olds tested in this experiment succeeded, suggesting that children have an abstract conception of nominal number by 36 months.

IPLP studies have shown agreement comprehension in several languages by around 3 years. In German, children appear to comprehend number on verbs around 2 to 4 years. This was seen in a study by Brandt and Höhle (2008) with audio stimuli like "Sie füttert einen Hund" (She feeds a dog) and "Sie füttern einen Hund" (They feed a dog). While 2- to 4-year-old children showed a plural bias in both their looking times and pointing scores, they overcame this bias and showed comprehension in the singular condition. In French, agreement comprehension seems to emerge around 30 months. Legendre et al. (2009) examined agreement comprehension in French-acquiring 24 and 30 month olds using audio stimuli like "Il apporte le mique" (He's bringing the mique) and "Ils apportent le mique" (They're bringing the mique). In these audio stimuli, the subject pronouns, analyzed as agreement markers, provide a single phonological cue to number ([ilapor lə mik] vs. [izapor lə mik]). In both looking times and pointing scores, 30 month olds, but not 24 month olds, showed comprehension of the number contrast. Furthermore, these results were not considered a simple reflection of the input, since instances of this kind of agreement are extremely low in speech a child might hear. Instead, these children seemed to have recognized which agreement forms are singular and plural and parsed the visual stimuli in a way consistent with the agreement and verbs. In these languages, children appear able to map between the meaning and form of agreement by 2 to 4 years.

While evidence of agreement comprehension in other languages appears between 2 and 4 years, comprehension of agreement in English may not emerge until as late as 5 years, despite the fact that children appear to have sophisticated knowledge of nominal number by 3 years. This late comprehension evidence comes from a picture selection task by Johnson, de Villiers, & Seymour (2005), where children were shown a singular-plural pair of pictures with one or two animals and heard sentences like "The duck swims on the pond" or "The ducks swim on the pond", where the number on the noun was phonologically masked. Forced to rely on verbal inflection, only 5 and 6 year olds, not 3 and 4 year olds, performed above chance. The same finding was replicated in Spanish, using sentences with pro-drop, like "Nada en el charco" ([it] swims in the pond) and "Nadan en el charco" ([they] swim in the pond) (Perez-Leroux, 2006). Such late comprehension, especially when contrasted with early production, is striking. It seems strange that language learners should be able to spend as many as three years producing utterances that they cannot entirely comprehend. As strange as it may seem, these studies suggest

that while English-acquiring children may learn the form of agreement early, they do not associate a meaning with it until much later. As discussed above, however, such a drastic asymmetry could exist if it were sufficiently difficult to map the meaning of agreement onto its form.

Poor performance on these comprehension tasks, however, may merely mask knowledge of meaning. Limited computational resources, for example, may be to blame if a task was too demanding. This is supported by an unpublished revised version of Johnson et al.'s English study, which showed comprehension of English subject-verb agreement at 4 (Perez-Leroux, 2006). This suggests that comparatively late comprehension in English may be tied to methodology. Any real production-before-comprehension asymmetry in agreement, then, may not be as drastic as Johnson et al. suggest.

Comprehension and Production

Several studies have attempted direct comparisons between agreement comprehension and agreement production, with mixed results. Fraser, Bellugi, & Brown (1963) compared comprehension of 37 to 42 month olds with two forms of production. Their comprehension, or C, task was a picture selection task where children were shown two pictures and were told their names, but not which name corresponded to which picture. The experimenter said one of the names (e.g. "The boy draws") and asked the child to select the corresponding picture. The experimenter then said the other name (e.g. "The boys draw") and again asked the child to select the corresponding picture. In their production, or P, task children were told the names that corresponded to a pair of pictures (e.g. "This is 'The boy draws' and 'The boys draw'"), but not which name corresponded to which picture. The experimenter then pointed to one of the pictures and asked the child to name it. In their imitation, or I, task children were asked to model the experimenter's utterances (e.g. "The boy draws"). Performance on the C task was slightly above chance, which was suggested to be around 25%. Overall, children performed best at the I task (58% correct), worse at the C task (29% correct), and worst at the P task (4% correct). It is not surprising that children performed worse at the P task than at the C task, since the P task required comprehension in addition to producing a response, but children's relative success at the I task and nominal success at the C task suggests that children 37 to 42 months old may have some knowledge of form, but they are still very limited in their ability to map meaning to that form.

Another pattern of results was found in a study by Keeney & Wolfe (1972) on children 36 to 59 months old. They tested production by gathering spontaneous speech from their subjects and found that these children had mastered agreement, using it correctly in 94% of obligatory contexts. They tested imitation by having children repeat sentences with both grammatical and ungrammatical agreement and found that children performed nearly perfectly on grammatical sentences and had a tendency to correct ungrammatical sentences. They tested comprehension using a picture selection task. When stimuli were full sentences, like "The birds hop" or "The bird hops", children were above chance of 50% correct, but still only around 60% correct. When stimuli contained the verb only, like "hop" or "hops", children did not perform above chance. From this, Keeney & Wolfe concluded that these children have a "syntactic, intralanguage rule for verb number inflection, but have no comprehension of the singular or plural meaning of such an inflection". These results are consistent with Brown's spontaneous production mastery between 26 and 46 months. Furthermore, children's success with imitation, and especially in correcting ungrammatical utterances, suggests knowledge of the form of agreement. Children's

poor performance with stimuli like "The bird hops" suggests again that at this age children are only beginning to map meaning to form.

While the different methodologies required for testing production and comprehension makes comparison difficult, these studies confirm the production data discussed above. Children may have knowledge of the form of agreement by around 3 years of age, evidenced through their near-perfect spontaneous production, their ability to imitate, and their tendency to correct ungrammatical agreement. And unlike Johnson et al.'s claim that children do not comprehend English agreement until they are 5 years old, these studies suggest some ability to comprehend between 3 and 4;11.

