Title

Multimodal Description of Instrument Events in Turkish and English

Permalink

https://escholarship.org/uc/item/31h4s3qp

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors

Moser, Christiana
Tarakçı, Bahar
Ünal, Ercenur
et al.

Publication Date

2024

Peer reviewed

Multimodal Description of Instrument Events in Turkish and English

Christiana Moser (christiana.moser@mail.utoronto.ca)[1]
Bahar Tarakcı (bahar.tarakci@ozyegin.edu.tr)[2]
Ercenur Ünal (ercenur.unal@ozyegin.edu.tr)[2]
Myrto Grigoroglou (m.grigoroglou@utoronto.ca)[1]
Department of Linguistics, University of Toronto[1] ; Department of Psychology, Özyeğin University[2]

Abstract

Daily experiences are conceptualized as events involving multiple participants and their relations (i.e., thematic roles). When describing events, speakers often do not include all event participants involved. Here, we explore how underlying conceptual requirements and language-specific encoding options influence the content of event descriptions in speech and gesture in two typologically different languages (English, Turkish). Focusing on conceptually peripheral instruments whose status is highly debated, we manipulated the conceptual status of event participants by including events that 'require' or 'allow' otherwise syntactically optional instruments. Results showed that the require-allow distinction did not manifest uniformly in Turkish and English in speech, gesture, or when both modalities were considered. However, mention of highly optional event participants (e.g., allowed instruments) was affected by language-specific syntactic encoding options. We conclude that, under more naturalistic elicitation conditions, planning descriptions of instrument events is more heavily affected by language-specific encoding than conceptual prominence of the roles.

Keywords: events; language production; gesture; argument structure; thematic roles; instruments; verbal semantics

Introduction

We make sense of the world by organizing our daily experiences in terms of structured representations referred to as *events,* which unfold in time and space and involve multiple participants (e.g., objects, people) and the relations between them. In linguistic theory, event participants are known as *thematic roles*, which express the semantic relationships between verbs and their syntactic constituents (e.g., Baker, 1997; Fillmore, 1968; Jackendoff, 1990) and serve as representational units for the organization of preverbal event conceptualizations (Levelt, 1989; cf. also Dowty, 1991; Jackendoff, 1990; Levin & Rappaport-Hovav, 2005; Pinker, 1989, a.o.). A central question for theories of language production is how such preverbal event representations map onto a linguistic message. The answer is not straightforward, as languages differ greatly in terms of how they encode events, and theoretical models of language production strongly support the possibility that conceptual units recruited for the purposes of language production are guided by language-specific encoding constraints (Levelt, 1989). Furthermore, language is a multimodal phenomenon, and speakers often use gesture along with speech to express information about event components (McNeill, 2005) in ways that align with language-specific encoding options

available in speech (Kita & Özyürek, 2003; Ünal et al., 2022). Finally, language production processes do not occur in a vacuum but are inherently embedded into communicative situations, which, although frequently ignored by language production research (Meyer et al., 2019), are shown to affect how event components are expressed in speech (e.g., see Grigoroglou & Papafragou, 2019a/b; Lockridge & Brennan, 2002). The goal of the current study is to investigate how underlying conceptual event representations map onto multimodal linguistic descriptions, elicited by speakers of two typologically distinct languages (English, Turkish), under more naturalistic experimental conditions.

Thematic Roles and Conceptual Prominence

Thematic structure plays a critical role in the organization of preverbal conceptualizations of events. For instance, an event where a woman is stirring soup with a ladle is organized in terms of three thematic roles: an Agent (woman), a Patient (soup), and an Instrument (ladle). However, thematic roles vary in terms of their importance to event structure and can be ranked in terms of a thematic prominence hierarchy (AGENTS > PATIENTS > GOALS/RECIPIENTS > INSTRUMENTS; Jackendoff, 1990; Baker, 1997). Thematic roles that are higher in the hierarchy (e.g., AGENTS) are central components of the event, and are more likely to be encoded as core syntactic arguments than thematic roles lower in the hierarchy (e.g., INSTRUMENTS), which tend to be syntactically more peripheral and can be omitted in speech without causing the sentence to be ungrammatical (Baker, 1997; Fillmore, 1968; Jackendoff, 1990; Rappaport-Hovav & Levin, 1998). Further, thematic roles with greater conceptual prominence tend to map onto syntactic positions with greater grammatical prominence (e.g., AGENTS tend to map onto subject NPs; Do & Kaiser, 2022; Ferreira, 1994).

