

# Speech-Gesture Relationship across Languages and in Second Language Learners: Implications for Spatial Thinking and Speaking

Aslı Özyürek  
Koç University

## 1. Introduction

Speakers frequently use hand gestures accompanying their speech while they speak. For example, a speaker can raise her hand in a climbing manner as she says *the cat climbed up the tree* in her speech. Some of these gestures are called iconic gestures, that is, they resemble the events they represent. This paper investigates the relationship between these iconic gestures and the content of the information speakers express in their speech. The nature of this relationship reveals the relations between speakers' spatial thinking and speaking during the speech production process.

Two opposing views have been proposed in the literature to account for this relationship. According to one view (Independence Hypothesis) gestures and speech are autonomous and separate systems. Gestures originate out of the spatial representations in working memory and independent of verbal representations during speech production. (e.g., Butterworth & Beattie, 1978; Krauss et al. 1996, Feyereisen & de Lannoy, 1991). However according to another view (Interface Hypothesis), speech and gesture are interdependent systems and gestures originate from an interface representation in which spatial and verbal representations interact (Kita, 2000; McNeill, 1985, 1992; Kita & Özyürek, under review; Özyürek & Kita, 1999).

One way to empirically test these two hypotheses is to compare representations in iconic gestures across typologically different languages that are about the same events. If speakers change their iconic gestures when they use different lexical and syntactic packaging of spatial information even though they depict the same events, this provides evidence for the view that there is interface between spatial and verbal representations through which gestures originate.

Two studies are conducted to test these hypotheses. Study 1 compares speakers' gestures in typologically two different languages, English and Turkish, as they talk about the same event. English and Turkish differ from each other in the way they lexicalize manner and path elements of the motion events.

Study 2 compares gestures of the same speakers as they learn a second language at different proficiency levels, that is, compares gestures of Turkish speakers using English as a second language. Independence Hypothesis would

predict that in both studies gestures would not differ when speakers use typologically different languages since in both cases gestures would originate from the same spatial representations and independent of the verbal representations. However Interface Hypothesis would predict that gestures differ in ways similar to the differences in lexical and syntactic packaging of spatial information since there would be interaction between the spatial and the verbal representations during speech production process.

## 2. Study 1: Gestures across typologically different languages

With regard to expressing motion events, such as describing a ball rolling down a hill, English and Turkish differs from each other in the way they lexicalize manner and path elements of a motion event. While English speakers use a verb and satellite construction that allows them to express both elements within one verbal clause, Turkish speakers have to use two verbs that can be connected (with connectors such as *-arak*; *-ip*) in a main and subordinate clause matrix.

### English:

- (1)      *rolls*      *down*  
           V            satellite  
           manner    path

### Turkish:

- (2)      *yuvarlan-arak*    *iniyor*  
           V-roll-Conn    V-descend  
           manner            path

Given these differences in lexicalization patterns between the two languages, the prediction with regard to gestures is that if there is interaction between gestures and speech, one could expect gestures to differ in similar ways to the differences to the lexicalization patterns. For example it is possible that Turkish speakers use separate gestures for manner and path as they use separate verbs to express both elements. However, English speakers might be able to express both manner and path within one gesture since they can express both elements within one verbal clause. On the other hand if there is no interaction between the two systems then one would expect gestures to be same regardless of the language spoken. In order test these hypotheses the following study was conducted. In this study Turkish and English speaking subjects were shown the same movie and asked to describe it to another native speaker and their speech and gestures were analyzed.

## 2.1. Sample and procedure

14 American English, 16 Turkish native speakers participated in this study. All the data was collected from monolingual speakers in Chicago and Istanbul. Each subject is asked to see an animated cartoon 'Canary Row' (8 minutes) and to narrate the cartoon story to an addressee who has not seen it. The narratives are videotaped.

## 2.2. Coding

One scene from the cartoon was selected for detailed analysis. In this scene Sylvester swallows a bowling ball and rolls down the street with this ball in his stomach. The aim of picking this scene is that both manner (rolling) and path (going down) are represented simultaneously. The part of narratives that refers to this event is analyzed. Gestures and speech that accompany the Rolling Event description are categorized into three types: a) Trajectory-only, b) Manner-only, and c) Manner-Trajectory Conflating.