It is conceivable, however, that this younger comprehension success was driven not by agreement comprehension but by nominal number comprehension, which was masked by Johnson et al. but not by Fraser et al. or Keeney & Wolfe. For Keeney & Wolfe, noun number alone is sufficient to pick the correct picture, since the singular pictures showed only one animal, while the plural pictures each showed two animals. If children relied on nominal number, however, it is surprising that children did not perform better than 60%, given that they were of an age when they should have had abstract knowledge of nominal number markings. Relying on nominal number may have been possible though more difficult in Fraser et al.'s study, where singular and plural pictures both depicted two animals, though in the singular picture only one animal was shown performing the relevant action. If children were relying on nominal number, their limited success indicates that it was not very helpful.

Children's poor performance may have been due to the visual representations of meaning they were given. Though only one picture may be the "best" match for a given sentence, there is more than one correct answer. Consider the example in figure 1 below.
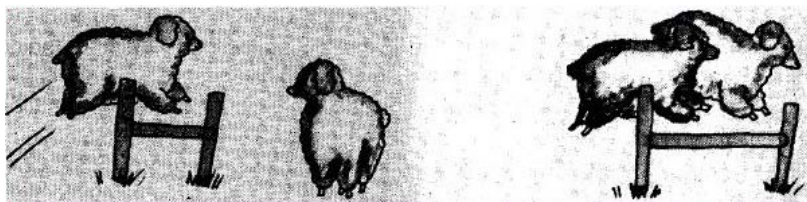

Figure 1: Example singular and plural pictures from Fraser et al. (1963).

If the child hears "The lamb jumps", he would ideally take it to refer to the picture on the left, which shows only one lamb jumping. He could, however, imagine that it refers to either of the jumping lambs in the plural picture on the right, such that "The lamb jumps" could potentially describe either or both pictures. Likewise, if the child hears "The lambs jump", he would ideally take it to refer to the picture on the right, which shows two lambs jumping. He could, however, imagine that it refers to any or all combinations of jumping lambs. Additionally, the child might imagine that the standing lamb could have recently jumped or may soon jump and therefore qualify as a jumping lamb. "The lambs jump" could therefore describe either or both picture. Children's difficulty picking the "best" picture, whether relying on nominal number alone or using agreement number, suggests that this methodology is not ideal for testing comprehension of agreement. This is especially poignant given evidence of agreement comprehension in French at 30 months and nominal number comprehension at 36 months in English using less ambiguous videos, where different nonce objects appeared in the singular and plural displays to discourage distributive readings in which plural audio could be represented by the sum of the singular and

the plural visual displays. It should also be noted that Legendre et al., unlike Johnson et al., Fraser et al., and Keeney & Wolfe, used dynamic displays, which may be important for verbal contrasts. Methodology, then, may be the cause of poor performance on comprehension tasks, not an authentic lack of agreement comprehension.

The large asymmetry between form, which may be acquired around 3 years, and meaning, which may not be acquired until around 5 years, might not be as great as proposed by Johnson et al. Instead, the methodologies used may have masked children's competence, giving the impression that children continued on for years producing forms they could not comprehend. Below, I ask whether English-acquiring children aged 24 to 40 months demonstrate comprehension of agreement. This is accomplished by asking whether these children display a form-to-meaning mapping between an auditory stimulus with singular or plural subject-verb agreement and a video representation of that stimulus showing one or two actors performing the indicated action. But first, Brown's claim of production mastery between 26 and 46 months is examined via replication.

Study 1

Study 1 seeks to replicate Brown's finding that production of third-person singular agreement is mastered between 26 and 46 months. Brown characterizes mastery as correct use in at least 90% of obligatory contexts. Determining obligatory contexts from corpora, however, can be quite difficult. For example, consider an utterance like "Linda see him ." (from Nina20, Suppes 1974). This could be considered an obligatory context, such that the child meant "Linda sees him." Alternatively, it could be a mis-transcribed command, "Linda, see him!", or the child could have intended "Linda saw him", "Linda will see him", "Linda can see him", or countless other utterances. Context can help disambiguate, but the choice is often unclear. A researcher's conservativity in considering contexts obligatory can greatly affect the age at which a child reaches Brown's 90% threshold, and unfortunately conservativity is often unclear within studies and inconsistent between studies. For example, Brown gives the following ages of mastery for three of the corpora he analyzed: Adam – 3;6, Sarah – 3;8, and Eve – 2;3. Wilson (2003) analyzed the data from the same corpora, but found that Sarah reached mastery much earlier and that Eve reached mastery much later. This discrepancy may be due to different conceptions of obligatory context. In an attempt to determine Brown's rough level of conservativity, the following study analyzes data from Adam's corpus according to three different levels of conservativity. Utterances from another child are then analyzed to see whether she reaches Brown's 90% threshold according to roughly Brown's level of conservativity between the ages of 26 and 46 months. Ramifications of conservativity choice are then considered.

Method

Data for this study was taken from corpora of spontaneous child speech in the CHILDES database (MacWhinney 2000). Files from corpora for two mono-lingual English-acquiring children were chosen: Nina (Suppes 1974, files 1-7, 9-20, 23, 30, 35, 37, 40-41, 44-45, 47, 49-50, 52-53, and 56) and Adam (Brown 1973, file 32). Data for Adam was taken when he was 3;6, the age when, according to Brown, he masters production of third-person singular lexical agreement. Data for Nina was chosen more or less randomly and was recorded between the ages of 1;11 and 3;3. Among the data, child utterances containing lexical verbs and third-person singular subjects were coded as agreeing or not, and the non-agreeing utterances were further

coded as a) being potentially mis-transcribed commands, b) as being utterances missing other tense, aspect, or mood markers (e.g. "is", "can", "will"), or c) neither. This further coding of non-agreeing utterances was done somewhat generously, such that, for example, any utterance that could have reasonably been interpreted as a command was marked as such. Nina's data was grouped into six-month bins. Results were calculated using three notions of obligatory context. The least conservative scores, which I will refer to as type 3, were calculated by dividing the number of agreeing utterances by all coded utterances (i.e. all utterances with lexical verbs and third-person singular subjects). More conservative scores, type 2, were calculated by dividing the number of agreeing utterances by all coded utterances minus those non-agreeing utterances that may have been miscoded commands. The most conservative scores, type 1, were calculated by dividing the number of agreeing utterances by all coded utterances minus both those that may have been mis-transcribed commands and those that may have been missing tense, aspect, or mood marking.

