The incremental effect of discourse-pragmatic sensitivity on referential choice in the acquisition of a first language

Mary E. Hughes a,*, Shanley E.M. Allen b,**

a Boston University, USA
b University of Kaiserslautern, Germany

Received 20 December 2012; received in revised form 7 March 2014; accepted 7 March 2014

Abstract

Previous research has demonstrated that children as young as 2;0 are sensitive to discourse-pragmatic context when selecting referring expressions. If a referent is present in the discourse context and/or jointly attended to by the listener, a child will be more likely to omit a referring expression or use a pronominal form. To date, most research has examined the effect of individual features in the discourse, whereas in reality the various features occur and work together. In this study, we explore children’s sensitivity to the incremental effect of six discourse-pragmatic features. This stepwise approach is a more nuanced approach to measure the cumulative effect of accessibility on argument realization in order to reveal the predictive patterns of accessibility. Videotaped data from four monolingual English-speaking children in spontaneous interaction with their caregivers are analyzed at two ages: 2;0–2;7 and 3;0–3;1. Caregivers and children at both ages are sensitive to incremental effects of accessibility. Their selection of linguistic forms follows a cline such that low information forms are utilized for accessible referents. As referents become increasingly inaccessible, they are more likely to be realized as high information forms. These results indicate another important dimension of children’s sensitivity to discourse-pragmatics and have implications for assumptions about children’s theory of mind.

Keywords: Argument realization; Referential choice; Discourse-pragmatics; Subject omission; Acquisition of reference

1. Introduction

It is by now well established that children as young as two years of age are sensitive to the knowledge of their interlocutors in selecting appropriate linguistic forms to realize subjects and objects in spontaneous speech (see review in Allen et al., 2008). In particular, they are sensitive to the conceptual accessibility of a referent, defined as “the ease with which the mental representation of some potential referent can be activated in or retrieved from memory” (Bock and Warren, 1985:50). This conceptual accessibility is determined by a set of discourse-pragmatic features including PRIOR MENTION (whether the referent is new to the discourse or has already been introduced), PHYSICAL PRESENCE (whether the referent is absent from or present in the physical context), and JOINT ATTENTION (whether the speaker and interlocutor are jointly attending to the referent at the time of speech or not) (e.g., Chafe, 1987; Du Bois, 1987; Givón, 1983; Gundel et al., 1993; Prince, 1985). Speakers then direct

* Corresponding author. Tel.: +1 15087889267.
** Corresponding author at: Psycholinguistics and Language Development Group, Faculty of Social Sciences, University of Kaiserslautern, Erwin-Schrödinger-Straße 57/409, 67663 Kaiserslautern, Germany.
E-mail addresses: mhughes@bu.edu (M.E. Hughes), allen@sowi.uni-kl.de (S.E.M. Allen).

http://dx.doi.org/10.1016/j.lingua.2014.03.001
0024-3841/© 2014 Published by Elsevier B.V.
their interlocutors’ activation or retrieval of the intended referents by using particular forms in speech to signal the level of accessibility of the referents (Ariel, 1994:99). Referents that are not accessible to the interlocutor – for example, referents newly introduced into the discourse – are typically realized with maximally informative forms like lexical noun phrases as in Speaker A’s utterance in (1). In contrast, arguments that are already accessible to the interlocutor through one or more discourse-pragmatic features – for example, referents mentioned in the preceding utterance – are typically realized with reduced forms like pronouns, as in Speaker B’s utterance in (1).

(1) Speaker A: Good morning, Bob. What a nice shirt you’re wearing!
Speaker B: It doesn’t fit me anymore.

This process of referential choice reflects sophisticated coordination of linguistic knowledge (which linguistic forms are permitted in a given language), discourse knowledge (how accessible referents are in the discourse and physical context), and social knowledge (what the interlocutor is attending to at the time of speaking).

In investigating this process, previous researchers have examined discourse-pragmatic features individually, compared the interaction of two or more discrete features, or looked at the effect of accessibility in general without differentiating individual features (e.g., Allen, 2000; Clancy, 1993, 1997; Guerriero et al., 2006; Hughes and Allen, 2013; Narasimhan et al., 2005; Serratrice, 2005; Skarabela et al., 2013; see Allen et al., 2008 for a complete review). In these studies, researchers code spontaneous speech or experimentally test one or more individual discourse-pragmatic features. They then assess children’s sensitivity to each feature in isolation and, if more than one feature is tested, often determine which one of the various features children are most sensitive to. However, it is very unlikely that children in fact attend to each individual discourse-pragmatic feature in isolation. Rather, they almost certainly attend to the various features in interaction and in combination with each other.

This sensitivity to the interaction between features has been much less well studied than sensitivity to the features in isolation even though the former is a better approximation of how things happen in the real world. Only one study thus far (Allen, 2007) has examined the incremental effect of accessibility by adding features together, one at a time in a stepwise manner, in order to investigate whether the resulting increase in “inaccessibility” determines the use of a high information forms, such as a lexical NP. Allen (2007) found positive results using this method in a null subject language, Inuktitut, but she used only a few features in a relatively small data set. The goal of the present study is to test this method more thoroughly, by including more accessibility features and using data from a language that does not permit null subjects. In this way, the present paper seeks to extend the literature on the children’s sensitivity to the interaction between features.

1.1. Previous approaches to studying the interaction of features

Numerous researchers who have investigated children’s acquisition of referential choice have found that preschool children are sensitive to a number of individual features of discourse and social knowledge. These include the three features mentioned earlier (PRIOR MENTION, PHYSICAL PRESENCE, JOINT ATTENTION) as well as several others including CONTEXTUAL DISAMBIGUATION (whether a particular referent has potential competitor referents in the physical context), LINGUISTIC DISAMBIGUATION (whether a particular referent has potential competitor referents in the linguistic context), ANIMACY (whether the referent is animate), TOPICALITY (whether the referent is the focus of conversation), QUERY (whether the referent is the subject of or the response to a question), EXPLICIT CONTRAST (whether a contrast is made explicit or emphasized by the speaker), and PERSON (whether the referent is first/second vs. third person) (see Allen et al., 2008). Further, this sensitivity to referent accessibility is not restricted to one language or one language typology. Similar results have been found in numerous languages1 including English (Campbell et al., 2000; Graf, 2010; Guerriero et al., 2006; Gundel et al., 2007; Hughes, 2011; Hughes and Allen, 2013; Matthews et al., 2006; Mishina-Mori, 2007; Serratrice et al., 2004), French (De Cat, 2004; Salazar Orvig et al., 2010a, 2010b), German (Schmitz, 2007; Wittek and Tomasello, 2005), Hindi (Narasimhan et al., 2005), Inuktitut (Allen, 2000; Allen and Schröder, 2003; Skarabela, 2007a, 2007b; Skarabela et al., 2013), Italian (Serratrice, 2005; Serratrice et al., 2004), Japanese (Guerriero et al., 2006; Mishina-Mori, 2007), Korean (Clancy, 1993, 1997), Mandarin (Huang, 2011; So et al., 2010), Spanish (Paradis and Navarro, 2003; Shin and Cairns, 2012) and Turkish (Demir et al., 2012; Gürcanli et al., 2007). Overall, preschool children are sensitive to a variety of individual discourse-pragmatic features in choosing whether to use an informative or reduced form to realize a referent in speech, and they are able to integrate that sensitivity with the language-specific possibilities for referential choice.

As research in the field progressed, it became increasingly clear that children were not only sensitive to these features individually, but they were sensitive to the interaction of these features. Several different approaches have been used so

---

1 All studies are with children aged 2–4 years except for Shin and Cairns (2012) who elicited data from Spanish-speaking children between the ages of 5:09–15:08.
far to gain insight into the sensitivity to this interaction between features. One approach has been to assume from the outset that the features are attended to in combination and thus to simply treat them as one feature. Serratrice (2005) follows this approach in her study of Italian-speaking children between the ages of 1;7 and 3;3. She assesses the feature ACTIVATION which measures the degree of “identifiability and accessibility” of a referent (p. 440) – essentially a combination of PRIOR MENTION, JOINT ATTENTION to the referent, and TOPICALITY. Clark and colleagues discuss the influence of common ground on referential choice for French-speaking children at 2;3 and 3;6 (Clark and Bernicot, 2008; Clark and Amaral, 2010). Common ground refers to the information that is “mutually known” between speaker and interlocutor (E. Clark, 2009:282; see also H. Clark, 1996), and thus is also essentially a combination of PRIOR MENTION, JOINT ATTENTION to the referent, and TOPICALITY. This approach has the shortcoming of not being very fine-grained – it is clear that the referents are accessible, but it is not clear what in particular makes them accessible.

A second approach has gone in the opposite direction, statistically assessing the independent contribution of each of several features (Allen, 2000; Hughes, 2011; Hughes and Allen, 2013; Skarabela, 2007a). For example, Allen (2000), who studied Inuktitut-speaking children from ages 2;0 to 3;6, found using logistic regression that an explicitly contrasted referent had a greater likelihood of being realized with an informative form (i.e., demonstrative or noun phrase) than a referent newly introduced into discourse, which in turn had a greater likelihood of being realized with an informative form than a referent with potential competitors in the physical context. This approach is somewhat problematic, however, given the difficulty of having enough data points in spontaneous speech data to meet the requirements of logistic regression.

