Joint Attention and the Development of the Use of Demonstrative Pronouns in Turkish

Aylin C. Küntay and Aslı Özyürek
Koç University

1. Introduction

Demonstrative pronouns such as this, that, here, and there provide ways to encode or grammaticalize features of the context of the utterance (Levinson, 1983; Lyons, 1977). Most traditional analyses of demonstratives propose that these forms provide primarily specifications about the position and the distance of the intended referent with respect to the location of the speech event participants (e.g., Anderson & Keenan, 1985; Diessel, 1999; Fillmore, 1971; Lakoff, 1974; Levinson, 1977). For example, Levinson (1977) pointed out that demonstrative pronouns depict the clearest paradigm for pragmatic variation along a proximal-distal dimension. However, more recent linguistic work on demonstratives has turned its attention to interactional factors that determine the speaker’s choice of specific forms (Hanks, 1990; 1992; Laury, 1996; Levinson, in press; Özyürek and Kita, 2000). Hanks (1992) states that cover labels such as “relative proximity” used for characterizing delictic reference “is not a matter of spatial contiguity. Instead, what is basic to deixis is the access (cognitive, perceptual, spatiotemporal) that participants have to objects of reference in the current speech event” (p. 60). Laury (1996), drawing on work on Finnish, shows that the meanings of demonstratives is not primarily based on concrete distance, but that instead, speakers use them to negotiate their mutual access to referents in conversation. Özyürek and Kita (2000) show that encoding joint attention of conversational participants is an essential semantic notion that is contrastively used in the demonstrative systems of Turkish and Japanese. In an overview of recent work on the pragmatics of deixis, Levinson (in press) proposes that the most important function of indexicality is directing the addressee’s attention to some feature of the spatio-temporal physical context or indicating the speaker’s presumption of the prior existence of that attention.

Similarly, most previous developmental investigations of how children acquire the meaning of demonstratives were conducted primarily with English with a focus on the acquisition of spatial semantic contrasts (Clark & Sengul, 1978; Tanz, 1976; Wales, 1986). Up to the present, there has not been any

* This paper reflects equal contribution from both authors. Sanem Günay, Şirin Duruk, Ibrahim Şenay, Engin Arık, and Pelin Kesebir helped in recording and transcribing the adult data and coding the entire corpus.

research on how children learn to use demonstratives in languages that encode social and interactive aspects of the context as a basic contrast such as in Finnish, Turkish or Japanese. However, how children use deictic forms in such languages might provide revealing data for efforts to determine whether factors based on distance or joint attention are at the core of the use of the demonstrative system. In fact, a recent development analysis of over here/over there carried out by Hallan (2001) suggests that the first uses by children do not necessarily encode strictly spatial distinctions, but serve to direct the attention of caregivers, closing the interpersonal "deictic gap" between first and second person.

Furthermore, the existing developmental studies have the following drawbacks: (1) they do not provide comparisons with the adult system, (2) they do not consistently include analysis of the relevant non-verbal actions of co-participants, except for inclusion of the pointing gesture in the Wales study, (3) they do not provide investigation of the social and interactional factors relevant for the use of demonstratives. In any developmental linguistic study, adult data are crucial for providing benchmarks for comparison with the younger age groups. An analysis of non-verbal actions and social-interactional factors is also crucial in any study of deictic terms. Hindmarsh and Heath (2000) point out that reference with deictic terms is usually underspecified without the contribution of bodily actions of co-participants (see also Hanks, 1990). Yet, non-verbal interactions are usually understudied in most analyses of deixis: "little research in linguistics has investigated how bodily gestures and actions are used in relation to talk in order to direct and encourage another to look at a particular (feature of an) object" (Hindmarsh & Heath, 2000, p. 1857).

Thus, the present study is an attempt to grapple with the problems of previous developmental accounts of the use of the demonstratives by using conversational data from Turkish speakers of different ages. The main questions addressed in this paper are: (1) how do children of different ages use demonstrative forms in a language such as Turkish that offers social and interactive factors as a semantic contrast in its demonstrative system in addition to spatial contrasts? and (2) how do Turkish children use demonstrative forms in comparison to adults?