Results and Discussion

      Results are shown in table 1. Adam's score is closest to Brown's 90% using type 1 obligatory contexts, suggesting that Brown was quite conservative. Nina's type 1 score reaches 90% by 2;6-2;11, as can be seen in figure 2. Using the most conservative conception of obligatory context above, these results support Brown's claim that mastery is reached between 26 and 46 months, or 2;2 and 3;10, as Nina reaches mastery by 2;6-2;11.

	1	2	3
Adam (3;6)	89% (25/28)	69% (25/36)	61% (25/41)
Nina (2;6-2;11)	90% (63/70)	72% (63/88)	66% (63/96)

Table 1: Production scores using three type of obligatory contexts, shown at Brown's age of mastery (90% in type 1).
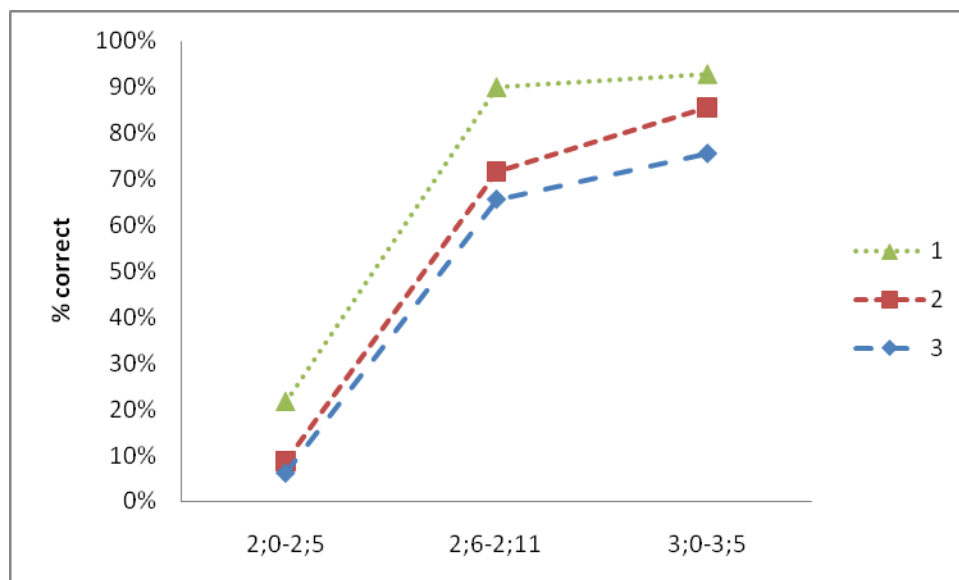

Figure 2: Nina's scores shown in 6-month bins across three types of obligatory contexts. Mastery in type 1 occurs by 2;6-2;11.

The question remains, however, as to whether it is appropriate to be so conservative with obligatory contexts. Notice that at mastery more of Nina's utterances containing third-person singular lexical verbs seemed to involve outright tense/aspect/mood errors (90%-72%=18%) than agreement errors (100%-90%=10%). This is somewhat surprising in light of claims that tense is acquired before agreement (Ingham 1998). If Nina's mastery of tense is still so weak as to allow such a potentially high number of tense errors, it would be strange to consider her a master of a later acquired skill, agreement production. This is a further indication that while children may have production 'mastery' by 26 to 46 month, they may not yet have mastered the form of agreement.

There are, however, more superficial explanations available for this relatively high number of errors. For one, children may have mastered inflection, but their grammar may still be subject to the Unique Checking Constraint (Wexler, 1998) such that they are unable to consistently check both tense and agreement features and therefore produce uninflected verbs. Alternatively, it could simply be that the limited computation resources at this age lead children to drop the third-person singular –"s", which is unnecessary from a communicative standpoint. Children may then have much greater competence than their performance reveals. Such superficial explanations are particularly appealing given that some form knowledge appears to be present from 19 months. Headturn Preference Procedure studies have shown that by 19 months children can distinguish agreeing (e.g. "The boy bakes bread", "The boys bake bread") from non-agreeing (e.g. "The boy bake bread") utterances (Soderstrom 2002). These children, however, do not distinguish agreeing utterances from ungrammatical doubly marked utterances (e.g. "The boys bakes bread", "The boy does bakes bread"), suggesting that whatever form they have learned is not adult-like. This data suggests that, though it may not be adult-like, children may have some knowledge about the form of agreement at a young age, well before they master production.

Study 2

Study 2 tests comprehension of subject-verb agreement by examining whether English-acquiring children aged 24 to 40 months demonstrate a form to meaning mapping between an auditory stimulus with singular or plural subject-verb agreement and a video representation of that stimulus showing one or two actors performing a given action. The auditory stimuli make no attempt to mask number on the noun, and IPLP was used with dynamic videos to minimize demands on the children in attempt to uncover knowledge which may have been masked in other more demanding studies.

Method

Participants

The participants in this study were 18 English-acquiring children (11 male, 7 female; age 24 months to 40 months, mean 32 months). An additional six children were tested but were excluded due to poor video quality (3) or experimenter error (3). Before coming into the lab, parents were asked to fill out a version of the MacArthur-Bates Communicative Development Inventories (CDI, Dale & Fenson, 1996) as an independent measure of the child's linguistic development. The return rate for these inventories was 83% (15/18). As compensation for their participation, each child received a small toy.