A third approach has been to mathematically assess the combinatory effect of the different discourse-pragmatic features. In this approach, the important factor is the number of discourse-pragmatic features that a particular referent is accessible or not accessible for. It is implicitly assumed that each of the features has an equivalent effect on referential choice, and so it is of less importance exactly which features a referent is or is not accessible for. As an example, take the referent “shirt” in (1). This referent is introduced in Speaker A’s utterance where it is “accessible” for four of the previously mentioned discourse-pragmatic features (PHYSICAL PRESENCE, LINGUISTIC DISAMBIGUATION, QUERY, EXPLICIT CONTRAST) and is “not accessible” for six of them (PRIOR MENTION, JOINT ATTENTION, CONTEXTUAL DISAMBIGUATION, TOPICALITY, ANIMACY and PERSON). Therefore, its inaccessibility score is 6 out of 10 – a relatively high score that leads to its realization as a noun phrase by Speaker B. Because this is a low score, the referent is realized as a pronoun in Speaker B’s utterance.

The simplest mathematical method is to assume a “threshold” of accessibility – a point below which children assume that a referent is more accessible, and above which they assume that it is less accessible (Allen, 2000; Narasimhan et al., 2005). In their study of Hindi-speaking children’s sensitivity to four discourse-pragmatic features (ages 3;0–4;0), Narasimhan et al. (2005) found that children differed in their referential choice for referents that were pragmatically prominent (i.e., “not accessible” for any one or more of the four features) vs. those that were not pragmatically prominent (i.e., “accessible” for all features). Thus, their threshold was set at the point of “not accessible for one feature” (i.e., 0 vs. 1+ features). Allen (2000) tested two thresholds: one set at “not accessible for one feature” (i.e., 0 vs. 1+ features), and one set at not accessible for two features (i.e., <2 vs. 2+ features). Children’s referential choice showed sensitivity to both of these thresholds.

A second mathematical method is to explore children’s sensitivity to the interaction between two particular features while controlling for the effect of other features. For example, Skarabela and colleagues (Skarabela and Allen, 2002; Skarabela et al., 2013) used Inuktitut spontaneous speech data from children ages 2;0–3;6 and coded for eight discourse-pragmatic features (PRIOR MENTION, PHYSICAL PRESENCE, JOINT ATTENTION, CONTEXTUAL DISAMBIGUATION, LINGUISTIC DISAMBIGUATION, ANIMACY, QUERY, EXPLICIT CONTRAST) to explore the interaction between two of the features: PRIOR MENTION and JOINT ATTENTION. To hold the effect of the other six features constant, only referents that were coded as “accessible” for all of those features were included in this analysis. When the effect on referential choice of each of PRIOR MENTION and JOINT ATTENTION was assessed separately, the two features had a similar magnitude of effect: lexical noun phrases were used to realize only 8% of referents coded as accessible for PRIOR MENTION and 21% of referents that were accessible for JOINT ATTENTION. In comparison, lexical noun phrases were used to realize only 7% of referents coded as “accessible” for PRIOR MENTION and only 4% of referents coded as “accessible” for JOINT ATTENTION. However, the children were very sensitive to the combinatory effect of the features: 64% of referents coded as “not accessible” for both of the two features were realized lexically compared to only 3% of referents coded as “accessible” for both features.

A more sophisticated mathematical method is determining the incremental effect on referential choice of a whole set of discourse-pragmatic features. As in the two other mathematical methods described, a referent that is “accessible” for all features is least likely to be realized with an informative linguistic form (i.e. demonstrative or noun phrase). The likelihood of being realized with an informative form increases with each additional feature for which the referent is not accessible. Thus, a referent that is coded as “not accessible” for four features is more likely to be realized with an informative form than a referent that is coded as “not accessible” for only three features, etc. Allen (2007) tested this premise in
spontaneous speech data from Inuktitut-speaking children ages 2;0–3;6 using four features which were found to have significant effects on referential choice in previous studies (PRIOR MENTION, PHYSICAL PRESENCE, EXPLICIT CONTRAST, and CONTEXTUAL DISAMBIGUATION). A clear and subtle relationship between accessibility and referential choice was found as shown in (2). (Note that informative forms here include demonstratives and noun phrases, and that no referents were coded as “not accessible” for all four features.)

(2) a. referents “accessible” for all four features → 18% realized with informative forms
   b. referents “not accessible” for one feature → 29% realized with informative forms
   c. referents “not accessible” for two features → 57% realized with informative forms
   d. referents “not accessible” for three features → 86% realized with informative forms

These results show that children are not only sensitive to a threshold of features or to the effect of two features in combination, but are also sensitive to the relatively more subtle incremental effect of features.

This incremental approach is very promising for determining just how subtle children’s sensitivity to accessibility is. However, it has only been tested in one language and should ideally be replicated in other languages of different typologies. Importantly, Inuktitut allows referents to be realized as null forms (i.e., the referent is omitted), and it does not allow pronouns to appear in either subject or object position. Therefore, the difference between a reduced form (null form) and a more informative form (demonstrative or lexical noun phrase) is maximally clear. It is not yet known whether children’s sensitivity would be so subtly revealed in a language like English where null forms are not grammatical and the typical reduced form is a pronoun. Further, Allen’s (2007) study did not assess the incremental effect of accessibility in caregiver data, and did not look for any effect of development.

1.2. The present study

The purpose of the present study is to fill these gaps in the literature. We assess the incremental effect of six discourse-pragmatic features in argument realization in a set of spontaneous speech data from four English-speaking children at two time points (2;0–2;7 and 3;0–3;1), as well as in data from their caregivers. The role of child-directed speech in the acquisition of referential forms has already been established in several previous studies. For example, Küntay and Özyürek (2006) examined the production of demonstratives in monolingual Turkish-speaking adults and children. Their results showed a pattern of demonstrative use when adults spoke to children as opposed to speaking to other adults, and the children mirrored this non-adult-like pattern. In another study, Guerriero et al. (2008) compared monolingual English-speaking mothers and children to monolingual Japanese-speaking mothers and children. The researchers found different patterns of development in the production of referents for English-acquiring and Japanese-acquiring children. These differences reflected a difference in input between the two groups of parents, both in choice of referents and in terms of interactional style. Rozendaal and Baker (2010) investigated English-acquiring children’s acquisition of determiners and pronouns in terms of the given vs. new distinction in discourse, and found that the children’s development of the determiners occurred at a faster rate than their pronoun use. This appeared to be a direct effect of the adult input data, which revealed inconsistency in pronoun usage. In light of these findings, we believe it is important to consider the shape of the input data in this study in order to investigate possible parallel patterns in terms of incremental effects of accessibility in the adult and child data.

We have already found in earlier studies that the children in the present study show sensitivity in their referential choice to each of six individual discourse-pragmatic features (Hughes and Allen, 2006, 2013). Here we build on that study by assessing the effect of these same features incrementally. We hypothesize that English-speaking children will show incremental sensitivity to discourse-pragmatic features in the same way as Inuktitut-speaking children, regardless of the differences in language typology. We also expect that caregivers will show high sensitivity to the incremental effect of discourse-pragmatic features. Finally, we expect that the three-year-old children will show a more sophisticated incremental sensitivity than will two-year-olds.

2. Method

2.1. Participants

Data for this study were taken from the Manchester-Max Planck Dense Database (Lieven et al., 2003, 2009). Participants were the same four monolingual English-speaking children studied in Hughes and Allen (2013) – Annie, Brian, Eleanor, and Fraser. The children lived with their families in a large metropolitan area in England and came from middle-class backgrounds. All four children were audi-taped for several hours each week and videotaped for one hour each week at ages 2;0–2;1 and 3;0–3;1. One child, Brian, was also taped throughout the entire year. For this study, we
used only the videotaped data (one hour per week per child) so that non-verbal aspects of the interaction and the physical context could be included in the analyses.

2.2. Data

The transcripts and videotapes, which were generously made available to us by Elena Lieven, recorded the children in spontaneous interactions with their mother and/or a familiar researcher (see Lieven et al., 2003, 2009 and Hughes and Allen, 2013, for further details about the recording, transcription, and coding process). In order to assess the children's development and based on previous findings that children's sensitivity to accessibility changes over time (e.g., Hyams, 1986; Guerriero et al., 2006; Serratrice et al., 2004; Salazar Orvig et al., 2010a,b), the children's utterances were assessed at the two available time periods. The ultimate goal was to analyze the children's utterances when they were all at similar linguistic levels. These two time periods will be referred to throughout the remainder of this article as Time 1 (2;0–2;1 for 3 children and 2;4–2;7 for 1 child²) and Time 2 (3;0 and 3;1 for all four children). From the available data, the following were excluded: utterances that were not fully intelligible; utterances that did not contain a verb; and utterances that were exact self-repetitions, exact imitations, recitation of poems or songs, frozen forms, and routines. This yielded a total 1901 utterances from 25 files at Time 1 and 3177 utterances from 21 files at Time 2. These 5078 utterances were then coded for grammatical and discourse-pragmatic information as described in the following sections.