Interestingly, however, more peripheral event roles such as instruments are not always tangential to event structure. In fact, semantic analyses find that verbs can be classified as conceptually 'requiring' or 'allowing' event participants that are otherwise syntactically optional (Levin, 1993; Koenig et al., 2003; Barbu & Toivonen, 2016; Barbu, 2020; Rissman, 2022; Rissman et al., 2015). For example, verbs that describe instrument events can require an Instrument (e.g., *stir* in 1a) or allow for one but not require it (e.g., *drink* in 1b).

(1) a. The woman *stirred* the soup <u>with a ladle.</u>
   b. The woman *drank* the milkshake <u>with a straw.</u>

These theoretical intuitions have been complemented with experimental evidence showing that participants (typically speakers of English) find required instruments to be more argument-like (thus, more important for conceptual event structure) than allowed instruments (Rissman et al., 2015; 2019; Barbu & Toivonen, 2016; Barbu, 2020). For instance, in self-paced reading tasks, participants were faster to recognize the involvement of an instrument when sentences included require-instrument verbs than allow-instrument verbs (Koenig et al., 2003; Koenig et al., 2008). Similarly, in grammaticality judgment and sentence completion tasks, participants were more likely to judge instruments as arguments when the verb was of the require type than the allow type (Rissman et al., 2015). More recently, these results were confirmed with a broader cross-linguistic sample of speakers of Spanish and Mandarin Chinese (Rissman et al., 2019), as well as Romanian, Spanish, and Turkish (Barbu, 2020). This evidence suggests that speakers have gradient event representations, but how such representations map onto linguistic messages remains unclear.

Cross-linguistic Event Description in Speech

Mapping non-linguistic event representations onto language requires the speaker to place event participants into specific syntactic positions (Levelt, 1989). However, languages vary in terms of how they encode events. For instance, languages differ in terms of the event components that are structurally required in an event description, and in the syntactic options available to encode certain event roles. How such language-specific encoding patterns affect event descriptions across languages has been extensively studied for motion events (e.g., Furman et al., 2014; Gennari, 2002; Kita & Özyürek, 2003; Papafragou et al., 2006; Ünal et al., 2021, a.o.), but their contribution to other events is relatively understudied. Here, we focus on the cross-linguistic expression of instrument events in English and Turkish, two typologically distinct languages with several structural differences.

One difference between English and Turkish concerns the obligatoriness of event components in the syntactic structure, with English requiring arguments to be mentioned overtly and Turkish permitting null arguments. More specifically, Turkish allows for subject drop (licensed by agreement morphology on the verb), as well as direct/oblique object drop (as long as there is a discourse accessible antecedent for co-reference); see (2) (Gürcanlı et al., 2007; Öztürk, 2005). Although, in both Turkish and English, instruments tend to be syntactically optional and can be omitted freely (Russo, 2021), it is possible that the tendency for argument drop in Turkish affects the mention of instruments as well.

(2) (Kadın)   (çorbayı)   (kepçe ile) karıştırdı
    (woman) (soup-ACC) (ladle with) stir-PST.3SG
    '(The woman) stirred (the soup) (with the ladle)'

A second difference between the two languages concerns the availability of encoding options for instruments. In English, instruments can be encoded in prepositional phrases (3a) or incorporated into the verb (3b).

(3) a. The woman stirred the soup with the ladle *(PP)*
    b. The woman whisked the eggs *(V)*

In Turkish, instruments can be encoded as case-marked noun phrases (4a), postpositional phrases (4b), or verbs (4c).

(4) a. Kadın çorbayı   kepçeyle karıştırdı *(DP)*
       woman soup-ACC ladle-COM stir-PST.3SG
       'The woman stirred the soup with the ladle'

    b. Kadın çorbayı   kepçe ile   karıştırdı *(PP)*
       woman soup-ACC ladle with stir-PST.3SG
       'The woman stirred the soup with the ladle'

    c. Yumurtaları   çırptı   *(V)*
       egg-PL-ACC   whisk-PST.3SG
       '(One) whisked the eggs'

These differences in encoding options are relevant to the issue at hand in that variability in encoding options can be associated with conceptual peripherality (Rissman & Majid, 2019; Ünal et al., 2021). As there are three regular options for encoding instruments within a single clause in Turkish (DP, PP, and V) but only two in English (PP and V)[1], this raises the possibility that instruments are more syntactically or conceptually peripheral in Turkish than in English. However, the "additional option" available in Turkish is a case-marked DP option, which is both syntactically and conceptually less peripheral than the PP option, as it is literally structurally closer to the verb (Baker, 1997; Brown & Dell, 1987). Given the tight link between the conceptual prominence of a thematic role and the grammatical prominence of its syntactic realization (Do & Kaiser, 2022; Ferreira, 1994), it is an open question how these differences in encoding options affect mention of instruments in event descriptions across languages.

