

Early Speech about Manner and Path in Turkish and English: Universal or Language-Specific?

Shanley Allen¹, Aslı Özyürek², Sotaro Kita³, Amanda Brown¹,
Reyhan Turanlı⁴, and Tomoko Ishizuka¹

¹Boston University, ²Koç University, ³Max Planck Institute for
Psycholinguistics, and ⁴BosphorousUniversity

1. Introduction

Manner and Path occurring simultaneously in a motion event can be encoded in speech within either one or two clauses, depending on the typology of the language (Talmy 1985). Specifically, satellite-framed languages such as English typically use one clause to encode Manner and Path as shown in (1). Verb-framed languages such as Turkish typically use two separate clauses, mostly in a matrix and dependent clause construction (2).

(1) Tomato Man rolled down the hill.

(2) yuvarlan-arak tepe-den iniyor
roll-Connective hill-Ablative descent:Present
'S/he/it descends the hill while rolling.'

There are two competing hypotheses in the literature concerning when children begin exhibiting these (and other) language-specific differences in their speech. The *Universal Hypothesis* (Bowerman 1982, Slobin 1985) claims that children begin with a universal default pattern that reflects universal cognitive representations, and only gradually learn the language-specific patterns as a result of exposure to their target language. In contrast, the *Language-Specific Hypothesis* (Slobin 1987, Bowerman & Choi 2000) claims that children follow language-specific patterns from their earliest productions.

The present study tests these hypotheses with motion event descriptions in speech in English- and Turkish-speaking 3-year-olds and adults when both manner and path of motion are salient and need to be expressed as part of the

* This research was funded by grant BCS-0002117 from the National Science Foundation to the first three authors, and by a Turkish Academy of Sciences Young Scientist Award to A. Özyürek. Substantial logistical support was provided by the Max Planck Institute for Psycholinguistics. Many preschools in both Boston and Istanbul generously allowed us to study the children under their care. We acknowledge with thanks all these contributions to our work.

event description. The specific research questions we address are: (1) What are children's early descriptions of motion events including both Manner and Path? (2) Does this early stage reveal a universal or a language-specific pattern?

Previous research by Özyürek & Özçalışkan (2000) and Özçalışkan & Slobin (1999, 2000) found language-specific differences with regard to motion event descriptions between Turkish and English as early as 3;0 in narrations of the wordless picture book *Frog, Where are You?* (Mayer 1968). However, the results of these studies may be an effect of the stimulus used, because there are few instances of "mandatory" expression of simultaneous occurrence of Manner and Path in the frog story. The present study looks at this question again using new elicitation stimuli, all of which contain simultaneous occurrence of Manner and Path.

2. Method

2.1 Participants

Participants in the study were 40 Turkish speakers and 40 English speakers. Twenty of each group were adults ranging in age from about 18 to 40, all university students in either Boston (English) or Istanbul (Turkish). The remaining 20 participants in each group were children; both Turkish-speaking and English-speaking groups had a mean age of 3;8, with similar ranges for each (Turkish 3;6-4;0, English 3;3-4;3). All children were tested in preschools in the same two cities.

2.2 Materials and Procedure

All data were collected by elicitation, using a set of 5 video clips depicting events involving simultaneous Manner and Path. Two manners and three paths were depicted, yielding the following combinations: JUMP + ASCEND, JUMP + DESCEND, JUMP + GO AROUND, ROLL + ASCEND, and ROLL + DESCEND. As indicated in (1) and (2) above, English typically represents these events within one clause, while Turkish typically represents them within two separate clauses.

Each video clip was about 6 seconds in duration, and had three salient components: an entry event, a target motion event, and a closing event. All clips involved a round red smiling character and a triangular-shaped green frowning character, moving in a simple landscape. Participants chose their own names for the characters; we refer to them here as Tomato Man and Green Man. The ROLL + ASCEND clip goes as follows. The initial landscape on the screen is a large hill ending in a cliff which drops off to the ocean; Tomato Man is located at the base of the hill. Green Man enters the scene from the left and bumps into Tomato Man [entry event], then Tomato Man rolls up the hill [target motion event], and finally Tomato Man falls off the cliff into the ocean and bobs up and down [closing event].