In the analysis, we focused on the effect of referent accessibility on referential choice. Therefore, we restricted the data set for analysis to contexts in which the accessibility of the referent in the discourse context could be determined to have an effect on the linguistic form produced by the speaker as opposed to the linguistic form being fixed for other reasons. This led to excluding the utterance types listed in (3):

(3) a. imperative utterances, because subjects of imperatives must be omitted for grammatical reasons in English (Time 1: 227; Time 2: 317)
b. utterances containing first and second person subjects, because first and second person referents are by definition virtually always fully accessible in the discourse (Time 1: 627 first person and 60 second person; Time 2: 1194 first person and 278 second person)²
c. utterances in which children and adults use proper names to refer to themselves or their interlocutors instead of first and second person pronouns (Time 1: 48; Time 2: 88; see Hyams, 2008, and Hughes and Allen, 2013 for more explanation)
d. utterances in which the existence of joint attention could not be determined due to the positioning of the video camera or movement out of frame by the participants (Time 1: 67; Time 2: 118)
e. utterances containing subjects that are explicitly contrasted with an alternative referent using stress or tone of voice (e.g., ‘This cup is mine and that cup is yours’) because this feature has been shown in previous studies to be categorical in English, necessitating rather than simply influencing the use of an overt subject (Time 1: 25; Time 2: 57)
f. utterances containing subjects that are the referent of or the answer to a question (e.g., referent of question: ‘What broke?’; answer to question: ‘The vase broke.’) because this feature has been shown in previous studies to be categorical in English, necessitating rather than simply influencing the use of an overt subject (Time 1: 124; Time 2: 97).

By including only utterances in which argument choice could truly be affected by a referent's accessibility in discourse, our analysis provides a clearer picture of the actual effect of the six discourse-pragmatic features in this study.

The numbers for the final set of data used for analysis are given in Table 1, in which the age, number and length of sessions, the MLU, and the number of possible subject contexts for each child at each of Time 1 and Time 2 are listed.

In order to examine the state and effect of the caregiver input, 15% of the utterances from the four mothers were randomly selected and coded using the same criteria as outlined above for the child data, which yielded a total of 1129 possible subject contexts for the four mothers.

² One child, Brian, had a somewhat lower mean length of utterance (MLU) than the other three children during these earliest recordings. Since we also had later recordings for him, we used a slightly later set of data for his Time 1 – from 2;4 to 2;7 – so that the linguistic level of all four children would be as comparable as possible.

³ At Time 1, the children omitted subjects in 37% of all first person utterances and 60% of all second person utterances. These percentages are significantly higher than the percentage of omitted third person subjects at Time 1 (24%). At Time 2, the children omitted far fewer subjects but still omitted 16% of all first person subjects. This is significantly higher than the proportion of second and third person omissions (3% and 4% respectively), which show a more adult-like rate of omission.

Please cite this article in press as: Hughes, M.E., Allen, S.E.M., The incremental effect of discourse-pragmatic sensitivity on referential choice in the acquisition of a first language. Lingua (2014), http://dx.doi.org/10.1016/j.lingua.2014.03.001
2.3. Accessibility features

All third person subjects from both the child and mother data that fit the above criteria were coded for accessibility using six discourse-pragmatic features: ANIMACY, CONTEXTUAL DISAMBIGUATION, PHYSICAL PRESENCE, PRIOR MENTION, LINGUISTIC DISAMBIGUATION, and JOINT ATTENTION. These are the same six features that were already assessed individually in Hughes and Allen (2013).

We limited our analysis to subject contexts for several reasons. First, there is a known asymmetry between subject and object omission in English that has been observed by researchers from different theoretical perspectives (e.g.,Processing: Bloom, 1990; Competence: Hyams and Wexler, 1993; Discourse-pragmatic: Allen, 1997). The issues involved in the realization of subjects vs. objects are substantially different in certain respects, such as the higher likelihood that objects represent new information. Second, there are very few null objects in our data (i.e., <2%), thereby making it difficult to analyze the full range of forms that we are investigating. Third, restricting the analysis to subject contexts is in line with much of the previous research on the acquisition of referential forms (e.g., Hyams and Wexler, 1993; Orfitelli and Hyams, 2012; Paradis and Navarro, 2003; Salazar Orvig et al., 2010a; Serratrice, 2005; Shin and Cairns, 2012). For all of the reasons stated above, we believe that restricting our analysis to subject contexts presents a clearer picture of the effects of discourse-pragmatics than would analyzing subjects and objects together.

Each subject in the data was assigned a binary value for each feature based on how accessible its referent was in the discourse context (e.g., Ariel, 2001; Clancy, 1993; Allen, 2000; Skarabela, 2007b). An accessible value indicates that the referent is easily identified in the discourse context, so the speaker is likely to use a reduced form to realize the referent in speech (e.g., null form, pronoun). An inaccessible value indicates that the referent is not easily identified in the discourse context, so the speaker is likely to use a more informative form to realize the referent in speech (e.g., demonstrative, noun phrase). Table 2 lists the six features and how accessibility vs. inaccessibility was defined for each.

---

Table 2

<table>
<thead>
<tr>
<th>Accessibility feature</th>
<th>Accessible value</th>
<th>Inaccessible value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANIMACY (AN)</td>
<td>Animate</td>
<td>Inanimate</td>
</tr>
<tr>
<td>CONTEXTUAL DISAMBIGUATION (CD)</td>
<td>Only referent in physical context</td>
<td>Multiple referents in physical context</td>
</tr>
<tr>
<td>PHYSICAL PRESENCE (PP)</td>
<td>Physically present</td>
<td>Physically absent</td>
</tr>
<tr>
<td>PRIOR MENTION (PM)</td>
<td>Given: mentioned within preceding 5 utterances</td>
<td>New: not mentioned within preceding 5 utterances</td>
</tr>
<tr>
<td>LINGUISTIC DISAMBIGUATION (LD)</td>
<td>No other possible referents in preceding 5 utterances</td>
<td>Other possible referents in preceding 5 utterances</td>
</tr>
<tr>
<td>JOINT ATTENTION (JA)</td>
<td>Referent focus of attention for child and interlocutor</td>
<td>Referent not focus of attention for child and interlocutor</td>
</tr>
</tbody>
</table>

---

In reality, these features are not binary but rather gradient in nature. For example, numerous hierarchies posit degrees of animacy in which people are considered more animate than animals, which are more animate than plants, which in turn are more animate than rocks (Silverstein, 1976). A similar gradient can be posited for the other five features as well. However, as it has been the practice in the literature to treat these features as binary in order to simplify the analysis and to allow for statistical analysis based on these values, the current study maintains the use of binary values.

---

Please cite this article in press as: Hughes, M.E., Allen, S.E.M., The incremental effect of discourse-pragmatic sensitivity on referential choice in the acquisition of a first language. Lingua (2014), http://dx.doi.org/10.1016/j.lingua.2014.03.001
The following example illustrates how each subject was coded. In the interaction in (4), the mother and child are playing with multiple blocks on the table. At the time of the final utterance, the mother is looking at and touching the block in question (Block B), and the child is also looking at it.

(4) MOT: *That block fell down.* [referring to Block A]
MOT: *Where does this block go?* [referring to Block B]
CHI: Ø go there. [referring to Block B] (Annie 2;0)

The (omitted) subject in the final utterance is coded as follows:

**ANIMACY.** A block is inanimate. Because animate entities are particularly salient to children in the discourse context, the subject that has an inanimate referent is coded as inaccessible.

**CONTEXTUAL DISAMBIGUATION.** At the time of the utterance, the child is playing with several blocks, so there are numerous other possible referents in the physical context that could fit the semantics of the verb and the grammatical elements of the intended referent. Therefore, the subject is coded as inaccessible.

**PHYSICAL PRESENCE.** The block is physically present, and therefore more easily identified than if it were absent from the physical context. The subject is coded as accessible.

**PRIOR MENTION.** Any referent that has been mentioned in the preceding 5 utterances is considered accessible. Because the block referred to was just mentioned in the previous utterance, it is coded as accessible.

**LINGUISTIC DISAMBIGUATION.** There are two referents mentioned in the preceding five utterances that fit the semantics of the verb and the grammatical elements of the referent, making it more difficult to identify. Therefore, the subject is coded as inaccessible.

**JOINT ATTENTION.** The block is the focus of attention for both the mother and the child, and they are aware of each other’s attention on the block. Thus the subject is coded as accessible.

In order to test inter-rater reliability, 9% of the data was blind-coded by two research assistants. These two coders achieved an average agreement of 85%. Moreover, every file was coded by one research assistant, and then subsequently checked by a second research assistant. The first author then reviewed all files in order to resolve any inter-rater differences.

