Early Language-Specificity in Turkish Children’s Caused Motion Expressions in Speech and Gesture

Reyhan Furman\textsuperscript{1}, Aslı Özyürek\textsuperscript{1,2,3} and Aylin C. Küntay\textsuperscript{3}
\textsuperscript{1}Radboud University, Nijmegen \textsuperscript{2}Max Planck Institute for Psycholinguistics and \textsuperscript{3}Koç University

1. Introduction

Caused motion events (e.g. a boy pulls a box into a room) are basic events where an agent (the boy) performs an Action (pull) that causes a Figure (box) to move in a spatial Relation or Path (into) to a Goal (the room). Languages differ in the way these semantic elements are mapped onto lexical and syntactic structures (Talmy, 2000). For instance, satellite-framed languages such as English encode Action in the verb (e.g. \textit{put} in (1)) and express Relation (Path) in the satellite (e.g. \textit{into} in (1)). Thus, these languages use verbs that are semantically general to express caused motion events. In contrast, speakers of verb-framed languages such as Turkish can encode both Action and Relation in the verb and also exercise the option of using semantically specific verbs to do so (e.g. \textit{sok-} ‘put in, insert’ in (2)).

\begin{enumerate}
  \item The girl put the book into her bag.
  \item Kız kitab-yı çanta-si-na sok-tu.
    \textquoteleft The girl put the book into her bag.\textquoteright
\end{enumerate}

Caused motion events are among the events that children understand and talk about early on (Slobin, 1985). Given these crosslinguistic differences, how do children start to talk about caused motion? Previous research has shown that

\begin{flushleft}
* This work has been supported by Dutch Science Foundation, ASPASIA award given to the second author and by the Turkish Academy of Sciences, in the framework of the Young Scientist Award Program to the third author (AK-TUBA-GEIP/2001-2-13).
\end{flushleft}

typological differences influence language acquisition. An early study compared how child speakers of Korean, a verb-framed language, and English, a satellite-framed language, lexicalized components of caused motion events and found their speech to be language-specific from as early as 17-20 months (Choi & Bowerman, 1991).

More recent studies have focused on placement events, a subtype of caused motion events, and investigated how children speaking different languages describe them. Slobin et al. (2008) compared how the 2-year-old speakers of four satellite-framed (English, Finnish, German, Russian) and four verb-framed (Hindi, Spanish, Tzeltal, Turkish) languages describe placement events. They found that children had already tuned in to the typological characteristics of their language at the age of 2. That is, children acquiring satellite-framed languages tended to use various sorts of directional locative markers and focused on the Relation element of the placement event. On the other hand, those acquiring verb-framed languages typically used verbs, focusing on the action of putting. Another cross-linguistic study examined the spontaneous speech of four children aged 1;8 to 2;8 speaking Hindi and Tzeltal, Narasimhan and Brown (2009) found that these children start to talk about placement events using semantically general verbs instead of specific ones, even though both languages are verb-framed.

The above-mentioned studies have shown that from 2 years onwards children are mostly sensitive to language-specific encoding of caused motion and placement. With the exception of Choi and Bowerman (1991), however, they have not examined caused motion events in general (i.e. focused on placement events only) and none of them have studied children’s cospeech gestures that accompany caused motion expressions.

Co-speech gestures are spontaneous and frequent accompaniments to speech and the expressions in the two modalities have been found to be tightly integrated pragmatically, semantically, and temporally (McNeill, 1992). Gesture is used as a tool by young children to enhance the information conveyed in their speech (Özçalıșkan & Goldin-Meadow, 2005). Children at the one-word stage use gestures to supplement their speech to produce a variety of constructions such as argument-plus-argument (e.g. by saying Mommy and pointing at a shoe to mean ‘Mommy’s shoe’) or verb-plus-argument (e.g. by saying Eat and pointing at an apple) (Özçalıșkan & Goldin-Meadow, 2009, 2005). Such supplementary gesture combinations also predict children’s later language development. For instance, the onset of supplementary speech-gesture combinations predicts the onset of two-word combinations (Iverson & Goldin-Meadow, 2005). Interestingly however, once children become adept at using a construction in speech, they cease using supplementary gestures while expressing this construction (Özçalıșkan & Goldin-Meadow, 2009).

