Strong indefinites in Turkish, referential persistence, and salience structure

Umut Özge[a], Duygu Özge[b], Klaus von Heusinger[a]

[a]**Universität zu Köln,** [b]**Harvard University**

(Version: September 2015)

Abstract

The paper investigates the discourse properties of "strong" indefinite noun phrases, which differ from their "weak" variants by possessing additional sentence- and discourse-level properties like specificity, presuppositionality, discourse-linking among others. The two forward-looking discourse properties we focus on are (i) the referential persistence of the referent introduced by the indefinite, which is measured in the frequency of anaphoric references targeting back this referent; (ii) and salience of the introduced referent, which manifests itself in the form of anaphoric expressions targeting the referent, where the more salient the referent is, the more reduced the form of the anaphor becomes. at optional use of accusative case marker on a direct object in Turkish where the presence of it indicates "strength". A sentence continuation study revealed no effect of case-marking (hence "strength") on referential persistence and salience.

1. Introduction[*]

Indefinite noun phrases come in different varieties: strong indefinites, regular indefinites and weak indefinites (Kamp 2014). In this paper we focus on strong indefinites, i.e. indefinites that come with a referential intention of the speaker and that show various effects at three semantic and pragmatic levels: (i) at the sentence level, for instance, strong indefinites tend to be directly referential and specific, they show wide scope behavior and escape scope islands; (ii) their backward-looking discourse properties include discourse-linking and presuppositional behavior; (iii) with respect to the upcoming discourse they introduce salient discourse referents that can be picked up by attenuated anaphoric expression and they signal that the introduced discourse referent will be the antecedent for extended referential and even topical chains. While the sentence-semantic properties of strong indefinites are fairly well understood, their backward- and forward-looking properties are still unclear and need further research. Our study aims at investigating their forward-looking properties.

Languages may provide lexical, morphological or syntactic means to mark strong indefinites. We assume that case marking of the indefinite direct object (Differential Object Marking or DOM) in Turkish indicates a strong indefinite. Case-marked indefinite direct objects constitute one prototype of specific indefinites (Enç 1991) or strong indefinites in our terminology. We employ Turkish strong indefinites as our empirical domain to investigate the forward-looking discourse properties of strong indefinites. The two properties we are presently interested in are (i) the referential persistence of the referent introduced by the indefinite, which is measured in the frequency of anaphoric references back to this referent; and (ii) salience of the introduced referent, which manifests itself in the form of anaphoric expressions targeting the referent. We report from a sentence continuation experiment investigating the competition between a referent introduced by a subject and a referent introduced by a direct object in a ditransitive construction. The results, however, do not show any effects of case marking of the direct object for (i) the referential persistence or (ii) the salience structure. But the results clearly provide a surprising bias towards the object with respect to (i) referential persistence, even if not with respect to (ii) the salience structure. This confirms recent discussions of the different morpho-syntactic effects of speaker planning (including referential persistence) and hearer expectations (including salience structure) as summarized in Kehler & Rohde (2013).

Section 2 introduces the notion of strong indefinites and argues on the basis of their sentence semantics that case-marked indefinite direct objects in Turkish are a clear instance of strong indefinites. Section 3 presents the classical analysis of Enç (1991), which addresses how strong indefinites relate to discourse in backward direction under the notion of discourse-linking. Section 4 introduces two forward-looking discourse properties associated with strong indefinites: (i) Discourse Structuring Potential (of which referential persistence is a part) and (ii) salience structure. Section 5 reports our sentence continuation experiment, which provides results pertaining to the two mentioned forward-looking discourse properties. Section 6 discusses the results and Section 7 concludes.