### 2.4. Argument forms

In order to achieve a better understanding of the effects of discourse-pragmatics on subject choice, all third person subjects in the analysis were categorized as one of four possible referential forms, based on Ariel’s accessibility theory (Ariel, 2001). The central claim of accessibility theory is that adult speakers will choose a referential form based on the accessibility of the referent. The four referential forms analyzed in this study are null forms, pronouns, demonstratives, and lexical noun phrases. **Table 3** demonstrates the hierarchy of these four forms. Lexical forms are low accessibility markers because they are situations when a referent is inaccessible, meaning that the speaker has to convey more information in order to identify the referent. Conversely, null forms are high accessibility markers because they mark that a referent is highly accessible in the discourse context, meaning that the speaker does not have to convey as much information to

---

5 In their corpus of Inuktitut child speech, Skarabela and Allen (2003) determined that few arguments have preceding references more than 5 utterances prior, showing that the difference between a cut-off of 5 versus 20 preceding arguments is minimal. They suggest that a threshold of 5 previous utterances to encode a referent as new rather than given may be an appropriate distance for child spontaneous speech (Allen et al., 2008).
identify the referent. In other words, the referential form signals the degree of accessibility of a referent. Demonstratives and pronouns fall in between lexical noun phrases and null forms in terms of accessibility; however, as in many previous studies, demonstratives are ranked slightly higher than pronouns in terms of informational status, as seen in Table 3 (see Ariel, 1990, 2001; Givón, 1983; Gundel et al., 1993 among others).

2.5. Measuring the incremental effect of accessibility

The general claim underpinning this study is that the effect of the accessibility features increases incrementally for both adults and children, so that a speaker is more likely to produce reduced forms when the majority of features are accessible and full lexical noun phrases when the majority of features are not accessible. To determine whether this is the case, each possible subject context in the data was assigned a binary value – “accessible” (0) or “not accessible” (1) – for each of the six accessibility features. The values of the six features were then summed for each subject, giving it a value between 0 (completely accessible in the discourse context) and 6 (completely inaccessible in the discourse context). A two-way contingency table analysis was then conducted for each participant with the summed value of accessibility as the independent variable and form (i.e., null, pronominal, demonstrative, lexical) as the dependent variable. Chi-square tests of independence were performed to determine whether there was a statistically significant relationship between the predictor and the form of the subject.

We maintain the following two hypotheses for our data. First, the cumulative effect of the discourse-pragmatic features will mean that both adults and children will be more likely to realize highly accessible referents with reduced linguistic forms and highly inaccessible referents with more informative forms. Second, children’s choice of linguistic forms will differ from adults most at Time 1, but will become more adult-like at Time 2, so that their choice of linguistic forms will more closely match those of their caregivers.

3. Results

What follows are the results of the analysis for both the children and the adults in this study. In Section 3.1, utterances from the four mothers as a group are analyzed in order to establish how adult caregivers handle the mapping from accessibility of a referent to choice of linguistic form in the input and also to determine whether the cumulative effects of accessibility can be seen in the adults’ referential choice. In Section 3.2, individual results are given for each of the four children at the two time periods to determine each child’s sensitivity to the incremental effect of accessibility on referential choice, and to investigate how this ability develops over time.

Finally, this approach takes as its starting point the assumption that all features have equivalent “weight” in determining argument realization. This is a necessary assumption in order to perform this kind of analysis; however, in reality, this is not exactly the case. In earlier work, we have investigated the relative weight and the interaction of specific features and found that certain features are stronger than others and affect the choice of referent differently (e.g., Hughes and Allen, 2006; Allen, 2007; Hughes, 2011; Hughes and Allen, 2013). Therefore, in Section 3.3, the relative strength and weight of features in determining referential form are explored for the children as a group at each time point using multinomial logistic regression analyses. These findings provide a more refined model of how the six features inform the interlocutor’s choice of referent.

3.1. Group analysis for the caregiver data

The prediction for the caregiver data is that there should be a strong connection between the accessibility of a referent and referential choice: accessible referents should be realized by reduced linguistic forms (i.e., null forms and pronouns) and inaccessible referents should be realized by informative linguistic forms (i.e., demonstratives and lexical noun phrases). When analyzed individually, the results for the caregivers followed very similar patterns, and no significant change or development occurred in their production of forms from Time 1 and Time 2. Therefore, the caregivers’ data was combined as shown in Fig. 1.

The graph in Fig. 1 depicts the cumulative effect of accessibility for six discourse-pragmatic features distributed across four linguistic forms. The accessibility values range from 0 to 6 along the x-axis, with 6 being most inaccessible. As expected for the caregiver data, there are very few null subjects overall; instead, the caregivers use pronouns for accessible referents. The caregivers show the predicted pattern in their data: when referents are completely accessible, they are realized with pronouns at a rate of 67%. The caregivers’ use of pronouns decreases as referents become less accessible, and their use of demonstratives and lexical noun phrases increases. When referents are completely inaccessible for all 6 features, the caregivers produce lexical noun phrases at a rate of 95%.

One thing that was not predicted is the relatively high proportion of lexical noun phrases produced by the caregivers for referents that were coded as “inaccessible” for only 0 or 1 discourse-pragmatic features. This phenomenon likely results
from the repetition of lexical noun phrases typical of adult–child discourse that has been documented in several studies such as Clark and Bernicot (2008). It will be discussed in more detail in Section 4.

3.2. Individual trends in the child data

Before investigating the hypotheses stated in Section 2.5, the overall distribution of linguistic forms for the four children as a group was determined. Fig. 2 represents this distribution for the children at Time 1 (Total N = 723) and Time 2 (Total N = 1028). Chi-square analyses were performed to compare the use of each form. At Time 2, there are significantly fewer null forms ($\chi^2 = 168.470$, $p < .0001$), significantly more pronouns ($\chi^2 = 173.103$, $p < .0001$), significantly more demonstratives ($\chi^2 = 7.690$, $p < .01$), and significantly fewer lexical NPs ($\chi^2 = 41.060$, $p < .0001$) than at Time 1. These findings are in line with previous studies.

It has been well-established that there is considerable variation between individual children in both the rate of language acquisition and the cognitive strategies that each child applies to the task of language learning (Goldfield and Snow, 1997). Therefore, results for the children are reported here individually, in order to exclude the possibility that one or two children are responsible for the tendencies seen in the data for the entire group. In Fig. 3a through d, the cumulative effect of accessibility across four linguistic forms is represented for each child at Time 1 and Time 2.

Although there are some differences between the children, they all follow a similar pattern that shows strong sensitivity to the incremental effect of accessibility. At both Time 1 and Time 2, the children generally employ a high proportion of reduced forms to realize referents that are fully accessible, and a high proportion of more informative forms to realize referents that are “not accessible” for 4, 5 or 6 features. Moreover, the proportion of reduced forms decreases, and the proportion of more informative forms increases, as the referents become increasingly inaccessible. As expected, the reduced forms used at Time 1 are largely null forms, although some pronouns are used by each child. At Time 2, however, null forms have largely been replaced by pronouns used for the same function. Chi-square tests were conducted comparing the distribution of forms at Time 1 vs. Time 2 for each child. They show that the distribution of forms differs
significantly between Time 1 and Time 2 when referents are "inaccessible" for 0–4 features (\(p < .001\) for all comparisons). However, when referents are "inaccessible" for 5 or 6 features, the differences were not significant: the children at both times primarily use lexical noun phrases to introduce referents that are fully inaccessible.

As in the adult data, the children produce a higher proportion of lexical noun phrases than expected for referents that are at the accessible end of the scale, especially at Time 1. This overuse of lexical noun phrases decreases at Time 2 for Annie, Eleanor, and Fraser, but Brian still relies strongly on lexical noun phrases to realize accessible referents at Time 2. This pattern will be discussed further in Section 4.

The graphs at Time 1 demonstrate that for all the children, the mapping from accessibility of the referent to the choice of argument form is not quite adult-like. For the most part, there is not the neat cline reflecting the accessibility-form relationship that is seen for the adults in Fig. 1, where reduced linguistic forms dominate the accessible end of the scale and there is a gradual increase in more informative forms as accessibility increases. However, even at Time 1, it is clear that the children are sensitive to the discourse context because when referents are inaccessible for 4, 5, and 6 features, all children rely heavily on lexical noun phrases. By Time 2, all of the children are producing pronouns, and a similar cline of accessibility that we saw in the adult data is seen in the child data, demonstrating that the children have honed their mapping skills, making obvious their increased sensitivity to the discourse context.

Individual differences are also apparent between the four children. Brian and Fraser use more null forms at Time 1, while Annie and Eleanor use fewer null forms and more demonstratives. This may indicate that the two girls are using...
more demonstratives instead of null forms and pronouns at Time 1 or that the constructions that they are producing require demonstratives; for instance, they may be using more copula constructions for identification than Brian and Fraser. The children's level of linguistic ability as measured by mean length of utterance also varies slightly, and this may have an effect. Although the children are age-matched and broadly language-matched, Annie has the highest MLUw at both Time 1 (MLUw: 2.45) and Time 2 (MLUw: 3.42) and the pattern of her data at both times looks the most adult-like. Fraser has the lowest MLUw at Time 1 (MLUw: 1.74) and the pattern of his data looks very erratic. Overall, even with these individual differences, the graphs in Fig. 3a through d look remarkably alike for the two age ranges.

