1.1. Discourse-pragmatics in argument realization: Stating the problem

Argument realization refers to the process of selecting linguistic forms (e.g. nouns, demonstratives, pronouns, argument omission) to realize verb arguments (i.e. subjects and objects of a verb). The choice of linguistic form to realize an argument is influenced by a number of different factors, including the discourse-pragmatic context. The general observation about how the context influences argument realization is that speakers select argument forms depending on how accessible the referent is to the interlocutor at any given point in the discourse (Ariel 2001). When a referent is easy to recover from the prior linguistic or discourse context and is accessible to the listener, the speaker is more likely to select a ‘light’ linguistic expression (e.g. pronoun or omitted argument), whereas when it is not easy to recover a referent from the context, the speaker is more likely to select a ‘heavy’ linguistic expression (e.g. lexical noun).

A number of different discourse-pragmatic factors have been shown to influence a referent’s accessibility (e.g. Chafe 1976, Givón 1983, Clancy 1980, 1993, Allen 2000). These, for example, include newness (i.e. whether or not a referent has been mentioned in the prior linguistic context), physical presence (i.e. whether or not a referent is physically present), joint attention (i.e. whether or not a referent is jointly attended to by the speaker and listener), contrast (i.e. whether or not a referent is in contrast with other referent(s)), animacy (i.e. whether or not a referent is human), disambiguation in physical context/discourse (i.e. whether or not a referent is in competition with other potential referents) (for further details, see e.g. Allen, Skarabela and Hughes 2008).


However, although there are significant differences in how accessibility of
individual features influences argument form, the effect is not absolute. While young speakers, for example, tend to omit given referents, they also omit a high proportion of new referents. They even realize new referents with omitted arguments more often than with lexical nouns (Allen 2000, Hamann and Plunkett 1998, Skarabela 2007). Children thus appear to follow discourse-pragmatic principles in some contexts, but ignore them in others. This has led some to claim that young children’s understanding of discourse-pragmatic principles is not very sophisticated (e.g. Grinstead 2004, Hamann and Plunkett 1998). Instead of interpreting this behaviour as a failure to show awareness of discourse-pragmatics and referent accessibility, here we would like to raise two points in favour of an alternative explanation.

The first point we want to make is conceptual and refers to the nature of argument realization as a linguistic phenomenon. Argument realization is a dynamic process that is determined by interactions between different features under changing discourse-pragmatic conditions. As a result, speakers’ selection of argument form in any given context is not an instance of a categorical rule. Instead, speakers just follow general tendencies in how they represent intended referents. This is an obvious, yet still commonly neglected point. It is thus not to be expected that once a referent has been mentioned, it will be, out of necessity, expressed with a pronoun or omitted argument on every subsequent occasion. It is possible that, for example, a previously introduced referent may be later contrasted with other referents, and will thus require the use of a full noun phrase. Or, a speaker may select a more explicit linguistic form because such referent may be in competition with other referents.

This brings us to the second – methodological – point, and how to approach argument realization in research. While researchers tend to focus on the workings of individual discourse-pragmatic features, given the nature of argument realization highlighted above, it is unlikely that speakers in actual discourse select argument forms based on one individual feature alone (Allen 2000, Allen, Skarabela and Hughes 2008). It is possible that a new referent may be realized with a light linguistic form, as, for example, when the referent is produced in the presence of joint attention and is thus particularly salient to both speaker and listener.

The following section explores whether and how prior research on early argument realization from a discourse-pragmatic perspective incorporates the two points – that the context in which arguments are realized changes and so do the requirements for how speakers may choose to realize referents.

1.2. Models of discourse-pragmatic features in argument realization

Several recent studies have begun to explore how features work together and how speakers encode referents with multiple features. Three models have emerged as a result of these efforts. While all three models acknowledge and consider the contribution of several features, they vary in their view of whether
individual features contribute to argument realization in the same way – to the same degree – or whether they are of different strength.