Cross-linguistic Event Description in Gesture

When describing an event, viewers do not exclusively rely on the modality of speech. Gesture is often used to reinforce mentions that occur in speech, or to express information not mentioned in speech (McNeill, 2005). For instance, children may use gesture to encode information that is otherwise "missing" in their descriptions of motion events (Furman et al., 2014). As event components may be mentioned only in gesture, considering both modalities of speech and gesture allows for a more complete estimate of an event representation when examining event descriptions (Ünal et al., 2022, 2023).

---

[1] Independent instrument DPs are possible in English for exceptional verbs (as in X draws chalk on the wall), when a PATIENT is in an embedded PP (X takes a knife to a box), or when the instrument is in a subordinate VP (using a knife).

Further, available evidence from cross-linguistic expression of motion suggests that gesture represents information that matches language-specific encoding patterns in speech. For instance, English speakers, who tend to express path and manner within the same clause in speech (i.e., manner in the main verb and path in a PP), tend to conflate path and manner into a single gesture, while Turkish and Japanese speakers who tend to use separate clauses to express path and manner in speech (i.e., path in the main verb and manner in a separate verb phrase), typically produce separate gestures for manner and path (Kita & Özyürek, 2003; Özyürek et al., 2008; Ünal et al., 2022; a.o.). Even though the variation in motion event expressions in gesture is typically driven by language-specificity, it is also possible to observe within-language variation driven by factors other than language, such as conceptual event structure (Ünal et al., 2023). Importantly, peripheral components such as manner of motion are more susceptible to such influences than core components of a motion event such as path.

Although these findings are robust for motion events, it remains an open question whether language-specific patterns and within-language variation surface in gesture for other types of events, such as instrument events. Given that instruments come with certain affordances for representation in gesture (as nouns with concrete referents that are often manipulable and hand-held; van Nispen et al., 2017), instrument events are a particularly interesting test bed for the relation between event conceptualization and message formulation across languages.

Present Study

In the present study, we investigate how adult speakers of English and Turkish describe instrument events in speech and gesture. For each event, Instruments were either semantically required (e.g., ...stirred the soup _with a ladle_) or allowed (e.g., ...drank the milkshake _with a straw_). Unlike prior research that examined the require-allow distinction mostly in metalinguistic tasks involving argument selection, sentence completion, and grammaticality judgments (Barbu & Toivonen, 2016; Rissman et al., 2015; Rissman et al., 2019), here we explore whether the require-allow distinction also manifests as an asymmetric preference for including conceptually required participants over conceptually allowed participants in free event descriptions. We hypothesized that (otherwise syntactically optional) instruments might be mentioned more frequently in speech and in gesture when they are semantically required than when they are allowed.

Unlike prior work on mention of instruments that has been conducted primarily with speakers of English or other typologically similar languages (e.g., German, Dutch; Rismann et al., 2022), here we elicit descriptions in two typologically distinct languages (English, Turkish) that differ in the availability of syntactic options to encode instruments (English typically encodes them in PPs, Turkish in both DPs and PPs) and the acceptability of argument omissions. We anticipate that these differences may affect how instruments

are expressed in each language in speech and co-speech gesture.

Finally, unlike prior research investigating event descriptions in a social vacuum, our task had a clear communicative motivation. Participants were asked to describe the events for a naïve familiar interlocutor (e.g., friend, relative) who did not have visual access to the events. Given the importance of listeners' communicative needs for the mention of optional event participants (e.g., Do et al., 2022; Grigoroglou & Papafragou, 2019a, 2019b; Lockridge & Brennan, 2002), we expected that this communicative setting would elicit more naturalistic responses that better represent speakers' linguistic behavior "in the wild".

Methods

Participants

We recruited adult participants who were native (L1) speakers of Turkish ($n = 30$, female = 27) and English ($n = 30$, female = 19). The L1 Turkish adults were Özyeğin University students, and the L1 English adults were University of Toronto students, all recruited through student participant pools in each University. Each adult participant was accompanied by a familiar listener who was a friend, sibling, parent, or roommate. In all cases, the listener was a naïve interlocutor who had no visual access to events.