Materials

Video stimuli. The videos in this study were taken from Legendre et al. (2009). Each of the videos showed two eight-year-old boys performing actions with nonce objects. For each action there was a singular video, in which one boy demonstrated that action with a nonce object while the other stands stationary, and a plural video, in which both boys demonstrated the action jointly with another nonce object. Nonce objects were used to prevent distributive readings. Having both boys in the singular as well as the plural videos prevents children from distinguishing singular and plural videos based on the number of boys on the screen. Ten singular-plural video pairs were chosen from the set of videos such that they corresponded to an English verb that was likely to be known by two year olds. These videos were arranged into a split-screen display (see table 2) using FinalCut. The videos were superimposed on a gray background, so as to illuminate the child's face for ease of coding, in an arrangement judged through piloting to make the videos maximally visible while encouraging eye movements that would be clear to a coder. Each was six seconds long and the singular-plural pairs were edited as best as possible to make the actions temporally coincide. It should be acknowledged, though, that the coincidence was not perfect for all pairs of videos. It is also important to note that actions across different pairs of videos were temporally different. In the catching videos, for example, the act of catching occurs rather quickly, while in the wiping videos, the act of wiping extended across nearly the full six seconds. This means that in videos like the catching ones, there is only a short window in which to distinguish the singular from the plural video. Likewise, these critical moments during which the singular and plural videos can be distinguished fall at different times in different videos (e.g. catching begins at approximately 2.5 seconds, while wrapping occurs at approximately .3 seconds). Videos of the boys dancing were used to mark the correct side at the end of a trial, and waving videos were used to introduce the boys and mark the beginning and end of the testing session. Each of these videos was 3 seconds long.

phase	audio	video	duration
	"Oh, look!"		~2sec
baseline	*ping*		6sec
	"The boys kiss the dite."		~2sec

test			6sec
	ding ding ding		3sec

Table 2: Trial organization showing phase, example audio stimuli, example video frames, and approximate durations.

Audio stimuli. The audio stimuli were grammatical sentences of English involving a transitive verb and a nonce object. The subject in these sentences was always either "the boy" or "the boys". All sentences were spoken in infant-directed speech using intonation that stressed the verb. This intonation pattern was chosen to draw attention toward the verb and, hopefully, away from the subject and its nominal number marking. This stress pattern, however, may be more natural with singular subjects, where it is the verb that bears the overt number marking (in sentences with plural subjects, it may be more natural to stress the subject, which bears the overt number marking). The verbs were chosen to correspond to the videos and are given in table 3.

Nonce objects were used in the videos to avoid a distributive reading, and their names were chosen to be similar to real monosyllabic words commonly produced by two-year-olds (these real words were produced by over 80% of 24-month-olds, as determined through the CDI lexical norms), in an attempt to avoid difficult phonemic sequences and unnecessarily high memory load. One exception is the name "dax", which was borrowed from Berko (2004). Also, "muk" was inadvertently mispronounced such that it did not rhyme with "book", but this pronunciation was not deemed problematic. The phonological form of the nonce names, along with the real word they are based on, can be seen in table 3. This table also contains the phonological form of third-person singular verb ending. Note that there is a roughly even distribution of plural phonological forms (s=3, z=4, ɪz=3). Before coming into the lab, parents were asked which of the ten verbs their child understood, and only these were used in the child's stimuli. This helped ensure that children heard verbs that they were familiar with and eliminate noise in the data due to verb unfamiliarity.

audio stimulus				real word	sg verb ending
The boy	catches	the	ɡiz	(cheese)	ɪz
The boy	kisses	the	naj	(eye)	ɪz
The boy	wipes	the	pus	(juice)	s
The boy	takes	the	nal	(ball)	s
The boy	pulls	the	daks	–	z
The boy	hangs	the	fu	(shoe)	z
The boy	ties	the	val	(doll)	z
The boy	throws	the	pʌn	(moon)	z
The boy	wraps	the	kut	(foot)	s
The boy	dresses	the	klam	(clock)	ɪz
The boys	catch	the	ɡæθ	(bath)	
The boys	kiss	the	dajt	(light)	
The boys	wipe	the	blɔf	(block)	
The boys	take	the	dʒʌp	(cup)	
The boys	pull	the	loz	(nose)	
The boys	hang	the	swɪp	(swing)	
The boys	tie	the	trʌl	(truck)	
The boys	throw	the	mʊk	(book)	
The boys	wrap	the	dar	(car)	
The boys	dress	the	læt	(hat)	

Table 3: Audio stimuli.

Procedure

Each child was tested individually while sitting on a parent's lap. During the test session the parent wore headphones streaming two children's songs simultaneously to mask the audio stimuli. Each child was presented two blocks of either six or eight trials. Each trial was organized approximately as shown in table 4. The trial began with the appearance of the sun, directing the child's gaze to the center, and the audio "Oh, look!" After a short attention-grabbing *ping*, singular and plural videos were then shown side-by-side in silence (baseline phase). When the videos ended, they remained frozen on screen as the sun reappeared, re-centering the child's gaze, and the audio stimulus was heard. The sun then disappeared and the videos played again (test phase). When the videos ended, a playful dinging sounded, and a dance was displayed on the side of the screen that matched the audio stimulus, i.e. the correct side.

This trial organization was seen by 10 of the 18 children. The other eight did not see the re-centering sun during the "Oh, look!", and three of these eight saw the dance in the center of the screen, not on the correct side. Finally, two of the ten children who saw the organization in table 4 had side and number conflated. This variety of presentations was due to methodological modifications made during the course of testing children in light of their ostensibly poor performance.

Each child saw two blocks of either six or eight trials, depending on the number of verbs they knew. The number and side arrangements were based on those used in Legendre et al. (2006). It has been recommended that left-right-right-left or right-left-left-right orderings be used for assigning the correct side to avoid strict alternations (e.g. left-right-left-right) or beginning trial with two of the same correct sides in a row (e.g. left-left-right-right) (Hirsh-Pasek and Golinkoff 1996). While this can be applied to side, it is not possible to apply this simultaneously

to side and number. Within each block there was an equal number of singular and plural stimuli, and the correct video appeared equally often on the left and on the right.

The second block was identical to the first block, except that singular trials became plural, and plural trials became singular. This means that the audio stimuli changed while the videos remained the same (except the dance, which was moved to the opposite side for children who saw the dance on the correct side). Children were generally less attentive in the second block, so only results from the first block will be presented here.