The first model, that we will call ‘THE INCREMENTAL MODEL’, builds on the assumption that children are sensitive to the incremental effect of features: they are less likely to omit an argument if it is inaccessible for 1 vs. 0 features, for 2 vs. 1 features, for 3 vs. 2 features, for 4 vs. 3 features, etc. (e.g. Allen 2007, Hughes and Allen 2008). The basic assumption of the second model, that we will call here ‘THE THRESHOLD MODEL’, is that children are sensitive to a threshold of features: children are more likely to omit an argument if it is accessible for a number of different features studied than if it is inaccessible for one or more of the features, and they are more likely to omit an argument if it is inaccessible for only one feature than if it is inaccessible for two or more features (e.g. Allen 2000; Narasimhan et al. 2005). The common feature of these two approaches – threshold and incremental – is their fundamental assumption that all features have equal import in determining accessibility of a referent.

Rather than assuming that each feature contributes to argument realization to the same degree, the third model, called here ‘THE INTERACTION MODEL’, proposes that features are, or can be, of different strength and contribute to argument realization to a different degree (e.g. Allen 2000, Serratrice 2002, Skarabela 2007a).

One way to begin to explore the assumptions and predictions of these three models in more detail is to focus on the interaction of two features that have been shown to yield significant effects in argument realization. In this paper, we will thus consider the contribution of newness and joint attention, both of which have been shown to play a role in child and adult argument realization (Matthews et al. 2006, Wittek and Tomasello 2005, Guerriero et al. 2006, Skarabela 2007a, b, cf Allen et al. 2008). As mentioned earlier, prior research shows that children distinguish between new versus given referents quite reliably, and they use different argument forms to do so: they are more likely to omit arguments to represent given referents, whereas they are more likely to use overt arguments to introduce new referents. Similarly, they use different forms depending on whether referents are produced in the presence or absence of joint attention: they are more likely to omit arguments in the presence of joint attention, whereas they are more likely to use overt arguments in the absence of joint attention. But, as we have already pointed out, these trends are not categorical and we find that children do omit new referents and they also omit referents that were produced in the absence of joint attention. It thus appears that the role of these features vary across different contexts. How do the two features interact? What contexts are more likely to license omission of new referents or of referents produced in the absence of joint attention? What are the predictions of the three models?

All three models predict that when two features are inaccessible (i.e. when new referents are produced in the absence of joint attention), speakers will avoid omitting arguments and they will be most likely to use overt arguments instead. The models yield different predictions when one of the features is accessible
(i.e. when given referents are produced in the absence of joint attention or when new referents are produced in the presence of joint attention) but also when both features are accessible (i.e. when given referents are produced in the presence of joint attention). While the incremental model clearly predicts the highest omission rate in the context with two accessible features and only some omission in the two conditions with only one accessible feature, the threshold model makes two different predictions, depending on the threshold level: if the threshold is set at the level of 0 vs. 1+ accessible features, then the omission rate in the context of two accessible features will be the same as in the two conditions with only one accessible feature (i.e. given referents produced in the presence of joint attention should be equally likely to be omitted as either new referents produced in the presence of joint attention or given referents in the absence of joint attention). Alternatively, if the threshold is set at the level of less than or equal to 1 vs. 2+ features, then the model predicts the highest omission rate in the condition with two accessible features and the use of overt arguments in the two conditions with only one accessible feature. Finally, the interaction model predicts interaction effects between the two features but the actual degree of omission across these three conditions depends on the individual strength of each feature. This means that the omission rate in the condition with two accessible features may be the same as in one of the other two conditions with only one accessible feature, or it may be higher. In comparison to the incremental and threshold model, the interaction model thus allows for a difference between the two conditions with one accessible feature, depending on the strength of each feature; it also allows that the condition with two accessible features may be the same as the other two conditions.

In what follows, we examine these predictions in argument realization in the spontaneous speech of four 2- to 3-year-old children acquiring Inuktitut. We first outline relevant points of the employed methodology.

2. Methodology

To test our predictions, we explored the realization of third person referents in a videotaped corpus of naturalistic spontaneous speech of four children aged 2;0-3;6 acquiring Inuktitut (Allen 1996).

2.1. Relevant structural properties

Inuktitut is a polysynthetic, morphologically ergative language of the Eskimo-Aleut family with SOV basic word order. It has rich inflectional morphology with nominal case-marking affixes and verbal cross-referencing affixes with frequent subject and object omission (Allen 2000). An inflectional suffix completes each verb stem, encoding information about the verb mood and the person and number of the referent (i.e. a transitive-verb suffix encodes information about the person and number of the subject and object, and an intransitive-verb suffix encodes information about the person and number of the
subject only). The following schema illustrates the verb formation process in the language (adapted from Swift and Allen 2002a):

verb base + postbase suffixes + portmanteau inflectional suffix

\textit{atjiliuq}– \textit{jau–guma–nngit–tunga}

\textit{film–PASS–want–NEG–PAR.1sS}\textsuperscript{2}

‘I don’t want to be filmed.’