Previous research has not, however, examined speech-gesture combinations in a specific event type or studied the types of semantic elements expressed in speech and gestures over development. An earlier study examined the development of linguistic and gestural expressions of caused motion events in
two typologically different languages—English and Turkish (Furman, Özyürek & Allen, 2006). The verbal and gestural descriptions of English- and Turkish-speaking adults differed in accordance with the typology of the language they spoke. However, child speakers of both languages were unadult-like in their verbal and gestural descriptions of caused motion and they started to display the adult patterns only after age 5. Another study recently investigated how English-speaking children aged 2;6 to 5 talked and gestured about a particular caused motion event elicited by description of a stimulus item where the experimenter pushed a ball across a small pool with the help of a stick (Göksun, Hirsh-Pasek & Golinkoff, in press). Children of all ages used supplementary gestures, although younger children were more likely to supplement their speech by location gestures whereas older ones tended to use instrument gestures as supplementary. Thus, supplementary gestures continued to augment children's speech even at age 5.

But a lot of questions relating to the development of speech and gestures about caused motion events remain unanswered. For instance, how do children start to talk and gesture about different types of caused motion in general? In what ways do their co-speech gestures contribute to their caused motion event expressions in speech? And what do co-speech gestures reveal in terms of learning the language-specific patterns of caused motion event expressions? Here we attempt to answer these questions by examining the spontaneous speech and gestures of seven Turkish-speaking children longitudinally from the age of 12 months to 36 months.

Turkish is an interesting language to track the development of caused motion expressions because of two reasons. First, although it is verb-framed, it uses both semantically specific verbs (e.g. sok ‘put in’ in (2)) and general verbs (e.g. koy ‘put’ in (3)) to encode caused motion and both verb types are found highly frequently in adult speech.

   girl book–Accusative bag–Possessive–Dative put–Past
   ‘The girl put the book in her bag.’

Second, arguments of a verb can easily be dropped in Turkish (Gürcanlı, Özyürek, Nakipoğlu, 2007). For instance, it is perfectly acceptable to utter only the verb, as in (4), to describe a situation where one throws a ball on the couch. In this case, the verb encodes Action and Agent (through the use of the person marker) and the remaining semantic elements can be recovered from the discourse context.

(4) At–ti–m.
   throw–Past–1sg.
   ‘(I) threw.’
In the present study, we investigate how Turkish children start to describe all types of caused motion events and investigate whether the language-specific trends found in the description of placement events apply to the broader category of caused motion. We also examine children’s co-speech gestures to find out which semantic elements they encode in relation to what is expressed in speech.

We predict that Turkish-speaking children will start to talk about caused motion events by encoding both Action and Relation using verbs, similar to child speakers of other verb-framed languages (Choi & Bowerman, 1991; Narasimhan and Brown, 2009; Slobin et al., 2008). However, there are two possibilities on what type of verbs they will use. In line with Choi and Bowerman (1991), we might expect their verb use to be language-specific from the outset and predict that they will use both semantically general and specific verbs. Alternatively following Slobin et al. (2008) and Narasimhan and Brown (2009), we might expect Turkish-speaking children to use general verbs earlier than specific ones. With regard to gestures, we predict that children will use gestures to supplement their speech and we also expect this supplementation to decrease with age, in line with Özcılışkan and Goldin-Meadow (2009, 2005). Finally, we explore the relationship between gesture use and the use of general and specific verb types.

2. Methods
2.1. Data

Using the Koç University Longitudinal Language Development Corpus, we sampled the spontaneous speech and co-speech gestures of seven Turkish-speaking children. The children were videotaped at home while engaged in daily activities such as eating, playing, and conversing with their caregivers (parents, relatives or nannies) or occasionally with the researchers. One-hour sessions were sampled for each child every three months between the ages of 12 and 36 months and approximately eight sessions were analyzed per child.