Design and Materials

Stimuli consisted of 36 silent video clips depicting people participating in different events. Target events depicted instrument events (n = 12), of which half required an Instrument (e.g. ...stirred the soup with the ladle; Figure 1a), and the other half allowed an instrument (e.g., ...drank the milkshake with a straw; Figure 1b). Filler events depicted caused motion events (half goal-oriented; half source-oriented; n = 12) and possession transfer events (half required; half allowed; n = 12). Events were selected based on verb classifications (Levin, 1993), instrument argument-hood judgments from English speakers (Koenig et al., 2003; Rissman et al., 2015), and conversations with Turkish-speaking informants. Table 1 presents sentences that describe instrument events in our stimuli. All event videos lasted between 7 and 10 seconds. The videos were filmed in ordinary settings where the events could realistically take place, with a plain background. Two presentation lists were created to counterbalance event order.

Procedure

Each participant completed the study via Zoom with a live experimenter who navigated a PowerPoint slideshow. The speaker sat in front of the computer while the listener sat at an angle that ensured that she could not see the screen but still allowed them to see each other.

During practice trials, participants described two static images to their listeners. In the free description task, participants watched event videos and described what was
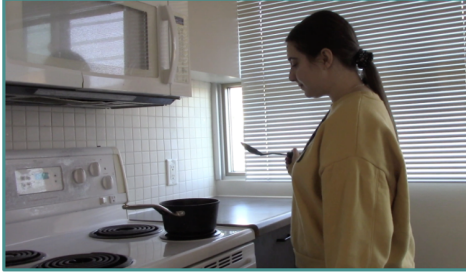
*Kadın çorbayı kepçeyle **karıştırdı.***
(a) 'The woman **stirred** the soup <u>with the ladle</u>'

*Kadın pipetle milkshake **içti.***
(b) 'The woman **drank** the milkshake <u>with a straw</u>'

Figure 1: Stills from the target video for two instrument events. (a) conceptually requires an instrument, (b) allows an instrument.

Table 1: List of instrument events used in the study.

Require Events	Allow Events
The girl **chops** the onion with a knife.	The girl **breaks** the vase with a hammer.
The girl **cuts** the fabric with scissors.	The girl **drinks** a milkshake with a straw.
The girl **draws** a flower with chalk.	The girl **eats** cake with a fork.
The girl **writes** in the notebook with a pen.	The girl **opens** the package with a knife.
The girl **stirs** the soup with a ladle.	The girl **scratches** the lottery card with a coin
The girl **beats** the egg with a whisk.	The girl **washes** the dishes with a sponge.

happening in each video to their listeners. If they described the first video with excessive peripheral detail (e.g., The woman wearing a yellow sweater is stirring soup with a spoon while standing in a white kitchen), the experimenter indicated that they could exclude details describing individuals and background, giving no other feedback. Participants' descriptions were video-recorded for transcription and coding. Sessions lasted around 30 minutes.

Coding

Multimodal event descriptions of target events were transcribed and coded for the presence of thematic roles in speech and gesture. Descriptions were coded by native speakers in each language, using ELAN (Version 6.6; Wittenburg et al., 2006). Event descriptions in speech were transcribed and segmented into clauses using conjunctions in English (e.g., *and, or*) and Turkish (*ve* 'and', *ama* 'but', *sonra* 'then'). Each clause was then coded for the explicit mention of Instruments. As the free description task allowed participants to describe the event in any number of clauses, for the present analysis, we consider only the clause that contains the main verb that described the event (e.g., *The woman stirs the soup <u>with the ladle</u>*). Instrument mentions were further coded for the syntactic structures in which they appeared. These included <u>independent noun phrases</u> (DPs, e.g. *Kadın çorbayı **kepçeyle** karıştırdı* 'A woman is stirring soup with a ladle'), <u>prepositional or postpositional phrases</u>

(PPs, e.g. *Kadın   çorbayı **kepçe ile** karıştırdı* 'A woman is stirring soup with a ladle'), or <u>conflated with a verb</u> (V, e.g., *yumurtalari **cirpiyor*** 'Someone is whisking eggs'). Although unexpected, there were some instrument mentions in English that were encoded as independent DPs. These included phrases such as *draws chalk on the wall* and *takes a knife to a box,* as well as embeddings in subordinate verb phrases such as *using X.*

In gesture, we first coded for the presence of Instrument gestures. Non-representational gestures that did not convey any meaning (e.g., beat gestures) were not coded. Instrument gestures were grouped into four categories by adapting the classification of iconic depictions developed by Ortega and Özyürek (2020) (see Figure 2): (1) *acting* gestures depicted how the instrument is being handled or manipulated, while the body represented the person handling/ manipulating it, (2) *representing* gestures depicted configurations recreating the form of an instrument, (3) gestures conflating *acting* and *representing* depicted how it would be handled or manipulated while the hand recreated the form of the instrument, and (4) *other* gestures. *Other* gestures (including pointing and drawing) were rare (occurring in 0.015% of mentions).

