The attention-grammar interface: eye-gaze cues structural choice in children and adults

Journal Article

Running Head: THE ATTENTION-GRAMMAR INTERFACE.

The Attention-Grammar Interface:

Eye-Gaze Cues Structural Choice in Children and Adults

Abstract

We investigated whether children (3- and 4-year-olds) and adults can use the active passive alternation – essentially a choice of subject– in a way that is consistent with the eye-gaze of the speaker. Previous work suggests the function of the subject position can be grounded in attentional mechanisms (Tomlin, 1995; 1997). Eye-gaze is one powerful source of directing attention that we know adults and young children are sensitive to; furthermore, we know adults are more likely to look at the subject of their sentence than any other character (Gleitman, January, Nappa & Trueswell, 2007; Griffin & Bock, 2000). We demonstrate that older children and adults are able to use speaker-gaze to choose a felicitous subject when describing a scene with both agent-focused and patient focused cues. Integrating attentional and grammatical information in this way allows children to limit the degrees of freedom on what the function of certain linguistic constructions might be.

Key words

Attention-Grammar Interface, Eye-Gaze, Subject, Argument-Structure Constructions, Social Cognition.

Introduction

Understanding the interface between attention and grammar is a central concern in many functional linguistic theories (e.g., Langacker, 1991; Talmy, 2007) and an important topic in psycholinguistic research (e.g., Henderson and Ferreira, 2004; Trueswell and Tanenhaus, 2005). Directing attention via referential and perceptual priming causes people to construe a scene in a particular way and typically this is reflected in the linguistic structures people use (Prentice, 1967; Turner and Rommetveit, 1968; Posner, 1980; Tomlin 1995, 1997; see also review by Myachykov et al., 2011). Here we investigate the extent to which eye-gaze influences speakers' construal of a scene in both adults and children. We offer a developmental perspective as children have to learn the particulars of the form-construal mapping in their language and using social-cognitive information such as eye-gaze might be one cue (among many others) as to how this works. First, we introduce the cognitive linguistic perspective on construal and the attention-grammar interface. Second, we discuss the hypothesis that the grammatical notion of subject can be grounded in the cognitive concept of attention. Third, we outline the development of subject knowledge, attention and eye-gaze sensitivity. Finally, we set out how this previous work relates the hypothesis that social-cognitive cues such as eye-gaze might scaffold construction learning.

Construal and the attention-grammar interface

The same state of affairs in the world can be encoded by the use of different linguistic devices to communicate a nuanced range of perspectives. For example:

1a. the roof slopes gently downwards

1b. the roof slopes gently upwards

The two scenes referring to the same roof can be mentally viewed or 'construed' from either above (1a) or below (1b) (cf. Langacker 1988: 62). Languages have of course evolved many different ways to alter how a particular concept is construed in the mind including different structural frames:

2a. the dog chased the cat

2b. the cat was chased by the dog

Note that the scenes in 1a-1b and 2a-2b are truth-conditionally equivalent in the sense that the state of affairs in the world which requires the statements to be true is the same for both. For example, in 1a and 1b there is a roof that exists such that it is angled at *x* degrees. The different expressions are therefore not describing different facts about the world rather they are conventionalized and prefabricated ways of expressing different perspectives. In some functional linguistic approaches (e.g., Goldberg, 2006) more abstract forms such as argument-structure constructions carry their own meaning independent of the items that appear in them; thus the hearer' s attention is still guided to what happened to *the gazzer* in the passive '*the gazzer was mibbed by the pubber'* regardless of the fact that we do not know what the words mean. Contrast this with theoretical frameworks that analyse 2b as the movement of *the cat* from the post verbal position into subject position (e.g., Radford et al., 1999). In the functionalist approach, "surface grammatical form does not conceal a 'truer', deeper level of grammatical organization; rather, it itself embodies the conventional means a language employs for the structuring and symbolization of semantic content" (Langacker 1987: 46-7).

Some of the most extensively investigated attention-directing devices are those components of dynamic events, including those that highlight the conceptual distinctions of

containment/support, path/manner, source/goal and figure/ground (Choi and Bowerman, 1991; Jackendoff, 1983; Lakoff, 1987; Talmy, 1985, 2000, 2007). The notion of figure/ground – whereby some information is highlighted with respect to relatively stable background– has been particularly well studied in relation to the English subject (for a similar conceptual division see Fillmore's (1976) "frame/highlighting" and Langacker's (1987) "base/profiling")

The English Subject and Attention

The subject of an English[1] clause is a 'mosaic' of prototypical coding and behavioural features (Keenan, 1976). For example, the subject typically comes before the verb and triggers agreement with it (e.g., *She smiles*, not *She smile*), has a special pronominal form (e.g., *She smiles*, not *Her smiles*) and entails certain structural properties (e.g., only the subject can leave in *She smiled at him and left*). The subject position is also associated with certain discourse properties[2]; most relevantly for this study, the foregrounding of items that appear in that position and, by definition, backgrounding other items in the clause (depending on the choice of linguistic terminology this function is also been variously referred to as figure-ground, perspective, theme, aboutness and prominence e.g., Talmy, 1985; MacWhinney, 1977; Langacker, 1991). Tomlin has shown that the foregrounding function of subject can essentially be re-described in terms of the cognitive concept of attention (1995, 1997). The key idea here is that attentional mechanisms privilege some forms of information over others by 'gating' perceptual input, sustaining focus on what is

[1] Cross-linguistically speaking, there have been about 30 different grammatical features that have been variously attributed to the concept of "subject" (for example controlling verb agreement, determining the actor in a subjectless second coordinate clause and so on) and any one language 'subject' is only a subset of these features which do not necessarily overlap.

[2] For example, the subject position is associated with given information whereas the object position is associated with new information (e.g., Halliday, 1985).

foregrounded. What is foregrounded then becomes the subject of the sentence (see Figure 1). This is most clearly exemplified by the active passive alternation in English, the function of which allows speakers to focus attention on what the agent did (active) in contrast to what happened to the patient (passive), examples 2a-2b and Figure 1. The utility of defocusing the agent's role in an action or state of affairs has not escaped many politicians, most famously realized in the non-apology apology "*...mistakes were made*…"

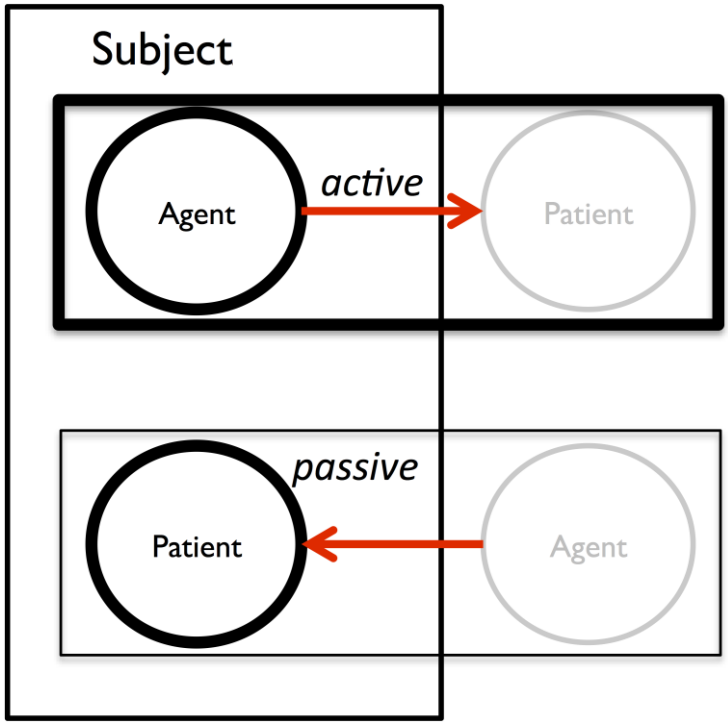

Figure 1. *A schematic of the relationship between the subject role and the active passive voice alternation. Participants that appear in the subject role receive attention. Active sentences are the unmarked form and are relatively more frequent transitive sentences than passives, as indicated by the heavier outline. Arrows indicate the direction of transitivity.*

To test the hypothesis that subject-is-theme-is-attention Tomlin (1995) asked participants to watch prototypical transitive scenes of two fish approaching each other until one swallows the other and swims away. Tomlin manipulated the attention of participants by placing a

flashing red arrow above one or other of the fishes 75 ms before the eating action was completed. Participants were asked to keep their eyes on the character the arrow pointed at and describe what they saw. The majority of adult speakers performed as the subject-is-theme-is-attention hypothesis predicted; on the cued agent animations (the arrow was above the fish that was doing the eating) the agent was assigned the subject position, and the clause was active. On the cued patient animations (the arrow was above the fish being eaten) the patient was assigned subject position and the overall clause was passive. Gleitman, January, Nappa & Trueswell, (2007) obtained similar results even when the cue was implicit (participants were largely unaware of the cue because it appeared so briefly), although with a decreased effect size of passivisation.