2. Strong indefinites in Turkish and their sentence semantics

"Indefinite noun phrase" is a cover term for quite different types of noun phrases. We follow Kamp (2014) and assume a three-way distinction comprising strong indefinites, regular indefinites, and weak indefinites. We call an indefinite "strong" if the speaker uses it with a referential intention, i.e. if the speaker can identify the referent, signaling to the hearer that that referent will be under discussion. Thus the hearer creates a discourse representation based on that of the speaker, which is permanent enough to be targeted by anaphoric expressions in the later stages of the discourse. A weak indefinite, on the other hand, is an expression that does not introduce a discourse referent with a long life, if it introduces one at all. The property expressed by the indefinite is generally included in the verbal component. Regular indefinites are indefinites that are neither strong nor weak, i.e. they do introduce discourse referents, but with no clear referential intention of the speaker. They constitute the main class of "regular" indefinites that show narrow scope behavior with respect to negation, modal operators and verbs of propositional attitudes, as already described by Karttunen (1969/1976). Languages offer a wide variety of lexical, morphological and syntactic means to mark different kinds of indefinites. For instance, English indefinite *this* (Prince 1981; Ionin

2006, Chiriacescu 2014) marks a strong indefinite, as *this*-indefinite noun phrases show referential readings. Differential Object Marking in Romanian also indicates a strong indefinite (Chiriacescu 2014). On the other side, bare nouns in English and other article languages indicate weak indefinites, while the use of the indefinite article is often unspecified with respect to the type of indefinite.

Turkish is an article language with the indefinite article *bir*, which is homophonous to the numeral for 'one', but differs in accent and distribution from the numeral (Kornfilt 1997). Turkish does not have a definite article, but case-marked direct objects without the indefinite determiner are definite. Other definite noun phrases like proper names, pronouns, demonstratives and demonstrative noun phrases obligatorily receive the accusative marker. The Acc(usative) case marker for the direct object is optional for indefinite noun phrases, which yields the following paradigm for direct objects:

(1) Formal marking of different referential properties
    a.   (ben)   *bu*   *kitab-ı*   okudum.   *demonstrative*
          I   this   book-Acc   read.1.sg
          'I read this book.'
    b.   (ben)   *kitab-ı*   okudum.   *definite*
          I   book-Acc   read.1.sg
          'I read the book.'
    c.   (ben)   *bir*   *kitab-ı*   okudum.   *indef. specific*
          I   a   book-Acc   read.1.sg
          'I read a certain book.'
    d.   (ben)   *bir*   *kitap*   okudum.   *indef. non-specific*
          I   a   book   read.1.sg
          'I read a book.'
    e.   (ben)   *kitap*   okudum.   *"incorporated"*
          I   book   read.1.sg
          'I did book-reading.'

The demonstrative determiner *bu* in (1a) makes the noun phrase definite and thus forces case marking of the direct object *bu kitabı*. The bare noun with case marker *kitabı* is interpreted as definite (since there is no definite article, the case and the lack of the indefinite article clearly signal definiteness). The form *bir kitabı* in (1c) with the indefinite article and the case suffix is interpreted as a specific indefinite noun phrase. The form *bir kitap* in (1d) with the indefinite article and without case is generally assumed to be a regular indefinite noun phrase, while the bare form *kitap* in (1e) is the incorporated form (see Dede 1986, Kornfilt 1997, von Heusinger & Kornfilt 2005).

We briefly review some evidence that the Acc-marked indefinites, as in (1c), receive strong interpretation. Case-marked direct objects show wide scope behavior with respect to commanding operators, while the non-marked alternates are confined to narrowest possible scope. Example (2) (from Özge 2011) presents the relevant evidence. In the Acc-marked variant (2a), the indefinite can get interpreted in all the available scopal positions, as in Readings (i) - (iii). The readings differ in the scope of the indefinite with respect to the other

two quantifier phrases *çoğu dilbilimci* ('most linguists') and *her makaleyi* ('every article'). The narrow scope reading (i) is true if most linguists read every article that solves one or another important problem. The intermediate scope reading (ii) is true if for most linguists there is one (potentially different) problem such that the linguist reads all the papers that solve that problem, and the wide scope reading (iii) is true if there is one problem such that most linguists read every solution for that problem. The non-marked variant (2b) on the other hand yields only one of these readings, namely the one where the indefinite has the narrowest scope with respect to two commanding operators *most* and *every*.