3.3. The effect of individual features

In the analyses thus far, the six features have been summed and treated as if they have equal weight in determining referential form. However, it is more likely the case that each feature does not contribute equally in determining the choice of referent. In fact, certain features may be stronger than other features in predicting particular outcomes (i.e., null forms vs. pronouns vs. demonstratives vs. lexical NPs).

In order to determine the contribution of each of the six features accurately, multinomial logistic regression analyses were performed for the children as a group at each of the two time points. Multinomial logistic regression is an advanced statistical technique that predicts the probability of a particular outcome when there are more than two levels of dependent or outcome variable (e.g., null form, pronoun, demonstrative, or lexical NP) based on a set of independent or predictor variables (e.g., the six discourse-pragmatic features under investigation). It also can help to determine which of the variables are strongest in determining the predicted outcome. In the following analyses, all independent variables are dichotomous: a value of ‘0’ was assigned if a referent was accessible for a particular feature, and a value of ‘1’ was assigned if a referent was inaccessible for that feature. The reference category for all analyses was 'lexical NP'. In other words, the analyses predict whether a speaker is more likely to choose a null subject vs. a lexical NP, a pronoun vs. a lexical NP, or a demonstrative vs. a lexical NP based on the values of the predictor variables – the six discourse-pragmatic features. As stated earlier, the general prediction is that accessible referents are more likely to be represented by low information forms (i.e., null forms and pronouns) and inaccessible referents are more likely to be represented by high information forms (i.e., demonstratives and lexical NPs).

3.3.1. Multinomial regression analysis for the children at Time 1

A comparison of the model with all six predictors for the child data at Time 1 against a model with no predictors (intercept-only6) showed significant improvement (\(df = 18, n = 825, \chi^2 = 313.202, p < .0001\)). Moreover, likelihood ratio tests showed that all features except for LINGUISTIC DISAMBIGUATION (\(p = .063\)) contributed significantly to the model (\(p < .0001\)). The model successfully predicted 49.9% of total referential forms overall. For each individual form, it predicted 76.1% of lexical NPs, 44.2% of null forms, and 36.2% of demonstrative forms correctly. However, it only predicted 8.7% of pronoun forms correctly. Overall, the model was successful enough in modeling the children's choice of referential forms at Time 1 that it is worth examining the findings more closely so that we can get a better idea of what features are important in driving these predictions.

Table 4 shows the effects of the six predictor variables on the choice of referential forms for the children at Time 1. The predictor variables for each category are listed in the leftmost column. In each case, the results indicate the likelihood of that category occurring as opposed to the reference category, which is lexical NP. The \( \beta \) values in the first column are the estimated multinomial logistic regression coefficients for the model. This is based on the accessible values of the predictor variables – a positive \( \beta \) value indicates that a referent is more likely to be realized as a null subject, a pronoun, or a demonstrative than a lexical NP, and a negative \( \beta \) value indicates that a referent is more likely to be realized as a lexical NP.7 The significance level of the Wald statistic is measured in the column labeled Sig. Anything with a \( p \)-value of less than .05 is considered significant. Finally, \( \exp (\beta) \) is the odds ratio for the predictors. This number describes how likely a certain outcome is: a value greater than 1 indicates that the outcome will fall into the comparison category (e.g., null rather than lexical), while a value less than 1 indicates the outcome will fall into the reference category (e.g., lexical rather than null). \( \exp (\beta) \) also predicts how many times more likely that outcome will be.

---

6 The intercept is the multinomial logit estimate for each outcome/dependent variable relative to the reference category (lexical NP) when the values for the six predictor variables in the model are set to zero; in other words, the intercept-only model shows the outcome when the model is run without the six discourse-pragmatic features as predictors.

7 The column labeled Std.Error contains the standard errors of the individual regression coefficients for the two models. The Wald Chi-Square statistic is the ratio of the coefficient (\( \beta \)) to its standard error squared. The column labeled \( df \) lists the degrees of freedom for each of the variables in the model. In this case, the degree of freedom for each variable is 1.

---

Please cite this article in press as: Hughes, M.E., Allen, S.E.M., The incremental effect of discourse-pragmatic sensitivity on referential choice in the acquisition of a first language. Lingua (2014), http://dx.doi.org/10.1016/j.lingua.2014.03.001
For the first category, lexical NP vs. null subject, all of the predictor values were found to be significant ($p < .05$ or lower) except for LINGUISTIC DISAMBIGUATION ($p = .103$). The prediction for three of these features, PHYSICAL PRESENCE, JOINT ATTENTION, and PRIOR MENTION, is that if a referent is accessible for any one of these (i.e., physically present, jointly attended to, and/or previously mentioned), it is more likely to be realized as a null subject rather than a lexical NP. The strongest predictor is PHYSICAL PRESENCE: according to the odds ratio ($\text{Exp} (b)$) a referent is 4 times as likely to be a null subject if it is accessible for this feature. A referent that is accessible for PRIOR MENTION or JOINT ATTENTION is more than twice as likely to be realized as a null subject. An accessible value for ANIMACY or CONTEXTUAL DISAMBIGUATION does not have the same predictive value: these referents are significantly more likely to be realized as lexical NPs showing that an accessible value for these two features does not cause the child to produce a null form. Perhaps, instead, the child relies more heavily on English grammatical constraints and supplies a subject. Furthermore, at Time 1, the child is also more likely to adhere to the convention of child and caregiver speech described above in which lexical NPs are produced to refer to animate third person entities instead of null forms or pronouns (see Sections 3.1, 3.2 and 4).

Four of the predictors were found to significantly affect the choice of pronoun vs. lexical NP. An accessible value for PHYSICAL PRESENCE or PRIOR MENTION predicts that a referent is almost three times as likely to be a pronoun rather than a lexical NP. An accessible value for ANIMACY OR CONTEXTUAL DISAMBIGUATION predicts that a referent is more likely to be realized as a lexical NP and not a pronoun, again demonstrating that accessible values for these features do not have a strong effect on the child’s production of low information forms.

Four of the predictors were also found to significantly affect the choice of demonstrative versus lexical NP, but not the same four. The strongest feature is PHYSICAL PRESENCE: an accessible value for this feature predicts that a referent is more than 14 times as likely to be realized as a demonstrative and not a lexical NP. A referent that is accessible for JOINT ATTENTION is over three times as likely to be realized as a demonstrative. These findings are not surprising because demonstratives are normally used for objects that are physically present; moreover, they are often a way to draw an interlocutor’s attention to a specific object. Accessible values for ANIMACY OR CONTEXTUAL DISAMBIGUATION predict that a referent is more likely to be realized as a lexical NP and not a demonstrative. As well as indicating that these features are not as strong, as we have seen earlier, it is also the case that demonstratives are not typically used for animate entities. They are, however, typically used to disambiguate an object in the physical context, so it makes sense that animate entities and unambiguous referents are less likely to be realized by demonstratives and more likely to be realized by another form, lexical NPs in this case.

Table 4
The effect of six predictors (ANIMACY, CONTEXTUAL DISAMBIGUATION, PHYSICAL PRESENCE, LINGUISTIC DISAMBIGUATION, JOINT ATTENTION, and PRIOR MENTION) on the choice of lexical NP vs. three other possible categories for children at Time 1.

<table>
<thead>
<tr>
<th>Predictor variables</th>
<th>$\beta$</th>
<th>Std.Error</th>
<th>Wald</th>
<th>$df$</th>
<th>Sig.</th>
<th>Exp ($\beta$)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Null subjects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AN***</td>
<td>-.626</td>
<td>.204</td>
<td>9.374</td>
<td>1</td>
<td>.002</td>
<td>.535</td>
</tr>
<tr>
<td>CD***</td>
<td>-1.246</td>
<td>.287</td>
<td>18.832</td>
<td>1</td>
<td>.000</td>
<td>.288</td>
</tr>
<tr>
<td>PP***</td>
<td>1.394</td>
<td>.344</td>
<td>16.419</td>
<td>1</td>
<td>.000</td>
<td>4.032</td>
</tr>
<tr>
<td>LD</td>
<td>.344</td>
<td>.211</td>
<td>2.655</td>
<td>1</td>
<td>.103</td>
<td>1.411</td>
</tr>
<tr>
<td>JA**</td>
<td>.695</td>
<td>.296</td>
<td>5.515</td>
<td>1</td>
<td>.019</td>
<td>2.004</td>
</tr>
<tr>
<td>PM***</td>
<td>.842</td>
<td>.220</td>
<td>14.669</td>
<td>1</td>
<td>.000</td>
<td>2.321</td>
</tr>
<tr>
<td><strong>Pronouns</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AN***</td>
<td>-1.519</td>
<td>.234</td>
<td>24.641</td>
<td>1</td>
<td>.000</td>
<td>.314</td>
</tr>
<tr>
<td>CD***</td>
<td>-1.547</td>
<td>.369</td>
<td>17.600</td>
<td>1</td>
<td>.000</td>
<td>.213</td>
</tr>
<tr>
<td>PP**</td>
<td>1.047</td>
<td>.346</td>
<td>9.154</td>
<td>1</td>
<td>.002</td>
<td>2.850</td>
</tr>
<tr>
<td>LD</td>
<td>.377</td>
<td>.227</td>
<td>2.762</td>
<td>1</td>
<td>.097</td>
<td>1.458</td>
</tr>
<tr>
<td>JA</td>
<td>.355</td>
<td>.316</td>
<td>1.262</td>
<td>1</td>
<td>.261</td>
<td>1.427</td>
</tr>
<tr>
<td>PM***</td>
<td>1.074</td>
<td>.244</td>
<td>19.386</td>
<td>1</td>
<td>.000</td>
<td>2.927</td>
</tr>
<tr>
<td><strong>Demonstratives</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AN***</td>
<td>-1.011</td>
<td>.266</td>
<td>17.068</td>
<td>1</td>
<td>.000</td>
<td>.333</td>
</tr>
<tr>
<td>CD***</td>
<td>-2.517</td>
<td>.549</td>
<td>21.027</td>
<td>1</td>
<td>.000</td>
<td>.081</td>
</tr>
<tr>
<td>PP***</td>
<td>2.675</td>
<td>.593</td>
<td>20.352</td>
<td>1</td>
<td>.000</td>
<td>14.517</td>
</tr>
<tr>
<td>LD</td>
<td>-.168</td>
<td>.245</td>
<td>.471</td>
<td>1</td>
<td>.493</td>
<td>.845</td>
</tr>
<tr>
<td>JA</td>
<td>1.167</td>
<td>.343</td>
<td>11.560</td>
<td>1</td>
<td>.001</td>
<td>3.214</td>
</tr>
<tr>
<td>PM</td>
<td>-.297</td>
<td>.245</td>
<td>1.471</td>
<td>1</td>
<td>.225</td>
<td>.743</td>
</tr>
</tbody>
</table>

