Using corpora to examine discourse effects in syntax

Shanley Allen, Barbora Skarabela and Mary Hughes

Introduction

A growing body of evidence shows that children’s acquisition of at least some aspects of syntax is affected by their understanding of information flow in discourse. Specifically, children choose whether and in what form to realize arguments in their speech depending on their determination of how cognitively accessible the referent is to the listener. In this chapter, we argue that studies of naturalistic corpora are essential for understanding this process. We discuss recently developed methods for these studies, recent findings, the complementary contribution of experimental studies, and directions for further research.

Argument realization entails the selection of a particular linguistic form - full noun phrase, demonstrative, pronoun, or zero anaphora (omission) - to express a referent as the subject or object of a verb. In English, for example, one can describe an event of hugging using any of the following utterances depending on the context: I hugged Marion, I hugged the girl, I hugged this one, or I hugged her. In null-subject or null-argument languages which allow ellipsis (e.g., Italian, Japanese, Inuktitut), the subject and sometimes the object can also be omitted.

It is well-documented that children omit arguments more frequently and in different contexts than adult speakers of their language, regardless of the typology of their language (e.g., Allen 2000; Bloom 1990; Gerken 1991; Greenfield and Smith 1976; Grinstead 2000; Hyams 1986; Kim 2000; Serratrice 2005; Skarabela and Allen 2003; Valian 1991; Valian and Eisenberg 1996). Accounts of this phenomenon abound from various theoretical perspectives, some claiming that early non-target-like argument omission is grammatical for the child (e.g., Hyams 1986; Hyams and Wexler 1993), and others claiming that it derives from early processing limitations (e.g., Bloom 1990; Gerken 1991; Valian 1991). This work has focused on the general conditions under which argument omission is permitted or likely to occur - i.e. why children omit arguments at all. However, although it is widely acknowledged that children do not omit all or often even most of the arguments that meet these general conditions, work from these theoretical perspectives does not tend to seek explanations for which arguments are actually
Spontaneous speech is a linguistically and cognitively inexpensive task, revealing their choice, to extend our understanding derived from theorizing and adult studies.

In the past 15 years, a different perspective on child argument realization has emerged, largely growing out of work on discourse-pragmatics in adult speakers of null-argument languages (e.g., Ariel 1990; Bock and Warren 1985; Chafe 1987; Du Bois 1987; Givón 1983; Gundel, Hedberg and Zacharski 1993; Prince 1985). This new perspective focuses on what factors children attend to in realizing an individual argument as lexical, pronominal, or null (e.g., Allen 2000, 2007; Cho 2004; Clancy 1993, 1997, 2003; Guerriero, Cooper, Oshima-Takane and Kuriyama 2001; Guerriero, Oshima-Takane and Kuriyama 2006; Hughes and Allen 2006; Narasimhan, Budwig and Murty 2005; Paradis and Navarro 2003; Serratrice 2005; Serratrice, Sorace and Paoli 2004; Skarabela 2006; Skarabela and Allen 2002). For example, speakers often omit arguments which have just been stated in the previous utterance, but tend to overtly realize arguments which are new to the discourse. Research identifying particular features of information flow and analyzing the effect of these in various sets of child data can allow us to come much closer to understanding and predicting the form in which a child will realize a particular argument, what factors children attend to most, and how children come to realize arguments in an adult-like way. This work has focused primarily on null-argument languages and whether children follow the same patterns as caregivers in which arguments they omit. Extending to non-null-argument languages such as English, we can further investigate which particular arguments that meet the general omission criteria established from grammatical and processing perspectives (or other criteria yet to be discovered) are likely to actually be omitted, based on children's understanding of information flow in discourse.

But what type of data is best suited for answering these questions? The growing sophistication of experimental paradigms might tempt one to believe that experiments are the obvious answer. As argued by Valian and Aubry (2005), experiments in this area can offer several advantages: high frequency of occurrence of variables which may be rare in spontaneous speech, control of the contexts of interaction to single out the effects of individual variables, systematic manipulation of variables, comparison of one variable against another in a constrained setting to test the relative influence of each, relatively large and representative sample of children, and systematic comparison of children with different levels of linguistic ability (e.g., by MLU). As discussed later, experiments clearly have made valuable contributions to research on argument realization.

We argue in this chapter, however, that naturalistic corpus research in this area is essential for three reasons. First, it allows us to identify individual factors that are relevant in natural child discourse, on the basis of children using language in contexts of their choice, to extend our understanding derived from theorizing and adult studies. Spontaneous speech is a linguistically and cognitively inexpensive task, revealing children's standard ability to select argument forms to realize intended referents, and allowing us to see which factors influence the way that children realize arguments under natural conditions. Relevant factors are obviously identified through experimental work as well, but arguably the naturalistic data allows identification of factors in more natural settings. In addition, information from the two approaches allows for triangulation by showing that the factors are relevant in a variety of situations. Second, naturalistic corpus research allows us to see how a variety of factors work together in influencing argument realization, rather than looking at only one or two carefully controlled factors in the typical experimental situation. Examples of complex interactions between the features of accessibility and argument form that occur in naturalistic communication situations can then be used as the basis for designing experiments to further explore these interactions. Third, it allows us to see how children learn and use the accessibility factors in the flow of long stretches of natural discourse as opposed to in the short interchanges typical of elicited production tasks and other experimental contexts. This includes following certain grammatical patterns across discourse (Du Bois 1985, 1987), building on patterns modelled in caregiver speech (Clancy 2003), and managing miscommunication (Skarabela 2006).

That said, the study of spontaneous speech corpora in this domain is not easy. Since the structures under investigation are often relatively rare, a large amount of data from several children is needed to provide enough power for statistical analysis. Transcript data available publicly is typically insufficient for looking at contextual and interactional determinants of cognitive accessibility such as presence/absence of the referent, contextual disambiguation of the referent, and joint attention; only videotaped data provides the necessary information. Finally, selecting which features to study, coding them in a consistent and replicable way within a study, and comparing across studies are all very challenging because of the large number of features involved and the myriad ways of defining them. Discussion in this chapter focuses on the latter, reporting on recently developed methods for coding and analyzing large corpora, and discussing the relative merits of corpus-based research for studying the effects of discourse on the acquisition of syntax.

The remainder of the chapter is structured as follows. After beginning with a brief outline of research on the effect of information flow on argument realization in adult speech to establish a context for the child studies, the most substantial part of the chapter provides an extended overview and (meta-)analysis of several aspects of the study of argument realization in naturalistic corpora. We first describe and define several individual features of information flow which have been identified in naturalistic studies, highlighting differences in definition or application across studies where relevant, and showing that children are sensitive to each of these features in their argument realization. We also touch on developmental results relevant to individual features, although this has not been a focus of naturalistic work for various reasons. Next, we examine the various ways in which the interaction between the features has been assessed, including studies which group features together as well as those which compare individual features to each
other. Finally, we present three examples of insight gained by looking at information flow across utterances within a stretch of discourse. In the last portion of the chapter we turn to an assessment of representative experimental studies, discussing their strengths and difficulties and showing how they provide complementary information about the role of cognitive accessibility in argument realization to that gleaned from naturalistic studies. We see how a reduced set of features have been studied in experiments, carefully controlled and typically not in interaction with each other, and also how studying development has been an important focus. We finish the chapter by summarizing the values of naturalistic corpus analysis in this area, and providing directions for further research.

2. The effect of information flow on argument realization in adult speech

Our understanding of the effect of information flow on argument realization is rooted in a by now ample literature based on adult spontaneous speech data in a variety of discourse contexts. The modern wave of research on this topic dates back some three decades (e.g., Ariel 1988, 1990, 2001; Arnold 1998; Arnold and Griffin 2007; Bock and Warren 1985; Brennan 1995; Chafe 1976, 1987, 1994, 1996; Clancy 1980; Clark and Haviland 1977; Clark and Marshall 1981; Du Bois 1985, 1987; Du Bois, Kumpf and Ashby 2003; Fretheim and Gundel 1996; Garrod and Sanford 1982; Givón 1983; Gordon, Groz and Giliom 1993; Groz, Joshi and Weinstein 1995; Gundel 1985; Gundel et al. 1993; Marslen-Wilson, Levy and Tyler 1982; Prince 1981, 1985). This work is broadly organized around the notion of accessibility of a referent within the flow of discourse. Bock and Warren (1985: 50) define conceptual accessibility as "the ease with which the mental representation of some potential referent can be activated in or retrieved from memory." A variety of discourse factors are investigated in the literature mentioned above and shown to feed into this accessibility. These include recency of prior mention of the referent in discourse, number of other potential competitor referents in the immediate discourse context, number of utterances that the referent persists in the discourse after its initial mention, extent to which the referent is the topic of current discourse, degree to which the referent is the focus of attention in the physical context, presence of the referent in the physical context, degree of animacy of the referent, degree to which the referent is uniquely identifiable for all in a given setting or social group (e.g., the sun, the floor, the queen, the boss, John), and degree of imageability of the referent. Several researchers have elaborated scales or hierarchies indicating how accessible a referent is on the basis of some or all of these characteristics. In general, a highly accessible referent tends to be recently mentioned, with no competitors, persistent after initial mention, the topic of current discourse, the focus of attention of the interlocutors, present in the physical context, highly animate, uniquely identifiable, and highly imageable. In contrast, a referent with a low degree of accessibility tends to be newly mentioned, with several competitors, not persistent after initial mention, not the topic of discourse or focus of attention, absent from the physical context, inanimate, not uniquely identifiable, and not easily imageable. Obviously few if any referents have all of these characteristics at once, so most referents lie at some intermediate point on an accessibility scale.

With respect to argument realization, Ariel (1994: 99) proposes that speakers "direct their addressees' retrieval of the intended referents by signalling to them how accessible those mental entities are" through use of particular forms in speech to express those referents. Again, several researchers propose scales of argument forms ordered by the degree of information about the referent provided in the form, indicating correspondence between each form and the level of referent accessibility it is used to express (e.g., Ariel 1990; Givón 1983; Gundel et al. 1993). At one end of the scale, the most accessible referents are realized by forms that provide very little information about that referent: zero anaphora, agreement markers, and unstressed or bound pronouns. The speaker signals the fact that the referent is known by providing little additional information in the linguistic expression. At the other end of the scale, the least accessible referents are realized by forms that provide substantial information about that referent: stressed or independent pronouns, demonstratives, definite and indefinite full noun phrases. Here the speaker signals that s/he does not expect the hearer to be able to identify the referent from the already-existing linguistic and physical context of discourse by providing more complete information in the current linguistic expression. Ariel also notes that speakers may sometimes choose a high information form (e.g., full noun phrase) for a highly accessible referent in order to combat the effect of natural decay or to reduce the possibility of interference from other possible referents. As Gundel et al. (1993) and others point out, the overall strategy of matching level of information in the linguistic form with referent accessibility corresponds with Grice’s (1975) Maxim of Quantity which states that the speaker should be as informative as required but no more informative than needed.