2.2. Speech Coding

Overall, 980 utterances that referred to caused motion events were transcribed and then coded for the type of verb, type of construction and the semantic elements represented. Children used both semantically general verbs which encode only Action and specific ones which encode Action and Relation. A comprehensive list of these verbs is given in the Appendix.

Two main construction types were distinguished: Verb Only and Verb plus Arguments. The Verb Only category denotes those utterances which included only a verb. Depending on the type of verb used, these constructions encoded either Action, as in (5a) or Action and Relation as in (5b).
The Verb plus Arguments category denotes constructions which included a verb and one or more of its arguments. Such constructions encoded a variety of semantic elements, such as Action, and Figure (6a), Action, and Goal (6b), or Action, Relation, Figure and Goal (6c).

(6) a. Bir mandal at-ti-m.
    one clothespin throw-Past-1sg.
    ‘(I) threw a clothespin.’ (İrem, 26 months)

b. Bura-ya koy-ahm.
    here-Dative put-Optative
    ‘Let’s put here.’ (Burcu, 36 months)

    hand-Possessive-Dative cream put.on-Past-1pl.
    ‘(We) put cream on my hand.’ (Can, 36 months)

2.3. Gesture Coding

When coding gestures, we focused only on those gestures that accompanied the utterances that encoded caused motion events in speech. The 246 gestures that accompanied such utterances were coded for type, the semantic element represented, and speech-gesture relationship. For type, gestures were categorized as deictic or iconic (McNeill, 1992). Deictic gestures included holding an object up to show it or pointing at an object or location (e.g. pointing at the carpet while saying Oraya dökütm ‘(I) poured there’). Iconic gestures represented characteristics or actions of entities (e.g. a child saying Tenis topunu böyle duvara attım ‘(I) threw the tennis ball to the wall like this’ while her hand, cupped as if holding a ball, moves from right to left).

Gestures were also coded for the semantic elements (e.g. Action, Agent, Figure, Goal/Relation or any combinations of these) they represented based on the framing of the co-occurring speech. For instance, a point at the armchair co-occurring with the utterance Buraya koyabilirsin ‘(You) can put here’ was coded as Goal/Relation.

---

1 Agents were coded only if they were mentioned as nouns or pronouns. Person marking on the verb was not coded as an Agent argument.
2 Gestures representing Goals and Relations were coded and analyzed as a single category.
Finally, gestures could be related to the co-occurring speech in three different ways (Özçalişkan & Goldin-Meadow, 2005; 2009). Reinforcing gestures expressed the same information as the utterance they occurred with (e.g. asking *Onu çöpe mi attın?* ‘Did you throw it to the garbage’ while moving a fist-shaped hand down). Disambiguating gestures clarified the referent of any deictic word in speech (e.g. saying *Anne şuna da su koysana* ‘Mommy put water in that one too’ and showing a water bottle). Supplementary gestures encoded semantic elements not conveyed in speech and added to the information to it (e.g. pointing at a notebook while saying *Kopardım* ‘(I) ripped’). These gestures could be both iconics and points.

3. Results

3.1. Speech

We first calculated the mean proportion of different construction types each child used in each session to determine whether construction use changed with age. Table 1 shows the relevant figures. Verb-Only constructions were used more frequently in the 14-27 month age period compared to the 28-36 month age period, $F(1, 49) = 73.19, p = .000$.

<table>
<thead>
<tr>
<th></th>
<th>Verb-Only</th>
<th>Verb-Plus-Arguments</th>
</tr>
</thead>
<tbody>
<tr>
<td>14-27 months</td>
<td>.75</td>
<td>.25</td>
</tr>
<tr>
<td>28-36 months</td>
<td>.25</td>
<td>.75</td>
</tr>
</tbody>
</table>

We then examined the mean proportion of utterances containing different semantic elements and found that children’s mention of elements changed with age, as indicated in Table 2. Children increased their mention of Agent, $F(1, 49) = 9.94, p = .003$, Figure, $F(1, 49) = 83.41, p = .000$ and Goal/Relation $F(1, 49) = 26.6, p = .000$, with age. There were no age differences in the mention of Action.