Reference category: lexical NP.

* Significant at $p < .05$.
** Significant at $p < .01$.
*** Significant at $p < .001$. 

For the first category, lexical NP vs. null subject, all of the predictor values were found to be significant ($p < .05$ or lower) except for LINGUISTIC DISAMBIGUATION ($p = .103$). The prediction for three of these features, PHYSICAL PRESENCE, JOINT ATTENTION, and PRIOR MENTION, is that if a referent is accessible for any one of these (i.e., physically present, jointly attended to, and/or previously mentioned), it is more likely to be realized as a null subject rather than a lexical NP. The strongest predictor is PHYSICAL PRESENCE: according to the odds ratio ($\text{Exp} (b)$) a referent is 4 times as likely to be a null subject if it is accessible for this feature. A referent that is accessible for PRIOR MENTION OR JOINT ATTENTION is more than twice as likely to be realized as a null subject. An accessible value for ANIMACY OR CONTEXTUAL DISAMBIGUATION does not have the same predictive value: these referents are significantly more likely to be realized as lexical NPs showing that an accessible value for these two features does not cause the child to produce a null form. Perhaps, instead, the child relies more heavily on English grammatical constraints and supplies a subject. Furthermore, at Time 1, the child is also more likely to adhere to the convention of child and caregiver speech described above in which lexical NPs are produced to refer to animate third person entities instead of null forms or pronouns (see Sections 3.1, 3.2 and 4).

Four of the predictors were found to significantly affect the choice of pronoun vs. lexical NP. An accessible value for PHYSICAL PRESENCE or PRIOR MENTION predicts that a referent is almost three times as likely to be a pronoun rather than a lexical NP. An accessible value for ANIMACY OR CONTEXTUAL DISAMBIGUATION predicts that a referent is more likely to be realized as a lexical NP and not a pronoun, again demonstrating that accessible values for these features do not have a strong effect on the child’s production of low information forms.

Four of the predictors were also found to significantly affect the choice of demonstrative versus lexical NP, but not the same four. The strongest feature is PHYSICAL PRESENCE: an accessible value for this feature predicts that a referent is more than 14 times as likely to be realized as a demonstrative and not a lexical NP. A referent that is accessible for JOINT ATTENTION is over three times as likely to be realized as a demonstrative. These findings are not surprising because demonstratives are normally used for objects that are physically present; moreover, they are often a way to draw an interlocutor’s attention to a specific object. Accessible values for ANIMACY OR CONTEXTUAL DISAMBIGUATION predict that a referent is more likely to be realized as a lexical NP and not a demonstrative. As well as indicating that these features are not as strong, as we have seen earlier, it is also the case that demonstratives are not typically used for animate entities. They are, however, typically used to disambiguate an object in the physical context, so it makes sense that animate entities and unambiguous referents are less likely to be realized by demonstratives and more likely to be realized by another form, lexical NPs in this case.
Multinomial regression analysis for the children at Time 2

For Time 2, a comparison of the model with all six predictors against a model with no predictors showed significant improvement (df = 18, n = 1171, \( \chi^2 = 493.393, p < .0001 \)). Moreover, likelihood ratio tests showed that all features contributed significantly to the model (\( p < .0001 \)). The model also successfully predicted 58.8% of the total referential forms overall: it predicted 73.0% of pronouns, 62.4% of demonstratives, and 42.0% of lexical NPs correctly. It failed to predict any null forms, but at Time 2, this only represents 4% of the children’s referents. Table 5 describes the effects of the six predictor variables on the choice of referential forms for the children at Time 2.

None of the predictor values were found to be significant in predicting null subjects except for PRIOR MENTION (\( p < .0001 \)). The children at Time 2 produce significantly fewer null subjects than at Time 1, and it seems that the only feature that influences this choice is PRIOR MENTION. In fact, if a referent has been previously mentioned in the last five utterances, it is over five times as likely to be a null subject rather than a lexical NP (Exp (\( b \)) = 5.329).

Five out of six predictors were found to significantly affect the choice of pronoun vs. lexical NP. An accessible value for PHYSICAL PRESENCE, LINGUISTIC DISAMBIGUATION, JOINT ATTENTION, or PRIOR MENTION predicts that a referent is more likely to be a pronoun and not a lexical NP. JOINT ATTENTION and PRIOR MENTION are the strongest predictors: referents that are accessible for either of these features are three times as likely to be realized as a pronoun. However, an accessible value for CONTEXTUAL DISAMBIGUATION predicts instead that a referent is more likely to be realized as a lexical NP and not a pronoun; this again indicates an accessible value for this feature does not play a strong role in determining the production of low information forms.

Four predictors had a significant effect on the choice of demonstrative vs. lexical NP. The strongest feature is PHYSICAL PRESENCE: an accessible value for this feature predicts that a referent is more than 11 times as likely to be realized as a demonstrative and not a lexical NP. A referent that is accessible for JOINT ATTENTION is almost three times as likely to be a demonstrative. Accessible values for ANIMACY or CONTEXTUAL DISAMBIGUATION predict that a referent is more likely to be realized as a lexical NP and not a demonstrative, most likely for the reasons given in Section 3.3.1.

To summarize, at Time 1, an accessible value for the features PHYSICAL PRESENCE, PRIOR MENTION, or JOINT ATTENTION is significant in predicting null subjects rather than lexical NPs, while an accessible value for the features PHYSICAL PRESENCE or PRIOR MENTION is significant in predicting pronouns rather lexical NPs. Lastly, an accessible value for the features PHYSICAL PRESENCE or JOINT ATTENTION is significant in predicting demonstratives rather than lexical NPs. Therefore, the accessibility value for these three features seems to have the greatest effect on the children’s production of lower information forms and demonstratives at Time 1.

Table 5
The effect of six predictors (ANIMACY, CONTEXTUAL DISAMBIGUATION, PHYSICAL PRESENCE, LINGUISTIC DISAMBIGUATION, JOINT ATTENTION, and PRIOR MENTION) on the choice of lexical NP vs. three other possible categories for children at Time 2.

