Towards Building a Corpus of Turkish Referring Expressions

Cengiz Acartürk, Murat Perit Çakır
Department of Cognitive Science, Informatics Institute
Middle East Technical University
Universiteler Mah. 06800 Çankaya, Ankara, Turkey
E-mail: acarturk@metu.edu.tr, perit@metu.edu.tr

Abstract

In this paper we report on the preliminary findings of our ongoing study on Turkish referring expressions used in situated dialogs. Situated dialogs of pairs of Turkish speakers were collected while they were engaged with a collaborative Tangram puzzle solving task, which was designed by Spanger et al (2011) in an effort to build a corpus of referring expressions in Japanese and English. The paper provides our preliminary results on the Turkish corpus and compares them with the findings of comparable studies conducted on Japanese and English referring expressions.

Keywords: Referring expressions, multimodal corpora, discourse annotation, Turkish language resources

1. Introduction

Referring expressions are linguistic resources that allow speakers to identify objects relevant to their ongoing interaction. Reference production and understanding of references involve the ability to think of and represent objects, to direct others' attention to relevant objects in the shared scene, and to identify what other speakers are talking about when they use such expressions (Gundel & Heldberg, 2008). Therefore, referencing practices in which such expressions are put into use are essential for understanding how language mediates cognition at the intra and inter-subjective levels (Hanks, 2009).

Referring expressions have gained increased attention from computational linguists due to the interest towards developing more natural and efficient human-agent interactions in the real-world context. Recently several corpora have been created to aid the analysis of referring expressions in English. For instance, the COCONUT corpus (Di Eugenio et al., 2000) includes a repository of referring expressions used during text-based interactions in the context of a 2-D interior design task. QUAKE (Byron & Fosler-Lussier, 2006) and SCARE (Stoia et al., 2008) corpora are based on interactions recorded in the context of a collaborative treasure hunting task in a 3-D virtual world.

The work on these corpora has led to the development of useful categorization schemes for English referring expressions. However, due to the restrictions imposed on participants at each task scenario, these characterizations usually apply only to a subset of the rich variety of uses referring expressions may have in situated dialogs. For instance, the COCONUT task posed limitations on the use of language by restricting participants to use a text-based interface and enforcing a strict turn-taking protocol. The corpus also did not include extra-linguistic features relevant for understanding the use of referring expressions. In contrast, the QUAKE and SCARE corpora were collected in a voice-enabled 3-D world, which models a more complicated and realistic context of interaction. However, participants were restricted to carry out limited set of actions such as pushing buttons and picking up or dropping objects in this virtual world. For that reason, the QUAKE and SCARE corpora were mainly used for studying location-based references (Byron et al., 2005).

Referring expressions are particularly important in the context of collaborative activity where interlocutors need to establish a mutual orientation towards relevant objects in the scene to coordinate and make sense of each other's actions (Goodwin, 1996; Hanks; 1992; Clark & Wilkes-Gibbs, 1986). Existing corpora of referring expressions lack a naturalistic situated dialog context, which may have an influence on the type and distribution of referring expressions identified based on such corpora. This motivated Spanger et al. (2011) to design a collaborative problem solving activity where pairs of participants coordinate their actions with talk in an unrestricted way. Spanger et al.'s work led to the culmination of the REX-J corpus, which includes referring expressions in Japanese and English. This corpus differs from the previously discussed ones in terms of its focus on the study of event or action based references.

Relevant work on the Turkish language primarily focuses on pronoun disambiguation and anaphora resolution in text. A synchronic corpus of 2 million words (METU Corpus, MTC), a morphologically and syntactically tagged subcorpus of MTC with 65,000 words (Say et al., 2002), and a 500,000-word subcorpus of MTC with discourse annotation (Zeyrek et al., 2009) are recently available tagged corpora in Turkish. On the other hand, previous work relevant to the study of Turkish referring expressions involves pronoun disambiguation and anaphora resolution in text with natural language processing techniques (Kılıçaslan et al., 2009; Tin & Akman, 1994), and a computational model of

contextually appropriate anaphor and pronoun generation for Turkish (Yüksel & Bozşahin, 2002).