<table>
<thead>
<tr>
<th></th>
<th>Action</th>
<th>Agent</th>
<th>Figure</th>
<th>Goal/Relation</th>
</tr>
</thead>
<tbody>
<tr>
<td>14-27 months</td>
<td>.97</td>
<td>.08</td>
<td>.10</td>
<td>.11</td>
</tr>
<tr>
<td>28-36 months</td>
<td>.98</td>
<td>.23</td>
<td>.43</td>
<td>.34</td>
</tr>
</tbody>
</table>

Last, we calculated the mean proportion of utterances that contained semantically general and specific verbs (out of all utterances with verbs) in each age period and found that both verb types were used equally frequently by children at different ages, as shown in Table 3.
These results reveal that Turkish-speaking children start to describe caused motion events using only verbs and only later talk about the semantic elements other than Action. Their use of semantically general and specific verbs does not change with age.

### 3.2. Gesture

We examined the mean proportion of different types of gestures produced by each child during their caused motion expressions. Regardless of age group, children used more deictic gestures than iconic ones, $F(1, 62) = 10.51, p = .002$. The use of different gesture types, however, did not change with age. That is, children in both age groups used deictic and iconic gestures at comparable proportions, as shown in Table 4.

**Table 4. Mean proportion of gesture types**

<table>
<thead>
<tr>
<th></th>
<th>Deictic</th>
<th>Iconic</th>
</tr>
</thead>
<tbody>
<tr>
<td>14-27 months</td>
<td>.68</td>
<td>.35</td>
</tr>
<tr>
<td>28-36 months</td>
<td>.61</td>
<td>.43</td>
</tr>
</tbody>
</table>

We next calculated the mean proportion of gestures encoding different semantic elements. Overall, there was a significant difference in the expression of semantic elements in gesture, $F(3, 124) = 20.47, p = .000$. Post-hoc tests revealed that Agent was encoded less frequently compared to Action, Figure and Goal/Relation in both age periods (Bonferroni, $p_s = .000$) (see Table 5).

**Table 5. Mean proportion of gestures containing each semantic element**

<table>
<thead>
<tr>
<th></th>
<th>Action</th>
<th>Agent</th>
<th>Figure</th>
<th>Goal/Relation</th>
</tr>
</thead>
<tbody>
<tr>
<td>14-27 months</td>
<td>.36</td>
<td>.00</td>
<td>.51</td>
<td>.51</td>
</tr>
<tr>
<td>28-36 months</td>
<td>.39</td>
<td>.03</td>
<td>.40</td>
<td>.56</td>
</tr>
</tbody>
</table>

We then investigated the relation between speech and gesture. Across the two age periods, children used reinforcing, disambiguating and supplementary gestures differently, $F(2, 93) = 3.69, p = .029$. Post-hoc analyses indicated that disambiguating gestures were used less frequently than reinforcing ones (Bonferroni, $p = .03$). The use of disambiguating, $F(1, 30) = 10.23, p = .033$, and supplementary gestures changed with age, $F(1, 30) = 17.50, p = .000$. Whereas supplementary gestures were used more frequently in the 14-27 month period, disambiguating gestures were used more in later ages. The relevant figures are shown in Table 6.
Finally, we examined whether gestures became tuned to verb types over development. That is, we focused on gestures that encoded Action or Goal/Relation and investigated which type of verbs these gestures co-occurred with. There were no significant differences between the use of Action and Goal/Relation gestures neither for semantically general verbs nor for specific ones- even though there was a trend for specific verbs to be used less frequently with Action gestures than with general verbs. There were no differences between the two age periods either. That is although older children were more likely to use Action and Goal/Relation gestures with specific verbs, these differences were not significant as shown in Table 7.

| Table 7. Mean proportion of Action and Goal/Relation gestures occurring with different verb types |
|-----------------------------------------------|-----------------------------------------------|
| **General Verbs** | **Specific Verbs** |
| Action Gesture | Action Gesture | Goal/Rel. Gesture | Goal/Rel. Gesture |
| 14-27 months | .52 | .63 | .20 | .45 |
| 28-36 months | .42 | .51 | .41 | .80 |

4. Discussion

This study investigated how child speakers of a verb-framed language, Turkish, start to talk and gesture about caused motion events. We focused on three specific questions. First, do children start with language-specific constructions and lexical items? Second, how do children’s co-speech gestures contribute to their caused motion event expressions in speech? And what do co-speech gestures reveal in terms of learning the language-specific patterns of caused motion event expressions?