In the above example, the verb base \textit{atjiliuq}, ‘film’, is followed by several optional postbase suffixes. The inflectional suffix –\textit{tunga} completes the word, representing a first person singular subject in the participial mood.

2.2. Data

The Inuktitut data used for this study were collected from four monolingual Inuktitut-speaking children—Elijah, Lizzie, Louisa, and Paul—ranging in age from 2;0 to 3;6 (Allen 1996). All the data were videotaped in naturalistic communication situations in the children’s homes while they were playing and talking with family and friends. Two hours of video footage per child at three 4-month intervals were selected and transcribed by native speakers in CHAT format. The total data set consists of 3,168 arguments.

2.3. Third person referents

Out of a total of 3,168 arguments, we selected third person referents only. This is for the following three reasons: (1) Only third person referents can be expressed with either overt or omitted arguments in Inuktitut. First and second person arguments are omitted except in contexts of extreme emphasis, which are very rare (Allen 1996, 2000). (2) Only third person referents can be either new or given. First and second person referents are always considered given, as they refer to the speech participants ‘I’ and ‘you’. (3) Only third person referents can be produced in the presence or absence of joint attention since we assume that first and second person referents are always jointly attended to.

Third person referents can occur as omitted or overt arguments, where OMMITTED refers to the null argument, while OVERT refers to an independent demonstrative, or an independent lexical NP (for examples and further details,

---

\textsuperscript{1} There are contexts where even the required suffixes can be dropped, however (for details, see Crago and Allen 2001, Swift and Allen 2002a, b).

\textsuperscript{2} The following abbreviations are used in the text: Verbal inflection: 1, first person; 3, third person; s, singular; S, subject; O, object. Verbal modality: IND, indicative. Word-internal morphology: PASS, passive; NEG, negative; PAR, participative.
see Skarabela 2007a, b).³ Note that there are no third person pronouns in
Inuktitut.

2.4. Coding

All subjects and objects of full verbal utterances were coded for
morphological form and for the following seven binary discourse-pragmatic
features: physical absence, contrast, differentiation in context, differentiation in
discourse, query, joint attention, and newness (see Allen 2000 and Allen et al.
2008 for a full justification of each feature). Each third person argument in the
data set received either an accessible or an inaccessible value for each
discourse-pragmatic feature. For example, if a third person argument was used
to refer to a referent that was physically present, it was coded as accessible for
physical absence; if the referent was not contrasted with other referents, it was
coded as accessible for contrast; if it was produced in the absence of joint
attention, it was coded as inaccessible for joint attention; if it introduced a new
referent, it was coded as inaccessible for newness etc. (for details, see Allen
2000, Skarabela 2007a, b).

Since we were interested in assessing the strength of two features only, and
thus needed to control for the contribution of the additional five features, out of
a total set of 1,890 third person arguments, we selected only those arguments
whose referents were accessible for the additional five features. That is, we
selected arguments whose referents: (1) were physically present in the context,
(2) were not contrasted with other referents, (3-4) were the only potential
referent in the physical and linguistic context, and (5) were not the subject of a
query. This yielded a total of 416 arguments for this analysis and each argument
varied in accessibility (i.e. either accessible or inaccessible value) for only
newness and joint attention.

2.5. Logistic regression

In order to assess the relative contribution of the two features to argument
realization in our dataset, we used a logistic regression analysis. Logistic
regression is a tool that allows us to evaluate whether both variables correctly
predict the form of an argument, but it also allows us to assess whether a certain
argument is likely to be omitted or overt. In other words, the results of logistic
regression give us the probability of a dependent variable (argument form)
occurring given known values of predictor variables (newness and joint
attention).

