

Turkish discourse anaphors: *bu* and *o*

Derya Çokal – University of Cologne
dcokal@uni-koeln.de

1 Introduction

In this study, we conducted a small-scale corpus analysis to examine: (1) accessibility of *bu* and *o* to an earlier and/or an immediately preceding clause; (2) how rhetorical relations *bu* and *o* are used; and (3) the role of type of verbs (i.e., dynamic verbs or stative verbs) in the use of *bu* and *o*.

Bu (i.e., *this*) and *o* (i.e., *that*) are accepted as linguistic/cognitive devices through which the addresser orients the addressee's attention to create joint attention, and through which the addressee constructs and modifies the mental representation of unfolding discourse (e.g., Cornish 2009, Çokal 2010, Webber 1991).

2 Methodology

The data for corpus analysis were retrieved from the METU Turkish Workbench. A total of 7849 paragraphs featuring *bu* were retrieved, as well as 4567 featuring *o*. 68 extracts featuring *o* or *bu* were taken randomly from fiction writing: 34 for *bu* and 34 for *o*. Texts were selected from the narrative genre. The number of words per text was between 22 and 115. The distance of an event or a proposition from *bu* and *o* was handled under two categories: an immediately preceding clause and non-adjacent clause. We used Marcu's (2000) rhetorical relations list. Schuster (1988) claims that the type of antecedent (i.e., events or action) determines the selection of the anaphoric expressions. In this study, stative verbs were annotated as a state or idea, while dynamic verbs were annotated as progressive actions on the part of the subject.

3 Results

Bu refers to the nearest clause itself (i.e., the right frontier; 90%), or a clause that is not adjacent to the clause containing the discourse anaphors (i.e., the left frontier; 10%). The antecedents of *o*, on the other hand, can only be the nearest clause 100%; extracts (1; 2).

Extract 1

(a) Hak veriyordu karısına; (b) çalışıyordu ve üstelik çok başarılıydı. (c) Zekiydi, (d) **çalışma diyemezdi**, (e) Cahide' yi dört duvar arasına hapsedmek olurdu **bu**.

(a) He understands his wife; (b) she studies and she is also very successful. (c) She is intelligent; (d) he has no right to tell her not to study. (e) **This** would be to imprison Cahide [between walls].

Extract 2

(a) Kadın çıldırması gibi. (b) Belli ki bir yakınına yıldırım çarptı. (c) Kim olduğunu çıkaramıyoruz. (d) **Kadına yardım etmek için gitmek istiyoruz** (e) ama selden geçilecek gibi değil ki. (f) Olanaksız **bu**.

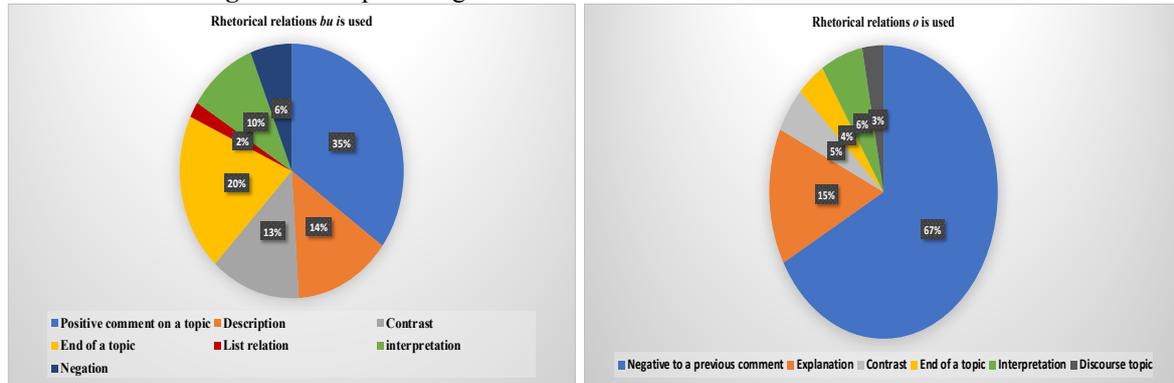
(a) The woman behaved as if she were mad. (b) Maybe one of her relatives had been struck by lightning. (c) We were not able to find out who she was. (d) We wanted to go to help her, (e) but it was impossible because of the flood. (f) **This** was impossible.

In Extract 1 *bu* refers to the proposition “çalışma diyemezdi/ he has no right to tell her not to study” in the immediately preceding clause. In Extract 2, however, *bu* refers to the proposition

in (d) “Kadına yardım etmek için gitmek istiyoruz/ We wanted to go to help her” which is not adjacent to (f).

In Figure 1, both *bu* and *o* are used in interpretation, contrast, and topic-end relations. However, these relations are less frequent than other relations which present significant semantic differences between *bu* and *o* in Turkish (e.g., negative vs. positive topic comment; end of topic, negation, and discourse topic).