Results

Data were analyzed with generalized binomial linear mixed effects modeling with crossed random intercepts for Subjects and Items. The models were fit with the *lme4* package (version 1.1.17; Bates, et al., 2015) in *R* (version 4.0.3; R Core Team, 2020). Significance levels for pairwise comparisons with corrections for multiple comparisons were obtained with *emmeans* (version 1.5.5-1; Lenth, 2021) and *multcomp* (version 1.4-16; Hotorn et al., 2008) packages. Figures were produced using the ggplot2 package (version 3.2.1, Wickham, 2016).

Instrument Mention in Speech

First, we looked at instrument mentions in speech across languages. The data was analyzed with a model that included mention of instruments in speech as the binary-dependent variable (1 = mentioned, 0 = not mentioned). Figure 3a summarizes the data. The model included fixed effects of Verb Type (Require, Allow) and Language (EN, TR) and
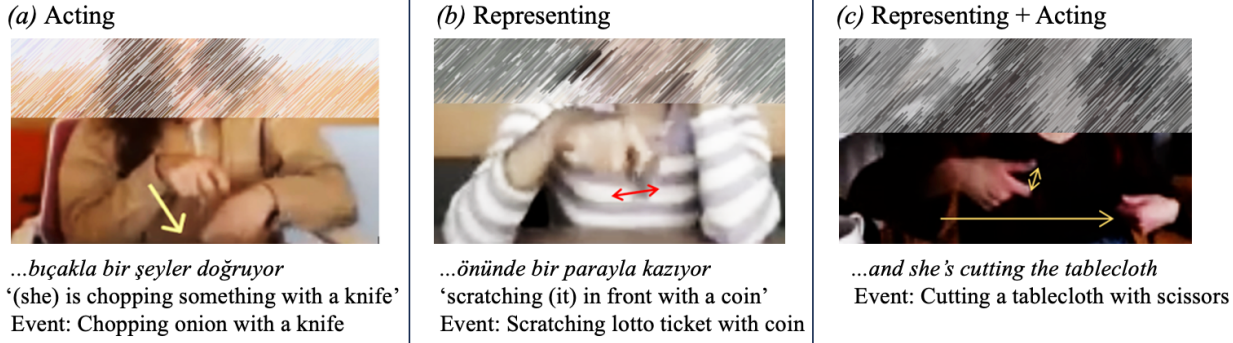
Figure 2: Examples of *(a)* acting, *(b)* representing, and *(c)* acting + representing gestures used to represent instruments.

their interaction, as well as random intercepts for Participants and Items. The fixed effects of Verb Type and Language were coded with centered contrasts (-.5, .5). The same coding strategy was followed in all subsequent analyses. The model revealed a significant interaction between Verb Type and Language (β = -0.803, SE = 0.376, z = -2.100, p = .033). Follow-up analyses showed that the interaction was due to the fact that speakers of Turkish were more likely to mention allowed instruments than speakers of English (β = -0.792, SE = 0.358, z = -2.211, p = .027; $M_{Turkish}$= 0.52, $M_{English}$= 0.37), but there was no such difference for required instruments (β = -0.786, SE = 0.366, z = 0.028, p = .98, n.s.; $M_{Turkish}$= 0.44, $M_{English}$= 0.46). No other effects were significant.

For completeness, we also looked at the syntactic structures in which instruments appeared in each language (Table 2). Observation of the data largely confirmed our predictions about the syntactic means used in each language to describe instruments, but also revealed some unexpected patterns. English speakers used primarily PPs to encode instruments, as well as some incorporations of the instrument into the verb. Surprisingly, English speakers also encoded instruments in DPs (i.e., as direct object of the verb, e.g., *draws chalk on the wall*; see also Coding).

As expected, Turkish speakers encoded instruments primarily in case-marked DPs, as well as a few verb incorporations. Although available as a syntactic option, in our data, Turkish speakers encoded instruments in PPs very rarely. Interestingly, in both languages, the verb-

Table 2: Percentage of instrument mention types (speech)

Mention Type	English		Turkish	
	Allow	Require	Allow	Require
PP	31.6	25.4	3.6	0.6
DP	5.7	4.0	48.5	34.1
V	0.6	16.9	0.0	10.6
No Mention	62.2	53.7	47.9	54.7

incorporation option was almost exclusively used for require-instrument events. This finding may be a potential reflex of the require-allow distinction: only conceptually prominent required instruments can be syntactically encoded through verb incorporation.