Items that appear at the start of an utterance occupy a salient slot and thus could trigger structural organization somewhat independently of grammatical status (MacWhinney, 1977). Because the sentence initial position is confounded with Subject role in English transitive sentences it is difficult to differentiate between a linear-ordering versus a grammatical-role account of the priming effects. To do so we must turn to languages that permit, under certain pragmatic contexts, more flexible word orders. Three recent studies analyzed perceptually primed structural choice in Russian (Myachykov and Tomlin, 2008), Finnish (Myachykov et al. 2010), and Korean (Hwang and Kaiser 2009). Overall the studies suggested that in flexible word-order languages the extent of perceptual priming is consistently weaker than in the fixed word-order languages. Myachykov et al (2010) propose that speakers universally attempt to employ the grammatical-role assignment mechanism in order to represent the perceptually salient referent but this interacts in complex ways with the availability and reliability of the linguistic resources of the particular language. In languages like Russian and Finnish, for example, passives are rare or largely

dispreferred and as a result, a linear-ordering mechanism is used to accommodate referential salience in terms of word order.

The exact nature of the attention-grammar interface is still uncertain and of course, subject position is one attentional cue among many others. For example, in unmarked cases, English tends to correlate theme with given information and subject position, and focus with new information and object position (and usually also prosodic stress). The famous Moses illusion takes advantage of this pattern: when asked "How many animals of each kind did Moses take on the Ark," most people respond "two," even though it can be independently established that they know that it was Noah, not Moses, who took the animals on the Ark (Erickson & Mattson, 1981). The fundamental role of attention in this process is underlined by the fact the illusion can be ameliorated when attention is focused on the incongruent item using structures such as clefts (3a) and there-insertions (3b) (Traxler, 2012)

3a. It was Moses who took two of each kind of animal on the Ark.

3b. There was a guy called Moses who took two of each kind of animal on the Ark.

While acknowledging the role that focus and other salience cues play at the attention-grammar interface, we concentrate here on the subject position for the theoretical reasons outlined above. Recognizing that these and other factors do influence structural choice, the counterbalancing procedure controls for these effects, see Design.

The development of subject knowledge, attention and eye-gaze sensitivity

The overall developmental picture is that children acquire the different features of the English subject gradually and at different times in a 'mosaic' fashion (Rispoli, 1991). Young children identify 'subject' as the most animate participant, or the first-mentioned

participant, or the agent, which suggests something less abstract than an adult-like notion of subject (see Corrigan, 1998 for a review). Experimentally Braine et al. (1993) have shown that mastery over the notion of English subject appears at around 5-6 years of age (see also Corrigan, 1998). Consistent with this, typically children do not produce full passives in spontaneous speech until about 4-5 years-of-age however performance can be significantly boosted when the passive form is supported with case marked pronouns (Ibbotson et al., 2009) with training (Pinker, Lebeaux & Frost 1987; Brooks & Tomasello, 1999) and when the passive form is more frequent in the ambient language (relative to Indo-European languages: Allen & Crago, 1996 for Inuktitut; Suzman, 1985 for Zulu; Pye & Poz, 1988 for K'iche' Mayan). Following Croft (2001) one explanation for this relatively late and piecemeal acquisition pattern is that in reality, abstract constructions such as intransitive, transitive, passive and there-constructions actually have their own subject. They may only be united by analogy later on in development under something like a highly schematic subject-predicate construction (Tomasello, 2003).

In terms of the development of attention, unsurprisingly, younger children are more easily distracted by task-irrelevant events (Enns & Girgus, 1985; Well, Larch, & Anderson, 1980; Carlson, 2005), they search less efficiently for a specified target (Day, 1978; Gibson & Yonas, 1966), they are less able to sustain attention on a given task (Corkum et al., 1995; Kannass and Oakes, 2008) and they are less able to switch efficiently from task to task (Guttentag, 1985; Pick & Frankel, 1974). The general developmental picture for infants is one of increasing cognitive control so that by around their fourth year they have begun to orientate their attention with volition and flexibility (Hughes, 1998; Ruff & Rothbart, 1996). Orienting of attention in this context refers to the alignment of some internal mechanism

with an external sensory input source that results in the preferential processing of that input. The external sensory input source we are most interested in here is eye-gaze.

Humans show a strong sensitivity to eye-gaze from birth (Farroni, Csibra, Simion, & Johnson, 2002). Neonates can follow gaze if the pupils are seen to move (Farroni, Massaccesi, et al., 2004) and at around 5-months-of-age they can discriminate between very small horizontal deviations (5°) of eye gaze (Symons et al., 1998). Like adults, infants process facial features in deeper way when gaze is directed towards them as compared with averted gaze (Farroni et al., 2002; Farroni, Johnson, & Csibra, 2004; Farroni, Massaccesi, Menon, & Johnson, 2007; Hood et al., 2003). Clearly, the capacity to use another person's eye gaze as a cue to attention develops very early in life, however, to begin with this might be achieved with rather low-level, low-mentalising mechanisms, for example, the perceptual geometry and luminance of the eye (Ando, 2002). Compared with other primates, humans have a relatively large white sclera surrounding a small dark pupil and iris making eye-gaze discrimination relatively easier in humans than in other animals (Kobayashi & Kohshima, 1997). Supporting the low-mentalising interpretation of eye-gaze sensitivity, a wide range of species have a very accurate ability to determine whether they are being looked at (e.g., Burger, Gochfeld, & Murray, 1992; Burghardt & Greene, 1990; Perrett & Mistlin, 1991) and nonhuman primates such as adult rhesus monkeys can discriminate between photographs depicting direct gaze and gaze averted by 5°, the same ability that has been reported in human infants (Campbell, Heywood, Cowey, Regard, & Landis, 1990; Symons, Hains, & Muir, 1998).

Although infants' eye-gaze sensitivities may be based on relatively simple mechanisms (gaze perception), young children soon begin to integrate eye-gaze information into a more sophisticated picture of how other people work including their future intentions

and mental states (Baron-Cohen, 1994; see Striano & Reid, 2006, for a recent review). The development of joint attention at 20 months predicts theory of mind abilities at 44 months (Charman et al., 2001) underlining eye-gaze as a key component in the development of social cognition in early life. Baron-Cohen, Campbell, Karmiloff-Smith, Grant, and Walker (1995) found that children ages 3 and 4 years old deduce the direction of gaze of a schematic face and they can ascribe mental states such as desires on the basis of the direction of gaze (see also Lee, Eskritt, Symons, & Muir, 1998). Thus, understanding that direction of gaze can indicate which objects a person knows exists, is currently attending to, and holds a mental state about can help a child infer much about the current visual world (although this understanding may not be as flexible as adults when cues conflict e.g., Friere, Eskritt, & Lee, 2004; Pellicano & Rhodes, 2003; Pellicano and Rhodes, 2003).

Eye-gaze following at 6 months has been shown to correlate with vocabulary size at 18 months (Morales et al., 2000; Morales, Mundy, & Rojas, 1998) and in noun learning, children can use eye-gaze, head posture and gesture to infer speakers' referential intention (e.g., Baldwin, 1991; Carpenter , Akhtar & Tomasello, 1998; Gergely, Bekkering & Király, 2002; Woodward & Sommerville, 2000). Nappa, Wessel, McEldoon, Gleitman & Trueswell (2009) showed that 3-, 4- and 5-year-olds used the eye-gaze of the speaker to infer the meaning of novel relational verbs (of the type *chase* vs. *flee*) in linguistically uninformative contexts (e.g., *He's mooping him*). Thus children who saw a speaker looking at the *chaser* when they uttered the novel verb were more likely to attribute 'chase-like' semantics to the novel verb. The opposite effect was found when a speaker looked at the *flee-er*.