In speech, Trajectory-Only and Manner-only refer to use of separate verbs for manner and path (e.g., *He is rolling; He goes down the street* or *He descends while rolling*), whereas Manner-Trajectory Conflating refers to conflating manner and path within one clause (e.g., *he rolls down*). In gesture, a Manner-only gesture represents the circular nature of the rolling and/or the repetitive aspect of rolling (e.g., a repetitive up and down movement of the hand), without representing change of location of the moving entity. A Trajectory-only gesture represents change of location without any manner representation. In a Manner-Trajectory Conflating gesture, the representation of trajectory and manner are superimposed (e.g., a hand sweeping horizontally, as it makes a small repetitive up and down movement).

## 3. Results

As predicted by the typological differences, all English speakers used one verbal clause (e.g., *he rolls down*) to express both the manner and path in the rolling event, whereas all Turkish speakers used two verbs (e.g., *he rolls and goes down the street*). In the English sample three speakers also used a manner-only (e.g., *he is rolling*) and one speaker used a trajectory-only speech (e.g., *he goes down*) in addition to their manner-trajectory conflating constructions.

The patterns of gestures also paralleled the differences in speech across languages. Percentage of subjects who used Manner-only, Trajectory-only and Manner-Trajectory Conflation gestures at least once in their description of the scene were calculated.

The first analysis showed that more Turkish speakers than English speakers used Manner-only gestures at least once in their to descriptions of the scene (Table 1).

**Table 1: Percentage of subjects who used Manner-only gestures at least once in their repertoire of gestures**

Language	Used at least once	Never used
English (N=14)	14%	86%
Turkish (N=16)	38%	62%

Furthermore it was found that more Turkish speakers than English speakers used Trajectory-only gestures to describe the scene (Table 2).

**Table 2: Percentage of subjects who used Trajectory-only gestures at least once in their repertoire of gestures**

Language	Used at least once	Never used
English (N=14)	35%	65%
Turkish (N=16)	75%	25%

However, there was no difference between the number of Turkish and English speakers who used Manner-Trajectory conflated gestures in their descriptions.

**Table 3: Percentage of subjects who used Manner-Trajectory conflation gestures at least once in their repertoire of gestures**

Language	Used at least once	Never used
English (N=14)	71%	29%
Turkish (N=16)	69%	31%

#### 4. Conclusion:

The findings show that speakers' gestures about the same motion events change with different lexical and grammatical encoding of information in different languages while also retaining some spatial elements of the events as observed. While more Turkish speakers used Manner-only and Trajectory-only gestures than English speakers, similar percentage of speakers used Manner-Trajectory Conflation gestures in the two languages.

These findings show that gestures are influenced both by the spatial representations in working memory independent of verbal representations (as seen by manner-trajectory conflated gestures used similarly across speakers of both languages) and also by the verbal representations used specifically for the type of language used (as seen by manner-only and trajectory-only gestures used

more frequently by Turkish than English speakers). Thus these findings support the Interface Hypothesis according to which there is interaction between speech and gestures that gestures are generated from an interface mechanism between spatial and verbal representations.

### 5. Study 2: Gestures in different levels of L2 proficiency

Even though speakers from typologically different languages use different types of gestures, this might be also due to cultural factors shaping gestural patterns rather than due to the interaction between spatial and verbal representations. For example Muller (2001) has shown that Spanish speakers use a larger gesture-space than German speakers in natural conversations. Thus the difference in the Manner- only and Trajectory-only gestures might be due to cultural differences between Turkish and American subjects independent of the language type, that is, does not reflect the influence of language specific verbal representations. However, if one can show that the same speakers change their gestures as they switch to another language as a function of their proficiency in their second language, then this would show that the iconic gestures differ due to dynamic interaction between verbal and spatial representations rather than due to cultural factors. Thus in Study 2 the relationship between speech and gesture is investigated by looking at whether and how the same speakers change their gestures as they use a typologically different language as their second language. For this purpose gestures of Turkish speakers using English as their L2 at different proficiency levels are investigated.