Instrument Mention in Gesture

Next, we examined whether instruments were encoded in gesture. We tested a model that included Instrument gestures (1 = represented in gesture, 0 = not represented in gesture) within the main clause of the description at the item level. The model included fixed effects of Verb Type (Require, Allow), Language (EN, TR), and their interaction, as well as random intercepts for Participants and Items. Data are summarized in Figure 3b. The analysis yielded a significant effect of Language (β = 1.431, SE = 0.641, z = 2.231, p =
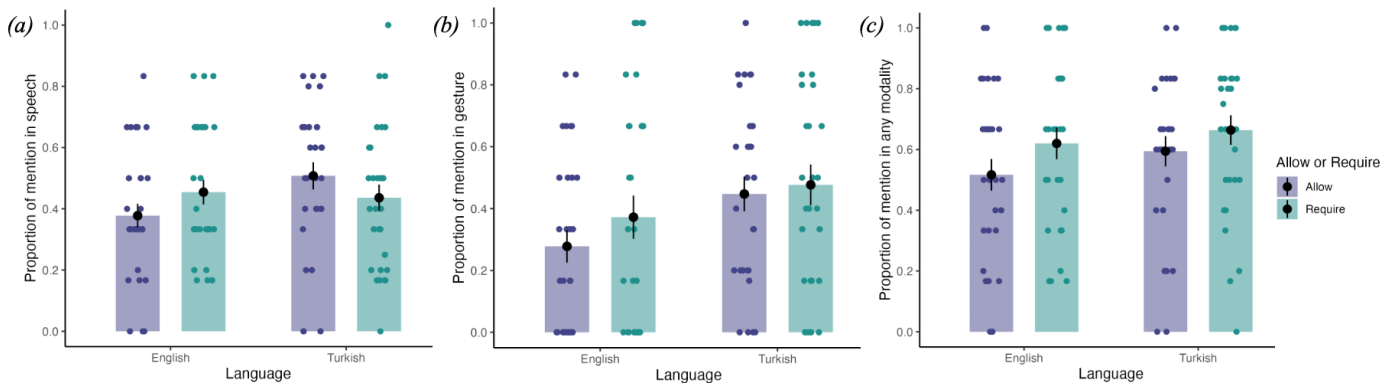

Figure 3: Proportion of instrument mentions in *(a)* speech, *(b)* gesture, and *(c)* any modality, across verb types and language groups. Error bars indicate standard error of participant means. Dots represent participant means.

.026), with Turkish speakers using, overall, more instrument gestures than speakers of English ($M_{Turkish} = 0.47$, $M_{English} = 0.33$). No other effects or interactions were significant.

Instrument Mention across Modalities

Finally, to determine whether there was an impact of the require-allow distinction on instrument mention regardless of modality, we investigated whether instruments are mentioned in either modality. Data are summarized in Figure 3c and Table 3. The model of instrument mention across modalities included fixed effects of Verb Type (Require, Allow), Language (English, Turkish), and their interaction, as well as random intercepts for Participants and Items. No effects or interactions were significant.

Table 3. Percentage of instrument mention across modalities

Mention Modality	English		Turkish	
	Allow	Require	Allow	Require
Speech Only	23.2	24.9	15.2	18.9
Speech & Gest.	14.7	20.9	36.0	25.4
Gesture Only	13.6	16.4	9.8	22.5
No mention	48.6	37.9	39.0	33.1

Discussion

When describing events, people must make choices about what to include in their descriptions. Here, we asked how these choices are shaped by the underlying conceptual requirements (i.e., the require-allow distinction) and language-specific encoding options. We investigated how adult speakers of English and Turkish use speech and gesture to describe instrument events, an event category typically not explored cross-linguistically. Using instrument events as a test case, we examined whether the require-allow distinction manifests as an asymmetric preference for including semantically required participants over semantically allowed participants in free multimodal event descriptions.

Contrary to our main prediction, we found that speakers of English and Turkish mentioned instruments equally frequently, independently of whether they were conceptually required or allowed, in both speech and gesture. However, speakers of English and Turkish presented several differences in their encodings of instruments, in line with the available encoding options in each language. In speech, speakers of English primarily used prepositional phrases to encode instruments, while speakers of Turkish primarily used case-marked noun phrases to encode instruments. Furthermore, speakers of Turkish were more likely to gesture about instruments than speakers of English, thus confirming prior findings showing differences in use of gesture across languages (Azar et al., 2020).