Change in gaze direction is one of several behavioral cues that individuals use in combination with changes in facial and vocal displays and body posture to mark the

intention to act on an object (Mumme et al., 2007). Crucially, just prior to speaking, adults are more likely to look at the subject of their sentence than any other character (Gleitman, January, Nappa & Trueswell, 2007; Griffin & Bock, 2000). This raises the possibility that children could use this cue in the input, probabilistically at least, to build a correspondence between the perspective of an event and how that perspective is expressed in their language (cf. Nappa et al., 2009). Just as verbs such as *chase* and *flee* can lead to different construals of the same (perceptual) scene so can argument-structure constructions like the active-passive alternation, which are basically perspective-taking devices.

In summary, the above evidence raises the possibility that young children could use the social-cognitive cue of eye-gaze – which directs attention – to infer the function of grammatical subject – which is grounded in attention. We investigated this by exploring if they can use the active passive alternation (essentially a choice of subject) in a way that is consistent with the eye-gaze of the speaker. The developmental hypothesis is that if the function of subject position is grounded in attentional mechanisms (Tomlin, 1995; 1997), then we would expect that developing attentional abilities should interact with developing linguistic ability to assign a subject. Thus we would expect different age groups to perform differently. Our second hypothesis is more of a general theoretical point and is relevant to adults and children. As the evidence we have reviewed suggests, some cognitive linguistic frameworks suggest a close relationship between attention and linguistic performance (e.g., Givon 1992; Landau and Jackendoff 1993; MacWhinney 1977; Osgood and Bock 1977; Talmy 2007). This general assumption leads to the specific hypothesis that a speaker's attention can be guided by the eye-gaze of their interlocutor and that this in turn affects the structural choice of their utterances, specifically whether to use an active or passive sentence.

Methods

Participants

A total of 91 participants were tested. For each age group we report the numbers in each group (*N*) the mean age in years (M) and the standard deviation of the group as calculated on months-of-age (SD). For the Test condition: 3-year-olds (*N*=21, M=3.55, SD±3.66), 4-year-olds (*N*=23, M=4.39, SD±2.79) and adults (*N*=11, M=23.31, SD±87.60). For the Control Condition: 3-year-olds (N=13, M=3.45, SD±3.63), 4-year-olds (N=12, M=4.38, SD±3.60) and adults (N=11, M=19.08, SD±41.25). 7 children were excluded from the analysis because of fussiness (4) or not producing a novel verb (3) – for example they persisted with "*the dog's hitting the cat*" even after correction, instead of the intended "*the dog's tamming the cat*".

Design and Materials

The between-subjects factors were Age (3-year-old, 4-year-old and adults) and Condition (Test, Control) and the within-subjects factor for the Test Condition was Gaze (Agent-Focus, Patient-Focus). Ideally, to increase the power we would have treated each child as its own control with a within-subjects design. We assumed that most children and adults would provide answers in the experiment with an active voice if not cued to do otherwise (as indeed they did). We reasoned that if we ran a cued and non-cued trial with the same individual there was a risk that we would structurally prime the active response to such a degree that the already weakly represented passive would not get produced at all. We were interested in the degree to which participants could switch between active and passive responses in a way that was congruent with eye-gaze. By the time a participant answered with 8 actives responses (as they most likely would have done in the test condition) it would

be much harder to elicit a passive than if they have just finished our training phase (see below). Likewise, running the control after the test might have artificially boosted passive responses in the control condition. Means and standard deviations of age between Test and Control were comparable (see Participants). Increasing the activation levels/ease of retrieval of the either active or passive was avoided by keeping control and test participants separate. Although this reduced the power, as it turned out the effect was still strong enough to detect.

If participants came to the experiment with no preference for describing a scene with either actives or passives and responded at random then we could expect the proportion of congruent answers to be 50%, that is, an active response to agent-focused question and a passive response to a patient-focused question. If participants came to the experiment with a preference for describing a scene with only actives then again we could expect the proportion of congruent answers to be 50% – participants score 4/4 on the active-focused questions but 0/4 on passive focused questions. The design therefore takes account of baseline preference with respect to congruency. Despite this methodical treatment of participant's preference we still ran a control to check there was nothing strange about the puppets used or experimental set-up and to measure the different age groups' 'default' response that might vary for whatever reason. We therefore present the control data alongside the test data in the results section as it impacts on how easily participants can switch from their 'default' response to an eye-gaze cued response.

There were 8 test questions per participant, with an equal number of Agent-Focused and Patient-Focused trials, the orders of which were randomised. Over the experiment, the order of appearance of model sentences (training phase, see below), the left/right appearance of the participant in focus and the participant acting as agent were counterbalanced. There were six hand puppets used, randomly assigned to either the training phase or the test phase.

Referential salience and conceptual accessibility affect the degree to which a participant might attend to a particular character. Care was taken to match the puppets on size, general colourfulness and animals that could be considered prototypical instances of their own kind, (a monkey, a pig, a horse, a dog, a cat, and a bear). All actions were reversible transitives and the experimenter took care to make the actions as comparable as possible between trials (as far is possible with the act-out method). Importantly however, the counterbalancing procedure controls for salience. For example, if participants considered the pig more salient than the cat and correspondingly chose this as the subject, the pig appeared both as the focus of attention in active and passive conditions, and the unfocused participant of active and passive conditions over the course of the experiment.

Procedure

Training Phase

The experimenter modelled novel verb, (*tamming*) using animal hand puppets. This was done twice in the active voice (e.g., "*look! the horse is tamming the goat*") and twice in the passive voice (e.g., "*look! the frog is getting tammed by the cow !*"). After each model the experimenter asked the child or adult to repeat the sentence (e.g., "*can you say the horse is tamming the goat?*"). The experiment did not progress until the participants had correctly repeated all four of the model sentences at least once. The action consisted of an agent hitting the top of a patient's head with a cardboard tube (see Figure 1). We used the cardboard tube to increase the distance between participants and thus make it easier to determine whether the experimenter was looking at the agent or patient (yet still remain a causal transitive scene). Throughout the training phase, the experimenter did not use eye-gaze to direct attention onto either of the characters.

Test Phase

The training puppets were abandoned and four new hand puppets were introduced. The experimenter asked the participant to name the animals, which all did correctly. The experimenter then said "*we're going to play a different game now, this time you have to tell me what the animals are doing*". The experimenter looked at the participant, engaged eye contact, looked at the target animal (either the agent or patient of the impending action) looked back at the participant and then finally looked back to the target animal again. Thus the idea was to establish triadic joint-attention between speaker, addressee and referent. The experimenter then performed the action while continuing to look at the target referent. Shortly after the action had begun the experimenter asked, "*what's happening*?" If the participant did not respond, the experimenter persisted for a while with the same question taking care not to linguistically cue them into an active or passive response (e.g., "*what is the cat doing*"? or "*what's happening to the dog*"?) – thus we tried to create an analogous situation to the Nappa et al. 'linguistically uninformative' condition for verb learning (2009). Children's and adult's responses were recorded and later coded for active or passive voice and the order in which participants were mentioned. We also recorded whether the subject of the sentence (either agent or patient) was congruent or incongruent with the eye-gaze of the speaker/experimenter. For example, if the speaker was looking at the agent of the action and the participant responded with the agent in subject position (i.e., active: *agent-V-patient*) it was coded as congruent. If the participant responded with the patient in subject position (i.e., passive: *patient-Aux-V-(agent))* it was coded as incongruent. It is important to note that either an active or a passive answer to the question "*what's happening*" is equally grammatical with respect to who is doing what to whom (assuming the arguments map onto what is actually happening in the scene). We are interested in what could be considered the

most pragmatically felicitous response (following adult performance in Tomlin, 1995; 1997 and anticipated adult performance in this study), namely those responses where the subject of the utterance maps onto where the attention has been directed, in this case via the eye-gaze of the speaker. Adults and children experienced the same experimental procedure with the same materials. While the hand-puppet procedure might have appeared strange to some adults it was necessary to make the results directly comparable – if we had not used hand puppets any difference between adults and children could be attributable to the different methodologies used.