Figure 1: The percentage of rhetorical relations in which *bu* and *o* is used



In Extract 3, the writer describes a “Havra”, or synagogue. *Bu* is used in a contrast relation with the discourse marker “Ama” to state that though the synagogue was surrounded by high walls and thus was very quiet, this did not affect the magical atmosphere visible on looking through the keyhole. In Extract 4, *o* is also used in a contrast relation, in a manner similar to *bu*. The writer claims that two teams would tie at the end of the game because of the penalty given at the last minute. Another character in the story contradicts that idea, saying “*o* eskidendi (‘That was in the past’).

Extract 3

- (a) Havra, ne de olsa, uzun upuzun duvarlarla çevriliydi. (b) Uçsuz duvarlar. (c) Sessiz
 (d) Ama **bu**, biz o delikten baktığımızda, içeriden bize renkli ışıklar süzülürken ortalığı sarıp sarmalayan o büyüü bozmadı hiç.
 (a) The synagogue was surrounded by very high walls. (b) Endless walls. (c) Quiet. (d) However, **that** did not spoil the magical atmosphere with coloured lights all around visible when we looked through the keyhole.

Extract 4

- (a) Tutar bir penaltı verir maçın son dakikasında, (b) adam da dengine getirir atar mı , (c) haydi al bakalım sahadan berabere ayrıl . . . (d) Yok artık abicim, (e) berabere yok, (f) **o** eskidendi şimdi yeneceğiz, yeniyoruz, yeneriz.
 (a) Now, what if the referee were to give a penalty in the last minute? (b) The player would score a goal, (c) then we would draw.... (d) No mate, no draw, (e) **that** was in the past, (f) this time we will win, are winning, are going to win.

While *bu* is mostly used in sentences in which a positive comment on a topic is given, *o* is used in a negative topic comment. In Extract 5, the writer describes the surroundings when the character leaves his/her flat. *Bu* in sentence (e) refers to the proposition in sentence (d), “Bakkal da kapamıştı”. The sentence with *bu* is a positive comment on the shop being closed.

Extract 5

- (a) Giyinip çıktı. (b) Merdivenleri hızla indi. (c) Sokak boştu. (d) Bakkal da kapamıştı ,
 (e) **bu** iyiydi. (f) Ağır ağır yürüdü.

(a) S/he put on his/her clothes and went out. (b) S/he went downstairs very quickly. (c) The street was quiet. (d) The shop was closed, (e) **this** was good. (f) He walked very slowly.

However, in Extract 6, the sentence with *o* presents a negative topic comment. The writer explains that when a politician loses the trust and love of a society, it is difficult to regain them. S/he provides his/her negative idea about winning back the trust and love of society. *O* is used more specifically in refuting or criticizing a given proposition.

Extract 6

(a) Ama milletin itimadını, milletin güvenini, milletin sevgisini kaybettin mi bir daha kazanamazsın. (b) **O** çok zor.

(a) However, if you lose the trust and love of society, you will never regain it. (b) **That** [regaining that trust and love] is very difficult.

The percentage of *o* is higher than that of *bu* in the use of dynamic verbs (*o*_dynamic verbs: 61%; *o*_stative verbs: 39%; *bu*_dynamic verbs: 50%; *bu*_stative verbs: 50%).

4 Conclusions

Bu and *o* refer to an entity/a proposition given in immediately preceding clause. Only *bu* refers to an entity/a proposition that is not close. However, the use of such cases is highly infrequent. *O* is mostly used in negative sentences and negative topic comment relations, whereas *bu* is preferred in positive topic comments and contrast relations. *Bu* is used more frequently than *o* (roughly twice as often), which indicates that if the writers do not want to give a negative meaning to their discourses, *bu* seems to be preferred over *o*. *Bu* is used to refer to both dynamic and stative verbs. *O* refers to dynamic verbs more frequently than to stative verbs. Larger-Scaled corpus studies need to be conducted to test these findings.

References

- Cornish, Francis. 2009. Indexicality by degrees: Deixis, “anadeixis” and discourse anaphora. Paper presented at the Symposium *Quel sens pour la linguistique?* April 2009, Université de Toulouse-Le Mirail.
- Çokal, Derya. 2010. The pronominal *bu-şu* and *this-that*: Rhetorical structure theory. *Dilbilim Araştırmaları* 1. 15–33.
- Marcu, Daniel. 2001. *The theory and practice of discourse parsing and summarization*. Cambridge: MIT Press.
- Schuster, Ethel. 1988. Anaphoric reference to events and actions: A representation and its advantages. In Vargha Dénes & Eva Hajičová (eds.), *Proceedings of the 12th International Conference on Computational Linguistics*, 602–607. Stroudsburg: ACL.
- Webber, Bonnie L. 1991. Structure and ostension in the interpretation of discourse deixis. *Language and Cognitive Processes* 6. 107–135.