Although it is not clear why the require-allow distinction, well-documented in speakers argumenthood judgments, did not surface in speakers' event descriptions, it is possible that

the preverbal event conceptualization for instrument events may not arise linguistically as straightforwardly as one might expect. In fact, despite our predictions that conceptually more prominent components should be more likely to be mentioned, the conceptual prominence of required instruments may have made them less likely to be encoded in linguistic descriptions of events than the highly peripheral allowed instruments. For instance, it is possible that speakers in our sample omitted required instruments because they thought these would be highly predictable for their listener (e.g., since *stirring* can only occur using an instrument, the instrument can be omitted as it is already highly accessible to the listener). By contrast, speakers may have considered allowed instruments as more important to mention, as their listener could typically not infer that an instrument was used for the performance of this event (e.g., drinking may not require an instrument, thus mentioning the instrument offers information that the listener would not be able to infer). This line of reasoning is in accordance with research showing that the choice to mention (or omit) instruments is guided by pragmatic factors, having to do with their degree of inferability in the event and the informational needs of the listener (Brown & Dell, 1987; Grigoroglou & Papafragou, 2019a, 2019b; Lockridge & Brennan, 2002).

Even though we found no independent effect of the require-allow distinction within each language, the require-allow distinction manifested in findings in other ways. First, we found that speakers of both English and Turkish were likely to incorporate instruments into verbs for events that require an instrument, but they very rarely – if at all – used verb-incorporation for allowed instruments. This finding is indicative of how instruments are more integral components of require-instrument events and are potentially generated closer to the verb in preverbal event conceptualizations for the purposes of language production.

Second, we found that the require-allow distinction interacted with language-specific constraints in rather unpredictable ways. Specifically, we found that Turkish speakers were more likely to mention allowed instruments in speech than English speakers, but no such difference arose for required instruments. Although it is not entirely clear why this higher rate of instrument mention in Turkish concerned only allowed but not required instruments, it is possible that the DP encoding option in Turkish may have disproportionately increased the mention of allowed instruments in speech within the same clause as the main event verb.

In sum, under the conditions of a more naturalistic free description task, we found that the conceptual require-allow distinction did not manifest uniformly across two typologically distinct languages, or across the modalities of speech and gesture. We conclude that planning descriptions of instrument events is more heavily affected by language-specific encoding options and communicative demands than the conceptual prominence of the roles.

References

Azar, Z., Özyürek, A., & Backus, A. (2020). Turkish-Dutch bilinguals maintain language- specific reference tracking strategies in elicited narratives. *International Journal of Bilingualism*, *24*(2), 376–409.

Baker, M. (1997). Thematic roles and syntactic structure. *Elements of Grammar*, 73–137.

Barbu, R.M., & Toivonen, I. (2016). Event participants and linguistic arguments. *Proceedings of the 38th Annual Meeting of the Cognitive Science Society*, 1961-1966.

Barbu, R.M. (2020). *On the psycholinguistics of argumenthood.* [Doctoral thesis]. Carleton University.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). lme4: Linear Mixed-Effects Models Using Eigen and S4. R package version 1.1-10.

Brown, P., & Dell, G. (1987). Adapting production to comprehension: The explicit mention of instruments. *Cognitive Psychology, 19,* 441–472.

Bunger, A., Papafragou, A., & Trueswell, J. C. (2013). Event structure influences language production: Evidence from structural priming in motion event description. *Journal of Memory and Language*, *69*(3), 299–323.

Do, M.*,* Kirby, S., & Goldin-Meadow, S. (2022). Regularization of word order in the verb phrase differs from the noun phrase: Evidence from an online silent gesture perception paradigm. *Proceedings of the 44th Annual Conference of the Cognitive Science Society*, 1211-1217.

Do, M., & Kaiser, E. (2022). Sentence formulation is easier when thematic and syntactic prominence align: Evidence from psych verbs. *Language, Cognition and Neuroscience, 37*(5), 648-670.

ELAN (Version 6.7) [Computer software]. (2023). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive.

Ferreira, F. (1994). Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language, 33*(6), 715–736.

Fillmore, C. (1968). The case for case. In E. Bach & R. Harms (Eds.), *Universals in Linguistic Theory*. Holt, Rinehart and Winston.

Furman, R., Küntay, A., & Özyürek, A. (2014). Early language-specificity of children's event encoding in speech and gesture: evidence from caused motion in Turkish. *Language, Cognition, and Neuroscience, 29*(5), 620-634.

Gennari, S. (2002). Motion events in language and cognition. *Cognition*, *83*(1), 49–79.

Grigoroglou, M. & Papafragou, A. (2019a). Interactive contexts increase informativeness in children's referential communication. *Developmental Psychology*, 55. 951–966.