³ In this paper, we do not distinguish between argument forms with or without inflection
since we have found in earlier research that the presence of inflection does not influence
argument realization (Skarabela 2007a).
2.6. Newness

Newness is defined here as a binary discourse-pragmatic feature that refers to whether a referent is new or given in the discourse. A referent was considered to be new if it had not been mentioned in the previous 20 utterances, and as given if it had been mentioned in the previous 20 utterances (Givón 1983, Du Bois 1987, Allen 2000). In practice, virtually all given referents had been mentioned in the previous five utterances (Skarabela and Allen 2002b).

2.7. Joint attention

We operationally defined joint attention as a triadic social activity wherein the child and the interlocutor are both focused on the same referent while aware of each other’s attention (Tomasello 1999). In the presence of joint attention the child sees that the interlocutor has visually registered and is focused on a particular referent, providing confirmation of the accessibility of the referent in question to the interlocutor. To code the video data for visual joint attention, we followed methodology introduced in Skarabela and Allen (2002a). We monitored eye gaze, body direction, head direction and gesture (including pointing) in order to determine whether both interlocutors were paying attention to the referent and that each was aware of the other’s attention. A typical example with joint attention in progress is illustrated by the following interaction as captured by the video footage (from Skarabela and Allen 2002a):

Two boys lie next to each other on the ground, playing and laughing. One of them, Elijah, stands up and throws a cap at a woman who is holding the video camera and recording them. When she gets hit with the cap, Elijah sits back down, laughing, he looks at the woman, then at the other boy who is meanwhile pretending to shoot the woman. Then Elijah says:

Ø milur-para!
Ø hit- IND.1sS.3sO
‘(I) hit (her)!’

(Elijah 2;9)

The whole scene lasts about 12 seconds. We determined that the boys were involved in joint attention: Elijah’s eye gaze is first focused on the woman, then on the other boy, who is meanwhile focused on the woman, pointing in her direction. The boys then glance at each other.

3. Results

In order to assess the relative contribution of the two features, we used a logistic regression analysis with argument form as the dependent variable and newness and joint attention as the independent variables. The analysis included a total of 416 cases. The full model was significantly reliable ($\chi^2(2) = 113.97$, df
The model accounted for between 24% and 36.5% of the variance in argument form (Cox and Snell $R^2 = .240$; Nagelkerke = .365), with 97% of the omitted arguments and 27% of the overt forms successfully predicted. Overall, 81% of predictions with the two predictor variables included were accurate in comparison to 77% of predictions in a constant-only model. Both variables were significant (both $p < .0001$) and reliably predicted argument form. The Wald statistic for joint attention was 63.171, while that for newness was 20.959, showing that joint attention was the stronger factor in argument realization in this dataset. The odds of an omitted argument being selected by a speaker were 5 times higher when the referent was given than when it was new and 20 times higher when the referent was produced in the presence of joint attention than when it was produced in the absence of joint attention.

To summarize the results of logistic regression, both features – newness and joint attention – significantly contributed to the selection of omitted and overt arguments in the sample of 416 examples in child Inuktitut. Importantly, a further result of the logistic regression analysis shows that joint attention had a stronger effect on argument form than newness.

So how exactly did the two features work together and what argument form choices were made in the four possible contexts? In order to assess our specific predictions about the interaction between newness and joint attention, we report the basic descriptive data in Table 1.

**Table 1.** The proportion of omitted arguments representing given vs. new referents in the presence or absence of joint attention

<table>
<thead>
<tr>
<th></th>
<th>Given</th>
<th>New</th>
</tr>
</thead>
<tbody>
<tr>
<td>JA-present</td>
<td>97% (165/169)</td>
<td>87% (45/52)</td>
</tr>
<tr>
<td>JA-absent</td>
<td>64% (102/160)</td>
<td>29% (10/35)</td>
</tr>
</tbody>
</table>

The results in Table 1 indicate that, overall, children omitted a large proportion of arguments (77%). They omitted more arguments in the presence rather than in the absence of joint attention (65%); they also omitted more arguments representing given rather than new referents (83%). The distribution of omitted arguments across the four conditions was significantly different ($\chi^2 (1) = 8.058, p < .005$). Given referents were omitted significantly more in the presence rather than absence of joint attention (97% vs. 64% respectively, $p < .001$). Crucially, while the children omitted new referents, they did so significantly more when they were involved in joint attention (87% vs. 29% of arguments in the presence vs. absence of joint attention respectively; $p < .0001$). Similarly, while they omitted arguments when they were not involved in joint attention, they did so significantly more with given rather than new referents (64% vs. 29% of arguments respectively; $p < .0005$). That is, the children were considerably less likely to omit arguments when the referent was new and produced in the absence of joint attention. Furthermore, the omission rate for new referents produced in the presence of joint attention was significantly lower.
higher than for given referents produced in the absence of joint attention (87% vs. 64% respectively; \( p < .005 \)). This shows a differential effect of the two factors in argument realization and highlights the finding that joint attention is a stronger factor than newness in argument realization in this dataset.