These relationships between referent accessibility and argument realization have been upheld in many studies of adult discourse across a wide range of languages of varying typologies, although languages differ in exactly which forms are used to realize referents at different levels of accessibility depending on the inventory of forms available in each language. In addition, Du Bois (1985, 1987) has shown that argument form and accessibility (as revealed by recency of mention of referent) have reflexes in the grammatical structure of discourse, which he labels Preferred Argument Structure. Only one lexical argument (i.e. full noun phrase as opposed to pronoun or zero anaphora) and one new argument typically appear in a verbal clause, regardless of the number of arguments present in the clause (intransitive, transitive, ditransitive). Further, both lexical and new arguments tend not to appear in A position (subject of transitive verb), and rather appear in S (subject of intransitive verb) or 0 (object) positions. Subsequent research across many languages of different typologies has upheld these observations, although languages differ in details such as how lexical and new arguments are distributed across S and 0 positions (e.g., Du Bois et al. 2003).

Given that it is well established that these features of information flow affect argument realization in adult speech, it is appropriate to ask how and when this develops...
in children. Are children sensitive to accessibility features from their earliest speech, or does this take time to develop? Are they sensitive to the same kinds of features as adults? Are they sensitive to the features in the same way as adults? Which of the features are most salient for children, and are reliably encoded in children’s speech? How do the features interact with each other, and with other aspects of grammar?

The answers to these questions will lead us to a deeper and more thorough understanding of what drives argument realization in children, and ultimately to a more complete theory of what drives the development of grammar in children, as hinted at in the following examples (see Allen 2006 for an extended discussion of this question). As noted earlier, explaining the child-adult differences in argument realization has been a major question in the field of language development over the past 30 years, and has served as a microcosm for exploring the various factors that feed into language development and how (if at all) they interact. The predominant theoretical explanations of child-adult differences in argument realization focus on identifying the general grammatical (e.g., Hyams 1986; Hyams and Wexler 1993) and processing (e.g., Bloom 1990; Gerken 1991; Valian 1991) conditions under which children are expected to omit arguments. But these approaches do not have the tools to explain which of the many arguments meeting those conditions are actually omitted. For example, versions of the grammatical approach claim on the basis of syntactic principles that subjects can be omitted in matrix but not subordinate clauses (e.g., Rizzi 1993/1994) or with non-finite but not finite verbs (e.g., Wexler, Schütze and Rice 1998), and versions of the performance approach claim that subjects are much more likely to be omitted when they occur with long VPs than with short VPs due to children’s limited processing resources (Bloom 1990). However, clearly not all subjects of non-finite verbs, subjects in matrix clauses, or subjects with long VPs are omitted. Understanding how cognitive accessibility contributes to argument realization can enable us to predict which subjects meeting these grammatical and processing conditions will be omitted. We may also find that the dynamics of information flow in fact underlie these grammatical and performance patterns: perhaps subjects of non-finite clauses are significantly more accessible overall than subjects of finite clauses, or subjects of long VPs are significantly more accessible than subjects of short VPs! In addition, the cognitive accessibility approach can serve to uncover new patterns of argument omission in children. For example, research on adult speech from this perspective predicts an asymmetry in child argument omission between subjects in the A vs. S positions, a pattern not yet revealed in the extensive literature from the grammatical perspective.

An understanding of when and how children hone their sensitivity to the relationship between accessibility features and argument realization could also help us understand how children retreat from early over-omission of arguments or why children increase their use of strong pronouns and lexical arguments, questions that have largely been left unexplained in literature from the grammatical perspective. Further, developmental data from corpora coded for accessibility features could help resolve an apparent conflict between the grammatical and performance approaches identified by Hyams and Wexler (1993). At least one version of the grammatical approach predicts that arguments omitted by young children in non-null-subject languages are those arguments that would be realized as pronouns by older children and adults. However, the performance approach predicts no such continuity since subjects are omitted on the basis of processing resources and not any discourse-based properties. A developmental assessment of argument form based on referent accessibility would clarify which of these predictions is consistent with the data.

With these ultimate motivations in mind, we now turn to a detailed overview and analysis of the effect of the dynamics of information flow on argument realization as revealed through research on naturalistic spontaneous child speech.

3. The effect of information flow on argument realization in child speech

Over the past 15 years, a growing number of studies have appeared in the language development literature using naturalistic data to investigate the relationship between information flow and argument realization (Table 1). As noted earlier, these studies have an advantage over experimental studies in that they view the workings of information flow in naturalistic contexts, typically in the child’s own home with familiar interlocutors (parents, siblings, friends, and/or familiar researchers). The situations of interaction are not contrived, and thus researchers can be more assured that children produce their utterances as they typically do, influenced by conditions of interaction that they encounter and use regularly in their daily life. In addition, the child may have a better knowledge of objects that are typically present in the home setting and thus be more free to show the full extent of their linguistic ability; it may be difficult to take in all the nonlinguistic information in an unfamiliar lab setting.

In most studies, data is collected in sessions from 20 to 60 minutes long, at a frequency of every two weeks to every four months, over a period of six months to two years. This allows one to observe development within a single child, and to check the stability of the use of information flow features over a period of time. Because of the intensity of data collection, however, usually only a few children are studied, thus compromising the generalizability of the work to some extent.

The first study along these lines, conducted by Greenfield and Smith (1976), shows that children at the one-word stage tend to select for production, based on their assumed target utterance, the one word that conveys the most essential information of their intended message. That word, in combination with contextual information, encodes the communicative intent of a full utterance in adult speech. Greenfield and Smith (1976) call this tendency the Informativeness Principle: children tend to omit presupposed information that can be taken for granted, and they tend to select the word that expresses what is new or changing in the situation, representing information that is not already shared in some way between the speaker and interlocutor.
### Table 1. Studies investigating referent accessibility and argument realization

<table>
<thead>
<tr>
<th>Study</th>
<th>Language</th>
<th>Age</th>
<th>Source of data</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allen and Schroeder (2003)</td>
<td>Inuktitut</td>
<td>2;0–3;6</td>
<td>Videotaped data (Allen 1996)</td>
<td>newness</td>
</tr>
<tr>
<td>Cho (2004)</td>
<td>Korean</td>
<td>1;8–3;0</td>
<td>Audiotaped data</td>
<td>newness</td>
</tr>
<tr>
<td>Clancy (1993, 1997)</td>
<td>Korean</td>
<td>1;8–2;8</td>
<td>Audiotaped data</td>
<td>absence, newness, contrast, animacy, person, query</td>
</tr>
<tr>
<td>Clancy (2003)</td>
<td>Korean</td>
<td>1;8–2;8</td>
<td>Audiotaped data</td>
<td>newness, animacy, person</td>
</tr>
<tr>
<td>Guerreiro, Oshima-Takane and Kuriyama (2006)</td>
<td>Japanese, English</td>
<td>1;9–3;0</td>
<td>Videotaped data (Oshima-Takane, Goodz and Derevensky 1996)</td>
<td>absence</td>
</tr>
<tr>
<td>Hughes and Allen (2006)</td>
<td>English</td>
<td>2;0–2;1</td>
<td>Videotaped data (Lieven, Behrens, Speares and Tomasello 2003)</td>
<td>newness, differentiation in context, differentiation in discourse, inanimacy, person</td>
</tr>
<tr>
<td>Narasimhan, Budwig and Murty (2005)</td>
<td>Hindi</td>
<td>2;10–4;3</td>
<td>Audio- and videotaped data (Budwig and Chaudhary 1996)</td>
<td>animacy, contrast, newness, query</td>
</tr>
<tr>
<td>Paradis and Navarro (2003)</td>
<td>Spanish, English, Spanish-English bilingual</td>
<td>1;0–2;0</td>
<td>Transcribed data (Deuchar and Quay 2000; Lopez Ornat 1994; Serrat Sellabona unpublished)</td>
<td>newness, contrast, absence, query, emphasis</td>
</tr>
<tr>
<td>Serratrice (2005)</td>
<td>Italian</td>
<td>1;5–3;0</td>
<td>Transcribed data with rich contextual notes (Cipriani, Pfanner, Chilosi, Cittadoni, Ciuti, Maccari, Pantano, Pfanner, Poli, Sarro, Bottari, Cappelli, Colombo and Veneziano 1989)</td>
<td>activation state, referential ambiguity, person</td>
</tr>
<tr>
<td>Serratrice, Sorace and Paoli (2004)</td>
<td>Italian, English, Italian-English bilingual</td>
<td>1;7–4;7</td>
<td>Transcribed data with rich contextual notes (Serratrice, no date given, for Italian-English bilingual child; Cipriani et al. 1989 for Italian monolinguals and Brown 1973; Sachs 1983; Suppes 1974 for English monolinguals)</td>
<td>person, absence, activation (=newness), contrast, differentiation in discourse, query</td>
</tr>
<tr>
<td>Skarabela (2006)</td>
<td>Inuktitut</td>
<td>2;0–3;6</td>
<td>Videotaped data (Allen 1996)</td>
<td>physical absence, linguistic newness, contrast, interference, lack of joint attention</td>
</tr>
<tr>
<td>Skarabela (2007)</td>
<td>Inuktitut</td>
<td>2;0–3;6</td>
<td>Videotaped data (Allen 1996)</td>
<td>joint attention</td>
</tr>
<tr>
<td>Skarabela and Allen (2002)</td>
<td>Inuktitut</td>
<td>2;0–3;6</td>
<td>Videotaped data (Allen 1996)</td>
<td>newness, joint attention</td>
</tr>
</tbody>
</table>
More recent studies in this domain have focused on identifying individual features of accessibility in information flow from adult discourse that are relevant to children, and determining how to code them in child data (which is often more challenging than adult data). Once the features are coded in naturalistic corpora, they are assessed both qualitatively and quantitatively to determine how they shape children’s realization of arguments. Table 1 lists many of these recent studies, including the languages and ages dealt with, and the accessibility features studied. We do not include studies based on narrative data since they do not involve large spontaneous speech corpora which is the theme of this volume. It should be noted, however, that several narrative studies have been published focusing at least partly on the relationship between referent accessibility and argument realization (e.g., Clancy 1992; Hickmann and Hendriks 1999; Karmiloff-Smith 1985).

As shown in Table 1, at least 11 studies to date (some reporting results in more than one paper) have investigated the effect of one or more accessibility features on argument realization across seven languages (English, Hindi, Inuktitut, Italian, Japanese, Korean, Spanish). All studies have found at least some evidence that children use relatively more informative linguistic forms (full noun phrases, demonstratives, stressed or independent pronouns) to realize referents which have a low degree of cognitive accessibility, and relatively less informative forms (unstressed or bound pronouns, omission) to realize referents which are highly accessible. Some studies have also looked at how this sensitivity develops in children, although this has not been a major focus of naturalistic work. In the next sections we focus on the different features used in these studies, how they are typically defined and coded, whether the different features are tapping into accessibility in the same way, and what developmental patterns are apparent.