Grigoroglou, M. & Papafragou, A. (2019b). Children's (and adults') production adjustments to generic and particular listener needs. *Cognitive Science*, *43*, e12790.

Gürcanlı, Ö., Nakipoglu Demiralp, M., & Özyürek, A. (2007). Shared information and argument omission in Turkish. In H. Caunt-Nulton, S. Kulatilake, & I. Woo (Eds.), *Proceedings of the 31st annual Boston university conference on language development*. Cascadilla Press.

Hothorn, T., Bretz, F., Westfall, P. & Heiberger, R. M. (2008) multcomp: Simultaneous Inference in General Parametric Models. http://CRAN. R-project.org.

Jackendoff, R. S. (1990). *Semantic Structures*. MIT Press.

Koenig, J.P., Mauner, G., & Bienvenue, B. (2003). Arguments for adjuncts. *Cognition, 89*(2). 67–103.

Koenig, J.P., Mauner, G., Bienvenue, B., & Conklin, K. (2008). What with? The anatomy of a (proto)-role. *Journal of Semantics, 25*(2), 175–220.

Kita, S. & Özyürek, A. (2003). 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, 48*, 16-32.

Lenth, R.V. (2021). emmeans: Estimated marginal means, aka least-squares means (R package Version 1.5.5-1).

Levelt, W. J. M. (1989). Speaking: From intention to articulation. The MIT Press.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation.* University of Chicago Press.

Lockridge, C. B., & Brennan, S. E. (2002). Addressees' needs influence speakers' early syntactic choices. *Psychonomic Bulletin & Review, 9*(3), 550–557.

McNeill, D. (2005). *Gesture and thought*. University of Chicago Press.

Meyer, A. S., Roelofs, A., & Brehm, L. (2019). Thirty years of Speaking: An introduction to the special issue. *Language, Cognition and Neuroscience*, *34*(9), 1073–1084.

Özyürek, A., Kita, S., Allen, S., Brown, A., Furman, R., & Ishizuka, T. (2008). Development of cross-linguistic variation in speech and gesture: Motion events in English and Turkish. *Developmental Psychology*, *44*(4), 1040–1054.

Papafragou, A., & Grigoroglou, M. (2019). The role of conceptualization during language production: Evidence from event encoding. *Language, Cognition and Neuroscience, 34*(9), 1117–28.

Papafragou, A., Massey, C., & Gleitman, L. (2006). When English proposes what Greek presupposes: The cross-linguistic encoding of motion events. *Cognition*, *98*(3), B75–B87.

Rappaport-Hovav, M., & Levin, B. (1998). Building verb meanings. In M. Butt & W. Geuder (Eds.), *The projection of arguments: Lexical and compositional factors*. CSLI publications.

Rissman, L. & Rawlins, K. (2017). Ingredients of instrumental meaning. *Journal of Semantics, 34*, 507-537.

Rissman, L., Rawlins, K., & Landau, B. (2015). Using instruments to understand argument structure: Evidence for gradient representation, *Cognition, 142,* 266-290.

Rissman, L., Rawlins, K., & Landau, B. (2019). Event Participants and Verbal Semantics: Non-Discrete Structure in English, Spanish and Mandarin. *Proceedings*

of the Annual Meeting of the Cognitive Science Society, 41, 960-966.

Rissman, L. & Majid, A. (2019). Thematic roles: Core knowledge or linguistic construct? *Psychonomic Bulletin & Review, 26*(6), 1850–1869.

Russo, L. (2021). *Arguments, adjuncts and instruments in English and Turkish.* [MA Thesis, Carleton University].

Ünal, E., Mamus, E., & Özyürek, A. (2023). Multimodal encoding of motion events in speech, gesture and cognition. *Language and Cognition*, 1–20.

Ünal, E., Manhard, F., & Özyürek, A. (2022). Speaking and gesturing guide event perception during message conceptualization: Evidence from eye movements. *Cognition, 225*, 105127.

Ünal, E., Richards, C., Trueswell, J., & Papafragou, A. (2021). Representing agents, patients, goals, and instruments in causative events: A cross-linguistic investigation of early language and cognition. *Developmental Science*, *24*(6), 1-13.

van Nispen, K., van de Sandt-Koenderman, W.M.E., & Krahmer, E. (2017). Production and comprehension of pantomimes used to depict objects. *Frontiers in Psychology, 8*(1095).

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: A professional framework for multimodality research. *Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation*.

Zheng, M., & Goldin-Meadow, S. (2002). Thought before language: how deaf and hearing children express motion events across cultures. *Cognition*, *85*(2), 145–175.