We found that language-specificity in caused motion event descriptions was evident from the start. Turkish children started to talk about caused motion events using only verbs, similar to children speaking other verb-framed languages (Choi & Bowerman, 1991; Narasimhan & Brown, 2009; Slobin et al., 2008). After the age of 28 months, Verb-only constructions were replaced by Verb-Plus-Argument constructions, suggesting that children had established the caused motion construction in their repertoire. However, children still used Verb-only constructions 25% of the time at the 28-36 month age period, indicating that their use of this construction was not solely due to developmental reasons. That is, children’s continued use of Verb-Only constructions might be due to the fact that arguments can be freely elided in Turkish and such constructions might be used to the same extent by adults. Future research should...
determine the rate of Verb-Only constructions in the caused motion expressions of older children and adults in order to find out whether the rate we report here is particular to the speech of 3-year-olds. A related issue is that argument ellipsis is linked to the existence of specific verbs in a language (Brown, 2008). For instance, semantically specific verbs in Tzeltal often encode information about the Figure (or patient) and both adults and 3;6-year-old children represent the object argument less in their speech when the verb they use is specific (Brown, 2008). Future research should also investigate whether such a link between the use of specific verbs and argument ellipsis exists in Turkish.

In addition to starting to describe caused motion events by using only verbs, Turkish children used semantically specific verbs as frequently as general ones even at very early ages. Interestingly, children’s use of specific verbs showed not only high token frequency but also high type frequency. That is, type frequency of specific verbs per child ranged between 7 and 13, showing that each child used many different specific verbs (see Appendix for a list of all verbs). In contrast to these results, 2-year-old Turkish children have previously been found to talk about placement events using the semantically general verbs koy- ‘put’ or at- ‘throw’ (Slobin et al., 2008). Likewise, child speakers of Hindi and Tzeltal, both verb-framed languages, rarely used specific verbs in their descriptions of placement events (Narasimhan & Brown, 2009).

There could be several reasons why our results are contradictory to those of previous research. First, the aforementioned studies have focused on only one type of caused motion event, i.e. placement, whereas we studied caused motion in general. The full pattern of language-specificity might be easier to spot when a more general domain of events is investigated. Support for this hypothesis comes from another study which examined the specificity of Tzeltal-speaking children’s verbs in their descriptions of intransitive and transitive events (Brown, 2008) and found that Tzeltal children use many specific verbs, in contrast to Narasimhan and Brown (2009). Second, intra-typological differences between languages might be influencing acquisition patterns. For instance although both Turkish and Hindi are verb-framed, Hindi does not typically use specific verbs which encode Action and Relation to express caused motion. Instead, it uses semantically general verbs (e.g. daal ‘put/drop’ or rakh ‘put/place’) encoding Action in conjunction with locative case-marked nominals encoding Relation (Slobin et al., 2008). In contrast, Turkish uses both patterns to express caused motion. Moreover, in Hindi specific verbs such as ghus-aa ‘insert’ are absent in caregivers’ input speech to children (Narasimhan & Brown, 2009) whereas they might be found more frequently in the Turkish input. Thus, as suggested by Slobin et al. (2008) languages exhibit the properties typical of the language type to differing degrees and such intricate intra-typological variation affects the development of caused motion expressions.