### 4. Individual accessibility features

A total of nine different accessibility features have been investigated across the different studies listed in Table 1, although they are labelled and defined somewhat differently across studies. They are all derived from the features used in the adult studies mentioned earlier, but not all the features studied for adults have been studied for children. Although each of the features is in reality both continuous and complex, in most studies each is treated as binary (accessible, not accessible) for purposes of the studies of argument realization discussed here. This enables assessment of the statistical relationship between the accessibility of the referent and the form of the argument in a way that using a continuum to code degree of accessibility would not. Further, it substantially simplifies the task of determining degree of accessibility of a referent for each feature. The "not accessible" value for each feature indicates that a referent would not be accessible to the interlocutor in a conversation solely on the basis of the discourse or situational information from that feature, such as a referent newly introduced into discourse or absent from the physical context of interaction. The "accessible" value, in contrast, indicates that a referent would be accessible on the basis of information from that feature, such as a referent already recently referred to in the discourse or present in the physical context. However, it is clear that some features are stronger than others in their influence on argument realization, and that most features are not symmetrical in their effect. Each of the nine features is discussed in some detail in the following paragraphs. Developmental trends with respect to individual features are noted after the features are discussed.

#### 4.1 Newness

The most widely studied feature is the recency of mention of the referent in the discourse. Indeed, this is the one feature that is regularly referred to as undoubtedly having some effect on child argument realization, even by researchers from theoretical perspectives not related to discourse-pragmatics (e.g., Bloom 1990, 1993; Hyams and Waxler 1993). The "not accessible" value for newness characterizes a referent which has not been mentioned in the prior discourse and is predicted to be realized by a high information form, while a referent that has been previously mentioned (i.e. "accessible") is predicted to be realized by a pronoun or null argument. This is one of the two features for which both binary values of accessibility are assumed to have a strong influence on argument form.

The criteria for evaluating whether a referent is new or given vary substantially across studies. Most studies restrict the definition to explicit linguistic mention in prior discourse. However, different cut-off points are used across studies to distinguish new vs. given, reflecting in part a poor understanding of the psychological mechanisms underlying the contribution of newness to overall accessibility. Based on criteria laid out in Chafe (1976) and Du Bois (1987) for adult research on elicited narratives, some child studies (e.g., Clancy 1993) recognize three categories of newness: non new (mentioned in the preceding utterance), accessible (mentioned in the preceding 2+ utterances), and new (never before mentioned in the interaction). This classification works well for elicited narratives which do not typically exceed 100-200 utterances, and in which participants are focused on the same storyline throughout the interaction. It seems much less useful, however, for spontaneous speech interactions in which the topic of conversation changes frequently and in which one session could easily include 1000 or more utterances. Therefore, most studies of newness in child spontaneous speech collapse the accessible and new categories together into one category leading to the application of a binary distinction between new and given referents. Studies following the earlier-mentioned adult work consider a referent as new when it has not been explicitly mentioned in the preceding 20 utterances, and given otherwise (e.g., Allen 2000; Guerriero et al. 2006; Hughes and Allen 2006; Mishina-Mori 2007; Skarabela 2006). However, some studies apply a much earlier cut-off point ranging from three prior utterances (Narasimhan et al. 2005) to 5-10 exchanges (Paradis and
Navarro (2003), while others do not specify any particular cut-off point (Serratrice et al. 2004). In an attempt to clarify the underlying motivation for the boundary between new and given in child speech, Skarabela and Allen (2003) analyzed in their corpus of Inuktitut child speech how frequently a target referent was mentioned one utterance prior, two utterances prior, and so forth, up to 20 or more utterances prior. They determined that very few arguments have preceding references more than 5 utterances prior, suggesting that the difference between a cut-off of 5 versus 20 preceding arguments is minimal, and that an earlier threshold between new and given information (i.e. a distance of 5 utterances to encode a referent as new) might be more appropriate for child spontaneous speech.

Another area of difference across studies is the relationship between newness and person of the referent. Following Chafe (1976) and Du Bois (1987), most studies consider first and second person referents as given because they always encode reference to speech participants ('I', the speaker, and 'you' the listener), while third person referents are classified as either new or given depending on the context (e.g., Allen 2000; Guerriero et al. 2006; Hughes and Allen 2006; Mishina-Mori 2007; Serratrice et al. 2004, Skarabela 2006). Some researchers are more cautious, classifying first introductions of first and second person referents as new (Clancy 1993; 1997, 2003; Narasimhan et al. 2005), although they acknowledge that the effect is slight since there are few first introductions of such referents.

4.2 Topicality

Topicality refers to whether the referent is or is not the focus of the current conversation (e.g., Givón 1983; also widely discussed in the generative literature), usually termed "accessible" and "not accessible" respectively. It overlaps substantially with newness to the extent that the two are often not easily distinguishable. One difference is that topicality takes as a reference point both extralinguistic and linguistic contexts, whereas newness is typically restricted to only linguistic contexts. Serratrice et al.'s (2004) feature activation is very similar if not identical to topicality; they consider a referent "accessible" for activation if it is associated with topic maintenance, and "not accessible" for activation if it has not been previously introduced in discourse or signals a shift of topic. Narasimhan et al. (2005) formulate their feature prior mention in terms of whether or not the referent is being talked about rather than in terms of purely linguistic mention of the referent — again consistent with the concept of topicality. This enables them to capture the tendency for referents to be talked about more frequently than they are explicitly mentioned. Cho (2004: 37) identifies a referent as topical if it "had been mentioned in the immediately preceding clause, or was present physically in the discourse context." Other researchers define topicality in terms of a referent's salience in the discourse context (Clark 2003; Kayama 2003). The coding of topicality is much less objective than that of some of the other features since it depends on one's impression of salience rather than on anything quantifiable. As is the case for

4.3 Absence

The feature absence refers to whether a referent is present in or absent from the physical context of the conversation, sometimes narrowly defined as the visual field/space of the discourse interaction (Paradis and Navarro 2003). Absent referents are considered less conceptually accessible to the hearer than present referents and are thus more likely to be realized by a high information argument form, as has been confirmed by all of the lie relevant studies listed in Table 1. There is much less strength in the reverse prediction; however: just because a referent is present in the physical context does not constitute grounds for realizing the associated argument with a low information form because mere presence of a referent does not make it particularly salient to the hearer. For this reason, Narasimhan et al. (2005) do not include absence as one of the indicators of accessibility (which they refer to as "pragmatic prominence") in their study. Note that several researchers have coded absence on the basis of contextual notes in transcript data (e.g., Paradis and Navarro 2003; Serratrice et al. 2004). This is less than ideal for obvious reasons.

4.4 Query

The feature query indicates whether or not a referent is the subject of or the response to a question. The underlying assumption is that (information about) the referent is either not yet identified or newly identified in an instance of query. As a result, the listener has less than complete knowledge of the referent, and thus the speaker is likely to realize it using a high information form. As with absence, the reverse is not true: the failure of a referent to be the subject of or response to a question does not make it salient to the hearer and thus use of a low information form is not expected on that basis alone.

The way that the definition of query plays out in actual coding differs substantially across studies. Allen (2000) and her colleagues (Hughes and Allen 2006; Skarabela 2006) define query quite narrowly such that in the interaction Who ate the cake? John did!, only who and John would be coded as "not accessible" for query. Serratrice et al. (2004) would have coded both cake (since information about the cake is queried) and John (they did not code question words for discourse-pragmatic features at all) (Serratrice personal communication, April 2007). Narasimhan et al. (2005) would have coded only John since they limited the definition of query to responses only. Paradis and Navarro (2003) coded "not accessible" for query "when the referent was being questioned (in an intonational interrogative) or used in response to a question" (note that they did not code any arguments in structurally interrogative utterances), so they would have coded only John as "not accessible" in the above example. Clancy (1993,
4.5 Disambiguation / contrast / interference

Several researchers assess whether a particular referent has potential competitor referents in the linguistic or physical context that could be easily confused with the target referent. Note that it is not necessary for the speaker to explicitly draw attention to this potential ambiguity; it is only necessary that the potential ambiguity exist. The prediction is that the speaker will more likely realize a potentially ambiguous referent with a high information form to make its identity clear. As with the previous two features, the reverse prediction does not hold. Just being unambiguously identifiable in either discourse or physical context does not make a referent salient, and thus does not by itself encourage use of a low information form to realize that referent.

This feature has appeared in the literature under several different names, and with different definitions. Both Clancy (1993, 1997) and Narasimhan et al. (2005) refer to it as contrast. They take the verb semantics as the main criterion, coding the target referent as "not accessible" for contrast if there is one or more other possible referents "bearing the same relation to the same predicate ... or bearing a parallel relation to a similar type of predicate" (Clancy 1997: 641). Paradis and Navarro (2003:379) similarly ascribe contrast to an argument "whose function [is] to disambiguate between two possible referents." Allen (2000) and Hughes and Allen (2006) separate the effects of linguistic and physical context into two features, termed disambiguation in discourse and differentiation in discourse. The former refers to a referent which has one or more potentially competing referents in the previous 5 utterances; the latter refers to a referent which has one or more potentially competing referents in the immediate physical context. Serratrice et al. (2004) and Serratrice (2005) restrict themselves to disambiguation in the linguistic context (termed differentiation in discourse and disambiguation respectively); they could not analyze the physical context since their analyses were based mostly on transcript data. Skarabela (2006) only investigates disambiguation in the physical context, using the term interference.

All of the mentioned studies confirm the relevance of disambiguation (as they construe it) to argument realization. Allen (2000) and Hughes and Allen (2006) show in separate studies of Inuktitut and English data that differentiation in context has a significant effect while differentiation in discourse does not when both are assessed together to determine their relative contribution to a logistic regression model. Allen (2000) suggests that this is due to confounding between these two features, as well as between these features and explicit contrast (see next section). Since other researchers either analyzed only one of these features or treated the two together as one feature, it is not surprising that the discrepancy in Allen (2000) and Hughes and Allen (2006) did not show up in those studies.

4.6 Explicit contrast / emphasis

Allen (2000), Paradis and Navarro (2003), Serratrice et al. (2004), and Skarabela (2006) investigate the effect on argument realization of a contrast made explicit or emphasized by the speaker. This is different from the factor disambiguation / contrast / interference just discussed because in addition to the ambiguous situation being present in the linguistic or physical context of discourse, it is actively resolved by the speaker through some type of emphasis. A referent that is explicitly contrasted with another potential referent is more likely to be realized with a high information form since there is more than one potential referent in the context. However, a referent which is not being explicitly contrasted would not pull for a low information form since it is not particularly salient. This feature is different from the others discussed so far in that it depends to some degree on assessing the speaker's intent rather than simply the situation presented by the discourse and physical context.