1.1. The demonstrative pronoun system in Turkish

Turkish has a three-way distinction in its demonstrative pronoun system, offering three forms, bu, şu, and o. Previous analyses of demonstratives in Turkish have indicated that the contrasts among these three forms encode the relative distance of the referent from either of the conversationalists (Anderson and Keenan, 1985; Kornfilt, 1997; Levinson, 1983; Underhill, 1993). In these accounts, bu is used for objects that are in proximity to the speaker, şu is used for objects at mid-distance or close to the addressee, and o is used for objects that are located at a far distance from the speaker. In challenging these past accounts of the Turkish demonstrative system, Özyürek and Kita (2000) have
provided a recent analysis of Turkish demonstrative system based on investigation of videotaped conversational data. This study has revealed that presence or absence of joint attention is a significant factor in determining the speaker's choice of demonstrative forms. As summarized in Figure 1, şu is used in utterances where there is no previous joint attention between the interlocutors regardless of the distance of the objects. For example, a speaker can refer to a wound on his arm as well as to a far away referent with şu as long as the addressee's visual attention is not on the object. On the other hand, bu and o are used in cases where joint attention had already been achieved. With bu used for proximal objects and o for distal objects with respect to the speaker. Figure 1 presents an analysis of the Turkish demonstrated system as proposed by Özyürek and Kita (2000).

Figure 1: Analysis of the Turkish demonstrative pronoun system (Özyürek and Kita, 2000)

In this paper, our aim is to reveal the developmental pattern of the use of the demonstrative şu in relation to bu and o, that is, how children learn to use a demonstrative that encodes presence or absence of joint attention in comparison to other pronouns that encode spatial contrasts. With this aim in mind, we conducted the following study based on conversational data with children and adults.

2. Method

Conversational data were collected from 4-year-olds, 6-year-olds, and college-age adults, including 3 pairs of participants from each group. All pairs were composed of individuals who reported having a previous "friendship" relationship. Participants were given a picture of a model made of lego pieces of different shapes and colors, and asked to reconstruct this model collaboratively with actual lego pieces. Each pair was videotaped for 12 minutes. Later all the utterances were transcribed. Relevant nonverbal actions accompanying the utterances containing demonstratives were also noted.

The setup in this study is a joint activity that calls for close coordination between pairs. In our task, the participants need to establish what Kendon (1992) calls a "joint transactional segment" in having to position their actions in such a
way as to establish an overlapping or joint interactive space. That is, even though the pairs might assign distinct responsibilities to each member, the task of responsibility allocation and management necessitates close collaboration. To be able to coordinate their activities with one another, the individuals in each pair have to establish shared orientation to the objects that are part of the assigned activity, i.e., the lego pieces, the pictured model, the parts of the model, etc.

2.1. Coding categories

Verbal and non-verbal actions that might be important for manipulation of joint attention and in the choice of the demonstrative forms were the basis of the coding categories chosen for the study. They were applied to utterances containing any of the three demonstratives.

The following features used in the utterances constituted the coding categories:

1. **Type of demonstrative**: which one of the demonstratives is featured in the utterance (i.e., *bu*, *yu*, or *o*).
2. **Referred object's location**: whether the object is close to speaker, close to addressee, or away from both (objects that were in the space between the participants or far away from both were coded in this latter category).
3. **Speaker's pointing during utterance**: whether the speaker was not pointing to the object at all, pointing with object in hand, or pointing with index finger,
4. **Addressee's eye-gaze after the use of the demonstrative form**: whether the addressee's eye-gaze moves to the object referred by the speaker or not, and
5. **Speech act function of the utterance**: whether the utterance constitutes an ideational act, a control act, or a question-answer act. These speech act categories will be explained below.

In regard to the speech act function of utterances containing demonstratives, the most relevant criterion seems to be about whether or not the utterance calls for explicit action from one or both of the participants. With this distinction in mind, some of the coding categories from INCA-A (Inventory of Communicative Acts) proposed by Ninio and Wheeler (1986) and developed by Ninio, Snow, Pan, and Rollins (1994) were adopted. The codes that were employed in this study are provided in the following table:
Table 1: Speech Act Coding Scheme (adopted from INCA-A by Ninio, Snow, Pan, and Rollins [1994])

<table>
<thead>
<tr>
<th>CS</th>
<th>Counter suggestion (of an act)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PF</td>
<td>Prohibit/forbid/protest hearer’s performance of an act</td>
</tr>
<tr>
<td>QN</td>
<td>Ask a product question (wh-question), querying for information</td>
</tr>
<tr>
<td>RD</td>
<td>Refuse to carry out act requested or proposed by others</td>
</tr>
<tr>
<td>RP</td>
<td>Request, propose, or suggest an action for hearer, or for hearer or speaker</td>
</tr>
<tr>
<td>SA</td>
<td>Answer a yes-no question with a statement</td>
</tr>
<tr>
<td>SI</td>
<td>State intent to carry out by speaker</td>
</tr>
<tr>
<td>ST</td>
<td>State or make a declarative statement</td>
</tr>
<tr>
<td>YQ</td>
<td>Ask a limited alternative yes-no question</td>
</tr>
</tbody>
</table>

Later, the categories listed in Table 1 were combined to reveal three broad types of illocutionary force:

1. **Control acts** that call for an action on the part of the listener or both the speaker and the listener (Ervin-Tripp, 1989). This category included CS, PF, RD, RP from the categories listed in Table 1.