Coding

Testing sessions were recorded using a digital camcorder situated directly above the television, and video recordings were coded offline in Supercoder (Hollich 2005). For each frame, children received a score of 100 if they were looking at the correct side, 0 if they were looking at the wrong side, and no score if they were looking at neither side.

Results and Discussion

Independent samples t-tests were used to determine whether subjects with different presentations could be grouped together. There were no significant differences between the scores of children who saw the dance in the center and those who saw the dance on the correct side (singular baseline: $t(16) = .876$; $p = .394$; singular test: $t(16) = -.802$; $p = .434$; plural baseline: $t(16) = 1.374$; $p = .188$; plural test: $t(16) = 1.064$; $p = .303$). Significant differences were found, however, between children with the side-number conflated presentation and the others (singular baseline: $t(16) = -.146$; $p = .886$; singular test: $t(16) = .425$; $p = .676$; plural baseline: $t(16) = -2.244$; $p = .039$; plural test: $t(16) = -2.865$; $p = .011$), so the data from the children with the conflated presentation is excluded from further analysis. The remaining subjects were compared to see whether those who saw the sun during "Oh, look!" could be combined with those who did not, and no significant differences were found (singular baseline: $t(14) = -.262$; $p = .797$; singular test: $t(14) = -.133$; $p = .896$; plural baseline: $t(14) = 1.865$; $p = .083$; plural test: $t(14) = .101$; $p = .920$). The following analyses, then, include 16 children (9 male and 7 female; ages 24 to 40 months, average 33 months).

A comparison of the average baseline scores and average test scores in singular and plural trails is shown in figure 3. Looking time to the correct side did not increase from baseline to test in singular or plural conditions ($t(15) = .688$; $p = .251$ and $t(15) = .501$; $p = .312$ respectively in one-tailed paired t-tests). For a closer look, frame-by-frame data is presented in figure 4. A trend can be observed in baseline for both singular and plural trials, namely a tendency to look toward the plural side. This is particularly apparent in the latter portions of the baseline, as highlighted in figure 5. This suggests that after children had acquainted themselves with the videos, they preferred to look at the plural video. The true baseline, then, should perhaps ignore the first few seconds of look-around time, which may mask a stronger plural bias. Similarly, it may be helpful to ignore the last few seconds of test. The moment in which the singular and plural videos can be distinguished (e.g. when something is caught by either one or two boys) occurs within the first four (or often even two) seconds of each video, so children may have, in their mind, completed the task within the first few seconds and subsequently allowed their gaze to wander. Though this trend is less apparent in figure 4, it is reflected in the average looking scores between the first four and last two seconds (decreases by 9 in singular and increase by 4 in plural).
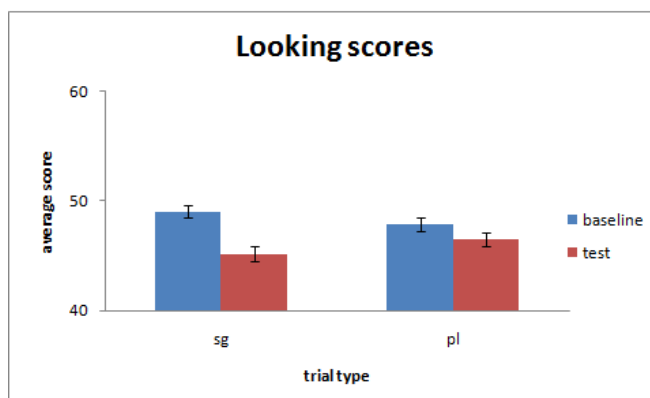
Figure 3: Looking scores in baseline and test from singular and plural trials, shown with plus and minus one standard error bars.
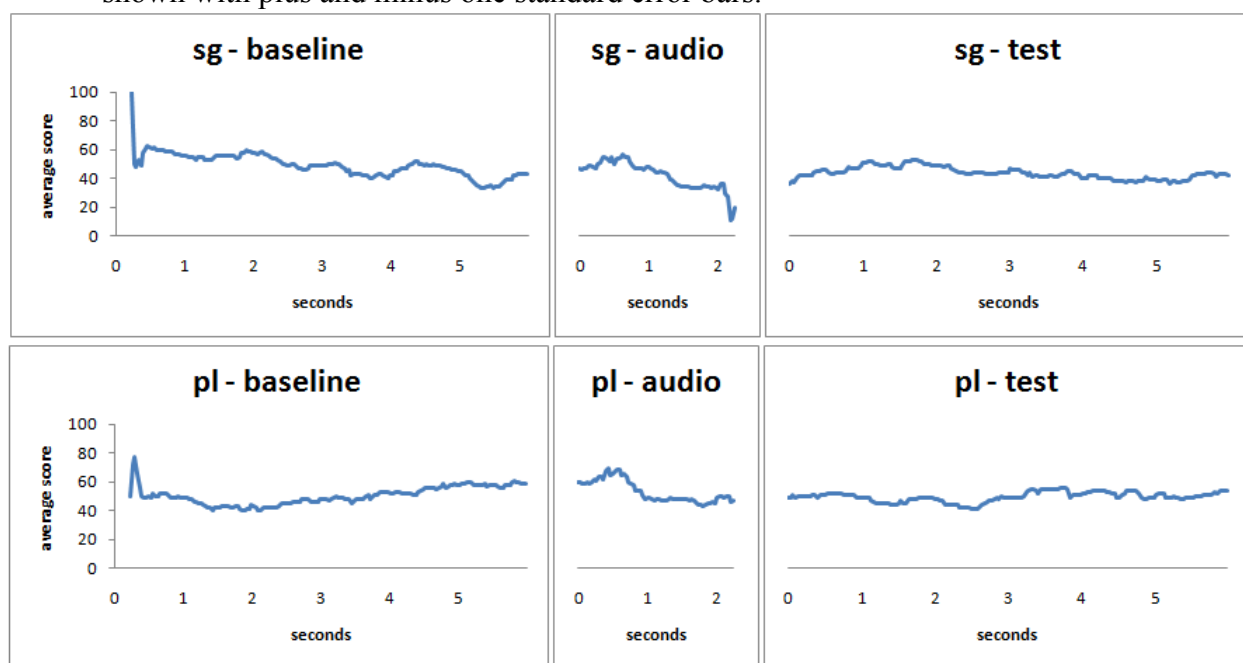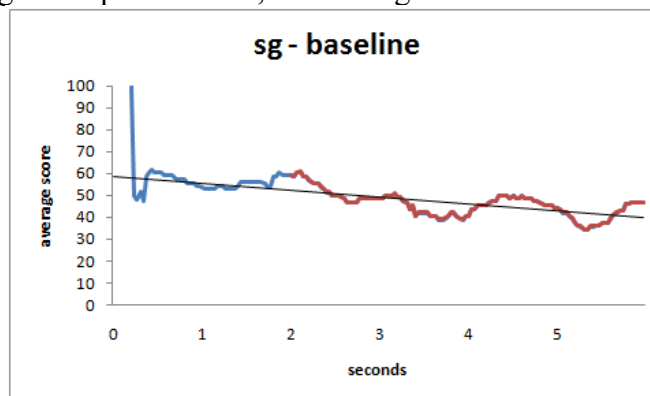