Allen (2000) and Skarabela (2006), both using the term contrast, employ both non-linguistic and linguistic means to assess whether the speaker is explicitly contrasting a referent, including stress, tone of voice and gesture. Allen (2000: 488) describes typical situations of explicit contrast as a child "wanting to prohibit others from doing something he or she is doing, or when a child wants to do something someone else is doing"; Serratrice et al. (2004) (also under the term contrast) use the preceding and following discourse context to determine the child's intention or assumption of contrast (since their data is mostly in the form of written transcripts, they have no access to intonation information). Paradis and Navarro's (2003) definition of contrast partly includes this idea of explicit contrast in that they use it to code an argument whose function is to focus on a particular referent. Their feature emphasis, used to code arguments that "could be read as if they had more prosodic prominence, in other words, the speaker seemed to intend to highlight that [argument]; also is similar to the notion of explicit contrast although not completely identical. All these studies find, as expected, that children use more informative forms to realize referents which they explicitly contrast with others.
Hughes and Allen (2006) purposefully leave explicit contrast out of the list of features coded in their study because it functions differently from the other features. Since it is most often indicated by prosodic prominence, the argument must be realized using a high information form (e.g., stressed pronoun, full noun phrase) for the feature to apply, and thus all arguments which are explicitly contrastive are realized with high information forms almost by definition. As noted earlier, the feature explicit contrast often overlaps with the various types of disambiguation coded across studies. The confounding between these features should be sorted out in future research.

4.7 Person

The person of referent is often included as a discourse-pragmatic feature since, following Du Bois (1987) and others, third person referents (e.g., he, she, it, they) have different discourse-pragmatic status than first and second person referents (e.g., I, we, you). The search space for the latter tends to be quite restricted, particularly in child discourse (Allen 2000), because there are relatively few participants in a typical conversation. In contrast, it is much more difficult for the hearer to identify the referent of a third person argument since there is a potentially unlimited number of third person referents in the search space. Thus, child speakers are predicted frequently to realize first and second person referents with low information forms. Most languages only allow first and second person referents to be realized as low information forms: pronouns, agreement markers or null forms depending on the language (e.g., Inuktitut does not permit first and second person pronouns in argument position). However, children are predicted to select the lowest form available in their language if more than one option is available. Third person referents can be expressed in most languages using a wide variety of forms - full noun phrases, proper names, demonstratives, independent and bound pronouns, verbal agreement markers, and null forms although the exact forms possible depend on the language (e.g., Inuktitut allows all but pronouns, English has no plural agreement markers and severely restricts null forms). Just the fact of being third person does not make the referent strikingly inaccessible to the hearer, however, so it does not command use of a particularly high information form. Thus, the effect of person is the converse of that for the previous four features discussed; here, the "accessible" value pulls for using a low information argument form because being first or second person makes the referent quite salient to the hearer, but the "inaccessible" value does not by itself pull for a high information form.

The tendency for first and second person referents to be realized with the lowest information form possible in a language, and to do so more frequently than third person referents, is confirmed by all the studies in typologically diverse languages listed in Table 1. Although no effect of person was found in data from one English-Italian bilingual child, the authors claim that that is due to interaction between person and other features (Serrattrice et al. 2004). Indeed, person clearly overlaps considerably with newness (first and second person referents in most studies are assumed to be "accessible" for newness by definition), absence (first and second person referents are virtually always present in the physical context of the conversation), disambiguation (it is almost always clear who the speaker and hearer are in a conversation), animacy (first and second person referents are always animate), and attention (first and second person referents are assumed to always be jointly attended to). Because of the powerful effect of person in argument realization, and because of the high degree of overlap between person and other features, some researchers have restricted some or all of their analyses to third person only (e.g., Allen 2000; Hughes and Allen 2006; Skarabela 2006, 2007; Skarabela and Allen 2002, 2003).

4.8 Animacy

The feature animacy has also been studied frequently in the context of argument realization in both child and adult language. Animacy refers to how alive or sentient an entity is, with first person humans at one end of the continuum and abstract entities at the other. Adult studies often elaborate detailed animacy hierarchies with 8 or more levels, which are crucial in many languages for understanding the intricacies of such phenomena as ergative marking, noun classification, and word order (e.g., Comrie 1989; Silverstein 1976). In the child studies discussed here, animacy is applied as a binary feature that differentiates between animate referents (human, animal) as "accessible" and inanimate referents as "not accessible". As noted earlier, there is substantial overlap between animacy and person because all inanimate referents are third person, and all first and second person referents are animate.

Although animacy is often used as an accessibility feature, it is distinct from such features since it is an inherent semantic property associated with individual referents. Unlike newness, absence and person which are determined by a particular discourse-pragmatic context, animacy is an inherent property that is stable across different discourse-pragmatic contexts. In certain contexts in child discourse, however, inanimate referents such as dolls, stuffed animals, and certain toys assume the characteristics of humans and are thus typically coded as animate.

What, then, is the relationship between animacy and argument realization? The logic is very similar to that just described for person. In typical child discourse, the number of inanimate entities (e.g., food, cup, television, furniture, clothes) outweighs the number of animate entities (e.g., mother, father, sibling, friend, dog). Thus, reference to an inanimate entity is potentially more ambiguous than to an animate entity, such that inanimate entities are predicted to be realized more frequently by high information argument forms than animate entities. Previous studies in this area yield controversial results, however. While animacy is found to influence argument realization in the expected way in some languages (e.g., Korean: Clancy 1993, 1997, 2003), it does not have a significant effect on the linguistic form in others (e.g., Inuktitut: Allen 2000; English: Hughes and Allen 2006). Although Narasimhan et al. (2005) include animacy in their analysis, its effect cannot be separated from that of others included in the
designations "pragmatically prominent": It is thus unclear to what extent animacy alone influences the choice of argument form. In addition, differences in the reported findings may originate in slightly different coding strategies (e.g., coding dolls as animate versus inanimate entities). Finally, the link between this inherent semantic property and argument realization may be too weak. Being animate does not make a referent particularly salient in discourse, and being inanimate does not make a referent particularly non-salient. Therefore, although there is a difference in relative salience between animates and inanimates, the effect is not especially strong, and probably not strong enough to have a significant effect on argument realization.

4.9 Attention

The feature attention assesses whether or not a speaker and listener are focused on the same referent while they are aware of each other's attention (Tomasello 1999). The basic prediction is that referents produced in the context of joint attention are more likely to be realized by low information forms than those produced in the absence of joint attention, because the former are particularly salient in discourse. Referents produced in the absence of joint attention are not particularly non-salient in discourse, however, so there is little reason to expect them to be necessarily realized by high information argument forms. To assess whether or not joint attention is in progress for a particular referent, videotaped data are examined for eye gaze, body direction, head direction and gesture (including pointing) of both interlocutors. First and second person referents are automatically considered instances of joint attention in progress since these refer to the speech participants. Further, referents not physically present in the context of the ongoing conversation cannot be coded for joint attention. However, the presence of the target referent in the physical context of the interaction does not automatically imply that the speaker and listener are involved in joint attention, i.e. a speaker and listener may be involved in "joint activity" (Clark 1996) without joint attention being established.

Analyses of data from 2- to 3-year-old learners of Inuktitut (Skarabela 2006, 2007; Skarabela and Allen 2002, 2004) show that the predictions are realized. In addition, a recent study of the role of joint attention in argument realization in Mandarin child-directed speech shows that Chinese caregivers are more likely to omit arguments of transitive verbs when they are jointly attending to the referents of those arguments with their children (Lee 2006). Lee suggests that joint attention thus facilitates children's identification of intended referents. Clark's (2001) study of English-speaking caregivers makes the similar suggestion that adults establish joint attentional scenes to facilitate early word learning.

Note that although the other factors are typically operationalized in the same way for naturalistic and experimental studies, the factor attention is operationalized somewhat differently (see later section on experimental studies). In experimental conditions that are categorized as "not accessible" for attention, the interlocutor is either out of the room or explicitly looking away when the child initially interacts with the referent or observes the referent on video. The interlocutor in the experiment also does not attend to the referent when the target utterance is spoken, either because s/he cannot see the referent because of an obstruction or because s/he purposefully does not look at the referent. These experimentally-induced manifestations of lack of joint attention are much stronger than in a typical naturalistic interaction where the interlocutor simply does not happen to be sharing attention with the child on some referent. Therefore, the "not accessible" condition for attention in experiments tends to pull much more strongly for a high information form than the similar condition in naturalistic interactions.

4.10 Developmental trends

Although all of the naturalistic studies of argument realization have assessed whether sensitivity to individual accessibility features is operative at one particular age or level of linguistic ability, few studies have explored the question of how that sensitivity changes over time. Most studies have taken as their focus establishing whether children are sensitive to the effect of accessibility features on argument realization at all, and how this interacts with the typological features of the language, leaving for later research the more complex question of when and how this sensitivity develops. Furthermore, given the constraints of naturalistic data, it is not easy to meet the criteria necessary to determine the timing of development. One would need either longitudinal data from one or more children over a sufficient time period to show developmental change or cross-sectional data from several children at each of two or more developmental stages. In addition, one would need sufficient data for each feature studied, at each time point, to do statistics or other analyses ensuring that the change from one time point to another was significantly different either quantitatively or qualitatively.

Several recent naturalistic studies are not suited to investigate development because they sample data from the target child(ren) at only one developmental point. For example, Hughes and Allen (2006) look at one child over only a 6-week period (2;0-2;1), Mishina-Mori (2007) examines four children each over only a two-month period (two older bilinguals 3;0-3;2, and two younger monolingual "controls" 2;6 and 2;9), and Narasimhan et al. (2005) investigate data from twelve children but at only one time point each (2;10-4;3) and not properly distributed to form groups according to age or linguistic ability. Some studies present data divided by age groups but no developmental trends are evident, either because no such trends are present or because the small number of children and/or relevant analyzable utterances obscures the trends (e.g., Cho (2004) with two Korean children aged 2;0-2;8; Clancy (1993) with two Korean children aged 1;8-2;8). Allen (2000) was originally intended to investigate development in nine months of longitudinal data from each of four children. However, the author decided not to publish the developmental portion of the study because it was not clear whether there were indeed no developmental effects, whether the study did not extend long enough for development to be evident, or whether there was not
sufficient data at each developmental point for unambiguous results in the domains investigated. In addition, because the four children ended up being at fairly different developmental levels and because only three data points were available for each child (months 1, 5, and 9 of the study), the developmental groups that could be formed on the basis of linguistic ability ended up being awkwardly unbalanced. The fact that one cannot determine the level of linguistic ability of participants in advance is one major drawback of naturalistic studies.