(2) a.

Çoğu	dilbilimci	önemli	bir	problem-i	çöz-en
most	linguist	important	a	problem-Acc	solve-Rel
her	makale-yi	oku-muş-tur.			
every	article-Acc	read-Ev.Cop-Aor			

'Most linguists have read every article that solves an important problem.'

    (i)   Reading (i): Most linguists are such that if an article solves some important problem they read the article ($most\forall\exists$).

    (ii)   Reading (ii): For most linguists it is the case that there exists an important problem *p*, such that s/he has read every article that solves *p* ($most\exists\forall$).

    (iii)   Reading (iii): There is a problem *p* such that most linguists have read every article that solves *p* ($\exists most\forall$).

    b.

Çoğu	dilbilimci	önemli	bir	problem	çöz-en
most	linguist	important	a	problem	solve-Rel
her	makale-yi	oku-muş-tur.			
every	article-Acc	read-Ev.Cop-Aor			

'Most linguists have read every article that solves an important problem.'

    (i)   The only reading: Most linguists are such that if an article solves some important problem they read the article ($most\forall\exists$).

The Acc-marked indefinites are also observed to receive referential readings in the context of referentially opaque verbs, as in (3) (from Dede 1986). In (3a) the Acc-marked direct object receives a referential or specific reading, i.e. the speaker is looking for a particular individual and the most appropriate description is that the individual is a secretary. In (3b) the unmarked direct object can only be interpreted non-specifically, i.e. the speaker is looking for one person or other with the property of being a secretary.

(3) a.

Bir	sekreter-i	arı-yor-um.	Bul-a-mı-yor-um.
a	secretary-Acc	look_for-Pr.Prog-1.sg	find-Abil-Neg-Pr.Prog.-1.sg

'I am looking for a secretary. I can't find him.'   [specific indefinite]

    b.

Bir	sekreter	arı-yor-um.	Bul-a-mı-yor-um.
a	secretary	look_for-Pr.Prog-1.sg	find-Abil-Neg-Pr.Prog-1.sg

'I am looking for a secretary. I can't find one.'   [non-specific indefinite]

To summarize, Turkish Acc-marking of indefinite direct objects makes them referentially strong, i.e. case-marked direct objects are free to take scope (narrow, intermediate or wide), i.e. they show referential readings in opaque contexts.

3. Strong indefinites in Turkish and their backward discourse semantics

In a seminal paper, Enç (1991) points to a correlation of Differential Object Marking with specificity and specificity with (implicit or discourse) partitivity, which she takes as an instance of discourse-linking (d-linking in short) following Pesetsky (1987). The d-linking effect of the Turkish accusative marker can be illustrated with the following well-known example:

(4)   Odam-a   birkaç   çocuk   girdi.
        my room-Dat   few   child   entered
        'A few children entered my room.'   (Enç 1991: ex. 16)

        a.   İki   kız-ı   tanıyordum.
                two   girl-Acc   knew.1sg
                'I knew two girls.'   (Enç 1991: ex. 17)

        b.   İki   kız   tanıyordum.
                two   girl   knew.1sg
                'I knew two girls.'   (Enç 1991: ex. 18)

Enç (1991: 6) observes that it is only with the accusative marked version *iki kız-ı* ('two girl-Acc) in (4a) that the girls are understood as belonging to the set introduced in the opening sentence of the discourse, while the unmarked indefinite *iki kız* ('two girl') in (4b) refers to two girls not included in the original set. Enç (1991) claims that her notion of d-linking is what underlies the more widely known "strong"/"weak" distinction of Milsark (1974), and that Turkish is a language where this distinction has an overt morphological realization. In this respect, Enç's (1991) proposal is an important one regarding the interpretation of strong indefinites across languages.