In this paper we appropriated Spanger et al.'s experimental setup in an effort to build a corpus of Turkish referring expressions. We aim to build and analyse a corpus that will guide subsequent work on a more general class of referring expressions used in *situated dialogs*. To the best of the authors' knowledge, our study is the first multimodal corpus that focuses on the use of referring expressions in situated, naturalistic dialogs in Turkish.

The rest of the paper is organized as follows. Section 2 provides an overview of the experimental setup and the annotation scheme used to build the Turkish referring expressions corpus. The next section summarizes the preliminary findings of our analysis on the Turkish corpus. The paper concludes with a comparison of our results with the findings of studies conducted on Japanese and English referring expressions, and with a discussion of some possible directions for future research.

2. Corpus Building

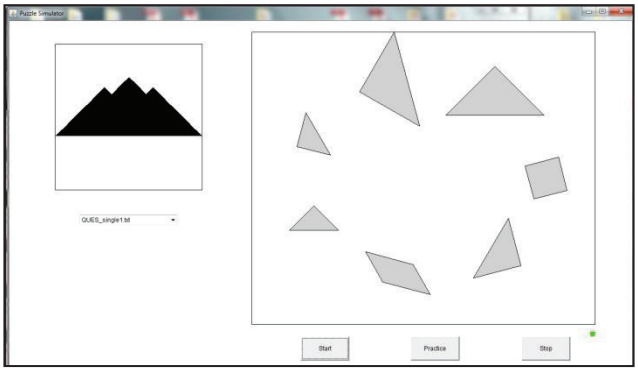

Figure 1: Screen shot of the Tangram Simulator software.

2.1 Experimental Setup

In the experiment, we employed the dual eye-tracking paradigm. Eight graduate students (2 female, 6 male) were recruited to participate in this study. The participants were grouped into 4 same-gender pairs. The pairs were located at different locations (two labs at METU campus) during the experiment. They coordinated their work through a screen sharing software called Team Weaver (www.teamweaver.com), which also enabled voice communication. Two non-intrusive eyetrackers (a Tobii T120 and a Tobii T1750) and the Tobii Studio software were used to record the eye movements, utterances and mouse gestures of both participants concurrently. All participants were native Turkish speakers.

During the experiment, each pair was asked to collaboratively solve four different tangram puzzles by using the Tangram Simulator software (Spanger et al., 2009, 2010; Tokunaga et al., 2010). Figure 1 above displays a screen shot of the Tangram Simulator. Tangram puzzles require solvers to construct a target shape by using seven pieces, which include two large triangles, two small triangles, one medium-size triangle, a square and a parallelogram. Participants used mouse gestures to move and rotate the Tangram pieces to construct the desired shape on their shared workspace. Before the experiment, each participant was asked to complete a short training task to get familiarized with the puzzle interface.

Participants were assigned to either the role of the *operator* or the *instructor* during each task. The operator had the control of the mouse, but had no access to the goal shape. Only the instructor could see the target shape, so it was the instructor's job to guide the operators' actions by uttering instructions. After completing the first two tasks, participants switched their roles. The operator's mouse pointer was not visible to the instructor. In other words, the instructor could only see a change on the shared space if the operator actually moves or rotates a specific piece.

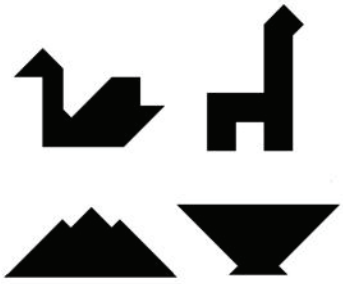

Figure 2: Swan, chair, mountain and vase constituted the four target shapes used during the experiment.

A total of four target shapes were used during the experiment (Figure 2). Pairs were allowed a maximum time of 15 minutes to work on each target. A hint was automatically provided by the software in every 5 minutes. The hint revealed the correct location of a single piece on the target description screen, so it was only visible to the instructor. The total duration of each experiment was approximately an hour.

In short, the experiment is particularly engineered in an effort to encourage participants to use referring expressions to coordinate their work. The roles assigned to the participants and the disembodied nature of the task were the two main constraints imposed by the activity. Hence, the task design eliminated the possibility of using cues such as pointing gestures and bodily orientations. The use of such interactional resources is beyond the scope of this corpus. A visual task that requires spatial arrangements of relevant objects was deliberately chosen to increase the chances of observing the use of referring expressions.