Figure 4: Average looking scores per frame for singular and plural trials in baseline, during audio presentation, and during test.
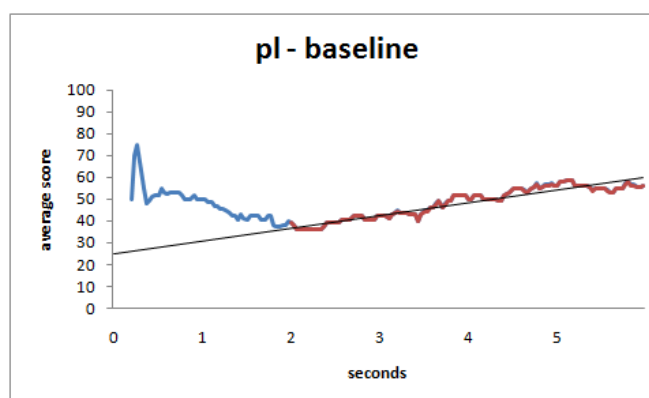
Figure 5: Average looking scores per frame for singular and plural trials in baseline with a linear trend line based on the last four seconds of data (red), exhibiting a trend toward plural.
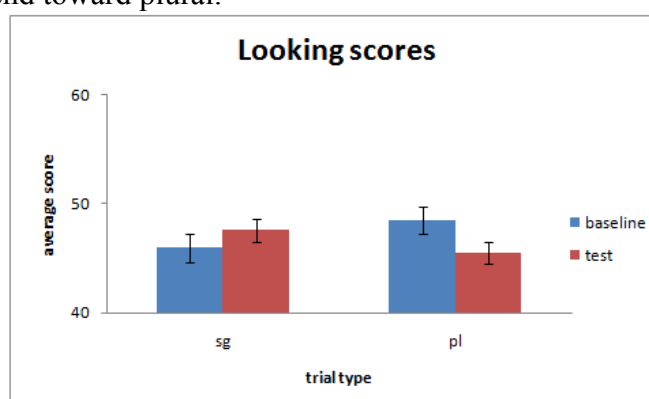

Figure 6: Looking scores in the last four seconds of baseline and the first four seconds of test from singular and plural trials, shown with plus and minus one standard error bars.

In light of this, figure 6 shows averages as in figure 3, but calculated from the last four seconds of baseline and the first four seconds of test. Still, a paired t-test reveals no significant differences (t(15) =-.307; p = .382 for singular, t(15) = .890; p = .194 for plural).

Additional paired t-tests were conducted on the frame-by-frame test data from figure 4, using five-frame bins. No significant increases in looks to the correct side were found either when looking scores in test were compared to the six-second baseline average or to the four-second baseline average. Frame-by-frame analyses were conducted on looking scores during audio, but paired t-tests comparing against baseline scores revealed no significant differences corresponding to meaningful points in the audio (onset of nominal number marking + .675 seconds for planning and executing eye-movements occurred around 1.255 seconds; onset of verbal number marking + .675 seconds for planning and executing eye-movements occurred around 1.742 seconds).

Figure 7 shows scatter plots of subjects' scores, with the difference between baseline and test scores for singular trials plotted on the x axis and the difference between baseline and test scores for plural trials plotted on the y axis. While some children seem to succeed at this task, demonstrated by their position in the first quadrant, they are in the minority. Difference scores

were additionally analyzed in terms of the child's age and CDI information, but no significant correlations emerged.
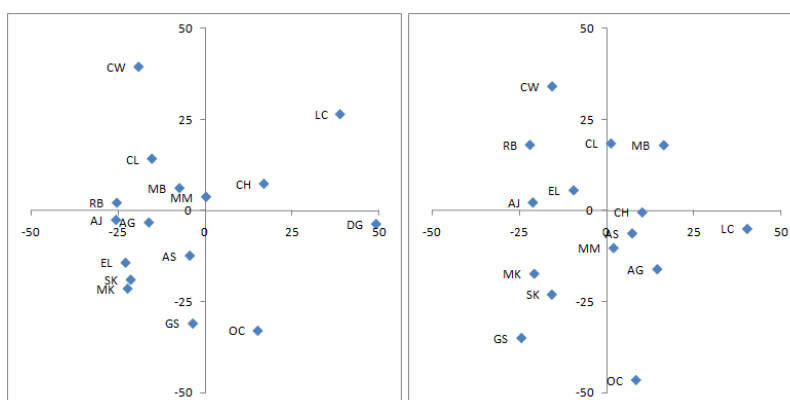

Figure 7: Scatter plot with subjects' average singular difference scores on the x axis and average plural difference scores on the y axis. Difference scores based on six-second averages are shown in the plot on the left, and difference scores based on four-second averages are shown in the plot on the right.