While Enç's (1991) analysis was highly influential, various authors like Farkas (1994), Zidani-Eroğlu (1997), Kelepir (2001), von Heusinger & Kornfilt (2005), Kılıçaslan (2006), Özge (2011, 2013a,b) among others noted that Enç's (1991) d-linking effect, while clearly attested for certain examples, fails to generalize to the full range of Turkish data. The objections are based on cases where the accusative marker is not needed for a discourse-linked reading and cases where the marker has a function other than discourse-linking. In addition to these, there are also cases where the presence of the marker does not have any observable effect regarding the relation of the indefinite to the backward discourse and there are no nominal or intensional operators that could interact with the indefinite. We call such cases "transparent contexts". As an example, take the following minimal pair, where there are no nominal or intensional operators that could interact with the indefinite.

(5) a. John   izleyiciler-e   bir   manken   tanıttı.
       John   audience-Dat   a   model   introduced
       'John introduced a model to the audience.'

    b. John   izleyiciler-e   bir   manken-i   tanıttı.
       John   audience-Dat   a   model-Acc   introduced
       'John introduced a model to the audience.'

Speakers do not seem to have reliable intuitions concerning the interpretative effect of the presence versus absence of the accusative marker in the above minimal pair, apart from the intuition that the indefinite object is somewhat more emphasized in the marked version (5b) in comparison to the non-marked version (5a). As far as we know, there has not been any attempt in the literature to experimentally investigate the contribution of the case marker to the discourse properties of the indefinite in sentences like (5a,b).

4. Discourse prominence: Discourse Structuring Potential and salience structure

Referential expressions not only introduce new discourse referents and link them to already established ones, they also activate these referents in the upcoming discourse in various ways. We consider two different types of forward-looking properties of discourse referents. One is the Discourse Structuring Potential (or DSP) and the other is the salience structure.

DSP is based on Givón's (1983) notion of topic continuity. In a series of papers, Chiriacescu and von Heusinger (Chiriacescu & von Heusinger 2010, von Heusinger & Chiriacescu 2013, Chiriacescu 2014) have further differentiated Givón's factors of topic continuity into two measurable parameters, which together represent the "forward-looking potential" or the Discourse Structuring Potential of a referential expression. The Discourse Structuring Potential of an expression introducing a discourse referent is defined as the information pertaining to the discourse status of the introduced referent in the subsequent discourse. The Discourse Structuring Potential is characterized by the two parameters "referential persistence" and "topic shift potential", which are defined as follows:[1]

(6)   Discourse Structuring Potential (DSP, Chiriacescu & von Heusinger 2010)
       (i)   *Referential persistence* measures the frequency with which a referent is anaphorically rementioned in the subsequent discourse (Givón 1983, Gernsbacher & Shroyer 1989, Arnold 1998, Kehler et al. 2008)
       (ii)   *Topic shift potential* measures the distance in sentences with which a non-topical referent is mentioned again as a topic for the first time in the subsequent discourse. (Note: Givón's topic continuity is different from this parameter, since it measures the duration of being a topic and not the first time it appears as a topic.)

The relation between DSP and strong indefinites has recently been studied in various languages in a number of studies. Chiriacescu & von Heusinger (2010), von Heusinger & Chiriacescu (2013) observed through a story continuation study that the Romanian DOM

---

[1] In the present paper we will be concerned only with the "referential persistence" component of DSP.

marker *pe* has a significant effect in increasing the DSP of both the definite and indefinite NPs it attaches to. Deichsel & von Heusinger (2011) report similar results for German indefinite *dies* ('this'), and Kim & Kaiser (2013) for case marking in Korean. Besides the DSP, there is a second type of forward-looking discourse activation: salience structure. Depending on various other parameters, certain forms of noun phrases promote their discourse referents to a high-ranked position in a salience structure (modeled as a ranked set of elements, see von Heusinger 2007, Bittner 2012). The higher a discourse referent in a salience structure the more easily it can be accessed. This high-ranked discourse referent is therefore very "accessible" and often a reduced element can be used to establish the anaphoric linkage.