2.2 Annotation Process

The screen recordings of pairs, with overlaid eye-gaze data, were synchronized and transcribed with the Transana data analysis software (www.transana.org).

Figure 3 shows a snapshot of the Transana transcription interface. Two native Turkish annotators (one annotator was one of the authors) independently annotated eight dialogues from two pairs following the annotation guidelines provided by Spanger et al. (2010) with minor modifications with respect to the classification of the annotation tags (explained below). Accordingly, we focused on annotating noun phrases that referred to the pieces in the working space of the puzzle interface. An inter-annotator reliability analysis of the annotation scheme was conducted on a sample of 5,620 tokens extracted from the corpus. Two annotators independently annotated this sample by identifying which tokens constitute referring expressions, and then selecting an appropriate label from the annotation scheme. Holsti's (1969) method found that the percent agreement among the two annotators was 0.847, where 1 corresponds to perfect agreement. This method takes into consideration the number of disagreements among annotators in terms of which tokens should be annotated, but it fails to correct for the role of chance agreement.

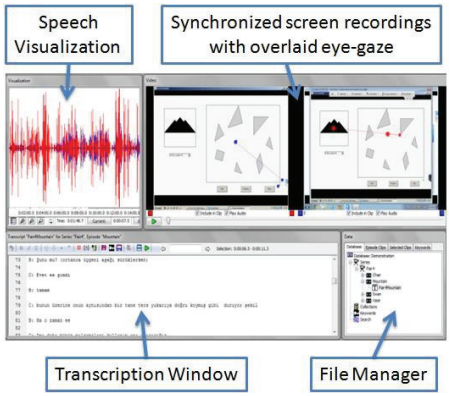

Figure 3: Interface of the Transana transcription tool, which displays a synchronized view of both participants' screens with overlaid eye gaze data.

3. Exploratory Analysis of the Corpus

In eight dialogues, we identified 1,109 tokens (844 produced by the solvers and 265 produced by the operators) with 170 different types of referring expressions. The collection of referring expressions involved 132 multiple-word referring expressions (418 tokens) and 38 single-word (691 tokens) referring expressions. Table 1 shows a partial list of referring expressions ordered by token frequency.

The data presented in Table 1 show that the participants produced the three demonstrative pronouns in Turkish (*o* 'it/that', *bu* 'this' and *şu* 'that') more frequently than others. As a consequence of the nature of the puzzle environment, they employed shape attributes (*parelelkenar* 'parallelogram', *üçgen* 'triangle' and *kare* 'square') in referring expressions. Finally, the participants used size attributes (*büyük* 'large', *küçük* 'small' and *orta boy* 'middle size'), demonstrative adjectives (*o* 'it/that', *şu* 'that' and *bu* 'this') and their combination (*o büyük üçgen* 'that large triangle') for modification of the shape attributes.

Table 1: Frequently used referring expressions in the corpus.

Referring Exp.	%	Referring Exp.	%
o 'it/that'	20.6	*küçük üçgen* 'small square'	3.9
bu 'this'	9.1	*orta boy üçgen* 'middle-size triangle'	2.9
şu 'that'	8.9	*o üçgen* 'that triangle'	2.6
paralelkenar 'parallelogram'	6.0	*şu üçgen* 'that triangle'	1.2
büyük üçgen 'big triangle'	5.7	*bu üçgen* 'this triangle'	1.1
üçgen 'triangle'	5.3	*şu paralelkenar* 'that parallelogran'	0.9
kare 'square'	5.0	*o büyük üçgen* 'that large triangle'	0.8

A more detailed analysis of syntactic/semantic properties of the referring expressions was conducted by a word-by-word based analysis of the identified referring expressions. For this, we annotated the single-word referring expressions and each word in the multi-word referring expressions according to their syntactic/semantic features. The feature list was prepared following the feature list identified by Spanger et al. (2009) for the Japanese corpus. We modified the feature list according to our findings peculiar to Turkish. The feature list is presented in Table 2.