The control condition contained a procedure that was identical to that described above except for the fact that the experimenter did not attempt to focus the attention of the child or adult onto any participant before the action began and looked straight ahead when asking "what's happening?" (Figure 2).

Agent-Focus Control Patient-Focus

Figure 2. *In this example 'the cat is tamming the dog' or, depending on your perspective, the 'dog is getting tammed'. The eye-gaze of the experimenter cues attention towards either the agent of the action (left), neither agent nor patient (centre) or the patient of the action (right).*

Results

First we present a comparison between the Control Condition and the Test Condition. We underline that the control condition was identical to the test condition except for the fact that the experimenter did not attempt to focus the attention of the child or adult onto any participant before the action began and looked straight ahead when asking "*what's happening*?" The results are presented in terms of Eye Gaze (Test Condition) and No Eye Gaze (Control Condition) as a proportion of active responses, displayed in Figure 3 below.

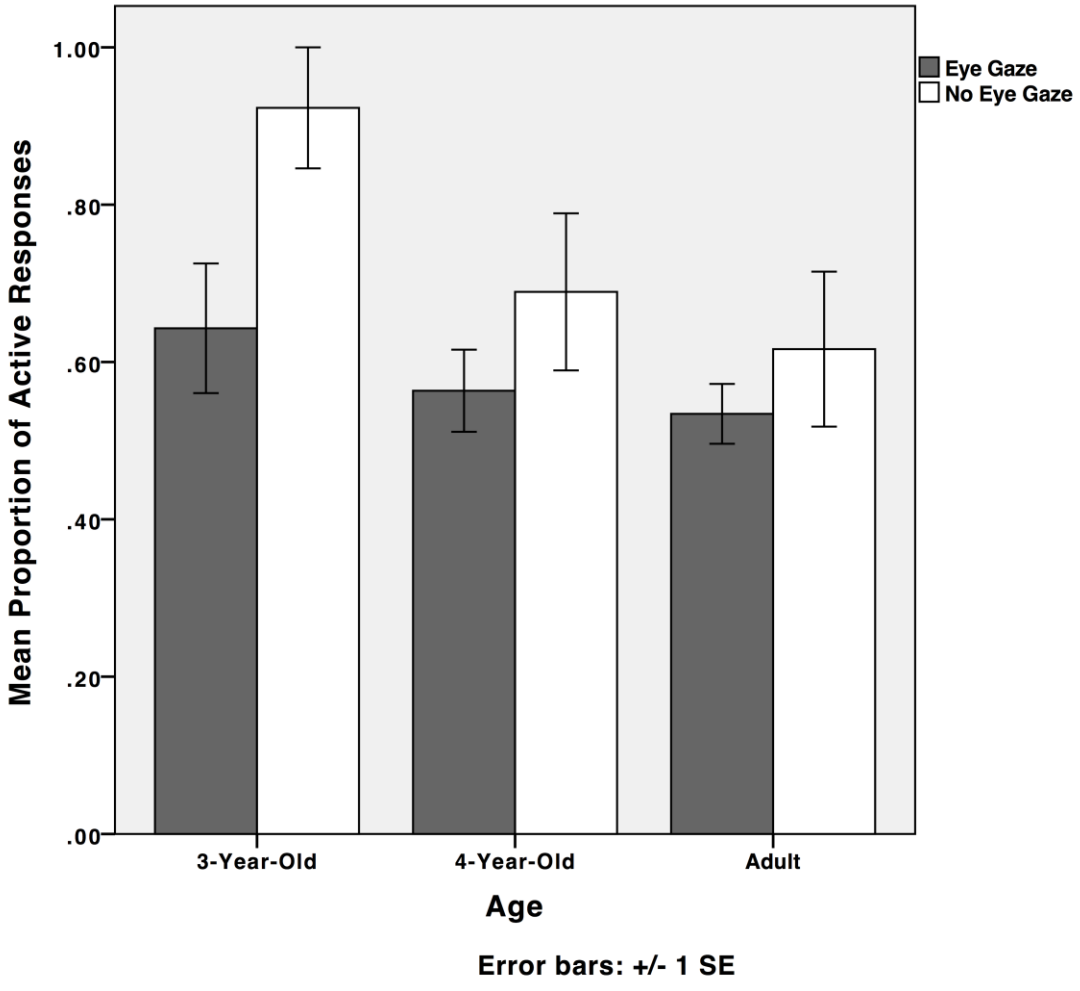

Figure 3. *Mean Proportion of Active Responses by Age Group.*

To determine whether there was a significant difference between Test and Control conditions with respect to the mean proportion of active utterances we conducted a 2 (Eye Gaze Focus, No Eye Gaze Focus) by 3 (3-Year-Olds, 4-Year-Olds, Adults) ANOVA with mean proportion of active responses as the dependent variable. There was a significant effect of Condition $F(1,85) = 6.07$, $p=.016$ $\eta_p^2 = .067$, Age $F(2,85) = 3.68$, $p=.029$ $\eta_p^2 = .08$ and no significant interaction between Age and Condition $F(2,85) = 0.85$, $p=.428$ $\eta_p^2 = .02$. This means Eye Gaze significantly affected the active/passive ratio for all age groups. What is most relevant for the discussion is that the youngest age group show the strongest preference for describing the scene (in the absence of eye-gaze focus) with active sentences. This means the youngest age group has the strongest 'default' to overcome if they are to switch to describing the scene with a passive.

We now present the analysis of the Test data with respect to the mean proportion of congruent responses to eye-gaze (obviously we cannot analyse the control data in this way as there is no eye gaze manipulation in the Control – note the change in y-axis from Figure 3 to 4). Figure 4 shows the mean proportion of congruent responses as a function of whether the experimenter was looking at the agent of the action or the patient of the action for different age groups.