Local effects of the salience structure are best accounted for in Centering Theory (Grosz, et al. 1995; Brennan 1995). This theory argues that the entities that the speaker assumes are in the hearer's center of attention tend to re-appear in the subsequent discourse segment. Moreover, the salient entities tend to be marked more with null pronouns than with overt pronouns and more with overt pronouns than with full referential expressions. The speaker tends to use full referential expressions when the entity in question is not easily identifiable by the hearer. Less salient entities, such as objects, are expected to be rementioned using fuller referential expressions than salient entities, such as subjects. Previous studies have shown that speakers use more pronouns when re-referring to the subject (i.e., the more salient entity), while they use full noun phrases when referring to the object that is less salient (Arnold 1998; Kaiser 2011). However, in contrast to the languages tested previously, Turkish frequently allows subject-drop, especially when the subject is an already familiar referent. In languages like Turkish (Turan 1996) or Spanish (Arnold 1998; 2003), it has been observed that the subjects, being prominent or salient in the discourse, are more likely to be dropped than pronominalized. Analyzing naturally occurring discourse utterances from the perspective of Centering Theory, Hoffman (1998) found that the subject referent in Turkish tends to be referred to in the discourse with null pronouns (see Runner & Ibarra 2016 for Spanish).

The parameters that determine the position of a discourse referent in the salience structure are various, including grammatical relation, semantic role, animacy, referential type, verb type, etc. It has been convincingly shown that in Turkish, grammatical relation/thematic role is among the most important ones. Turan (1998) provides evidence that subjects are usually the most salient element, with the exception of psychological predicates, where non-subjects are more salient (see also Schumacher, et al. 2016 for German). Turan (1998) observes that in (7) (adapted from her (11)), the null pronoun in the second sentence of (7a) can only target the subject referent *Murat* in the first sentence, which is confirmed by the ungrammaticality of (7b) where the null pronoun is intended to refer to the direct object *Ali* in the first sentence.

(7) a. Murat_i Ali-yi_k davet etti. *pro_i* on-a_k içki ikram etti.
          Murat   Ali-Acc   invited   him-Dat   drink served
          'Murat invited Ali.'   'He$_i$ served him$_k$ a drink.'

     b Murat_i Ali-yi_k davet etti. *#pro_i* on-a_k içki ikram etti.
          Murat   Ali-Acc   invited   him-Dat   drink served

7

'Murat$_i$ invited Ali$_k$.' Intended reading: 'He$_k$ served him$_i$ a drink.'

One important observation we can add to this example is that even if we change the order of the subject and object, and alter the pragmatic plausibility in favor of the object referent, the pattern observed by Turan (1998), namely that it is the subject referent that can be targeted by a reduced (null) form, stays intact. Here is our variation on (7):[2]

(8) a. Ali-yi$_k$ Murat$_i$ davet etti.   *pro$_i$*   on-a$_k$ çiçek   getirdi.
         Ali-Acc   Murat   invited   him-Dat   flowers   brought
         'Murat invited Ali.'   'He$_i$ brought him$_j$ flowers.'

     b   Ali-yi$_k$ Murat$_i$ davet etti.   *#pro$_k$*   on-a$_i$ çiçek   getirdi.
         Ali-Acc   Murat   invited   him-Dat   flowers   brought
         'Murat$_i$ invited Ali$_k$.'   Intended reading: 'He$_k$ brought him$_i$ flowers.'

These observations show that subjecthood is a strong determinant of salience in Turkish (see Turan 1996; 1998 and Hoffman 1998 for more data and arguments).

To summarize: Turkish Acc-marked direct objects are referentially strong as the standard tests show. They are often, but not always linked to the previous discourse. The salience structure in Turkish is primarily determined by grammatical relations, i.e. the referent introduced by a noun in the subject position is the most salient and typically realized as a null pronoun in the next sentence (as Turkish is a pro-drop language). In the next section, we investigate the effect of Acc-marking of the direct object with respect to the salience structure and referential persistence.