Participants were tested individually in a quiet space at their university (adults) or preschool (children). All interactions were videotaped for later

coding and analysis. The procedure had two parts. During the warm-up phase, the experimenter showed participants a typical scene from a clip and introduced them to the characters and the landscape. She explained that they would see a series of clips depicting adventures of Tomato Man and Green Man, and after each one they should narrate the clip to a listener who had not seen it. She then gave them two practice rounds with clips involving motion events like those in the test clips, and particularly encouraged them to give information about the target event if they did not spontaneously do so.

In the testing phase, the experimenter presented the 5 test clips for the participant to narrate. She played each clip twice, and then asked the participant to tell the listener what happened in the clip. If participants did not mention the target event in their narration, either the experimenter or the listener encouraged them to do so with a question like “What happened after Green Man bumped into Tomato Man?” Crucially, this question did not focus explicitly on either Manner or Path. Each of the 5 clips was narrated in turn.

Note that the procedure described here is part of a larger study involving a total of 10 video clips and two follow-up tasks. Only the 5 clips relevant to this paper are discussed here. Data from the two follow-up tasks have not yet been analyzed, so they are discussed only tangentially in section 4 with respect to their role in helping to interpret the results of the narration task. In addition to the data discussed here, we have collected data on all these tasks from English-speaking and Turkish-speaking 5-year-olds and 9-year-olds, as well as Japanese speakers in all four age groups. These data are also not yet analyzed, and will be reported in future publications.

2.3 Coding and Analysis

For purposes of analysis, we transcribed all utterances from the narrations describing the target event, and divided each into “sentences.” We define a sentence here as a matrix clause plus its subordinates; two matrix clauses separated by a coordinating conjunction are considered two sentences. We then coded each sentence for the structural patterns of speech relating to Manner and Path. Four coding categories were used, as exemplified in (3) with data from English. The first category, “C”, denotes a tight combination of Manner and Path within the verbal clause, as in the verb-satellite construction typically found in English. “M” denotes use of only a Manner element in the clause (i.e. no Path), while “P” denotes use of only a Path element (i.e. no Manner). Finally, “MP” denotes use of both Manner and Path in one sentence where each of these is a separate verbal element (e.g. tensed verb, gerund).

- (3) C: Tomato Man rolls up the hill.
 M: And the Red Guy twirled.
 P: He goes up a hill.
 MP: The red tomato was spinning when he was going up the steep incline.

The analysis was item-based (i.e. clip-based) rather than sentence-based. We determined the number of items (out of a total of 5) for which a given participant used a given construction (C,M,P,MP) at least once in their description of the clip. For example, an item would be coded as both "P" and "C" for the following description of the target event in the ROLL + ASCEND clip: *The Red Guy went up; he twirled up*. We then calculated the percentage use of each construction for each participant, and determined the mean percentage use for each groups of participants. As indicated in section 2.1, four groups of participants were compared: Turkish-speaking children (TC), Turkish-speaking adults (TA), English-speaking children (EC), and English-speaking adults (EA). We then compared the means across the groups using ANOVAs and independent t-tests.

3. Results

Table 1 shows the mean proportion of uses of each construction in describing the five target event clips for each of the four groups.

Table 1: Mean proportion of uses of each construction in describing the five target event clips

	C	M	P	MP
Turkish-speaking adults	.00	.08	.20	.82
English-speaking adults	.72	.11	.20	.12
Turkish-speaking children	.00	.46	.58	.22
English-speaking children	.55	.31	.29	.00

A 2*2 ANOVA analysis with language and age as factors was significant for the four constructions types. There were main effects of language for three of the four construction types ($p=.00$ for C, $p=.01$ for P and $p=.00$ for MP) and of age for four types ($p=.01$ for C, $p=.00$ for M, $p=.00$ for P, and $p=.00$ for MP). The interaction between language and age was significant for C ($p<.05$), P ($p=.015$) and MP ($p=.00$), and marginal for M ($p=.059$). Significant effects at this level allowed us to proceed with post-hoc comparisons of individual groups.

3.1 Adults

We first compared results across the two adult groups to determine whether there are indeed linguistic differences for the children to acquire. Table 2 gives the relevant figures. Independent t-tests showed that the Turkish-speaking adults used significantly fewer C constructions ($p=.00$) and significantly more MP constructions ($p=.00$) than the English-speaking group. However, no significant differences were found between the groups for use of M-only and P-only constructions. These results reveal a clear language-specific effect in choice of construction when both Manner and Path are expressed within a sentence.