Table 2: Syntactic/semantic features of the referring expressions

Feature	Example
Demonstrative	
Adjective	*bu üçgen* 'this triangle'
Pronoun	*bu* 'this', *şu* 'that', *o* 'it/that'
Nominalized form	*küçükler* 'small-PLU'
Partitive	*-den biri* 'one of …'
Determinative	*diğeri* 'other', *aynısı* 'same'
Pronominal Quantifier	*bu şey* 'this thing'
Attribute	
Size	*büyük üçgen* 'large triangle'
Shape	*büyük üçgen* 'large triangle'
Direction	*sola bakan* 'the one facing to left'
Spatial relation	
Projective	*sağdaki* 'the one on the right'
Topological	*dışarıdaki* 'the one outside'
Overlapping	*üstündeki* 'the one on the top'
Action mentioning	*çevirdiğin* 'the one you turned'
Time adverbial	*deminki* 'the one a moment ago'

The difference between Spanger et al.'s (2009) feature list and ours is the addition of a set of features (*nominalized form*, *partitive*, *pronominal quantifier* and *time adverbial*) in the Turkish feature list. The *determinative* class is involved in the 'other' category in the feature list for the Japanese corpus. We found it necessary to identify those features separately because we observed that the token frequency of those features in Turkish was higher than some of the common features in the Turkish corpus and in the Japanese corpus.

A frequency analysis of syntactic/semantic features of referring expressions revealed that the two types of attributes (shape and size) and the two types of demonstratives (adjectives and pronouns) were more frequently produced by the participants compared to other features. Those four types constitute approximately 84% of all the referring expressions produced by the participants.

The analysis revealed similarities between Turkish and Japanese referring expressions, as well. The major finding for the similarity between the two languages is that those four types of referring expressions were more frequent in the Japanese corpus, as well, constituting 85% of all the referring expression tokens. Table 3 gives a complete list of Turkish referring expressions, as well as Japanese referring expressions ordered by the percentage of token frequency.

Table 3: The syntactic/semantic feature distribution of the referring expressions (TR: Turkish, JAP: Japanese by Spangler et al., 2009)

Syntactic/Semantic Feature	**TR %**	**JAP %**
Shape (Attribute)	34.4	32.0
Pronoun (Demonstrative)	26.0	29.2
Size (Attribute)	14.4	14.1
Adjective (Demonstrative)	10.5	10.4
Determiner (Demonstrative)	4.69	1.59
Projection (Spatial Relation)	2.65	0.76
Action Mentioning	2.04	4.50
Partitive (Demonstrative)	1.56	NA
Pronominal Quantifier (Demonstrative)	1.02	NA
Nominalized Form (Demonstrative)	0.09	NA
Topological (Spatial Relation)	0.09	0.01
Direction (Attribute)	0.04	0.03
Time adverbial	0.04	NA
Overlap (Spatial Relation)	0.00	0.01

A comparison of the Japanese corpus with an English corpus of referring expressions was performed by Tokunaga et al. (2010), suggesting that Japanese participants use more *projection* (spatial relation) expressions and more *action mentioning* expressions compared to English participants. Our results suggest that Turkish participants exhibit a similar pattern with Japanese rather than English. However, a comparative investigation of the three languages will be performed after the completion of the analysis of all the recorded dialogues.

4. Conclusion and Future Work

In this paper, we presented the initial findings of an ongoing research on the construction of a corpus of Turkish referring expressions that employed a situated dialog environment. In its recent form, the data have been partially annotated.

Our preliminary results reveal that demonstrative pronouns, shape attributes and size attributes are the frequently employed features in referring expressions in the described situated dialogue environment of the Tangram puzzle solving task. The results also indicate that there are similarities between the syntactic/semantic feature distribution of Turkish and Japanese referring expressions. Like Japanese speakers, Turkish speakers also tend to use more projection and action mentioning referring expressions as compared to English speakers. We also identified additional features that are peculiar to Turkish referring expressions used in situated dialogs. Nevertheless, our findings are limited to the part of our corpus that has been annotated. A more thorough comparative investigation of the three languages will be performed once the annotation of all the recorded dialogues in our corpus is complete.

In the future we plan to expand this work across various dimensions. First, we will investigate whether the distribution of referring expression types differ across pairs, roles and task types (e.g. symmetric versus asymmetric target shapes). Second, we will focus on the eye tracking data to investigate how eye-gaze patterns are aligned with the referring expressions used by the participants. Finally, we will focus on the sequential organization of utterances that contain referring expressions to identify their communicational roles for the establishment and management of common ground for collaborative work. In particular, we aim to observe how different types of referring expressions are used in repair sequences to address problems of referential understanding.