A few naturalistic studies, however, do focus on development. Serratrice et al. (2004) investigated subject and object realization in one English-Italian bilingual child, six Italian-speaking monolingual children, and four English-speaking monolingual children. They divided the data from each language into four developmental stages according to mean length of utterances assessed in words (rather than morphemes): I (MLU 1.5-2.0), II (MLU 2.0-3.0), III (MLU 3.0-4.0), and IV (MLU 4.0+). Right from Stage I, a significantly higher percentage of the children’s null arguments realized referents with accessible features than referents with inaccessible features, separately assessed for each of absence, activation (= topicality), contrast, differentiation in discourse, and query. The monolinguals in each of the two languages produced over 80% of their null arguments to realize accessible referents (except for query at Stages II and IV in English), and the bilingual child over 70% (except for activation at Stages II and IV in Italian and absence at Stage IV in English). However, the bilingual child used third person pronouns in pragmatically inappropriate contexts in 9% of instances at Stages III and IV in Italian (personal subject pronouns were not used virtually at all before this point). Pronouns in Italian normally signal contexts of focus or topic shift, whereas the bilingual child used them in fully accessible contexts where pronouns would be appropriate in English. Note that Paradis and Navarro (2003) found the same pattern of overuse of pronouns to realize accessible referents at 1;9, 2;6, and 3;0, as well as data from the children’s mothers. MLUs for the English data were 1.09-2.25 and 2.90-4.25 respectively; MLUs for the Japanese data were 1.07-1.91 and 2.86-4.29 respectively. The English-speaking children already used more non-lexical than lexical forms to realize given referents at 1;9, with this tendency increased at 3;0. Within the non-lexical forms, null forms predominated at 1;9, but pronominal forms took over by 3;0, similar to the mothers’ speech and as expected for the language typology. As for new referents, the children used slightly more non-lexical than lexical forms to realize them at 1;9, but reversed this pattern by 3;0 consistent with their mothers’ performance. This suggests that it takes more time for children to learn to introduce new referents as opposed to maintaining existing referents. The Japanese-speaking children produced very few arguments at 1;9. By 3;0 these children behaved like their mothers in producing more non-lexical than lexical forms for given referents, but produced an equivalent number of lexical and non-lexical forms to

Skarabela (2006) investigated development in children’s sensitivity to the relationship between argument form and attention, comparing this with caregiver data from the same data set. Her data were divided into three groups according to the mean length of utterance in morphemes of all utterances containing a verb (see Allen (1996) for a justification of this grouping): I (MLU 3.25-3.99), II (MLU 4.0-4.74) and III (MLU 4.75-5.49). Note that values for MLU would be expected to be considerably higher than values for MLU taking all utterances into account, so these numbers do not represent unusual linguistic ability for children aged 2;0-3;6. Chi-square tests assessing the relationship between attention (accessible, inaccessible) and argument form (overt, null) were not significant at Stages I or II, but children at Stage III produced significantly more overt forms in the context of inaccessible referents than in the context of accessible referents (p<.001). When the variable argument form was differentiated into three categories - lexical, demonstrative, and null - the Chi-square results were significant at all three stages. As expected, children produced more null forms when attention was "accessible" (i.e. when the referent was jointly attended to by speaker and hearer), and more lexical forms when attention was "not accessible." Children at Stages I and II were equally likely to produce demonstratives for both values of attention. However, children at Stage III followed the adult pattern of producing more demonstratives when attention was accessible. Skarabela (2006) also looked at development in each child individually with data grouped according to the just-mentioned stages. Two stages were represented in the data of each child. For three of the four children, the degree of association between attention and argument form increased from the earlier to later stage. As Skarabela states, these results might be interpreted as suggestive of an increasing role of attention in the three different argument forms (lexical, demonstrative, null) with children’s increasing MLU.

Guerriero et al. (2006) is the only naturalistic study we are aware of which was explicitly designed to investigate the effect of accessibility on argument realization at predetermined developmental periods. In the first study they report (Study 1), they looked at the effect of newness in six English-speaking and six Japanese-speaking children at each of ages 1;9 and 3;0, as well as data from the children’s mothers. MLUs for the English data were 1.09-2.25 and 2.90-4.25 respectively; MLUs for the Japanese data were 1.07-1.91 and 2.86-4.29 respectively. The English-speaking children already used more non-lexical than lexical forms to realize given referents at 1;9, and this tendency increased at 3;0. Within the non-lexical forms, null forms predominated at 1;9 but pronominal forms took over by 3;0, similar to the mothers’ speech and as expected for the language typology. As for new referents, the children used slightly more non-lexical than lexical forms to realize them at 1;9, but reversed this pattern by 3;0 consistent with their mothers’ performance. This suggests that it takes more time for children to learn to introduce new referents as opposed to maintaining existing referents. The Japanese-speaking children produced very few arguments at 1;9. By 3;0 these children behaved like their mothers in producing more non-lexical than lexical forms for given referents, but produced an equivalent number of lexical and non-lexical forms to
realize new referents. The mothers used slightly more non-lexical than lexical forms to realize new referents at 1.9, but reversed this pattern by 3.0. Thus, it is possible that the Japanese-speaking children at 3.0 were simply following the earlier model in their input. In Study 2, Guerriero et al. (2006) replicated the results of Study 1 with two additional children learning each language, this time with four time points per child: I (MLU 1.0-1.99), II (MLU 2.0-2.99), III (MLU 3.0-3.99) and IV (MLU 4.0 +). The English-speaking children used primarily null forms to realize given arguments at Period I, but largely switched to pronominal forms by Period II for one child and Period III for the other consistent with language-specific patterns. Neither child introduced many new referents at Period I, but they correctly preferred to realize new referents with lexical forms from Period II where new referents started to appear in their speech. Interestingly, virtually all new referents realized with non-lexical forms in both child and mother data also co-occur with some non-linguistic indicator of the referent (e.g., gesture, touch, eye gaze). Both Japanese-speaking children correctly used non-lexical forms to realize given referents right from Period I. However, they only began to prefer lexical forms to realize new referents at Period III for one child and at Period IV for the other. As in Study 1, the Japanese-speaking children's patterns for new referents mirrored their mothers' input at the previous period. In contrast to the English speakers, the Japanese speakers used non-linguistic indicators to identify new referents expressed non-lexically only about half the time. Guerriero et al. (2006) suggest that the Japanese mothers are following a typical interaction style between familiar interlocutors whereby the burden is on the hearer to guess the referents from shared knowledge, rather than the English pattern of the burden being on the speaker to be as clear as possible. Children are then mirroring this pattern modelled in the input.

The developmental data taken together reveals some interesting patterns. First, children appear to be already sensitive to the relevance of accessibility to argument realization from around 2.0 or an MLU of 2.0. Second, children increase significantly in this sensitivity somewhere between MLU 2.0 and 3.0 with differences depending on the language and feature studied. Third, children seem to take somewhat longer to produce adult-like forms for inaccessible referents than for accessible referents, perhaps because the pattern in the input is often stronger for accessible referents. Fourth, children's growing sensitivity to accessibility features may be revealed more clearly by treating argument forms as having three or more levels (e.g., lexical, pronominal, null) rather than two (either overt vs. null or lexical vs. non-lexical). As noted earlier, however, it is much more difficult to study development in naturalistic data than in experimental studies; the exploration of developmental trends is one of the main contributions of experimental literature to our understanding of argument realization.

4.11 Summary

For each of the accessibility features discussed earlier, at least one study has shown a relationship between the accessibility value of a referent for that feature and the argument form with which that referent is realized in child speech. The effect of newness has been shown the most robustly, both with the largest effects and in the most languages. Other features such as query and animacy are less well studied and show less obvious effects across studies. Overall, however, it is clear that children show substantial sensitivity to referent accessibility in determining the form in which they realize arguments. In studies which have compared child data to caregiver data, it is also clear that children show very similar patterns of argument realization to those of their caregivers (Clancy 1993, 1997; Guerriero et al. 2006; Paradis and Navarro 2003; Skarabela 2006). And in studies which have investigated development, it is clear that children are sensitive to the effect of accessibility factors on argument realization from fairly early on and improve significantly in showing this sensitivity at some point between MLU 2.0 and 3.0 (roughly ages 2.0-3.0).

The preceding section has focused on the effect of the accessibility features individually to determine their unique role in argument realization. This is largely because most of the studies cited have analyzed the features individually, either using simple percentages (e.g., a higher percentage of "not accessible" than "accessible" referents is realized by a high information form), or using Chi-square to measure the statistical significance of the difference. However, there are at least three reasons to analyze the accessibility features in combination with each other. First, speakers are unlikely to consider each feature in an isolated fashion during actual discourse. The dynamics of discourse indicate that speakers rather consider each factor in light of the contribution it makes to the overall accessibility of the referent, and make decisions about argument realization based on that overall assessment. Second, we noted earlier that there is considerable overlap between some of the features - for example, animacy and person, disambiguation and contrast - so it is clear that these features should not be analyzed independently of each other.

Finally, we saw earlier that although many features have quite strong effects for one of their binary values, most do not function symmetrically and thus are very open to the influence of other features for the less strong binary value. This point is elaborated nicely by Serratrice (2005: 444-45) in her discussion of the relationship between disambiguation and other features. She notes that referents that are "not accessible" for disambiguation (i.e., that have competitor referents in the preceding discourse) should be realized with high information forms regardless of their accessibility status for other features. However, referents that are "accessible" with respect to disambiguation might nevertheless be realized with high information forms if they are "not accessible" for other features. In the latter case, the "not accessible" value of the other feature(s) would outrank the "accessible" value of disambiguation to result in the use of a high information form.

In the next section of the chapter, then, we turn to discussing the ways in which researchers have attempted to understand how accessibility features work in combination with each other.
5. Accessibility features working in combination

As just discussed, it is clear that interlocutors engaged in actual discourse attend to accessibility features in combination with one another rather than to each feature individually. However, it is not clear exactly how this happens. There are several possible ways in which speakers could be attending to the interaction between features, some of which we discuss next.

One simple possibility is a sort of threshold model. The idea here is that the crucial factor that the child attends to for purposes of argument realization is the distinction between zero and one feature coded as "not accessible," or perhaps between one and two features, and that additional features coded as "not accessible" do not add to the likelihood that a referent will be realized with a high information form. A second possibility would be that all features contribute equally to a final outcome, with all features essentially having identical "weight" with respect to one another. The effect of the features would then be incremental such that, for example, being inaccessible for two features means that a referent is twice as likely to be realized in an informative form as if it were inaccessible for only one feature. Yet a third possibility would involve features working together in a more complex way that should best be assessed by a regression model.

Possible models for attentiveness to feature interaction have not been much investigated in the domain of child argument realization. However, this is a crucial potential contribution of naturalistic corpora to this field because experimental paradigms are usually too constrained to investigate the complicated relationships of several features at once. Therefore, we outline next the attempts so far to study the ways in which accessibility features work in combination to affect argument realization in child speech.

5.1 Several features in one coding category

One possible way of acknowledging the interaction of features is to assume up front that certain features work so closely together that it makes no sense to separate them for individual analysis. Serratrice (2005) has done exactly this. She groups three of the most central features into one, labelled \textit{activation}, which measures the degree of "identifiability and accessibility" (Serratrice 2005:440) of referring expressions in discourse. The definition of \textit{activation} includes \textit{newness}, \textit{attention}, and \textit{topicality}. Serratrice finds, as expected, that referents are more frequently realized with high information forms when they are not activated — i.e. not jointly attended to, newly introduced into discourse, and not topical. Conversely, referents are more frequently expressed with low information forms when they are activated — i.e. jointly attended to, already introduced, and topical. This approach is also followed by Clark's (2001) work on common ground and its role in word learning.