2. **Ideational acts** that merely serve the function of information provision, performing merely on the plane of ideation. This category included SI and ST from the categories listed in Table 1, and

3. **Question-Answer acts** that prompt the listener to answer a question with information or that provide replies to such questions. This category included QA, QN, SA, and YQ, from the categories in Table 1. These three types of categories are exemplified below in (1), (2), and (3) respectively.

1. Control Act: Proposal (RP)
   Şimdilik sunulmuş başlalalım
   ‘Now, let’s start with this one.’ (Adult data)

2. Ideational Act: Statement (ST)
   Kapı değil ki bu
   ‘This is not a door’ (6-year-old data)
(3) Question-Answer Act: Limited alternative question (YQ)
Bu uçak mı helikopter mı?
‘Is this a plane or a helicopter?’ (4-year-old data)

3. Results

In the first analysis we calculated the frequencies of demonstrative usage at different age groups. Overall, adults use more demonstratives per utterance than both of the child groups. Specifically, 39% of the total adult utterances contained a demonstrative, whereas this figure is 22% for 6-year-olds and 20% for 4-year-olds.

The distribution of the three demonstratives also varied across age groups. Table 2 summarizes the number and percentage of the usage of each of the demonstratives by different age groups.

<table>
<thead>
<tr>
<th></th>
<th>Bu</th>
<th>Şu</th>
<th>O</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adults</td>
<td>36%</td>
<td>39%</td>
<td>25%</td>
<td>298</td>
</tr>
<tr>
<td>6-year-olds</td>
<td>63%</td>
<td>24%</td>
<td>13%</td>
<td>206</td>
</tr>
<tr>
<td>4-year-olds</td>
<td>53%</td>
<td>22%</td>
<td>25%</td>
<td>138</td>
</tr>
</tbody>
</table>

It appears that, in utterances containing demonstratives, adults use şu most often, followed by bu and then o. The distribution of the preference seems to be quite different for children. In both child groups bu is the most preferred demonstrative followed by şu and o. The amount of usage of şu in the child groups is less than that of adults, while the usage of bu is more. The frequency of usage of o in adults and 4-year-olds do not differ much. Even though there is a drop in the use of o in 6 year olds, this does not change the general ordering in the preference of demonstratives by children, i.e. bu, followed by şu, followed by o.

Among nonverbal factors accompanying deictic utterances, we examined the referred object’s location, pointing patterns, and the addressee’s eye-gaze right after the use of the demonstrative. Table 3 presents the distribution of preferred object location during the use of the three demonstratives in all the age groups.
Table 3: Distribution of preferred object location during the use of three demonstratives in different age groups (in percentages; total usage of each demonstrative in parentheses)

<table>
<thead>
<tr>
<th></th>
<th>Close to speaker</th>
<th>Close to addressee</th>
<th>Away from both</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bu</td>
<td>Şu</td>
<td>O</td>
</tr>
<tr>
<td>Adults</td>
<td>57%</td>
<td>39%</td>
<td>26%</td>
</tr>
<tr>
<td>(109)</td>
<td>(122)</td>
<td>(68)</td>
<td></td>
</tr>
<tr>
<td>6-year-olds</td>
<td>59%</td>
<td>46%</td>
<td>21%</td>
</tr>
<tr>
<td>(139)</td>
<td>(41)</td>
<td>(26)</td>
<td></td>
</tr>
<tr>
<td>4-year-olds</td>
<td>70%</td>
<td>23%</td>
<td>67%</td>
</tr>
<tr>
<td>(73)</td>
<td>(31)</td>
<td>(35)</td>
<td></td>
</tr>
</tbody>
</table>

In adult usage, *bu* was preferred for objects close to speaker, and *o* for objects away from the speaker. *Şu*, on the other hand, was equally used for all the coded spaces. This finding confirms the analysis by Özürek & Kita (2000) that *şu* is employed for other communicative reasons than to encode the spatial distance of the object to the speaker. Children’s patterns of usage did not seem to differ from those of adults in the case of *bu* and *şu*. However, in 4 and 6-year-olds we found that *o* was used predominantly for spaces close to speaker and addressee but not much for the space away from both participants. It is possible children in this age group made reference only to objects in spaces close to speaker or addressee in this particular task, and thus, this task is not very suitable for determining how children learn to encode distal spaces with demonstratives.