If it is the cultural patterns that shape gestural patterns, one would expect Turkish speakers not to change their gestures at all when they speak English at different proficiency levels. This would be in line with the Independence Hypothesis in that it would suggest that gestures and speech are independent systems. Another prediction in line with Independence Hypothesis would be that gestures compensate for the lack of proficiency in L2. In fact previous research by Gullberg (1998) has shown that Swedish speakers learning French use abstract deictic gestures to locate referents in the gesture space more frequently when they speak French than when they speak Swedish in cases where they can not use nominal pronouns but use only full NPs in French. That is, abstract gestures compensate for the lack of proficiency in using pronouns by Swedish speakers in speaking French. However, Interface Hypothesis would predict that speakers' gestures would change when they speak L2 and changes in gestures would parallel the lexicalization patterns in L2 speech that can be expressed at each level of proficiency rather than compensate for the lack of grammatical competency.

In order to test these hypotheses the same cartoon used in Study 1 were shown to Turkish speakers who spoke English as L2 at different levels of

proficiency. Later they were asked to narrate the story once in Turkish and once in English to different listeners. Their gestures that accompany their English descriptions were analyzed.

### 3.1. Sample and procedure

18 native Turkish speakers who are at different levels of English proficiency participated in this study. Among these 6 were at Beginner (Mean age =20,3), 6 were at Intermediate (Mean age= 22,2) and 6 were at Advanced levels (Mean age= 35,4). Beginner group has started learning English when they were 19 years old and has attended intensive English education classes for a year. Intermediate level group has been educated in English speaking schools in Turkey since they were 11 and were at 3<sup>rd</sup> year of university education at the time of testing. However, they have never lived in an English speaking country before. The Advanced group has also been educated in English speaking schools since they were 11 years old and has spent the last 10 years of their life in the US before they returned back to Turkey. Beginner and Intermediate group consisted of students at Koc University in Istanbul where the medium of education was in English. Subjects at the Advanced group were lecturers at the same university.

Data collection took place at Koc University, Istanbul. Each subject told the same cartoon, "Canary Row " used in the Study 1. Afterwards they told it once in English and once in Turkish to two different Turkish listeners. Their order of narration was counter balanced across subjects. Their narratives were videotaped.

### 3.2. Coding

The speech and gestures used to depict the Rolling Event in English were analyzed and coded using the same criteria as in Study 1.

### 3.4. Results

For the first analysis the percentage of speakers who used Manner-only speech and gestures in their descriptions at least once were calculated (Table 5).

**Table 5: Percentage of Turkish speakers who used Manner-only descriptions in their speech and gesture at different proficiency levels in English**

Level of proficiency	Gesture		Speech	
	Used at least once	Never used	Used at least once	Never used
Beginner (N=6)	55%	60%	55%	45%
Intermediate (N=6)	100%	0	80%	20%
Advanced (N=6)	30%	70%	0	100%

The results first of all showed a parallel trend in speech and gestures at different proficiency levels. More speakers at the Intermediate level used Manner-only speech (e.g., *he is rolling*) than at the Beginner level in their descriptions. However, this level dropped to zero at the Advanced level since speakers could use Manner-Trajectory Conflating constructions (e.g., *he rolled down*) in similar ways that are used by native English speakers. The fact that Beginner Level speakers did not use Manner-only speech as much can be explained by the fact that they did not know enough manner verbs. At the Intermediate level they have learned the manner verbs but expressed it as separate verbs and did not yet combine it with satellite particles (i.e., *down*). A similar pattern was observed with regard to gestures since more speakers at the Intermediate level used Manner only gestures than speakers at Beginner level. However, this percentage dropped to 30% at the Advanced level paralleling the drop in speech

In the second analysis the percentage of speakers who used Trajectory-only speech and gestures in their descriptions at least once were calculated (Table 6).

**Table 6: Percentage of Turkish speakers who used Trajectory-only descriptions in their speech and gesture at different proficiency levels in English**

Level of proficiency	Gesture		Speech	
	Used at least once	Never used	Used at least once	Never used
Beginner (N=6)	100%	0	80%	20%
Intermediate (N=6)	60%	40%	60%	40%
Advanced (N=6)	50%	50%	30%	70%

The results again showed similar patterns between speech and gestures. There was a decrease in the percentage speakers who used Trajectory-only speech (eg., *he is going down the street*) as speakers became more Advanced. In the Advanced level speakers expressed path elements in satellite constructions while they preferred to do so in separate verbs at the Beginner and Intermediate levels. Similarly the percentage of speakers who used Manner only gestures also decreased as they became more advanced in English.