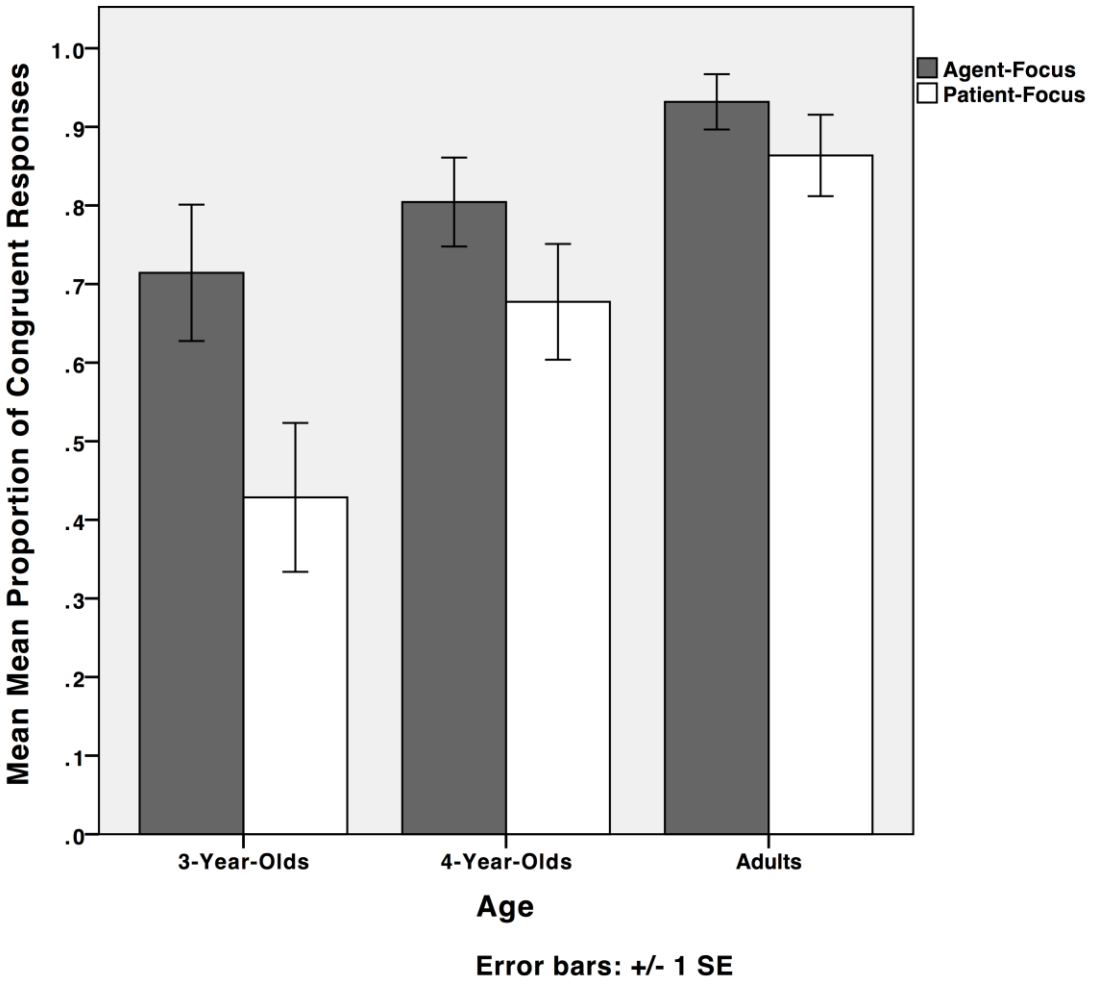

Figure 4. *Mean Proportion of Congruent Responses as a function of Eye-Gaze and Age. Note the change in y-axis from Figure 3.*

Overall Congruency Comparisons

To determine the overall effect of eye gaze on congruent responses, we first conducted a 2 (Congruent, Incongruent) by 3 (3-Year-Old, 4-Year-Old, Adults) mixed model analysis of variance (ANOVA). This showed a significant main effect of Congruency $F(1,52) = 97.94$, $p<0.001$, $\eta_p^2 = .653$ (M .737 > .263) and a significant interaction between Congruency and Age $F(2,52) = 14.36$, $p<0.001$, $\eta_p^2 = .356$. This means that overall there were more

congruent responses than incongruent. The interaction suggests that the difference between congruent and incongruent responses was not the same across all age groups. To analyze this interaction further we conducted an ANOVA by Age. We found significant differences between congruent and incongruent responses for 4-Year-Olds $F(1, 22) = 36.59$, $p<0.001$, $\eta_p^2 = .625$ and Adults $F(1, 10) = 306.25$, $p<0.001$ $\eta_p^2 = .968$ but not for 3-Year-Olds $F(1, 20) = 3.50$, $p = 0.07$ $\eta_p^2 = .149$. This means that regardless of the type of eye-gaze focus (agent/patient) only the 3-Year-Olds were unable to consistently match a congruent response. We now analyze how this overall effect of eye-gaze on congruency is influenced by whether the gaze is focused on the agent or patient.

Agent-Focused

We analyzed Agent-Focused trials with a 2 (Congruent, Incongruent) by 3 (3-Year-Old, 4-Year-Old, Adults) mixed model analysis of variance (ANOVA). There was a significant main effect of Congruency $F(1, 52) = 52.49$, $p<0.001$ $\eta_p^2 = .502$ (M .817 > .183) and no interaction between Congruency and Age $F(2, 52) = 1.82$, $p=0.172$ $\eta_p^2 = .066$. This means all age groups consistently produced active sentences when the experimenter was looking at the agent.

Patient-Focused

Second, we analyzed Patient-Focused trials with a 2 (Congruent, Incongruent) by 3 (3-Year-Old, 4-Year-Old, Adults) mixed model ANOVA. There was a significant main effect of Congruency $F(1, 52) = 9.25$, $p <0.001$ $\eta_p^2 = .151$ (M .657 > .343) and a significant interaction between Congruency and Age $F(2, 52) = 5.72$, $p<0.001$ $\eta_p^2 = 180$. The interaction suggests that the difference between congruent and incongruent responses for Patient-Focused trials was not the same across all age groups. To analyze this further we

conducted an ANOVA by Age. We found significant differences between congruent and incongruent responses for 4-Year-Olds $F(1, 22) = 5.80$, $p=0.025$, $\eta_p^2 = .209$ and Adults $F(1, 10) = 49.23$, $p<0.001$ $\eta_p^2 = .831$ but not for 3-Year-Olds $F(1, 20) = 0.56$, $p=.460$, $\eta_p^2 = .028$. This means only the 3-Year-Olds could not consistently produce passive sentences when the experimenter was looking at the patient.

Finally we report that for both Control and Test conditions the vast majority of passive responses from children and adults were truncated (e.g., "*the cat got tammed*") rather than full passives (e.g., "*the cat got tammed by the dog*"), Table 1.

Table 1. Proportion of Full Passives by Age Group and Condition (%)

	Test	Control
3-Year-Olds	4/168 (2.38)	0/104 (0)
4-Year-Olds	11/184 (5.97)	2/96 (2.08)
Adults	19/88 (21.59)	12/88 (13.63)

This did not have a significant impact on the main analyses as we are primarily interested in whether the (correct, i.e., congruent) agent/patient appeared in the subject role as a function of speaker's eye gaze. In principle there are other ways of describing the scene, where the subject would be congruent with eye-gaze but the structure would not be active or passive, such as *Dog is hurting from cat tamming it*. We did not witness any examples of these types of sentences, perhaps because the training phase primed/reduced the degrees of freedom on what participants thought is an appropriate response to the question "what's happening?", i.e. an active or passive was modeled by the experimenter in the training phase (and participants were asked to repeat them) and once participants had begun to answer with these options they continued to do so. In sum, all ages gave more congruent responses than

incongruent responses to both agent-focused and patient-focused trials, except for the 3-Year-Olds with patient-focused trials.

Discussion

We set out to investigate whether children can use the active passive alternation – essentially a choice of subject/focus – in a way that is consistent with the eye-gaze of the speaker – a task that requires integrating social and syntactic knowledge. We established that 4-Year-Olds and adults are able to use speaker-gaze to choose a felicitous subject when describing a scene; 3-Year-Olds can do so consistently only when the focus is on the agent. Our interpretation of how participants achieve this performance and of the relatively impaired performance of the 3-Year-Olds on the passive follows. To succeed at this task is not trivial. We don't know for sure all the steps needed to produce the appropriate response but at the very least, the following processes are all credibly involved.

First, both children and adults need to understand that following gaze establishes reference. Subcomponents of this ability are recognizing that looking is intentional behavior directed to external objects and events; that looking results in the mental experience of seeing an object or event; and that others share in the capacity to see things (D'Entremont, et. al., 2007). Second, participants need to coordinate where their attention has been focused with a linguistic representation. This involves selecting the construction that best serves the function of foregrounding a participant, which in this case is the subject position (perhaps something like Fig. 1). Part of this ability requires suppressing the preferred information structure and most heavily entrenched form (the active) when one needs to describe a scene from the perspective of the patient. There is evidence that this maybe more of a challenge for the 3-Year-Olds because they have the strongest preference for describing the scene with an active in the absence of social cues (see Control Condition). In addition, success on the

experiment not only requires attention for a given trial but the ability to switch attention between trials. Thus inhibitory control, attentional flexibility, and working memory are all implicated in giving the correct response, which as noted in the introduction, are still in development around 4 to 5 years-of-age (Hughes, 1998). Finally, participants need to produce a string of nouns, verbs and auxiliaries that not only satisfy the grammatical requirements of who did what to whom but also conform to the social-pragmatic demands of the context.