A variety of filters were also applied to the data in an effort to uncover a pattern of comprehension. The reasoning behind these filters was children do not always look at the screen for the entire time the videos are playing, which may cause them to miss the critical moments when the singular and plural videos can be distinguished. Failure to look longer to the correct side during test is less interesting in these cases, and it would be preferable if they could be removed. Finding and removing these trials would be rather impractical, but this can be approximated by removing trials where children did not look at the videos for some minimum amount of time. Several filters were applied selectively to trials where the child looked at least .5 seconds at each in baseline and to trials where no more than 80% of the child's looks were to one video in baseline), but none revealed significant effects. Analyses of first look and longest look were also considered, but did not appear promising.

An analysis was also conducted to see whether the information in the frozen videos during audio may have affected children's looking scores. The videos for "hang", "take", and possibly "throw" and "tie" appeared ambiguous in their frozen form, such that either could have been singular or plural. The other videos, on the other, had indications (e.g. one vs. two boys holding the nonce object) as to whether they were singular or plural. Analyses were redone with only the six remaining non-ambiguous verbs, but did not reveal any signs of comprehension. Rather, the English-acquiring children aged 24 to 40 months in this study did not demonstrate comprehension of subject-verb agreement through a form to meaning mapping between an auditory stimulus with singular or plural subject-verb agreement and a video representation of that stimulus showing one or two actors performing the indicated action.

Failure to find evidence of comprehension may be methodological. Children may not have found the task sufficiently clear or engaging. Also, the videos may have been confusing since the same two boys were shown on both sides on the screen. When children heard audio like "The boys kiss the dite" they may have seen two boys kissing something, but the same two boys were also on the other side and one of them was kissing something else, so children may have been drawn to this contrasting instance of the boys. Given the comprehension success cited by

Fraser et al. and Keeney & Wolfe in their picture selection tasks, it may have also been the case that the videos in this task were too difficult to process, require too much of children's computational resources. These issues seem unlikely, however, given that 30-month-old French-acquiring children showed comprehension with the same videos.

Still, there are considerable differences between the French and English versions of this study. In the English study the videos stay on screen during the audio and in the French version they disappear. The still-present videos in the English version may have distracted children from the audio, causing them to not perform the intended task. It is also possible that having the continuous visual effect of leaving the videos on the screen during the audio somehow lulled children into an idle TV-watching mood, or the disappearance and reappearance may otherwise be important in re-engaging children's attention. Also, the audio in the English and French versions was different in several ways. The French stimuli had pronouns, while the English stimuli had full noun phrases. The full English noun phrases with their higher functional load may have made the task more difficult for reasons of memory or processing, or they may have made the audio less felicitous, since the boys were established early in the discourse and it may be pragmatically odd to continue to refer to them with full noun phrases. The nature of French and English agreement marking may have also affected the results. The agreement marker occurred early in the French audio streams. In the English audio streams, children had to wait longer for the agreement and may not be able to hold the subject in memory long to check the relevant verb features, or children may have chosen to map the noun phrase to the visual display right after hearing the subject, not having heard enough information to properly distinguish the videos.

Failure to find comprehension in this task, however, may run deeper than methodology. English-acquiring 24 to 40 month olds may simply not be able to map from form to meaning for number contrasts in agreement, as suggested by Johnson et al. This possibility is discussed below.

General Discussion

This study found no evidence of an ability to map the form of agreement to its meaning in English-acquiring 24 to 40 month olds. These children, then, do not appear to comprehend subject-verb agreement. As mentioned above, this lack of comprehension may not reflect children's linguistic competence but may instead reflect methodological challenges. There are, however, a number of possible explanations which suggest a lack of linguistic competence.

First and perhaps least interesting is the possibility that the children tested were just slightly too young. For example, 40 to 46 month olds may have been able to succeed at this task. After all, Brown suggests that robust knowledge of form may not emerge until as late as 46 months, so while comprehension may emerge earlier than 5 years, it may not be quite as early as 24 to 40 months. The paucity of individual success as shown in figure 7, however, suggests that this is not the case.

Comprehension of subject-verb agreement may then emerge quite late. Johnson et al. propose a large comprehension/production asymmetry and suggest that production trouble may stem from difficulty extracting singularity from a bundle of person-number features or from determining meaning from an uninterpretable feature. These explanations would be surprising, however, given that a similar study found comprehension in three-year-old German-acquiring

children (Brandt & Höhle, 2008), and there is no apparent reason why German-acquiring children should be generally more meta-linguistically precocious than English-acquiring children. Study 1, however, suggested that adult-like production may be later than claimed. This English comprehension-production asymmetry, then, might not be so large, and late comprehension may simply reflect a general delay in the acquisition of English inflection. This delay may be related to shape of English inflection.

English subject-verb agreement uses phonologically null morphemes in all person-number combinations except third-person singular. Perez-Leroux suggested that null morphemes may be more difficult to learn than overt morphemes. This is supported by results from Johnson et al., where singular inflection (e.g. swim<u>s</u>) was comprehended better than plural inflection (e.g. swim_), and Perez-Leroux, where plural inflection (e.g. nadar → nada<u>n</u>) was comprehended better than singular inflection (e.g. nadar → nada_). The ubiquity of null morphemes in English subject-verb agreement may make it more difficult to learn. For example, null morphemes may be more difficult to detect in the input, making it harder for children to learn the form. Also, the ubiquity of null morphemes leads to a great deal of phonological-form overlap among morphemes which could potentially make the child's task of drawing the correct generalizations more confusing and difficult.

Bittner et al. (2003) provide a similar explanation, whereby the acquisition of verbal morphology is affected by richness, uniformity, and transparency of inflectional paradigms. This puts learners of morphologically impoverished languages like English at a disadvantage, predicting an overall delay in acquisition of verbal morphology. This is consistent with late comprehension in English and suggests that form might not be acquired until later than 26 to 46 months. Partial knowledge of form may emerge early enough to allow the productivity with novel forms seen in Theakston et al., but it may not be fully generalized. Furthermore, this analysis is consistent with the cross-linguistic results presented above. French is morphologically richer than English, particularly if clitics are considered agreement, so it is not surprising that comprehension in French emerges earlier than in English. The same holds for German. And while Spanish may be morphologically richer than German and French, late comprehension evidence in Spanish may be due to methodology, since the German and French studies used an IPLP task while the Spanish study used a picture-selection task. This analysis may allow masking of form knowledge if pressure on computational resources can force children to rely on item-based knowledge instead of non-robust morphological paradigms they have recently generalized. But while Bittner et al.'s account appears plausible, it, like the other accounts discussed, has difficulty distinguishing masked form from an absence of form in English.