Lastly, the percentage of speakers who used Manner-Trajectory Conflating speech and gestures in their descriptions at least once were calculated (Table 7).

**Table 7: Percentage of Turkish speakers who used Manner-Trajectory descriptions in their speech and gesture at different proficiency levels in English**

Level of proficiency	Gesture		Speech	
	Used at least once	Never used	Used at least once	Never used
Beginner (N=6)	10%	90%	0%	100%
Intermediate (N=6)	0	100%	0%	100%
Advanced (N=6)	50%	50%	80%	20%

This analysis shows that speakers began using conflated constructions (e.g., *he rolls down*) in speech mainly at the Advanced level but this was rarely the case at the Beginner or Intermediate level. Similarly speakers used Manner-Trajectory conflating gestures only at the Advanced levels.

### 3.5. Conclusion

Study 2 shows that the same speakers change their gestures when they speak a second language and these changes parallel how information is expressed grammatically at each level of L2 proficiency. These findings show first of all that when speakers' gestures change in different languages as shown in Study 1, this is not due to cultural factors. They also show that In L2 learning, gestures do not only reveal compensation strategies of speakers (Gullberg, 1998) but also interactions between spatial and linguistic representations of speakers at different proficiency levels in the target language. It is important to note here that the results of this study do not suggest that speakers never use compensation strategies in L2. The compensation gestures observed Gullberg's study might be those types of gestures that are "meant to be seen" for the addressees in cases



of incompetency in L2 language, such as lexical access. That is, these gestures might be generated specifically to be seen by the addressees in communication and thus reveal a different type of conceptualisation during speaking. Further study is needed to pull out the effects of different influences on gestural representations. If this is true it would suggest that there is another influence on the representations that gestures originate from, that is, the influence of environment (Özyürek (in press); see Kita and Özyürek, under review for different types of influences on gestural representations). However, I argue that the manner and path gestures investigated in this study reveal the interaction of representations between spatial and verbal representations in the production of L2 speech.

Finally the speech results reveal that the verb and satellite constructions used dominantly by native speakers of English is a hard construction to master for Turkish speakers in L2 and needs years of practice in a country where the L2 is spoken. That is, typologically distinct and different constructions across languages are hard to learn in L2.

## 6. General Conclusion

The results of both studies in this paper show that speakers convey information in their gestures in ways similar to the syntactic-lexical encoding of information of the language they speak. This effect is demonstrated both across languages and during second language acquisition. These findings support Interface Hypothesis according to which gestures and speech are interdependent systems and gestures are generated from an interface mechanism between spatial thinking and speaking (McNeill, 1985;1992; Kita, 2000; Kita & Özyürek (under review); Özyürek, (in press)).

## References

- Butterworth, B. & Beattie, G. (1978). Gesture and silence as indicators of planning in speech. In Campbell & Smith (eds.). *Recent advances in the psychology of language*. pp. 347-360. New York: Plenum.
- Feyereisen, P. & de Lannoy, J. D. (1991). *Gestures and Speech: Psychological Investigations*. Cambridge: University of Cambridge Press.
- Gullberg, M. (1998). *Gesture as a communication strategy in second language discourse*, PhD. Dissertation, Lund, Lund University Press.
- Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.), *Language and gesture* (pp.162-185) Cambridge: Cambridge University Press.
- Kita, S. & Özyürek, A. (under review). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*.
- Krauss, R. (1998). Why do we gesture when we speak? *Current Directions in Psychological Science*, 7-2, 54-60.
- McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review*, 92-3, 350-371.

- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.
- Muller, C. (2001). Gesture-space and culture. In C. Cave, I. Guaitella et S. Santi (eds.), *Oralite et Gestualite: Interactions comportements multimodaux dans la communication* (pp.565-571). Paris: L'Harmattan.
- Özyürek, A. (in press). Do speakers design their co-speech gestures for their addressees? The effects of addressee location on representational gestures *Journal of Memory and Language*.
- Özyürek, A. & Kita, S. (1999). Expressing manner and path in English and Turkish: Differences in speech, gesture, and conceptualization. In M. Hahn & S. C. Stonnes (eds.), *Proceedings of the Twenty-first Annual Meeting of the Cognitive Science Society* (pp. 507-512). Amsterdam: Erlbaum Publishers.