Children’s gestures also played an important role in the expression of caused motion events. As expected, gestures were used to supplement speech and at the Verb-Only stage they revealed more elements of the caused motion event than speech. That is, Figure and Goal/Relation were encoded in gesture as frequently
as Action in all ages. This pattern of representation was different than the one in speech where the semantic elements other than Action emerged in time. As children learned the Verb-Plus-Argument construction, the use of supplementary gestures decreased, in line with Özçalışkan and Goldin-Meadow (2009, 2005). However, 28-36 month-olds still used gestures to supplement their speech 25% of the time, suggesting that some semantic elements continue to be encoded exclusively in gesture even after children are able to express all elements in speech. In addition, children increased their use of disambiguating gestures over time, as documented in previous research (Özçalışkan & Goldin-Meadow, 2009, 2005). We surmise that the increase in disambiguating gestures might be related to children’s increased use of deictic words with age and the development of their comprehension of the discourse constraint that the referents of deictic words have to be clarified. These possible links have to be tested in future studies.

We also found that overall children used deictic gestures more compared to iconic gestures when describing caused motion events. This is in line with Göksun, Hirsh-Pasek and Golinkoff (in press) who also found that English-speaking children between the ages of 2;6 to 5 produce more deictics than iconics in their narrations of causal events. In contrast to previous research (Göksun, Hirsh-Pasek & Golinkoff, in press) which found that children start to use iconics in their causal event descriptions after the age of 4, age did not influence the type of gesture used. That is, iconics were used with the same frequency across the two age periods in our study.

Early language-specificity was also evident in several aspects of children’s gestures. First, children of all ages represented Actions in their gestures as frequently as Goals/Relations. This phenomenon can be likened to children’s use of both general (encoding Action) and specific (encoding Action and Relation) verbs from early on. That is, Action and Relation were represented equally frequently in both speech and gesture- a reflection of the typological properties of Turkish as a verb-framed language.

Second, even at the age of 3 when they had learned a complex construction, children still used some supplementary gestures. We believe that the continued use of supplementary gestures might be related to argument ellipsis in Turkish. That is to say, arguments representing Figure and Goal/Relation can be easily dropped since they can both be recovered from discourse context and verb semantics (Brown, 2008) and thus continue to pop up in gesture even in later years of development.

Third, language-specificity is apparent in gesture types that children produce. Namely, Turkish children use deictic and iconic gestures equally frequently from the beginning and around 30-40% of gestures produced by 14-36 month old children are iconics. Compared to English-speaking children who used almost no iconics before age 4-5 (Göksun, Hirsh-Pasek & Golinkoff, in press), Turkish children’s early and frequent use of iconic gestures might appear precocious at first glance. However, we believe that it is related to the typological properties of Turkish. Since Turkish is a verb-framed language,
children’s early speech contains a lot of verbs thus many co-speech gestures represent Actions, which can only be depicted by producing iconics, from early on. In contrast, learners of satellite-framed languages might produce few iconics early on in development since their early descriptions of caused motion events do not include verbs (Choi & Bowerman, 1991).

Finally, gesture type might have also interacted with speech-gesture relation as a result of language-specificity. That is, in a satellite-framed language such as English, supplementary gestures tend to be deictics (Özçalışkan, personal communication) whereas half of all supplementary gestures in our data are iconics. Thus, the verb-framed nature of Turkish might lead young children to produce many verbs in their early speech about caused motion, which in turn, allows them to use iconic Action gestures that encode supplementary elements such as Goal/Relation and/or Figure. An example of a supplementary iconic gesture is a child talking about wearing a backpack and saying Öyle takyorum ‘(I) put.on/wear like that’ and moving both of her hands, cupped as if holding the straps of the backpack, from the sides of her body to her shoulders. In this case, while her speech encodes Action and Relation through the use of the specific verb tak- ‘put.on/ attach’, her gesture represents Action, Figure and Goal. Testing the link between language typology, gesture type and speech-gesture relation is a task for future crosslinguistic research.

In sum, we have investigated how Turkish-speaking children between the ages of 1 and 3 represent caused motion events in their speech and gestures. We found that children’s speech about caused motion- both in the choice of constructions and in the type of verbs used- was language-specific from the start. Consequently, the gestures children produced also mirrored this robust language-specificity. Further research is necessary to clarify the effects of language typology on caused motion event speech and gestures.

Appendix
List of AH Verbs Used

Semantically General

Semantically Specific
References