1.2 Threshold approach

A \textit{threshold approach}, as noted earlier, takes as its starting point the assumption that all features have equivalent "weight" in determining argument realization, and that the \textit{threshold} is discriminating only up to a certain threshold in his/her sensitivity to the indicated by those features. For example, the child may attend only to the distinction between zero and one feature coded as "not accessible," and any additional coded as "not accessible" do not add to the likelihood that a referent will be with a high information form.

Narasimhan \textit{et al.} (2005) follow such an approach, coding four individual features separately (\textit{animacy}, \textit{contrast}, \textit{newness}, \textit{query}) but analyzing them all together under the general rubric of \textit{pragmatic prominence}. They consider a referent pragmatically prominent if it is tagged as "not accessible" for any one of the four features, and non-prominent if it is tagged as "accessible" for all of them, thereby assessing in their study whether general pragmatic prominence is linked to argument form. By implication, they assume that all referents contribute similarly to the speaker's overall assessment of accessibility, and that \textit{level} of accessibility for any one of the features is enough to lead to an effect on argument realization. They indeed find that lexical noun phrases are pragmatically prominent more frequently (95% of the time) than are pronominal and null arguments (64%).

Allen (2000) uses a method similar to that of Narasimhan \textit{et al.} (2005). Recall that die coded all arguments in her data set for eight discourse-pragmatic features — \textit{newtvu}, \textit{absence}, \textit{contrast}, \textit{query}, \textit{differentiation in discourse}, \textit{differentiation in context}, \textit{animacy}, and \textit{person}. For the analysis assessing features together, she considered only the first six features since the final two do not tap into accessibility in the same way. She licit coded each argument for whether it was "not accessible" for no features, one feature, or two or more features. This strategy essentially assumes that each feature contributes equivalently to overall accessibility. Allen then analyzed the data to find evidence for a threshold either between 0 and 1 features coded as "not accessible," or between 1 and 2 features. Through a logistic regression analysis, she found that the odds of an argument being realized with a high information form (here, overt as opposed to null form) were almost four times as large if the argument was coded as "not accessible" for two or more features than if it was coded as "accessible" for all features.

A second logistic regression showed that the odds of an argument being realized with a high information form (again, overt as opposed to null form) were almost three times as large if the argument was coded as "not accessible," or between 1 and 2 features, than if it was coded as "not accessible" for one or fewer features. This indicates that, licit being "not accessible" on the basis of one feature clearly has an effect on argument realization, being "not accessible" for two features has an additional effect. However, it is not clear evidence for a threshold at either point.
5.3 Incremental contribution

"Maxing out" at a threshold is not the only possibility if accessibility features indeed have equal "weight" in determining argument realization. Another logical possibility is that they could be added to each other, leading to an incremental effect whereby the likelihood of a referent being realized with a high information form increases with each feature for which it is "not accessible". Allen (2007) pursued this question. She considered only the four features which were found to be significant in Allen (2000): newness, absence, contrast, and differentiation in context. As in Allen (2000), any overt form was taken to be a form of high information, in comparison with zero anaphora and verbal agreement markers which were taken to be low information forms. She found that referents which were "accessible" for all four features were realized as overt in 18% of cases, referents coded as "not accessible" for only one feature were overt in 29% of cases, referents coded as "not accessible" for two features in 57% of cases, and for three features in 86% of cases. No referents were coded as "not accessible" for all four features, at least partly because no referent could both be absent and have competitor referents in the physical context. This result strongly suggests a cumulative effect of accessibility features in predicting argument realization. In further research, it would be interesting to investigate whether the particular features involved make a difference, or whether such an incremental effect could be obtained with any features.

5.4 Independent contribution

The most sophisticated way to determine the relative contribution of the different accessibility features is to assess them together simultaneously through a statistical model. This allows one to determine the strength of contribution of each feature given the previous or simultaneous contribution of the others. Prediction overlap between features should thus be revealed, either rendering one feature not significant while its competitor is significant, or reducing the significance of each of the competitors. Allen (2000) used logistic regression to assess the independent contribution of each of the eight features she assessed individually (newness, absence, contrast, query, differentiation in discourse, differentiation in context, animacy, and person). Logistic regression is the equivalent of multiple regression for categorical rather than continuous variables. It first evaluates the relationship between a particular outcome (here, argument realization as either overt or null) and the predictors taken together as a set. If a relationship is found, then one can assess the contribution to the outcome of each of the individual predictors. Allen found that the predictors as a set reliably distinguished between overt and null arguments (p<.001), and that the features person, contrast, newness, absence, and differentiation in context each contributed significantly to that prediction. Finally, she was able to determine the relative effect of each feature through the odds ratios generated by the logistic regression. For example, the odds of an argument being overt were higher for a explicitly contrasted referent ($e^{B}=2.0387$) than for a new referent ($e^{B}=1.3566$). Serratrice (2002) also used logistic regression in analyzing the independent contribution of each of the eight features she coded for: newness, activation, contrast, absence, query, disambiguation, person, and transitivity. She found that the predictors as a set reliably distinguished between overt and null arguments (p<.001), and that all features except newness and query contributed significantly.

Skarabela (2006) found similar results looking at newness, contrast, disambiguation in physical context, absence, and attention. The predictors as a set reliably distinguished between overt and null arguments (p<.0001), and all features except absence contributed significantly. In her analysis, contrast had a stronger effect than newness and disambiguation, which in turn were stronger than attention. This is similar to findings for caregiver data using the same procedure, except that disambiguation does not contribute significantly for caregivers. A multinomial logistic regression showed that the same five predictors as a set distinguished reliably between three levels of the dependent variable (lexical, demonstrative, omitted). All predictors contributed significantly to the selection of a lexical vs. omitted argument, while only four (not disambiguation) contributed significantly to the selection of a lexical vs. demonstrative argument (recall that Inuktitut does not allow pronouns as arguments). This is similar to findings for caregiver data using the same analysis, except that disambiguation contributes significantly to distinguishing between lexical and demonstrative forms in the caregiver data.

These logistic regression results are particularly interesting because other analyses of the same data using Chi-square analyses for each individual feature showed that every feature was significant (Allen (1997) assessed all features except attention used in Allen (2000) and Skarabela (2006); Serratrice et al. (2004) assessed all features in Serratrice (2005), except newness because it overlaps almost completely with activation). The comparison between features assessed individually and comparatively reveals the clear importance of more sophisticated statistical techniques in determining the effect of referent accessibility on argument realization. The more subtle and comprehensive assessment permitted through logistic regression clearly allows one to take into account that the features interact with one another and may not be significant once the effect of other features is factored out.

5.5 Case study of interaction between two features

In addition to considering all the possible accessibility features in one analysis to determine their relative strength and contribution to argument realization, it is also fruitful to more narrowly compare two features. Although this might normally be considered the domain of experimental studies, it is actually extremely difficult to set up an experimental paradigm uniquely testing two features in each of the four required conditions because the contexts come across as much too contrived given the short
time period and restricted dynamics of an experimental setting. Spontaneous speech interactions allow each of these conditions to naturally occur, although some conditions are obviously much more common than others.

Skarabela and Allen (2002) pursued a narrow comparison of this sort in their study of the interaction between newness and attention. From the original data set of 3168 arguments in their Inuktitut child data, they selected only those which were "not accessible" for person (i.e. all third person), and which were "accessible" for absence, query, explicit contrast, and both linguistic and physical components of disambiguation. To maximize the distinction between "accessible" and "not accessible" for newness, they only included referents which were newly referred to (i.e. "not accessible") and those which had been referred to in the immediately preceding two utterances (i.e. "accessible"). Thus, they excluded from analysis all referents which were most recently referred to in the preceding 3 to 20 utterances. The remaining referents were categorized according to the four possible conditions: both "not accessible", both "accessible", the first "not accessible" and the second "accessible", and the reverse. Skarabela and Allen then assessed which referents were expressed using lexical forms, i.e. full noun phrases, versus null forms. Expressions using demonstratives were not considered. A total of 347 referents were included in the analysis once all the relevant criteria were controlled for.

Results showed the following. When each feature was treated separately without taking the other into account, about 21% of "not accessible" referents were realized lexically (23% for newness, 20% for attention), as compared with about 5% of "accessible" referents (7% for newness, 4% for attention). The picture changed substantially once both features were treated together. Referents which were "not accessible" for both features were lexical in 64% of cases, while referents which were "accessible" for both features were lexical in only 3% of cases. Being "not accessible" for both of these features dramatically increases the extent to which a child provides a high level of information about the referent in his or her speech. All of the situations in which a referent not accessible for both features was expressed using a low information form led to breakdown in communication between the child and his/her interlocutor; 75% of those miscommunications were repaired. Referents which were "not accessible" for attention but "accessible" for newness were lexical in 12% of cases, while the reverse situation resulted in lexical expressions in 6% of cases. This indicates that, although children may be slightly more influenced by attention than by newness, accessibility on the basis of only one of these features is generally enough for a child to use a low information form. Accessibility for both features increases the likelihood of choosing a low information form, but not drastically.

This kind of direct comparison of different features is very helpful for advancing our knowledge of which features are more important to argument realization. Spontaneous speech allows for such comparison, but experiments typically do not, or at least have not yet been successful in this regard, for the reasons mentioned earlier.

5.6 Summary

Studies just reviewed show definite evidence of children's sensitivity to accessibility features working together to determine argument realization. Both the threshold incremental models show promise for further research, as well as constrained comparisons of two features and large scale statistical analyses that assess the effect of features in comparison with each other. There is clearly room for much further research in this area using much more sophisticated statistical analyses and other techniques.

6. Usefulness of extended stretches of discourse

The previous sections have highlighted two advantages of naturalistic over experimental studies of the effect of information flow on argument realization: identification of discourse-pragmatic factors influencing argument realization in natural settings, and understanding how the child attends to the interaction of these factors. We now turn to a third advantage of assessing argument realization in naturalistic corpora: it allows us to see how children are learning and displaying their knowledge across a series of related utterances in an extended stretch of discourse. Observing how a child realizes the same referent over various turns in conversation, for example, can show that child's, understanding of how the accessibility of a referent changes over time, and how this is modelled by their caregivers. We elucidate next three ways in which the interaction between referent accessibility and argument realization over extended stretches of discourse has been investigated in the literature: Preferred Argument Structure, conversational sequences, and managing miscommunication.