Acknowledgements. The authors would like to thank Dr. Takenobu Tokunaga for making the Tangram Simulator software available to us, to Emine Eren and Semra Küçük for their help with corpus annotation, to Dr. Kürşat Çağıltay, Özge Alaçam and the METU HCI Research and Application Laboratory for granting us access to their eye tracking facilities, and to the anonymous reviewers for their comments and suggestions.

5. References

Byron, D., Mampilly, T., Sharma, V. & Xu, T. (2005). Utilizing visual attention for cross-modal coreference interpretation. In *Proceedings of CONTEXT 2005*, pp: 83-96.

Byron, D. & Fosler-Lussier, E. (2006). The OSU Quake 2004 corpus of two party situated problem solving dialogs. In *Proceedings of the Fifth Language Resources and Evaluation Conference (LREC 2006)*.

Clark, H. H. & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22:1-39.

Di Eugenio, B., Jordan, P. W., Thomason, R. H. & Moore, J. D. (2000). The agreement process: An empirical investigation of human-human computer mediated collaborative dialogues. *International Journal of Human-Computer Studies*, 53(6): 1017-1076.

Goodwin, C. (1994). Professional Vision. *American Anthropologist* 96(3): 606-33.

Gundel, J. K. & Hedberg, N. (2008). *Reference: Interdisciplinary Perspectives.* New York, NY: Oxford University Press.

Hanks, W. (1992). The indexical ground of deictic reference. In A. Duranti & C. Goodwin (Eds.), Rethinking context: Language as an interactive phenomenon (pp. 43-76). New York: Cambridge University Press.

Hanks, W. (2009). Fieldwork on deixis. *Journal of Pragmatics,* 41, pp. 10-24.

Holsti, O. R. (1969). *Content analysis for the social sciences and humanities*. Reading, MA: Addison-Wesley.

Kılıçaslan, Y., Güner, E. S., & Yıldırım, S. (2009). Learning-Based Pronoun Resolution for Turkish with a Comparative Evaluation. *Computer Speech and Language*, 23(3): 311-331.

Say, B., Zeyrek, D., Oflazer, K., Özge, U. (2002). Development of a Corpus and a Treebank for Present-day Written Turkish. In *Proceedings of the Eleventh International Conference of Turkish Linguistics*.

Spanger, P., Yasuhara, M., Iida, R., Tokunaga, T. (2009). A Japanese Corpus of Referring Expressions Used in a Situated Collaboration Task. *In Proceedings of the 12th European Workshop on Natural Language Generation (ENLG 2009)*. pp: 110-113.

Spanger, P., Yasuhara, M., Iida, R., Tokunaga, T., Terai, A. Kuriyama, N. (2010). REX-J: Japanese referring expression corpus of situated dialogs. *Language Resources and Evaluation*. (Dec). Online First, DOI: 10.1007/s10579-010-9134-8

Stoia, L., Shocley, D. M., Byron, D. K., & Fosler-Lussier, E. (2008). SCARE: A situated corpus with annotated referring expressions. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC 2008)*.

Takenobu, T., Masaaki, Y., Asuka, T., Morris, D., & Belz, A. (2010). Construction of bilingual multimodal corpora of referring expressions in collaborative problem solving. In *Proceedings of the Eighth Workshop on Asian Language Resources*. pp.38 – 46.

Tın, E. & Akman, V. (1994). Situated processing of pronominal anaphora. In *Proceedings of Second Conference for Natural Language Processing (KONVENS'94)*. pp. 369–378.

Yüksel, Ö. & Bozşahin, C. (2002). Contextually Appropriate Reference Generation. *Natural Language Engineering*, 8: 69-89.

Zeyrek, D. Turan, Ü. D., Bozşahin, C., Çakıcı, R., Sevdik-Çallı, A., Demirşahin, I., Aktaş, B., Yalçınkaya, I., Ögel, H. (2009). Annotating Subordinators in the Turkish Discourse Bank. In *Proceedings of the ACL-IJCNLP, LAW Annotation Workshop III*. Singapore, August 6-7, 2009, pp: 44-48.