Table 2: Mean proportion of uses of each construction in describing the five target event clips, TA vs. EA

	C*	M	P	MP*
Turkish-speaking adults	.00	.08	.20	.82
English-speaking adults	.72	.11	.20	.12

* denotes significant difference between groups for construction type indicated

3.2 Turkish (children and adults)

A comparison between the two Turkish-speaking groups (3-year-olds and adults) also shows some similarities and differences as illustrated in Table 3.

Table 3: Mean proportion of uses of each construction in describing the five target event clips, TA vs. TC

	C	M*	P*	MP*
Turkish-speaking adults	.00	.08	.20	.82
Turkish-speaking children	.00	.46	.58	.22

* denotes significant difference between groups for construction type indicated

The Turkish-speaking children resemble the adults in their preference for expressing the Manner and Path separately (M,P,MP). However, they differ from the adults in their use of the different separation strategies. Turkish-speaking children used more M-only and P-only sentences than the adults, but fewer MP sentences, all according to independent t-tests significant at the $p=.00$ level.

3.3 English (children and adults)

A comparison between the two English-speaking groups (3-year-olds and adults) reveals a similar pattern as indicated in Table 4.

Table 4: Mean proportion of uses of each construction in describing the five target event clips, EA vs. EC

	C*	M*	P	MP*
English-speaking adults	.72	.11	.20	.12
English-speaking children	.55	.31	.29	.00

* denotes significant difference between groups for construction type indicated

Like the English-speaking adults, the English-speaking children prefer a tight combination of Manner and Path (C) over constructions which separate the two (M,P,MP). However, they also differ from the adults. According to independent t-tests, English-speaking children use significantly fewer C ($p=.00$) and MP ($p=.00$) constructions, and significantly more M constructions ($p=.00$), than do the adults.

3.4 Children

Finally, we compare the two groups of children, as shown in Table 5.

Table 5: Mean proportion of uses of each construction in describing the five target event clips, TC vs. EC

	C*	M	P*	MP*
Turkish-speaking children	.00	.46	.58	.22
English-speaking children	.55	.31	.29	.00

* denotes significant difference between groups for construction type indicated

Comparison between the two child groups reveals language-specific differences at an early age. English-speaking children use C constructions at least once in describing 55% of the clips. Turkish-speaking children use MP at least once in describing 22% of the clips, but English-speaking children never use this pattern. Finally, Turkish-speaking children use significantly more P constructions than English-speaking children ($p=.00$).

4. Discussion

This study used a new and highly focused set of stimuli to test whether children show language-specific or universal patterns in their earliest uses of speech about motion events. In contrast with stimuli used in other similar studies, our stimuli strongly encouraged production of both Manner and Path in one sentence to describe the target motion events. Unlike other studies, we found both universal and language-specific patterns in children's motion event speech. We discuss each of these patterns in turn.

4.1 Universal Patterns

Three-year-old children speaking both English and Turkish produced more clauses than their counterpart adults with just one element of the motion event represented – either Manner or Path. Turkish-speaking children produced a significantly higher proportion of both M-only and P-only constructions than Turkish-speaking adults. English-speaking children also used more of both construction types than English-speaking adults, but only the difference for M-only constructions was significant. Given that children are less skilled users of language with fewer resources for cognitive processing, this difference between children and adults is not surprising. We consider several potential sources for it here.

First, children may not have the cognitive resources to focus on or take in information regarding more than one aspect of the event at once (i.e. only Manner or Path but not both), and therefore represent only the one aspect in their speech. Second, children may only be able to remember one aspect of the event at once during narration, even though they are cognizant of both during

their viewing of the clip. Third, children may find it morphologically or syntactically more difficult to represent both Manner and Path in one sentence than to just represent either one alone. Fourth, children may follow a narrative style which prefers to represent only one new element per clause; this would correspond with the preference for representing only one new argument per clause found in Du Bois's (1987) Preferred Argument Structure (see Clancy 2003 and Allen & Schröder 2003 for information on this pattern in Korean- and Inuktitut-speaking children respectively). A fifth potential factor is vocabulary knowledge: children may not either know or be able to access quickly the relevant vocabulary items for each of the manners and paths they wish to describe. While lack of appropriate vocabulary is almost certainly relevant for a few scattered omissions, it may also be the source of a systematic pattern if 3-year-olds overall have particular difficulty with the words required to describe the manners or paths depicted in the clips in this study.