6.1 Preferred argument structure

As noted earlier, Du Bois (1985, 1987) has observed a powerful pattern in discourse by which the accessibility of a referent is linked not only to the form in which that referent is realized in speech, but also to the argument role held by that referent as well as the number of referents of low accessibility which can appear in one utterance. This Preferred Argument Structure shows that new and lexical (i.e. full noun phrase) referents tend to avoid appearing in A position (subject of transitive verb), and rather appear in S (subject of intransitive verb) and/or O (object) position depending on the language at hand. Further, only one lexical and one new argument typically appear within one verbal clause. While these patterns are observed at the level of the utterance, they play out as part of the dynamics of the overall flow of discourse and thus are far less likely to be observed in isolated utterances. Only looking at extended stretches of discourse in naturalistic corpora will allow us to see these patterns.
The patterns of Preferred Argument Structure have been substantiated in a large number of adult languages (e.g., Du Bois 1985, 1987; Du Bois et al. 2003). In recent years, research with child data from at least six different languages - Korean (Clancy 2003), Inuktitut (Allen and Schröder 2003), Japanese, English (Guerriero et al. 2001; Guerriero et al. 2006), Hindi (Narasimhan et al. 2005), and Venezuelan Spanish (Bentivogli 1996) - reveals that children are also following this pattern in their discourse. The Inuktitut-speaking children in Allen and Schröder’s study, for example, produced more than one lexical or one new argument in only 0.04% of utterances in their data (only 1 out of more than 2500 utterances). Further, only 1.1% of lexical arguments and 0.7% of new arguments appear in A position in the Inuktitut child data. This shows that children are highly sensitive not only to individual links between referent accessibility and argument realization, but also to the broader discourse patterns this entails.

6.2 Conversational sequences

Clancy (1996, 1997, 2003), building on the work of Vygotsky, has long been drawing attention to the importance of the structural properties of adult-child discourse for children's acquisition of grammar. She reviews evidence from conversational sequences between children and their parents to show how structured interactions provide children with sufficient information about the relationship between discourse and grammar. In her corpus of spontaneous speech of two-year-old children learning Korean, she observes that early parent-child interactions frequently involve sequences with one predicate which occurs with the same referent, and this referent is expressed by different linguistic forms across several utterances as the referent accessibility changes. For example, a new referent is introduced with a lexical noun but subsequently the same referent is most likely expressed with a null argument. Children are thus exposed to a range of argument forms that are used with few but frequent individual predicates and, simultaneously, to a range of discourse factors associated with particular referential forms.

Similarly, Clancy argues for this iconic relationship between discourse and grammar in early stages of language acquisition in her work on Preferred Argument Structure. She shows that the nature of children's early activities is the underlying source of their early sensitivity to Preferred Argument Structure. Namely, children tend to participate in activities that involve a human agent that acts on an inanimate object. In discussing these activities they tend to produce a small number of verbs, each used frequently, that represent the human agent in A position. The inanimate object tends to be encoded in O position. While the A argument is relatively constant across different contexts, the O argument frequently changes. Clancy argues that children use the O position to direct adults' attention to new information. She finds that children also use the S position to encode new information that frequently changes, although to a lesser degree. Preferred Argument Structure therefore serves an attention-focusing function in child discourse.

Clancy’s account shows that argument realization provides a powerful connection between the levels of discourse and acquisition of grammar, i.e. grammar and linguistic structures arise from discourse interactions between children and their caregivers. Similarly, however, these types of insights can be gained only from detailed analyses of stretches of naturalistic interactions and not from experimental studies.

6.3 Managing miscommunication

Naturalistic corpora also allow us to see what happens when a speaker selects a linguistic form that does not provide enough information about the intended referent. Children may misjudge the knowledge of their hearer, or may wrongly assume that their hearer possesses the same knowledge that they do about a specific referent. As a result, they may realize a referent with a low information form when it is not in fact conceptually accessible to the listener. This leads to a situation of communication breakdown. Observing extended discourse allows us to see what a child naturally does in such a situation.

Factors associated with these problematic referents. In addition, they also provide information about whether or not children are able to recover from these "errors" and how.

Skarabela (2006) studied miscommunication and its effects in 780 omitted arguments from four two-year-old Inuit children. She found that although the children tended overwhelmingly to omit arguments that represented accessible referents, they also omitted arguments representing inaccessible referents. This was particularly the case in the context of children's introduction of new referents. Interestingly, however, these omitted arguments were used to realize new referents primarily in the context of joint attention. In other words, although not using the default adult form to introduce a new referent (i.e. a lexical noun), the children did show sensitivity to whether or not a referent is conceptually accessible by taking advantage of the "low cost" of contextual information rather than by linguistic means.

In a few instances, children selected omitted arguments to represent new referents in the absence of joint attention, leading to a communication breakdown. In some cases, the conversation deteriorated and led to a different topic. In others, however, children either spontaneously repaired the miscommunication with a full lexical noun phrase or, after being explicitly asked for clarification by the listener, they used verbal or non-verbal means to disambiguate the intended referent. Overall, the children were successful at identifying the sources of ambiguity and subsequent communication repair.

These instances of miscommunication can be interpreted as episodes of useful feedback when children are provided with explicit information about what a listener considers an accessible referent and what is the best way to realize it. In this respect, "errors" in argument realization contribute to children's developing understanding that the speaker's world is not always identical with that of the listener.

This type of information about children's abilities to realize arguments and their abilities to cope with communication breakdown is again not elicited in experimental...
studies. Instead, it requires naturalistic interactions between the speaker and the listener who negotiate meaning together in the flow of conversation. In addition, this method is likely to be much more useful in discovering differences in argument realization between children and adults since controlled experiments require as a prerequisite the knowledge of “the rules of the game.” In this respect, using naturalistic speech data rather than experiments provides particularly insightful information about how children use different discourse-pragmatic factors.

6.4 Summary

The three areas just reviewed show how particular aspects of children’s understanding of the relationship between referent accessibility and argument realization are revealed across stretches of extended discourse in ways that aren’t usually visible through the isolated utterances and limited interactions characteristic of experiments in this domain. Through the lens of Preferred Argument Structure, we see that children are sensitive to the effect of discourse-pragmatic factors not only on the form in which arguments are realized but also on the syntactic role in which they appear. We see that caregivers scaffold this behavior by realizing arguments in different forms and syntactic roles as their accessibility changes through discourse, and that children’s typical early interactions also lend themselves to this patterning. Finally, we see that children are largely able to identify and correct their miscommunications resulting from speaker-hearer differences in referent accessibility.

Up to this point, we have discussed at length the values of studying argument realization in naturalistic data as opposed to experiments. However, we fully acknowledge that carefully designed experiments in this area are very helpful for triangulation of findings and can contribute important information that naturalistic data is not well suited to study. In particular, experiments allow one to assess the unique effect of one individual feature, or the relative effect of each of two features, while carefully controlling the effect of the others. Because the topic and form of interactions are controlled in experiments, they permit a better understanding of what children’s intended utterances are (and thus what they are omitting) than in naturalistic studies where one must often guess what the child intended. Experiments are also ideal for illuminating the trajectory of development because they allow for testing many children at once in different groups determined by age or linguistic ability, and because the larger number of children and data points in each group permits more powerful statistical analysis and generalizability than is generally possible in naturalistic studies. Finally, experiments enable tests of comprehension effects related to argument realization, which is not possible using naturalistic studies. In the next section we review experimental literature on children’s comprehension and production of arguments in different forms to illustrate the different ways in which experimental and naturalistic corpus studies contribute to our understanding of argument realization.

7. Experimental studies

Experimental studies of argument realization have focused on both comprehension and production of referential expressions. Only a small number of studies on children’s comprehension of reference have been reported in the literature (e.g., Arnold, Brown-Schmidt and Trueswell 2007; Arnold, Brown-Schmidt, Trueswell and Fagnano 2005; Kayama 2003; K.S. Shin 2006; Song and Fisher 2005, 2007; Tyler 1983; Wykes 1981, 1983). They have focused on features which are known to be particularly powerful, as well as features which lend themselves well to experimental situations: newness, disambiguation in the linguistic context, and topicality. In a typical experiment, the child is presented with one to three sentences of a story context which exhibit one or more of the accessibility features, and then assessed to see how they determine the reference of an ambiguous form (usually a pronoun) in the final target sentence. On the basis of response times in a mispronunciation task, for example, Tyler (1983) suggested that 5-year-olds could interpret a pronoun as coreferent with the subject of a preceding story. Wykes (1983) showed that 5-year-olds could use the semantics of a preceding sentence to correctly interpret the antecedent of a later pronoun. Arnold et al. (2005) predicted that gender information would be a stronger cue to pronoun reference than order of mention in a preceding sentence or appearance of the antecedent in subject position, and showed the strength of gender as a cue in an off-line task with 3.5- to 4.5-year-olds. Song and Fisher (2005, 2007), in contrast, showed in a series of preferential looking studies that children aged 3.0 and 2.5, respectively, could identify referents which were made prominent by virtue of being mentioned first in the story, appearing in subject position, being mentioned more often than other referents, and being pronominalized once. Song and Fisher (2005) also succeeded in carefully comparing the effect of disambiguation vs. topicality in 3-year-olds. Use of different methodologies across studies - including act-out tasks, looking-preference paradigms and eye tracking - allows for triangulation of findings and investigations at ever-younger ages. In general, these studies have confirmed the attentiveness of children to all of these features in assessing reference.

At least five experimental production studies have recently investigated the effect of information flow in argument realization (Campbell, Brooks and Tomasello 2000; Gürçanh, Nakipoglu and Özyürek 2007; Matthews, Lieven, Theakston and Tomasello 2006; N.L. Shin 2006; Wittek and Tomasello 2005). Most of these studies focus on either newness or attention, although a couple also investigate disambiguation and absence. In the typical methodology, a referent is presented and engaged in some action, either physically or on video. Then the participant is asked to request the referent from or recount the action to an experimenter who either did or did not observe the initial action. The linguistic form for the referent produced by the participant is assessed to determine whether it conforms to expectations based on the accessibility features present in the study. Since production (rather than comprehension) is the focus of this
We discuss these studies in some detail next to illustrate the strengths and difficulties of experimental studies in comparison with studies based on naturalistic data.

7.1 Strengths of production studies

An excellent example of a study assessing the unique effect of one feature comes from Matthews et al. (2006), who studied argument realization in one hundred English-speaking children, a third each aged 2, 3, and 4 years. The first part of the study focused on the role of attention, holding all other features constant (all referents were "not accessible" for newness and person, and all were "accessible" for absence, query, disambiguation, explicit contrast, and animacy). Participants viewed 10 short video clips (e.g., clown jumping, fairy eating an apple). For one block of 5 clips, the experimenter was watching the screen with the child (i.e. "accessible" for attention); for the other, the experimenter was not able to view the screen (i.e. "not accessible" for attention). After viewing each clip, participants were asked to recount the clip to the experimenter with the request: What happened? What did you see? Results showed that the 3- and 4-year-olds, but not the two-year-olds, chose different linguistic forms (noun vs. pronoun) to realize the referents depending on whether the interlocutor shared attention to the video or not. The second part of the study held attention constant (= "not accessible") and varied newness. The same participants narrated another block of 5 similar clips to an experimenter who could not see them, and whether the experimenter mentioned or did not mention the target referent prior to asking the child what happened in the clip (e.g., What is the clown? Oh! What happened? vs. That sounds like fun! What happened?). All three age groups differentiated their choice of referring expression depending on whether the referent had just been expressed by the experimenter or not. The authors concluded that both attention and newness influence children’s argument realization, but that children are sensitive to the effect of newness earlier than to attention. Campbell et al. (2000) found an identical pattern of results in a study with very similar design but where the English-speaking participants viewed real-life events (rather than videos) involving inanimate (rather than animate) toys moving in different ways (e.g., being pulled on a train, sliding down a chute).