Overall, the results from Study 2 suggest that meaning may be associated with subject-verb agreement in English quite late. Form, as suggested in Study 1 and by English's morphological impoverishment, may be acquired later than generally thought, though it remains difficult to determine the abstractness of children's knowledge. For a clearer picutre of whether methodological issues were to blame for the late comprehension seen in study 2 and to better assess the abstractness of children's knowledge of agreement, a further study is currently in preparation.

This new study involves a pointing task, where children first participate in training trials to make clear the task and its reliance on the verb to differentiate videos, which pointing can reinforce. The videos show nonce objects performing nonce actions, with one nonce object performing one nonce action (e.g. one dite spimming) in the singular videos and five nonce

objects performing another nonce action (e.g. five hotes sibbing) in the plural videos. This should eliminate a distributive reading and avoid the problem mentioned above whereby children may have been confused or distracted by seeing the same two actors on both sides of the screen. The audio stimuli will mask the nominal number, as in Johnson et al., so that children will not be able to succeed by relying on nominal number (e.g. "The dite spims"). Furthermore, by preventing reliance on item-based knowledge, the use of nonce objects will offer a view of children's abstract knowledge of form and its associated meaning.

References

Berko, J. (2004). The Child's Leaning of English Morphology. In B. Lust, & C. Foley, *First Language Acquisition: The Essential Readings* (pp. 253-273). Blackwell Publishing.

Bittner, D. Dressler, W. U., & Kilani-Schoch, M. editors (2003). *Development of Verbal Inflection in First Language Acquisition: A Cross-Linguistic Perspective*. Berlin: Mouton de Gruyter.

Brandt, C., & Höhle, B. (2008). Sensitivity to cerbal inflection by German toddlers: Evidence from eyetracking. *IASCL*.

Brown, R. (1973). *A first language: The early stages.* Cambridge, MA: Harvard University Press.

Clark, E., & Hecht, B. F. (1983). Comprehension, Production, and Language Acquisition. *Annual Review of Psychology (34)* , 325-349.

Dale, P. S. and L. Fenson (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers 28*, 125–137.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, 44, 491-505.

Fraser, C., Bellugi, U., & Brown, R. (1963). Control of grammar in imitation, comprehension, and production. *Journal of Verbal Learning and Verbal Behaviour*, 121-135.

Hirsh-Pasek, K. and R. Golinkoff (1996). *Methods of Assessing Children's Syntax*, Chapter The Intermodal Preferential Looking Paradigm: A Window onto Emerging Language Comprehension. MIT Press.

Höhle, B., Schmitz, M., Santelmann, L. M., & Weissenborn, J. (2006). The Recognition of Discontinuous Verbal Dependencies by German 19-Month-Olds: Evidence for Lexical and Structural Influences on Children's Early Processing Capacities. *Language Learning and Development* , 277-300.

Hollich, G., K. Hirsh-Pasek, and R. M. Golinkoff (1998). Introducing the 3-D intermodal preferential looking paradigm: A new method to answer an age-old question. *Advances in Infancy Research 12*, 355–373.

Ingham, R. (1998). Tense without agreement in early clause structure. *Language Acquisition 7*(1), 51-81.

Johnson, V. E., de Villiers, J. G., & Seymour, H. N. (2005). Agreement without understanding? The case of third person singular /s/. *First Language* , 317-330.

Keeney, T., & Wolfe, J. (1972). The acquisition of English agreement. *Journal of Verbal Learning and Verbal Behavior (11)* , 698-705.

Kouider, S., Halberda, J., Wood, J., & Carey, S. (2006). Acquisition of English Number Marking: The Singular-Plural Distinction. *Language Learning and Development* , 1-25.

Legendre, G., Barriere, I., Goyet, L., & Nazzi, T. (2009). Comprehension of infrequent subject-verb agreement forms: Evidence from French-learning children. *DRAFT* .

Levelt, W. M. J. (1989). *Speaking: From intention to articulation*. MIT Press, Cambridge, MA .

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk. Third Edition.* Mahwah, NJ: Lawrence Erlbaum Associates.

Naigles, L. R. (2002). Form is easy, meaning is hard: resolving a prardox in early child language. *Cognition* , 157-199.

Naigles, L., Bavin, E., & Smith, M. (2002). Generalizing novel verbs to different structures: evidence for the importance of understanding meaning. In B. Skarabela, S. Fish & A. H. J. Do (Eds.), *Proceedings of the 26th Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.

Perez-Leroux, A. (2006). Interpreting plurality. *GLOW.* Barcelona.

Phillips, C. (2009). Acquisition at the syntax-semantics interface. *GLOW.* Nantes, France.

Pinker, S. (1991). Rules of Language. *Science, 253,* 530-535.

Soderstrom, M. (2002). T*he Acquisition of Inflection Morphology in Early Perceptual Knowledge of Syntax*. Ph.D. thesis, Johns Hopkins University.

Suppes, P. (1974). The semantics of children's language. *American Psychologist 29*, 103–114.

Theakston, A., Lieven, E., & Tomasello, M. (2003). The role of input in the acquisition of third person singular verbs in English. *Journal of Speech, Language, and Hearing Research*, 863-877.

Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua 106*, 23-79.

Wilson, S. (2003). Lexically specific constructions in the acquisition of inflection in English. *Journal of Child Language 30*, 75–115.

Winitz, H., Sanders, R., & Kort, J. (1981). Comprehension and production of the /-əz/ plural allomorph. *Psycholinguistic Research (10)* , 259-272.