Salamo, Lieven and Tomasello (2010) note that children have more difficulty giving pragmatically appropriate responses to sentence-focus questions of the type "*what's happening*?" than either argument-focus questions "*who is VERB-ing?*" or predicate-focus questions "*what is AGENT doing*?" This pattern corresponds with sentence-focus questions being relatively less frequent to the other types in child directed speech. In this experiment we deliberately chose an unbiased question (linguistically-speaking) to assess the role of social cueing in isolation (i.e., "*what's happening*?"). To the extent that sentence-focus questions are more pragmatically difficult, this seems to affect the youngest age group in this study the most. Again, one possible explanation for this is in terms of the development of underlying domain-general capabilities – the idea is that argument-focus questions and predicate-focus questions help to anchor the relevant piece of information (either verb or agent) in short term-memory, from which the appropriate response to the question is constructed (an advantage which is not present in sentence-focus questions). Although more infrequent in child directed speech, sentence-focus questions may benefit the most from the support of social cues precisely because they are linguistically uninformative. Indeed Nappa et al., (2009) found the strongest effect of social cues on verb learning in the linguistically uninformative condition. In reality children face a complex matrix of social and linguistic

cues and they have to work out how reliable these are across multiple situations. Using the design in this study, one might predict an even stronger effect of subject choice alternations (and perhaps at younger age) where linguistic anchoring (e.g., "*what's happening to the cat*?") acts in coalition with social cues (e.g., speaker looks at cat).

As noted in the introduction, in English the active is the preferred information structure for describing a causal transitive scene: actives outnumber passives 9:1 in written text and the ratios are even higher for spoken discourse. Not only this, the token frequency for passives is low. In one analysis of child directed speech, the passive-per-utterance rate is 0.36% (Gordon and Chafetz, 1990). Thus all age groups have had relatively less experience with the passive than the active, and the results show that there are more congruent active responses than passive responses for all age groups. The active construction also aligns with a non-linguistic preference for conceptually framing events as source-to-goal or instigator-patient rather than the reverse (Lakusta and Landau, 2005; Lakusta, Wagner, O'Hearn & Landau, 2007). Nappa et al. (2009) note that in many languages, verbs that align with this preference (e.g., *chase*) outnumber their asymmetric partners (e.g., *flee*). They also note the instigator-patient bias is present pre-linguistically (Csibra, Biro, Koos, & Gergely, 2003; Woodward, 1998) and continues through to adulthood (Gleitman et al., 2005). Increased experience (use) of the passive with time, particularly with a literate undergraduate population from which the adult sample was mainly drawn (see Dąbrowska, 2006), acts in opposition to this tendency. This experience would make it relatively easier to retrieve the passive construction from memory when it is required (i.e., on the patient-focused trials). Also, in the design of this experiment, all participants were explicitly shown in the training phase (and perhaps reminded in the case of adults) that actives *and* passives were legitimate ways to describe the scene. This part of the design may have increased the conceptual

accessibility (Osgood & Bock, 1977) of the passive and boosted the level of passives description from that which we would normally see in spontaneous speech.

Experimentally it has been shown children around their second year of life can use the attentional and emotional states of speakers as cue to their communicative intentions (e.g., Baldwin, 1991; Tomasello and Farrar, 1986). One might wonder why it takes participants in this study until their fifth year of life to link the linguistic and the social domains in a way that allows them to succeed on both agent and patient focused trials. As discussed above, the complexity of the task and the relative frequency of the passive construction hints at one answer but another possibility, as mentioned in the introduction, is that it takes time for young children to integrate eye-gaze information (as a basic perceptual cue) into a more sophisticated picture of how other people work including their future intentions and mental states (Baron-Cohen, 1994; see Striano & Reid, 2006, for a recent review). The difference between the 3 and 4 year olds in our experiment might also be partly attributable to development in this social-cognitive capacity. Without further tests that can disassociate the lower-level perceptual cues from intentional eye-gaze cues (for example a non-voluntary eye-gaze shift) it is not clear to what extent the children in this experiment are using in the intentions of the speaker to assign a subject. For that reason we use the term the attention-grammar interface rather than the social-grammar interface.

It is also worth noting that the experiments that have focused on the relationship between attentional states of speakers and communicative intentions (e.g., Baldwin, 1991; Tomasello and Farrar, 1986) have mainly focused of word learning. The adult-like function of subjecthood makes it a much more abstract and less concrete learning challenge than learning words, which, in the usage-based framework at least, means that mastery of subjecthood requires more evidence and more experience with using it. While the 3-year-

olds in this study probably possess many of the social-cognitive foundations that the 4-year-olds do, success on the task needs competence in linguistic and executive control domains as well as effective connections between these domains.

Priming studies show that children already have some representation of the passive younger than 4 years-of-age (e.g., Huttenlocher, Vasilyeva & Shimpi, 2004). One might expect that the 3-Year-Olds in this study might be able to succeed on both agent-focused and patient-focused trials. However, the factors we have reviewed above – the steps needed to succeed at this task, the relative frequency of the passive, the conceptual bias for instigator-patient, the relative difficulty with sentence-focus questions, the developing executive control, the abstractness of 'subject' relative to word learning – could all disproportionally affect the youngest age group.

The performance of the 3-, 4-year-olds and the adults in this study provides further support for the subject-is-theme-is-attention attention hypothesis (Tomlin 1995; 1997). What is more, we have demonstrated this using a more ecologically valid cue than a red arrow hovering above a participant, namely, eye-gaze. Our interpretation of these results is that eye-gaze plays a similar role to the flashing arrow in the Tomlin studies, namely, it is an attention-directing cue that foregrounds one character and, by definition, backgrounds the other. Importantly, we know adults are more likely to look at the subject of their sentence than any other character (Gleitman, January, Nappa & Trueswell, 2007; Griffin & Bock, 2000). This raises the possibility that young children could be using the social cue of eye-gaze in situ (which directs attention) to infer the function of grammatical subject (which is grounded in attention).

More generally, the methodology we have used here advocates exploring linguistic cues in combination with the social-pragmatic context. By using eye-gaze we have been

able to consider a broader range of cues than a traditional corpus-based approach to the development of language. By doing so, we have been able to get closer to reconstructing the rich social-pragmatic-linguistic world in which the child actually grows up. The challenge is to explore ways in which social-pragmatic skills interact with prodigious pattern-finding abilities in a way that which explains the emergence of linguistic knowledge. Yu & Ballard (2007) found that a computational model of word-learning performed better when it used social information (joint attention and prosody) in combination with statistical cues (cross-situational learning) than when it relied on purely statistical information alone. There is no reason to doubt that a combination of social and linguistic cues also helps in learning syntactic relations (e.g., Nappa et al., 2009).

We noted in the introduction that, typically, children do not produce full passives in spontaneous speech until about 4-5 years-of-age however performance can be significantly boosted when the passive is supported with case marked pronouns (Ibbotson et al., 2009) with training (Pinker, Lebeaux & Frost 1987; Brooks & Tomasello, 1999) and when the passive form is more frequent in the ambient language (relative to Indo-European languages) (Allen & Crago, 1996; Suzman, 1985; Pye & Quixtan Poz, 1988). To this list we might add that felicitous passive use is also boosted when supported with the kind of social-cognitive cues that children might actually experience, such as eye-gaze.

Our understanding of children's social cognition has progressed significantly, however, it has yet to be worked through in any detail how this knowledge interacts with emerging syntactic representations. This paper is a step in that direction and investigates whether eye-gaze influences the choice of grammatical subject for young children and adults. The linguistic notion of subjecthood is in principle an abstract one. Previous work suggests the function of the subject position can be grounded in terms of attention and

information structure. One powerful source of directing attention that we know young children are sensitive to is eye-gaze. We have demonstrated that 4-year-olds and adults (but not conclusively 3-year-olds) are able to use speaker-gaze to choose a felicitous subject. It has been shown before that social-cognitive cues help children learn words but this is the first demonstration that eye-gaze could be important in learning something as abstract as subject role. Integrating attentional and grammatical information in this way allows children to limit the degrees of freedom on what the function of certain linguistic constructions might be (cf. Nappa et al., 2009) and allows linguistic theory to ground abstract functions in deeper cognitive and communicative principles (Tomlin, 1995,1997; Goldberg, 2006; Langacker, 1991).

References

Allen, S. & Crago, M. (1996). Early passive acquisition in Inuktitut. Journal of Child Language 23(1). 129-155.

Ando, S. (2002). Luminance-induced shift in the apparent direction of gaze. *Perception, 31,* 657–674.

Baldwin, D. A. (1991). Infant contribution to the achievement of joint reference. Child Development, 62, 875–890.