Wittek and Tomasello (2005) again found similar results in tests of newness and contrast within a similar design (both features tested but not assessed relative to each other, attention always "not accessible"). In their study, German-speaking children were asked to retrieve inanimate objects from a shelf where the children had placed them earlier, using the requests What happened to the broom? ("accessible" for both newness and contrast), What do we need to get? ("not accessible" for newness, "accessible" for contrast), and Did the clown have a vacuum cleaner? ("not accessible" for both newness and contrast). Children aged 2,5 and 3,5 used low information forms for the first question and high information forms for the other two. Children aged 2,9 produced overwhelmingly high information forms for the final question suggesting sensitivity to the effect of contrast even at this young age. However, they did not differentiate in their use of forms between the first two questions, suggesting that children become sensitive to newness sometime between 2,0 and 2,6.

All three of these studies are quite carefully designed for the most part, succeeding well in being as natural as possible and eliminating the effect of confounding features. I hey also illustrate well how age effects can be shown from experimental studies, since a large number of children within a narrow age range can be studied.

7.2 Difficulties with production studies

Other studies illustrate some of the difficulties of a less naturalistic interaction and of controlling for the effects of features other than the one under investigation. Gürçanh et al. (2007) conducted a study close in design to the first part of Matthews et al. (2006), but with only one age group of Turkish-speaking children (3;0-4;11), as well as a group of adults, who experienced the two values of attention as a between-subjects rather than within-subjects condition (i.e. half the participants recounted the video clips to an experimenter who viewed the clips with them, the other half to an experimenter who was out of the room when the video was shown). Although the children produced more high information forms in the condition where the experimenter did not share attention during the video, the adults showed no difference in the two conditions. The adult participants seem to have interpreted this as a "test question" rather than as a real interaction situation because they responded using high information forms for all questions rather than following patterns that would be characteristic of a real conversation with natural dynamics of information flow. This may have happened because it is quite strange for someone to ask you to recount to them something they have just seen, and because this task was very easy for adults.

N.L. Shin (2006) study testing the sensitivity of children to disambiguation also suffers from unnaturalness as well as difficulty in controlling the relative effects of the other features. Participants were 181 monolingual speakers of Mexican Spanish aged 5,9 to 15,8 and 30 adults. They were told 12 two-sentence stories acted out by animate figures (e.g., Maria and Jose sing songs. Maria sings a ranchera.), with the third and final sentence acted out but not spoken (e.g., one of the characters sings a children's song). Under the rationale of helping the non-native researcher better understand how to speak Spanish, the participants were then asked to select from two orally-provided options which third sentence would best complete the story (e.g., Later, he/she sings a children's song. or Later, she/he sings a children's song.). In the example given, the null subject option would be preferred if Maria sings the children's song since there are no intervening competitors. In contrast, the pronoun option (here he) would be preferred if Jose sings the children's song because the competitor referent Maria intervenes between the current and previous mention of the target referent. All features other than disambiguation are controlled for, with a "not accessible" value for person and "accessible" values for all other features. The results revealed that the younger children tended to overuse null forms in situations where there was a competitor referent, indicating
that they found the referent accessibility provided by the other features - probably particularly recency of mention of the referent and sharing attention with their interlocutor - far more powerful than the lack of accessibility incurred by disambiguation. Further, the experiment must have been somewhat repetitive and confusing since all 12 experimental items as well as 9 of 12 filler items involved the same 2 characters, Jose and Maria. In addition, a group of older children produced pronouns in many contexts where there was no competitor referent, implying that they were probably treating the questions as "test questions" rather than abstracting to natural interactional patterns.

Finally, the third study in Wittek and Tomasello (2005) illustrates the difficulty of controlling for the effect of other features. German-speaking children aged 2;6 and 3;6 again played games with an experimenter and placed the toys on the shelf after use. The features absence and disambiguation (in physical context) were manipulated, but not in comparison with one another; all other features were held constant with newness, attention, animacy and person "not accessible" and the other features "accessible": This time, each toy was placed on a shelf either out of sight in a box ("not accessible" for absence, not relevant for disambiguation), right next to another toy ("not accessible" for disambiguation, "accessible" for absence), or in a separate location on the shelf where it could be individuated by a point ("accessible" for both features). Children were instructed to ask a second experimenter (who did not witness placing of the items on the shelf) to get particular items. Although the items differed in whether they were visible and easily individuated, both groups of children typically asked for all items using noun phrases rather than pronouns, null references, or simply pointing. The authors concluded that children essentially ignored relative physical location of the objects in this task, responding much more powerfully to the lack of shared knowledge of the objects (both newness and attention) in any of the conditions.

7.3 Summary

Experimental assessments of the effect of accessibility features on argument realization have some advantages over naturalistic ones. In particular, they allow for singling out the effect of a particular feature holding all others constant, and they allow careful testing of the effect of features by age. However, no experiments have yet succeeded in comparing the effect of two features within one design, much less assessing the relative effect of all nine features taken together. This is a clear limitation in determining how argument realization works in the actual child who is attending to all of these features at once. Further, the studies cited above reveal the difficulty of assessing the effect of less powerful features such as disambiguation and absence in the face of the much more powerful features newness and attention. The latter essentially outweigh the effect of the former in the sorts of contrived situations that are the necessary setting for experiments. Even some of the authors of the experimental studies themselves point out that the experimental situations were somewhat extreme in that the information provided (or not) by the discourse was quite explicit, lacking much of the subtlety and influence from other factors that would be typical in natural interaction (e.g., Matthews et al. 2006). In addition, asking questions about simple events that both participant and experimenter had just witnessed is quite unnatural since these sorts of questions would be unlikely to occur in natural conversation, and clearly cannot really be asking for information. Thus, it is not clear how much the children's responses reflect their natural speech, or whether they have simply learned the expected formula to answer test questions (N.L. Shin 2006; Gürcanli et al. 2007). Nevertheless, the children in most of the studies showed clearly their knowledge of the factors tested, and altered their realization of arguments accordingly. This indicates at minimum that children use the same kinds of discourse-pragmatic knowledge in answering test questions that they do in real conversation.

Discussion and conclusion

We have seen throughout this chapter that naturalistic studies have made an important and unique contribution to our understanding of the relationship between discourse and syntax. Naturalistic studies are essential for identifying the accessibility features that affect argument realization in every-day discourse, for investigating the ways in which children attend to those features in combination with each other, and for uncovering how children learn about the application and misapplication of those features through extended stretches of discourse with familiar interlocutors. Experiments also provide a valuable contribution to this understanding. They allow for singling out the effect of individual features, controlling the contexts of interaction to constrain the effects of competing features, systematically comparing children of different ages and linguistic ability, and amassing a large sample of participants to facilitate generalizability and powerful statistical analysis. However, the interactions are usually more contrived, and experiments are limited in the number and type of features they can investigate at once. Information from both sources - experiment and naturalistic data - clearly have their place. But without the dimensions and understanding that naturalistic data provide, we lose an essential piece of the puzzle of how children's knowledge of the principles of discourse affects their syntactic production.

Studies of argument realization from the perspective of cognitive accessibility have looked at nine main features: newness, topicality, absence, query, disambiguation, explicit contrast, person, animacy, and attention. Each of these has been defined slightly differently in different studies, but most have been found to be significant factors in children's argument realization. The strongest factors are likely newness, explicit contrast, topicality, and absence. These have been found significant in several studies looking at individual factors, as well as in studies using logistic regression and incremental approaches to assess the combined or comparative effect of these factors on argument realization. Query, disambiguation, and animacy have generally been found less consistent in effect across studies. Although attention has not been studied enough
Using corpora to examine discourse effects in syntax

Shanley Allen, Barbora Skarabela and Mary Hughes

to be sure of its effect, it seems very promising in both naturalistic and experimental studies. Person is also powerful, but is less clearly a feature of accessibility than the others, and is more strongly affected by language typology (e.g., whether a language allows arguments of a certain person to appear in a particular form). These features will provide a strong base for further exploration of gaps in grammatical and performance-based claims discussed earlier. In particular, further research should use them to investigate which of the arguments shown by those perspectives to be candidates for omission - for instance, subjects of non-finite clauses, subjects of matrix clauses, subjects with long VPs - are actually omitted.

Developmental results from both naturalistic and experimental studies discussed here also bear on theoretical questions raised in the literature. The results of several studies converge to show development in children's sensitivity to several features. Experimental results suggest that contrast is one of the earliest features to be attended to, followed by newness and attention sometime between 2;0 and 2;6. Naturalistic data also show that newness, person, and activation are probably attended to early on but also are attended to more fully with age. Children's decrease in use of null arguments during the period of 2;0 to 3;0 could thus be at least partially explained by an increase in their attention to accessibility and in their sensitivity to the effect of accessibility features on argument form. The timing could also help to explain children's increase in use of strong pronouns and lexical forms towards the end of this time period. Finally, developmental data from Guerriero et al. (2006) can help to resolve the debate between the generative and performance perspectives concerning the continuity of null to pronominal forms. Guerriero et al. clearly showed that pronouns increased as null forms decreased from ages 1;9 to 3;0, while lexical forms remained fairly steady. The evidence seems to support strong continuity, consistent with the prediction of Hyams and Wexler (1993). Although the developmental results obtained to date are very useful, much research remains to be done in assessing the trajectory of development of other features not yet studied, in assessing how and when children become sensitive to the interaction of features, and in determining the mechanisms for children's learning.

As Valian (1991) first pointed out, children are quite sensitive to the typology of their language from very early on in terms of the proportion of arguments omitted in their speech. While speakers of both non-null-subject and null-subject languages overomit arguments at the earliest stages, the proportion of omissions is much higher for speakers of null-subject languages. As mentioned in the introduction, virtually all studies of the effect of accessibility on child argument realization have been conducted in null-subject languages. Only three studies reported here investigate a non-null-subject language (English): Guerriero et al. (2006), Hughes and Allen (2006) and Serratrice et al. (2004). Each of these show that English-speaking children attend to accessibility features according to the same patterns as their counterparts speaking null-subject languages. However, no study to date has investigated particular grammatical contexts in English such as subjects of finite vs. non-finite verbs to determine how grammatical and accessibility constraints interact. This would be a very fruitful direction for future study.