Next we analysed the pointing gesture patterns used with the demonstrative *şu* to find out whether *şu* is used to encode presence or absence of joint attention by different age groups. Table 4 summarizes the results of this analysis. Adults were most often using index-finger pointing, followed by pointing with the object when *şu* was used. Children, in both age groups, did not use pointing gestures to achieve joint attention in adult-like ways. In both age groups, many uses of *şu* were not accompanied by pointing. When they pointed, both younger age groups preferred to carry out pointing with the object rather than with the index finger.

Table 4: Distribution of types of pointing gestures during the use of *şu* in different age groups (in percentages)

<table>
<thead>
<tr>
<th></th>
<th>No pointing</th>
<th>Pointing with object</th>
<th>Pointing with index finger</th>
<th>Total <em>şu</em>'s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adults</td>
<td>11%</td>
<td>47%</td>
<td>42%</td>
<td>122</td>
</tr>
<tr>
<td>6-year-olds</td>
<td>25%</td>
<td>65%</td>
<td>10%</td>
<td>41</td>
</tr>
<tr>
<td>4-year-olds</td>
<td>26%</td>
<td>51%</td>
<td>22%</td>
<td>31</td>
</tr>
</tbody>
</table>
Another indicator of joint attention manipulation that we looked at was whether or not addressee’s eye gaze moved to the referent after the demonstrative was used. Table 5 provides the percentage of cases where the addressees’ eye-gaze turns toward the direction of the referent after the speakers’ usage of each demonstrative. For adults, addressee’s eye-gazes turned towards the object referred to by the speaker after the use of ʂu. Children did not exhibit the same patterns of recipient reaction. In fact, their eye gaze moved towards the referred object most frequently after a bu rather than after a ʂu.

Table 5: Distribution of addressee’s eye-gaze after the demonstrative use in different age groups in percentages (in percentages; total usage of each demonstrative in parentheses)

<table>
<thead>
<tr>
<th></th>
<th>Bu</th>
<th>ʂu</th>
<th>O</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adults</td>
<td>61% (109)</td>
<td>74% (122)</td>
<td>49% (68)</td>
</tr>
<tr>
<td>6-year-olds</td>
<td>85% (139)</td>
<td>65% (41)</td>
<td>46% (26)</td>
</tr>
<tr>
<td>4-year-olds</td>
<td>58% (73)</td>
<td>36% (31)</td>
<td>44% (35)</td>
</tr>
</tbody>
</table>

The final analysis concerns the speech act functions of utterances containing demonstratives in adult and child conversations. Table 6 summarizes the percentages of types of demonstrative usage in relation to utterances serving one of the speech act functions (i.e., ideational acts, control acts, and question-answer acts) in adults’ and children’s conversations.

Table 6: Percentages of types of demonstratives in utterances with different speech act functions (in percentages; total usage of each speech act in parentheses)

<table>
<thead>
<tr>
<th></th>
<th>Ideational Act</th>
<th>Control Act</th>
<th>Question-Answer Act</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bu</td>
<td>ʂu</td>
<td>O</td>
</tr>
<tr>
<td>Adults</td>
<td>52% (133)</td>
<td>22% (133)</td>
<td>26% (133)</td>
</tr>
<tr>
<td>Children</td>
<td>72% (163)</td>
<td>10% (163)</td>
<td>19% (163)</td>
</tr>
</tbody>
</table>

In adult data, most demonstrative-containing proposals calling for joint or addressee action (i.e., control acts) include ʂu while demonstrative-containing utterances with purely informational content (i.e., ideational acts) feature bu. The demonstrative o is found in utterances serving an informational function
slightly more frequently than those that carry out control acts. But the difference between the two functions is not as sharp as for bu and su. Question-Answer acts that include demonstratives feature more bu's than the other two demonstratives.

For children, the picture we see in Table 6 is different. For all of the three speech acts characterizing utterances including demonstratives, children of both age groups prefer bu over the other two forms. The frequency of bu is higher, even for control acts, where adults predominantly use su.