Baron-Cohen, S. (1994). How to build a baby that can read minds: Cognitive mechanisms in mindreading. *Cahiers de Psychologie Cognitive* [*Current Psychology of Cognition*]*, 13,* 513–552.

Baron-Cohen, S., Campbell, R., Karmiloff-Smith, R., Grant, J., & Walker, J. (1995). Are children with autism blind to the mentalistic significance of the eyes? *British Journal of Developmental Psychology, 13,* 379–398.

Braine, M.D.S., Brooks, P.J., Cowan, N., Samuels, M.C., & Tamis-LeMonda, C. (1993). The Development of categories at the semantics/syntax interface. Cognitive Development, 8, 465-494.

Brooks, P. J. and Tomasello, M. (1999). Young children learn to produce passives with nonce verbs. Developmental Psychology, 35(1), 29-44.

Burger, J., Gochfeld, M., & Murray, B. G. (1992). Risk discrimination of eye contact and directness of approach in black iguanas (*Ctenosaura similis*). *Journal of Comparative Psychology, 106,* 97–101.

Burghardt, G. M., & Greene, H. W. (1990). Predator simulation and duration of death feigning in neonate hognose snakes. *Animal Behaviour, 36,* 1842–1844.

Campbell, R., Heywood, C. A., Cowey, A., Regard, M., & Landis, T. (1990). Sensitivity to eye gaze in prosopagnosic patients and monkeys with superior temporal sulcus ablation. *Neuropsychologia, 28,* 1123–1142.

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. Developmental Neuropsychology, 28, 595-616.

Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen- through 18-month-old infants differentially imitate intentional and accidental actions. Infant Behavior and Development, 21(2), 315–330.

Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Cox, A., & Drew, A. (2001). Testing joint attention, imitation, and play as infancy precursors to language and theory of mind. *Cognitive Development, 15,* 481–498.

Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition, 43,* 83–121.

Corkum, V., Byrne, J. M., & Ellsworth, C. (1995). Clinical assessment of sustained attention in preschoolers. Child Neuropsychology, 1, 3–18.

Corrigan, R. (1988). Children's identification of actors and patients in prototypical and non-prototypical sentence types. Cognitive Development, 3, 285-297.

Croft, W. (2001). Radical Construction Grammar: syntactic theory in typological perspective. Oxford: Oxford University Press.

Csibra, G., Biro, S., Koos, O., & Gergely, G. (2003). One year old infants use teleological representations of actions productively. Cognitive Science, 27, 111–133.

Dąbrowska, E. and Street, J. (2006). Individual differences in language attainment: Comprehension of passive sentences by native and non-native English speakers. Language Sciences 28, 604-615.

Day, M. C. (1978). Visual search by children: The effect of background variation and the use of visual cues. Journal of Experimental Child Psychology, 25, 1–16.

D'Entremont, B., Yazbeck, A., Morgan, A. & MacAulay, S. (2007). Early gaze-following and the understanding of others. In R. Flomm, K. Lee & D. Muir (Eds.), Gaze-following: It's development and significance (pp. 77–94). Mahwah, NJ: Lawrence Erlbaum Associates.

Enns. J. T., & Girgus, J. S. (1985). Developmental changes in selective and integrative visual attention. Journal of Experimental Child Psychology, 40, 319-337.

Erickson, T. A. and Mattson, M. E. (1981). From words to meaning: A semantic illusion. Journal of Verbal Learning and Verbal Behavior, 20: 540-552.

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. Proceedings of the National Academy of Sciences, 99, 9602–9605.

Farroni, T., Johnson, M. H., & Csibra, G. (2004). Mechanisms of eye gaze perception during infancy. Journal of Cognitive Neuroscience, 16, 1320–1326.

Farroni, T., Massaccesi, S., Menon, E., & Johnson, M. H. (2007). Direct gaze modulates face recognition in young infants. Cognition, 102, 396–404.

Farroni, T., Massaccesi, S., Pividori, D., & Johnson, M. H. (2004). Gaze following in newborns. *Infancy, 5,* 39–60.

Fillmore, C. (1976). Frame Semantics and the Nature of Language. In "Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech". Vol. 280, 20-32.

Friere, A., Eskritt, M., & Lee, K. (2004). Are eyes windows to a deceiver's soul? Children's use of another's eye gaze cues in a deceptive situation. *Developmental Psychology, 40,* 1093–1104.

Gergely, G., Bekkering, H., & Király, I. (2002). Rational imitation in preverbal infants. Nature, 415, 755.

Gibson, E. J., & Yonas, A. (1966). A developmental study of the effects of visual and auditory interference on a visual scanning task. Psychonomic Science, 5, 163–164.

Givon, T. 1992. The grammar of referential coherence as mental processing instructions. Linguistics 30. 5–55.

Gleitman, L. R., January, D., Nappa, R., & Trueswell J. C. (2007). On the give and take between event apprehension and utterance formulation. Journal of Memory and Language, 57(4), 544–569.

Gleitman, L., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. (2005). Hard words. Language Learning and Development, 1(1), 23–64.

Goldberg, A. (2006). Constructions at Work: The nature of generalisations in language. Oxford: Oxford University Press.

Gordon, P. & J. Chafetz. (1990). Verb-based versus class-based accounts of actionality effects in children's comprehension of passives. Cognition, 36, 227-254.

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. Psychological Science, 11, 274–279.

Guttentag, R. E. (1985). Memory and aging: Implications for theories of memory development during childhood. Developmental Review, 5, 56-82.

Halliday, M. A. K. (1985). An introduction to functional grammar. London: Edward Arnold.

Henderson, J., and F. Ferreira. 2004. The integration o and f language, vision, and action: eye movements and the visual world. New York: Psychology Press.

Hood, B. M., Macrae, C. N., Cole-Davies, V., & Dias, M. (2003). Eye remember you: The effects of gaze direction on face recognition in children and adults. *Developmental Science, 6*, 67–71.

Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. British Journal of Developmental Psychology, 16, 233-253

Huttenlocher, J., Vasilyeva, M. & Shimpi, P. (2004). Syntactic priming in young children. Journal of Memory and Language, 50, 182-195.

Hwang, H., and E. Kaiser. (2009). The effects of lexical vs. perceptual primes on sentence production in Korean: an on-line investigation of event apprehension and sentence formulation. Talk presented at the 22nd CUNY conference on sentence processing, Davis, CA.

Ibbotson, P. (2010). A Cognitive Constructivist Approach to Early Syntax Acquisition. PhD thesis (unpublished), submitted to the University of Manchester.

Ibbotson, P., Theakston, A., Lieven, E., & Tomasello, M. (2011). The Role of Pronoun Frames in Early Comprehension of Transitive Constructions in English. Language, Learning and Development. 7: 24–39.

Jackendoff, R. (1983). Semantics and cognition. Cambridge, MA: MIT Press.

Kannass, K. N., & Oakes, L. M. (2008). The development of attention and its relations to cognitive skills in infancy and toddlerhood. Journal of Cognition and Development, 9, 222–246.

Keenan, E. (1976). Towards a universal definition of "subject". C.N. Li (ed.), Subject and topic, 303 – 33. New York: Academic Press.

Kobayashi, H., & Kohshima, S. (1997, June 19). Unique morphology of the human eye. *Nature*, *387*, 767–768.

Lakoff, G. (1987). Cognitive models and prototype theory. In U. Neisser (Ed.), Concepts and conceptual development: Ecological and intellectual factors in categorization. Emory symposia in cognition (Vol. 1, pp. 63–100). Cambridge, England: Cambridge University Press.

Lakusta, L, Wagner, L., O'Hearn, K., & Landau, B. (2007) Conceptual foundations of spatial language: Evidence for a goal bias in infants. Language Learning and Development, 3(3), 179–197.

Lakusta, L., & Landau, B. (2005). Starting at the end: The importance of goals in spatial language. Cognition, 96, 1–33.

Landau, B., and R. Jackendoff. 1993. ‘‘What’’ and ‘‘where’’ in spatial language and spatial cognition. Behavioral and Brain Sciences 16. 217–65.