<table>
<thead>
<tr>
<th>Predictor variables</th>
<th>( \beta )</th>
<th>Std.Error</th>
<th>Wald</th>
<th>df</th>
<th>Sig.</th>
<th>Exp (( \beta ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null subjects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AN</td>
<td>-.008</td>
<td>.344</td>
<td>.001</td>
<td>1</td>
<td>.982</td>
<td>.992</td>
</tr>
<tr>
<td>CD</td>
<td>-.479</td>
<td>.385</td>
<td>1.547</td>
<td>1</td>
<td>.214</td>
<td>.619</td>
</tr>
<tr>
<td>PP</td>
<td>.817</td>
<td>.631</td>
<td>1.678</td>
<td>1</td>
<td>.195</td>
<td>2.264</td>
</tr>
<tr>
<td>LD</td>
<td>.544</td>
<td>.372</td>
<td>2.141</td>
<td>1</td>
<td>.143</td>
<td>1.723</td>
</tr>
<tr>
<td>JA</td>
<td>.662</td>
<td>.540</td>
<td>1.500</td>
<td>1</td>
<td>.221</td>
<td>1.939</td>
</tr>
<tr>
<td>PM***</td>
<td>1.673</td>
<td>.409</td>
<td>16.711</td>
<td>1</td>
<td>.000</td>
<td>5.329</td>
</tr>
<tr>
<td>Pronouns</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AN</td>
<td>.297</td>
<td>.174</td>
<td>2.926</td>
<td>1</td>
<td>.087</td>
<td>1.346</td>
</tr>
<tr>
<td>CD**</td>
<td>-.1245</td>
<td>.213</td>
<td>34.075</td>
<td>1</td>
<td>.000</td>
<td>.288</td>
</tr>
<tr>
<td>PP**</td>
<td>.671</td>
<td>.289</td>
<td>5.382</td>
<td>1</td>
<td>.020</td>
<td>1.956</td>
</tr>
<tr>
<td>LD**</td>
<td>.817</td>
<td>.179</td>
<td>20.751</td>
<td>1</td>
<td>.000</td>
<td>2.263</td>
</tr>
<tr>
<td>JA**</td>
<td>1.208</td>
<td>.262</td>
<td>21.214</td>
<td>1</td>
<td>.000</td>
<td>3.346</td>
</tr>
<tr>
<td>PM***</td>
<td>1.251</td>
<td>.167</td>
<td>56.148</td>
<td>1</td>
<td>.000</td>
<td>3.493</td>
</tr>
<tr>
<td>Demonstratives</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AN**</td>
<td>-.937</td>
<td>.228</td>
<td>16.855</td>
<td>1</td>
<td>.000</td>
<td>.392</td>
</tr>
<tr>
<td>CD***</td>
<td>-.2379</td>
<td>.375</td>
<td>40.225</td>
<td>1</td>
<td>.000</td>
<td>.093</td>
</tr>
<tr>
<td>PP***</td>
<td>2.438</td>
<td>.439</td>
<td>30.834</td>
<td>1</td>
<td>.000</td>
<td>11.454</td>
</tr>
<tr>
<td>LD</td>
<td>-.186</td>
<td>.203</td>
<td>8.36</td>
<td>1</td>
<td>.361</td>
<td>.831</td>
</tr>
<tr>
<td>JA***</td>
<td>1.045</td>
<td>.275</td>
<td>14.436</td>
<td>1</td>
<td>.000</td>
<td>2.843</td>
</tr>
<tr>
<td>PM</td>
<td>-.068</td>
<td>.200</td>
<td>.116</td>
<td>1</td>
<td>.733</td>
<td>.934</td>
</tr>
</tbody>
</table>

Reference category: lexical NP.

* Significant at \( p < .05 \).
** Significant at \( p < .01 \).
*** Significant at \( p < .001 \).
By Time 2, the children are more adult-like in their choice of referents, and only one feature is significant in predicting the occurrence of a null subject rather than a lexical NP — PRIOR MENTION. However, an accessible value for four features — PHYSICAL PRESENCE, LINGUISTIC DISAMBIGUATION, JOINT ATTENTION, and PRIOR MENTION — predicts that a referent is more likely to be a pronoun rather than a lexical NP. Finally, an accessible value for the features PHYSICAL PRESENCE or JOINT ATTENTION is significant in predicting demonstratives rather than lexical NPs, with PHYSICAL PRESENCE being the strongest predictor. At Time 2, children’s production of lower information forms is predicted by an accessible value for the same three features shown to be significant at Time 1, with the addition of LINGUISTIC DISAMBIGUATION. However, these low information forms are primarily null forms at Time 1 and pronouns at Time 2.

4. Discussion

In the present study, we examined English-speaking children’s and caregivers’ sensitivity to the incremental effect of six discourse-pragmatic features on referential choice at two developmental time periods. Each of the six features had already been shown to individually influence referential choice in these same speakers in a previous study (Hughes and Allen, 2013). Here we showed that both the children and their caregivers are sensitive in their referential choice not only to the individual features but also to the cumulative effect of these features.

Speakers produced a high proportion of reduced forms (i.e., null forms and pronouns) when referents were maximally accessible, and a high proportion of informative forms (i.e., lexical noun phrases) when referents were maximally inaccessible. More importantly, all speakers showed sensitivity to the number of features for which a referent was inaccessible. The proportion of reduced forms decreased incrementally, and the proportion of informative forms increased incrementally, as the inaccessibility score of the referent increased. Although there was some variation in the magnitude of the effect, every participant showed a pattern of this sort. Furthermore, the effect was evident at both developmental stages that we examined — Time 1 (2;0–2;7) and Time 2 (3;0–3;1). The children at Time 1 showed this effect most clearly as referents became increasingly inaccessible. The children at Time 2, like the caregivers, showed a gradient pattern across the scale of accessibility.

The results of this study extend Allen’s (2007) previous findings about incremental sensitivity in Inuktitut to English, a language of a very different typology. We further show that there is some development in this sensitivity, with three-year-olds performing more like adults than two-year-olds. Finally, we show that caregivers are sensitive to the incremental effect of accessibility as well.

All groups, but especially the children at Time 1, also displayed an interesting anomaly in the gradient pattern: an unexpectedly high proportion of lexical NPs for maximally accessible referents and referents with one inaccessible feature. This anomaly is largely due to repetition of a part of the previous utterance, a phenomenon that is quite common in child–adult interaction. Some examples from our data are in (5).

(5) a. CHI: Butterfly has gone.
   MOT: Where has the butterfly gone?[about a puzzle piece] (Fraser, 2;01)

b. MOT: Is Diesel happy?
   CHI: Diesel happy. (Fraser, 2;00)

Because these examples represent true third person subjects, they were not removed from the analysis.

The high number of noun phrases realizing accessible referents that we found in both child and caregiver data is not surprising, especially for younger children. In a study of 978 French child–adult exchanges involving repetitions, Clark and Bernicot (2008) show that speakers use repetitions for a number of strategic purposes. Adults repeat to check on and confirm children’s communicative intentions, often in cases where children have not provided complete enough information in their initial utterances to make their meaning clear, as well as to correct errors in what children have said. Both adults and children also use repetitions to signal that they are attending to the other’s utterances, and to place the repeated information in common ground as a basis for further interaction by acknowledging new information from the preceding speaker’s utterance and showing that they are now treating this as given information (Clark and Bernicot, 2008). Because children’s speech is often less clear than that of adults, repetition of the full noun phrase rather than just a pronoun can be used to clarify what exactly has been said. Repetitions of full noun phrases are also less cognitively demanding for children to both comprehend and produce than are pronouns and demonstratives; because less information about the referent is provided in the latter, the interlocutor must infer more information about the referent from

---

8 Repetition is common in adult-to-adult speech as well, and for the same purposes (e.g., Gundel et al., 1993; Walker, 1996). However, noun phrases are usually reformulated as pronouns (or null forms) for the second or later mention of a referent.
the discourse and physical environment. In addition, repeating the full noun phrase offers children a way to participate in conversation and assume a role in turn-taking before they have fully mastered the more complex system of pronouns and demonstratives typically used to maintain continuity of reference.

An additional finding related to development involves the use of null vs. pronoun forms for highly accessible referents. Although pronouns are typically used in English to realize accessible referents, the younger children very often produced null forms instead (e.g., “need hat” instead of “he needs a hat”). They appropriately select a low information form in this context, but pick an inappropriate or non-adult-like form. This phenomenon of high use of null subjects by English-speaking children has been well-studied for several decades (see review in Hyams, 2011). Again, the complexity of the pronoun system could explain this deviation from the target language. Alternatively, use of null forms could also reflect processing difficulty (e.g., Bloom, 1990; Gerken, 1991) or differences in the grammar between children and adults (e.g., Hyams, 2011; Wexler, 1998).

The present study also shows that pronoun use replaces omitted subjects once the children reach a later developmental stage. This finding substantiates the assumption in the literature that null forms are used early in acquisition in the place of pronouns (Hyams and Wexler, 1993). The children at Time 1 only produce pronouns for completely accessible referents 9% of the time. Instead they use null forms (25% of cases) and lexical noun phrases (64% of cases). The children at Time 2 rarely use null forms to realize accessible referents (only 7% of cases); instead, they now use pronouns (60% of cases). This mirrors the reduced forms used by the caregivers to realize accessible referents: 49% pronouns and only 1% null forms. While it is true that the production of null forms is largely replaced by the production of pronouns, some null forms might also be taking the place of other forms not fully acquired, such as demonstratives. These two phenomena in the speech of the younger children – overuse of lexical NPs and overuse of null forms – indicating that the mapping from accessibility of referent to linguistic form is not completely straightforward for these children.

In a second analysis, we established that the six features did not actually have equal weight in determining referential form, and by using logistic regression analysis, certain features were demonstrated to be stronger than other features in predicting particular referential forms.