4. Conclusions and discussion

This study contributes to research on the development of the use of the demonstratives by including data from a non-English language, namely, Turkish that encodes interactional factors as a semantic contrast in its system. In addition, it provides data from children in comparison to adults, providing a preliminary developmental profile of how children's use of demonstrative systems might differ from those of adults. Finally it provides additional evidence for the claim made by Özyürek and Kita (2000) on the analysis of the Turkish demonstrative system.

In summarizing the results of this study, the following similarities and differences between children's and adults' patterns of usage of the Turkish demonstrative system have been observed:

1. Adults use more demonstratives per utterance than children. No differences were observed with respect to the frequency of demonstrative usage between 4-year-olds and 6-year-olds.

2. Adults use su most frequently, followed by bu in their utterances containing demonstratives. Children, on the other hand, prefer bu over su. The percentage of the use of o does not appear to differ in children and adults.

3. Adults use bu mostly for objects close to speaker and o for objects in between the speaker and the addressee. Su is used by adults equally for all spaces. Children's use of demonstratives with respect to the distance of the referred object matches that of adults, except for o.

4. Children in both age groups did not accompany their use of su with adult-like pointing gestures.

5. Children in both age groups did not respond to their partners' use of su with adult-like orientation as reflected in the direction of their eye-gaze.

6. Adult preferred su over bu for utterances with a control act function, and bu over su for utterances with an ideational act function. Children did not use su as frequently as adults when their utterances served a control act function. They preferred bu irrespective of the speech act category of their utterances.

These results indicate that Turkish-speaking adults employ demonstratives more frequently and in different usage patterns than Turkish-speaking children in a task that calls for mutual orientation. Differences of children's data in
comparison to adults are especially apparent with the use of $yu$, a form used for conversational management of joint attention by adults. Children do not use this form as frequently as adults, and seem to default to the employment of $bu$ where $yu$ is called for according to adult standards. Furthermore, when they use this form they do not seem to deploy it for its joint attention getting function. Thus the amount and the nature of referring to and communicating about external objects and locations develop beyond 6-years-of age to emulate adult-like usage of the demonstrative system in Turkish.

This study also indicates that preschool children are not as competent as adults in the use of the Turkish demonstrative system for conversational management of joint attention. As Hanks (1990) suggests, demonstrative reference is a mode of engagement between the speech participants and the ostensible world. Children do not appear to be as conversationally flexible as adults in exhibiting such a mode of engagement, where the mutual orientation of the conversational participants with respect to the external world is smoothly managed.

These results might sound surprising in light of research indicating that joint attention is a very early communicative process that appears in infancy (Trevathan, 1998). Butterworth (1998), for example, finds that, at 6 months of age, babies look in the direction of their caregivers’ eye-gaze and can direct their focus on external objects and events that their caregivers call to their attention. More relevantly to the development of linguistic competence, Tomasello and Kruger (1992), among others, show that novel words are learned earlier in ostensive contexts where joint attention is achieved than in non-ostensive contexts. Thus, given such previous research, and in a linguistic system that encodes joint attention as a basic contrast, we would expect for children to have mapped $yu$ onto this function quite early. However, we have not found this to be the case.

We can speculate on the reasons why Turkish children use $yu$ to manipulate joint attention later and in different patterns than adults. One reason might be that $bu$ is more frequent than $yu$ in child-directed speech, leaving children with inadequate opportunities to figure out the subtleties of the usage of $yu$. This claim can be tested in Turkish child-directed speech. This explanation, even if child-directed data ends up warranting it, is unlikely given that absence of maximum frequency in the use of a certain linguistic form (in contrast to forms in the same paradigm) does not necessarily preclude the learner from figuring out the functions of this relatively infrequent form. Even if $yu$ turns out to be less preferred than $bu$ in speech directed to children, it definitely must play a role in adult-child conversations.

Another reason we can propose is that the integration of nonverbal factors with verbal expressions is a protracted developmental process (Goldin-Meadow, Alibali & Church, 1993), and needs to develop further beyond 6 years of age. Especially when this integration is called for in a conversational task, the challenge for preschool children might be larger than in an experimental task in which often a single utterance is called for. Work on development of
conversational competence demonstrates that mapping of conversational functions onto transparent and adult-like conversational cues is a skill that develops well beyond preschool years (Ninio & Snow, 1996). Thus, it is probably not the case that young Turkish children have an across-the-board incompetence about some of the functions of şu that adults indicate awareness of in their conversational interactions. However, in a task where they need to conversationally take into account and manipulate joint attention, and demonstrate this conversational competence by integrating nonverbal cues with verbal utterances, preschool children do not appear as flexible as adults.

References