Further study of the data is clearly necessary to tease apart these five alternatives. In addition, two follow-up tasks conducted with the participants but not yet analyzed will help to interpret the narrative results. In the first follow-up task, the experimenter asked the participant to silently act out the clip with a set of toy models, purportedly to show the listener more clearly what happened. Each clip was shown once and then acted out. (Additional showings and attempts at acting out occurred frequently with the children.) The results of this act-out task should offer insight into the participants' cognitive understanding of and memory for the event, independent of their language production.

In the second follow-up task, the participant's vocabulary knowledge was tested for each clip where they failed to name the manner. The experimenter showed the participant the target motion event portion of the original clip, and asked them to name the action. Results from this task will help to rule out simple vocabulary knowledge as a factor in performance.

4.2 Language-Specific Patterns

The second major finding of this study is that English- and Turkish-speaking children also pattern like adult speakers of their respective languages. English-speaking children predominantly use a tight combination of Manner and Path (C) within one clause, and they never use constructions which have two verbal elements in one sentence (MP). In contrast, Turkish-speaking children have already started to use constructions typical of their target language in which Manner and Path are expressed in two verbal elements in one sentence (MP).

It is important to note that input patterns cannot account for all of the child results reported here: children cannot just be producing what they hear most in the input (insofar as the patterns of adult speech found in this study correspond to patterns in the input). Uses of M and MP occur with nearly identical frequency in our English-speaking adult data (0.11 and 0.12 respectively). However, M and MP are used in quite different proportions by the English-

speaking children ($M=0.27$, $MP=0.0$). If children were merely mirroring the input they receive, we would expect a much closer correspondence between these proportions across children and adults.

The finding of language-specific patterns in the child data corresponds broadly with findings from previous similar studies using the wordless picture book *Frog, Where are You?* as a stimulus (Özyürek & Özçalışkan 2000, Özçalışkan & Slobin 1999, 2000). However, there is at least one important difference between those studies and ours that make our findings an important advance in understanding this phenomenon. Specifically, the previous studies elicited relatively few clauses in which both Manner and Path were produced. For instance, English-speaking adults reported in Özyürek & Özçalışkan (2000) produced a much lower proportion of C constructions than in our study (0.32 vs. 0.72), and Turkish-speaking adults produced no MP constructions at all (0.00 vs. 0.82). As in our study, English-speaking adults also produced virtually no MP constructions, and Turkish-speaking adults produced no C constructions. Therefore, the “frog story” did not serve well to elicit constructions where both elements are expressed. As we have seen in section 3, the stimuli used in our study are much more suitable for targeting the adult usage and developmental pattern of simultaneous Manner and Path expressions.

Finally, it is another important finding of our study that while English-speaking children used the most common English adult pattern (C) in 55% of the clips, Turkish-speaking children used the most common Turkish adult pattern (MP) in only 22% of the clips. This difference points out the fact that in general when two elements need to be expressed simultaneously it might be easier for children to acquire constructions in which two separate elements can be expressed within one clause than in a matrix-subordinate clause construction.

4.3 Future Directions

Although the findings from this study represent a significant advance in our understanding of the universal and language-specific tendencies of young children related to motion event speech, several questions remain to be resolved. The first obvious question is to tease apart the various factors influencing children’s early tendency to produce only one of the components of the motion event in their narration of the event, in comparison with the adult propensity to produce both Manner and Path.

A second related question asks why English-speaking children seem to produce both Manner and Path together in one clause more frequently (or earlier developmentally) than Turkish-speaking children: in 55% of event narrations for English-speakers vs. only 22% for Turkish-speakers. Here a linguistic complexity explanation seems plausible: it might be easier syntactically to produce a verb-satellite construction (e.g. *rolled down*) than a matrix-subordinate construction (e.g. *descended rolling*). Since the former is virtually never used in adult Turkish, the Turkish-speaking children would then avoid both structures

involving a combination of Manner and Path, and instead produce either M alone or P alone, which is what our results show.

Finally, examination of data from the same tasks with older children would provide information about the developmental pattern of use of Manner and Path, which might help to shed light on reasons for the 3-year-olds' systematic preference for representing only one aspect of the motion event in their narratives. The data we have collected from 5- and 9-year-olds will be useful in this regard.