4. Discussion and conclusion

This corpus study examines the role of two important factors – newness and joint attention – that had been claimed to affect children’s choice of argument form in spontaneous speech. Our aim was to determine, using the tool of logistic regression, the degree to which the two features independently predict overt versus omitted arguments in the corpus of spontaneous speech in four monolingual 2- to 3-year-old children acquiring Inuktitut (Allen 1996). We used the results in order to evaluate the assumptions and predictions of three recently proposed models of the role of discourse-pragmatics in argument realization: the incremental model (Allen 2007, Hughes and Allen 2008), the threshold model (Allen 2000, Narasimhan et al. 2005) and the interaction model (Allen 2000, Serratrice 2002, Skarabela 2007a).

The results of logistic regression revealed that both factors significantly contributed to argument realization in our corpus. The children were most likely to omit arguments representing given referents produced in the presence of joint attention (i.e. when both features were accessible); they were least likely to omit arguments representing new referents in the absence of joint attention (i.e. when both features were inaccessible). Furthermore, however, the results of logistic regression also showed that the two factors contributed to argument realization to a different degree. Specifically, joint attention was shown to be a stronger predictor than newness. This became obvious from a differential rate of argument omission in the two conditions with only one accessible feature: the proportion of arguments representing new referents that were omitted in the presence of joint attention was significantly larger than the proportion of arguments representing given referents that were omitted in the absence of joint attention. While the results of the first two conditions are at least partially compatible with all three models we aimed to evaluate in this paper, the interaction effects for the two conditions with only one accessible feature were not predicted by either the incremental or the threshold model. The overall results thus best support the interaction model.

What are the differences between the three models? As reviewed earlier, the incremental and threshold models both build on the same assumption that features are of equal strength and contribute to argument realization to the same degree. For the threshold model, this leads to the prediction that when the threshold is set between 0 vs. 1 accessible feature, there would be no difference between the three conditions with at least one accessible feature. This, as we have seen, was not the case in our study. Alternatively, if the threshold is set for between 1 vs. 2 accessible features, the model makes the prediction that there would be some omission in the condition with two accessible features, but
speakers would be more likely to use overt arguments – and to the same degree – in the two conditions with only one accessible feature. This again was not the case in this dataset. The incremental model, in comparison, predicts that there would be a difference between argument form choice in referents with one vs. two accessible features, a prediction that was borne out. This model, however, fails to predict that there may also be a difference between the two conditions with only one accessible feature; instead it predicts that, as long as one feature is accessible, regardless of which one, the omission rate would be the same.

To explain these results we appeal to how newness and joint attention affect accessibility and how accessibility gets encoded and signalled by speakers’ use of linguistic forms. While the general assumption is that argument omission in the context of new referents indicates a failure in signalling accessibility properly – and lack of children’s awareness of discourse-pragmatics – the results of this study point out that this is not necessarily the case. Instead, we argue, there are different ways to make a new referent accessible. The speaker has, on one hand, the choice of using an overt argument (e.g. lexical nouns). This, we have seen, is particularly the case when speakers are not involved in joint attention. But, as this study shows, alternatively, new referents can also be omitted when the speaker and interlocutor are involved in joint attention. That is, joint attention enhances accessibility of new referents. In contrast, it does not appear to affect accessibility of given referents, or at least not to the same degree, as given referents are already accessible via prior discourse. It remains to be seen whether or to what extent these trends are observable in later stages of development and in adult spontaneous speech.

In conclusion, the results of this study demonstrate that when newness and joint attention are considered together, children in fact realize arguments in accordance with subtle discourse-pragmatic principles and based on the referent’s accessibility to the listener. These findings thus advance our understanding of the role of discourse-pragmatics in early argument realization and further highlight children’s sensitivity to subtle interactions of multiple discourse and socio-cognitive cues.

References