Langacker, R.W. (1991). Foundations of cognitive grammar: Vol. 2–Descriptive application. Stanford: Stanford University Press.

Langacker, Ronald (1987). *Foundations of cognitive grammar: Theoretical prerequisites* (vol. 1). Stanford: Stanford University Press.

Langacker, Ronald 1991. Foundations of Cognitive Grammar, vol. 2.: Descriptive Application. Stanford: Stanford University Press.

Langacker, Ronald W. (1988). Autonomy, agreement, and cognitive grammar. Agreement in Grammatical Theory, ed. by Diane Brentari, Gary Larson, and Lynn MacLeod, 147-180. Chicago: Chicago Linguistic Society.

Lee, K., Eskritt, M., Symons, L. A., & Muir, D. (1998). Children’s use of eye gaze for “mind reading”. *Developmental Psychology, 34,* 525–539. MacWhinney, B. (1977). Starting points. Language, 53, 152-168.

MacWhinney, B. 1977. Starting points. Language 53(1). 152–68.

Morales, M., Mundy, P., & Rojas, J. (1998). Following the direction of gaze and language development in 6-month-olds. *Infant Behavior and Development, 21,* 373–377.

Morales, M., Mundy, P., Delgado, C. E. F., Yale, M., Neal, A. R., & Schwartz, H. K. (2000). Gaze following, temperament, and language development in 6-month-olds: A replication and extension. *Infant Behavior and Development, 23,* 231–236.

Mumme, D., Bushnell, E., DiCorcia, J. & Lariviere, L. (2007). Infants' use of gaze cues to interpret others' actions and emotional reactions. In R. Flomm, K. Lee & D. Muir (Eds.), Gaze-following: It's development and significance (pp. 143–170). Mahwah, NJ: Lawrence Erlbaum Associates

Myachykov, A, and R. S. Tomlin. (2008). Perceptual priming and structural choice in Russian sentence production. Journal of Cognitive Science 6(1). 31–48.

Myachykov, A, S. Garrod, and C. Scheepers. (2010). Perceptual priming of structural choice during English and Finnish sentence production. Language & Cognition: state of the art, ed. by R. K. Mishra and N. Srinivasan,54–72. Munich: Lincom Europa.

Myachykov, A., Thompson, D., Scheepers, C., & Garrod, S. (2011). Visual attention and structural choice in sentence production across languages. Language and Linguistic Compass, 5(2), 95-107. DOI: 10.1111/j.1749-818X.2010.00265.x

Nappa, R., Wessell, A., McEldoon, K.L., Gleitman, L.R. & Trueswell, J.C. (2009). Use of Speaker's Gaze and Syntax in Verb Learning. Language Learning and Development, 5(4), 203-234.

Osgood, C. E., and Bock, J. K. (1977). "Salience and sentencing: some production principles," in Sentence Production: Developments in Research and Theory, ed. S. Rosenberg (Hills- dale, NJ: Erlbaum), 89–140.

Papafragou, A., Cassidy, K., & Gleitman, L. (2007). When we think about thinking: The acquisition of belief verbs. Cognition, 105, 125–165.

Pellicano, E., & Rhodes, G. (2003). The role of eye-gaze in understanding other minds. *British Journal of Developmental Psychology, 21*, 33–43.

Perrett, D. I., & Mistlin, A. J. (1991). Perception of facial characteristics by monkeys. In W. C. Stebbins & M. A. Berkley (Eds.), *Comparative perception* (pp. 187–215). New York: Wiley.

Pick, A. D., and Frankel, G. W. (1974) A developmental study of strategies of visual selectivity. Child Dev, 45: 1162-1165.

Pinker, S., Lebeaux, D. and Frost, L. (1987). Productivity and constraints in the acquisition of the passive. Cognition 26. 267.

Posner, M. I. 1980. Orienting of attention. Quarterly Journal of Experimental Psychology 32. 3–25.

Poulin-Dubois, P., & Forbes, J. (2002). Toddlers' attention to intentions-in-action in learning novel action words. Developmental Psychology, 38, 104–114.

Prentice, J. L. 1967. Effects of cuing actor vs. cuing object on word order in sentence production. Psychonomic Science 8. 163–4.

Pye, C. & Poz, P.Q. (1988). Precocious passives (and antipassives) in Quiché Mayan. Papers and Reports on Child Language Development 27. 71-80.

Radford, A. et al. (1999) Linguistics: an introduction. Cambridge:CUP.

Rispoli, M. (1991). The mosaic acquisition of grammatical relations. Journal of Child Language, 18, 517-551.

Ruff, H. A., & Rothbart, M. K. (1996). Attention in early development: Themes and variations. New York: Oxford University Press.

Salomo, D., Lieven, E., & Tomasello, M. (2010). Young children's sensitivity to new and given information when answering predicate-focus questions. Applied Psycholinguistics, 31(1), 101-115.

Striano, T., & Reid, V. M. (2006). Social cognition in the first year. *Trends in Cognitive Sciences, 10,* 471–476.

Suzman, S. (1985). Learning the passive in Zulu. Papers and Reports on Child Language Development 24. 131-37.

Symons, L. A., Hains, S. M. J., & Muir, D. W. (1998). Look at me: Five-month-old infants' sensitivity to very small deviations in eye-gaze during social interactions. *Infant Behavior and Development, 21,* 531–536.

Talmy, L. (1988). Force dynamics in language and cognition. Cognitive Science, 12, 49–100.

Talmy, L. (2000). Toward a cognitive semantics. Cambridge, MA: MIT Press.

Talmy, Leonard. (1985). Lexicalization patterns: semantic structure in lexical forms. In T. Shopen (ed.), Language typology and syntactic description III, 57-149. New York: Cambridge University Press.

Talmy,L. (2007). Attention phenomena. In Dirk Geeraerts & Hubert Cuyckens (eds.). Oxford Handbook of Cognitive Linguistics Oxford University Press, 264-293.

Taylor, J. (2002). Cognitive grammar. Oxford, England: Oxford University Press.

Tomasello, M. (2003). Constructing a Language: A Usage-Based Theory of Language Acquisition. Harvard University Press.

Tomasello, M., & Barton, M. (1994). Learning words in nonostensive contexts. Developmental Psychology, 30, 639–650.

Tomasello, M., & Ferrar M. J. (1986). Joint attention and early language. Child Development, 57, 1454–1463.

Tomlin, R. S. (1995). Focal attention, voice, and word order: an experimental, cross-linguistic study. In P. Downing & M. Noonan (Eds.), Word order in discourse (pp. 517-554). Amsterdam: John Benjamins.

Tomlin, R. S. (1997). Mapping conceptual representations into linguistic representations: The role of attention in grammar. In J. Nuyts & E. Pederson (Eds.), Language and conceptualization (pp. 162-189). Cambridge: Cambridge University Press.

Traxler, Matthew. 2011. Introduction to Psycholinguistics: Understanding Language Science. Wiley-Blackwell: Chicester, UK.

Trueswell, J. C., and M. K. Tanenhaus. 2005. Approaches to studying world-situated language use. Cambridge, MA: The MIT Press.

Turner, E. A., and R. Rommetveit. 1968. Focus of attention in recall of active and passive sentences. Journal of Verbal Learning and Verbal Behavior 7(2). 543–8.

Well, A.D., Larch, E.P., & Anderson, D.R. (1980). Developmental trends in distractibility: Is absolute or proportional decrement the appropriate measure of interference? Journal of Experimental Child Psychology, 30, 109-124.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. Cognition, 69, 1–34.

Woodward, A. L., & Sommerville, J. A. (2000). Twelve-month-old infants interpret action in context. Psychological Science, 11, 73–77.

Yu, C. & Ballard, D. (2007). A unified model of early word learning: Integrating statistical and social cues. Neurocomputing 70, 2149-2165.

Acknowledgements

Thank you to St Vincent's Infant School, Freshfields Nursery and all the children who took part in this study; to Anna Coates and Mickie Glover for also assisting in this experiment; to Andriy Myachykov, Ewa Dabrowska and one anonymous reviewer who provided helpful comments on the manuscript. This research was supported by a postdoctoral fellowship from the Max Planck Institute for Evolutionary Anthropology, Leipzig.