Each of these questions requires further research beyond the scope of this paper to reach a satisfactory answer. We hope to provide more insight into these and other questions in further publications.

In summary, we have shown that children exhibit both universal and language-specific tendencies in their earliest speech about motion events involving Manner and Path. Our findings extend and strengthen previous related findings of language-specificity in the literature because our stimuli virtually require use of both Manner and Path to describe each event (stimuli used in other studies with motion events did not have this feature). The particular stimuli we used also enabled us to uncover a universal tendency across the two child groups that previous studies did not find. Specifically, children represent only one of the two elements of the motion event in their description of the clips significantly more often than the adults. In general, we expect that children's early cognitive representations of the motion event elements (e.g. ability to remember or represent one or two elements at the same time) as well as the degree of complexity inherent in the syntactic structure required to express these elements play major roles in how children learn to talk about motion events. The particular stimuli we used seems to illuminate these two effects on children's acquisition of motion event expressions in ways not shown previously with other stimuli.

References

- Allen, Shanley E. M. and Schröder, Heike. (2003). Preferred Argument Structure in early Inuktitut spontaneous speech data. In John W. Du Bois; Lorraine E. Kumpf; and William J. Ashby (Eds.), *Preferred Argument Structure: Grammar as architecture for function*. Amsterdam: John Benjamins.
- Bowerman, Melissa. (1982). Starting to talk worse: Clues to language acquisition from children's later speech errors. In S. Strauss (Ed.), *U-shaped behavioral growth*. New York: Academic Press.
- Bowerman, Melissa & Choi, Soonja. (2000). Shaping meanings for language: Universal and language specific in the acquisition of spatial semantic categories. In Melissa Bowerman & Stephen C. Levinson (Eds.), *Language acquisition and conceptual development*. Cambridge: Cambridge University Press.
- Clancy, Patricia M. (2003). The lexicon in interaction: Developmental origins of Preferred Argument Structure in Korean. In John W. Du Bois, Lorraine E. Kumpf & William J. Ashby (Eds.), *Preferred Argument Structure: Grammar as architecture for function*. Amsterdam: John Benjamins.

- Du Bois, John W. (1987). The discourse basis of ergativity. *Language* 63, 805-855.
- Mayer, Merce. (1969). *Frog. Where are you?* New York: Dial Press.
- Özçalışkan, Şeyda & Slobin, Daniel I. (1999). Learning how to search for the frog: Expression of manner of motion in English, Spanish, and Turkish. In Annabel Greenhill, Heather Littlefield, and Cheryl Tano (Eds.), *Proceedings of the 23th annual Boston University Conference on Language Development* (pp. 541-552). Somerville, MA: Cascadilla Press.
- Özçalışkan, Şeyda & Slobin, Daniel I. (2000). Climb up vs. ascend climbing: Lexicalization choices in expressing motion events with Manner and Path components. In S. Catherine Howell, Sarah A. Fish, and Thea Keith-Lucas (Eds.), *Proceedings of the 24th annual Boston University Conference on Language Development* (pp. 558-570). Somerville, MA: Cascadilla Press.
- Özyürek, Aslı & Özçalışkan, Şeyda. (2000). Differences in learning how to encode motion events in Turkish versus English: Evidence from speech and gesture. In Eve V. Clark (Ed.), *Proceedings of the 30th Stanford Child Language Research Forum* (pp. 77-86). CSLI: Stanford.
- Slobin, Daniel I. (1985). Crosslinguistic evidence for the language-making capacity. In Daniel I. Slobin (Ed.), *The crosslinguistic study of language acquisition, Vol. 2: Theoretical issues* (pp. 1157-1256). Hillsdale, NJ: Erlbaum.
- Slobin, Daniel I. (1987). Thinking for speaking. *Proceedings of the 13th Annual Meeting of the Berkeley Linguistics Society* (pp. 435-445). Berkeley, CA: Berkeley Linguistics Society.
- Talmy, Leonard. (1985). Lexicalization patterns: Semantic structure in lexical forms. In Timothy Shopen (Ed.), *Language typology and syntactic description. Vol. III: Grammatical categories and the lexicon* (pp. 57-149). Cambridge: Cambridge University Press.