Results showed that the features PRIOR MENTION, PHYSICAL PRESENCE, and JOINT ATTENTION are strong predictors of a speaker’s use of low information forms and demonstratives: at Time 1, an accessible value for PHYSICAL PRESENCE is the strongest predictor of null subjects vs. lexical NPs; accessible values for PRIOR MENTION and PHYSICAL PRESENCE are the strongest predictors of pronominal subjects vs. lexical NPs; and accessible values for PHYSICAL PRESENCE and JOINT ATTENTION are the strongest predictors of demonstratives vs. lexical NPs. At Time 2, an accessible value for PRIOR MENTION is the strongest predictor of null subjects; accessible values for PRIOR MENTION and JOINT ATTENTION are the strongest predictors of pronominal subjects; and accessible values for PHYSICAL PRESENCE and JOINT ATTENTION are the strongest predictors of demonstratives. These findings are in line with our linguistic knowledge of these forms. Demonstratives are normally used for objects that are physically present in order to draw the interlocutor’s attention to a specific object without regard to its status as discourse-old or discourse-new, while the use of null subjects or pronouns is more likely to reflect whether a referent has been previously mentioned in the discourse.

As for the other three features, there are varying degrees of significance. LINGUISTIC DISAMBIGUATION does not seem to have any real predictive power in the model and is rarely significant. However, when it was removed from the analyses, there was no real improvement or change in the models. It may be the case that, although this feature is quite weak when compared to the other features, it adds predictive strength in combination with other features.

An accessible value for ANIMACY is predictive at Time 1 for null subjects vs. lexical NPs and for pronouns vs. lexical NPs, but it predicts membership in the reference category (i.e., lexical NPs) and not the comparison category. This is most likely due to the fact that only third person pronouns were used in the analysis and the children at Time 1 are using a convention of child and caregiver speech in which they employ lexical NPs instead of null forms or pronouns to refer to animate third person entities. In the category of demonstratives vs. lexical NPs, ANIMACY also predicts membership in the reference category at Time 1 and Time 2 and for the adults. Again, this is as expected because demonstratives are not typically used for animate referents.

Finally, an accessible value for CONTEXTUAL DISAMBIGUATION does seem to have predictive power for the children at both times; however, like ANIMACY, it predicts membership in the reference category – lexical NPs. This prediction makes sense in the case of demonstratives as they are often used to disambiguate an object in the physical context; therefore, unambiguous referents are less likely to be realized by demonstratives and more likely to be realized by lexical NPs. However, for the case of the other categories, it merely suggests that this feature does not have strong predictive power for the production of low information forms.

It is important to note that, while these features do not have equal weight, the model with six predictors showed significant improvement over the model with no predictors at both Time 1 and Time 2 and, in most cases, all features contributed significantly to the model.

The analyses also show an important developmental change from Time 1 to Time 2. The model at Time 1 was able to correctly predict 76.1% of lexical NPs, 44.2% of null forms, and 36.2% of demonstrative forms, but only predicted 8.7% of...
pronoun usage and are overcompensating by using many lexical NPs. In contrast, the model at Time 2 correctly predicts 73% of pronouns, 62.4% of demonstratives, and 42% of lexical NPs, but does not predict any null forms because the children at Time 2 are producing very few null forms and instead are producing pronouns in more consistent and adult-like manner. The jump between Time 1 and Time 2 reflects the fact that children at Time 1 are less consistent in their choice of referential forms than the children at Time 2. As the input becomes more consistent, the model has more predictive power.

While it is true that all six features are not of equal weight and strength, the overall pattern of incremental sensitivity that we found in this study shows, even more strongly than previous studies based on individual features, that children have some sense of the knowledge of their interlocutors as evident from discourse and contextual features. Although it might be controversial to label this skill in children as young as 2;0 or 3;0 as a type of ‘mind-reading’ of their interlocutors’ intent, there is precedent in recent studies (e.g., Gundel et al., 2007; Gundel and Johnson, 2013; Matthews et al., 2006; O’Neil, 1996, 2005; Wittek and Tomasello, 2005). As Gundel et al. (2007;Gundel and Johnson, 2013) point out, children are able to use the full range of referential expressions fairly accurately by the age of 3;0. This ability reflects the fact that children at this early stage are able to gauge the attentional states of their interlocutors. Although the original Theory of Mind (ToM) approach is more concerned with the ability of children to understand the belief systems of others, the rudimentary elements of ToM are apparent in the ability of very young children to appropriately use the referential system of the language that they are acquiring. While children may not be able to express the beliefs of their interlocutors at earlier ages, they are still able to gauge others’ attention by monitoring expressions, eye gaze, and gesture (e.g., Clements and Perner, 1994; Repacholi and Gopnik, 1997; Robinson and Whitcombe, 2003). In a series of experiments, O’Neil (1996, 2005) asserts that very young children’s choice of referent appeals to certain aspects of ToM. In one experiment, she demonstrates that typically developing children as young as two were able to grasp the given-new distinction based on the co-presence or absence of their interlocutors when requesting an object that had first appeared in an earlier event. Children used more informative referential forms and more explicit gestures to indicate the object when they were aware that their interlocutors were not co-present in some way, either by being out of the room or by shutting both of their eyes at the time of the first presentation of the object (O’Neil, 2005). The children in this study, some as young as 22 months, were able to judge the attentional states of others, which in turn gave them an understanding of what constitutes new vs. given in the discourse context, “leading to a general sense of ‘what people talk about’” (p. 92).

The present study supports and adds to the growing consensus that very young children have rudimentary aspects of ToM in place. Gundel et al. (2007) maintain that three-year-old children’s adult-like use of referring expressions indicates that they have “the ability to appropriately assess what cognitive status the intended interpretation has for the addressee at a given point in the discourse” (p. 18). In other words, children can determine their interlocutor’s knowledge of the linguistic expression in context. This skill appears at a younger age than previous accounts suggest (see Baron-Cohen, 1995 and Tomasello and Haberl, 2003). However, children at 2;0 and even 3;0 might still might still not be able to take the interlocutor’s point of view, and therefore, they might not provide the appropriate information that is sufficient or necessary for their interlocutor to identify a referent in context (Grice, 1975). The ability to take another’s point of view seems to develop at a slightly later stage, after 4;0 or even later for certain types of complex deictic referential expressions and contexts. This can give rise to mapping problems, which are evident in the early use of referential forms. Importantly, the acquisition of discourse pragmatics, in particular the acquisition of referential forms, involves not only cognitive skills, but also linguistic knowledge. By studying the acquisition and development of referential forms, the interface between a child’s developing linguistic system and his or her emerging cognitive system becomes evident. In this study, we see development along both of these lines from Time 1 to Time 2, marking a distinctive change in the children’s production of referential forms by approximately age 3. In terms of linguistic ability, as the children’s MLUs increase so does their production of more adult-like forms, possibly via constructional learning mechanisms (e.g., Tomasello, 2003; Lieven et al., 2003) and/or the maturation of innate language mechanisms (e.g., Hyams and Wexler, 1993; Wexler, 1998). However, as pointed out above, the children’s correct use of referents is just as dependent on their developing cognitive abilities – only this interface between linguistic and cognitive knowledge allows the children to correctly gauge the attentional states of their interlocutors and then choose the appropriate linguistic form to communicate their intentions.

Overall, the findings here constitute and confirm strong evidence against previous claims that children have a deficient pragmatic system (e.g., De Cat, 2013; Deutsch and Pechmann, 1982; Hamann and Plunkett, 1998; Hyams and Wexler, 1993; Schaeffer, 2000; van Hout et al., 2010). Instead, children appear to possess a well-grounded though still developing pragmatic system – a system that three-year-olds continue to fine-tune on their road to adulthood and that involves not only linguistic knowledge, but also the coordination of logical, psychological, and social knowledge. This knowledge is not simply a binary differentiation between “accessible” and “not accessible” as we might find in a threshold model. Rather, looking at the incremental effects of accessibility as we have done here reveals a quite
subtle sensitivity that enables children to detect small differences in referent accessibility and integrate that with their language production. This sensitivity is already evident in children aged 2;0–2;1, and continues to increase with age.

Acknowledgements

We thank Elena Lieven and Jeannine Goh of the Max Planck Child Study Center in Manchester, UK, for providing access to the transcripts and videotapes of the data. This study was funded by National Science Foundation grant BCS-0346841 to Shanley Allen.

References


Please cite this article in press as: Hughes, M.E., Allen, S.E.M., The incremental effect of discourse-pragmatic sensitivity on referential choice in the acquisition of a first language. Lingua (2014), http://dx.doi.org/10.1016/j.lingua.2014.03.001


Mary E. Hughes completed her Ph.D. in 2011 in Applied Linguistics at Boston University. Her dissertation focused on the interaction of syntactic and discourse-pragmatic effects in the acquisition of reference in an attempt to shed light on the early development of referential choice in dyadic conversation. Currently, she is a Lecturer at Boston University. Her research interests include the integration of competence-based and usage-based linguistic approaches in order to show the range of factors at work in L1 acquisition. She has also done research on the acquisition of Irish Gaelic, and is interested in minority language revitalization.

Shanley E.M. Allen received her Ph.D. in Linguistics from McGill University in Montreal, Canada. She has published extensively on the acquisition of argument realization, the acquisition of morphosyntax and narrative by Inuktitut-speaking children, the coordination of speech and gesture in motion events across typologically different languages, and code mixing in bilingual preschool children. She is Co-Editor of the book series Trends in Language Acquisition Research published by Benjamins, and is on the Editorial Board of Bilingualism: Language and Cognition, International Journal of Language Sciences, and Journal of Child Language.