5. A sentence continuation experiment

5.1 Objective

Our main question in the present study is whether the Turkish differential object marker (i.e. Acc case), in transparent contexts like (5) contributes to the Discourse Structuring Potential of the indefinite as previously observed for other languages such as English (Prince 1981, Ionin 2006), German (Deichsel 2014) or Romanian (Chiriacescu & von Heusinger 2010; von Heusinger & Chiriacescu 2013, Kim & Kaiser 2013). We concentrate on the referential persistence component of DSP (i.e. the increase in the likelihood of a particular referent being rementioned in the upcoming discourse), leaving the investigation of the topic-shift dimension for future research.[3] If the accusative case increases the referential persistence of the object entity, we predict that the object entity would be more likely to occur to again in the upcoming discourse. We also examine the salience structure and analyze the types of referring expressions used to refer to the entities that are rementioned in the upcoming discourse. If saliency is reflected via shorter referential expressions as predicted by

---

[2] In the absence of any intonational or aspectual devices that would alter the discourse segmentation, the continuation in (8b) results in a slightly odd situation, as one would expect the guest rather than the host to bring flowers.
[3] Nilsson (1985) and Erguvanlı & Zimmer (1994) hint at the possibility that the accusative marker might be a signal of topicality in the upcoming discourse.

Centering Theory, the referential expressions should appear in reduced forms for the salient entity, such as pronouns or null pronouns rather than full NPs.

5.2 Method
5.2.1 Participants
Thirty-three participants took part in this study. Turkish was their native/first language. They were either university students or graduates. They reported that they did not have any language disorders or neurological problems.

5.2.2. Materials
We used two-sentence discourses that followed a consistent pattern in terms of (i) number of sentences, (ii) length of sentences, (iii) discourse genre (all were made-up news scripts), (iv) similar syntactic structures, and (v) the number of discourse referents.

The first sentence acted as a discourse initiator, as in (9a) and (10a), where an event (e.g. a discussion on health or a fashion day), a location (e.g. a TV show or Safran Palace Hotel), and a time (e.g. the whole week or last night) were introduced. The second sentence acted as a referent introducer, as in (9b) or (10b), where a sub-event (e.g. the first program or the first show of the night) of the event mentioned in the discourse initiator (i.e. health discussion or fashion day) was introduced. The main function of the second sentence was to introduce a discourse-new subject and a discourse-new object in a ditransitive construction with Agent-Patient arguments. The case marking on the object was manipulated between zero and accusative (indicated by '(-yl)' in bold).[4]

(9) a. Sağlık medyasının ana gündem maddesi haline gelen diyabet hastalığı geçen hafta boyunca Sağlık TV kanalında incelemeye alındı.
'Diabetes, which is currently in the focus of health media, has been analyzed on the Health TV channel last week.'

      b. Serinin ilk programında   tanınmış   bir   diyetisyen
in the first program of the series   well-known   a   dietician
program-a   bir   diyabetli**(-yi)**   konuk etti.
program-Dat   a   diabetic**(-Acc)**   hosted
'In the first program of the series, a well-known dietician hosted a diabetic in the show.'

(10) a. Yeni markaların tanıtımına fırsat veren Nişantaşı moda günlerinin açılışı dün gece Safran Palas Oteli'nde yapıldı.
'Nişantaşı fashion days, which serve for the introduction of new brands, started last night at Safran Palace Hotel.'

      b. Etkinliğin ilk defilesi-nde   başarılı   bir   modacı
first show of the event-Loc   successful   a   designer
izleyenler-e   bir   manken**(-i)**   tanıttı.

---

[4] The Turkish accusative marker is realized as *i* or *ı* (abstracted as *I*) depending on the requirements of vowel harmony. It requires the buffer consonant *y* when attached to a word ending with a vowel.
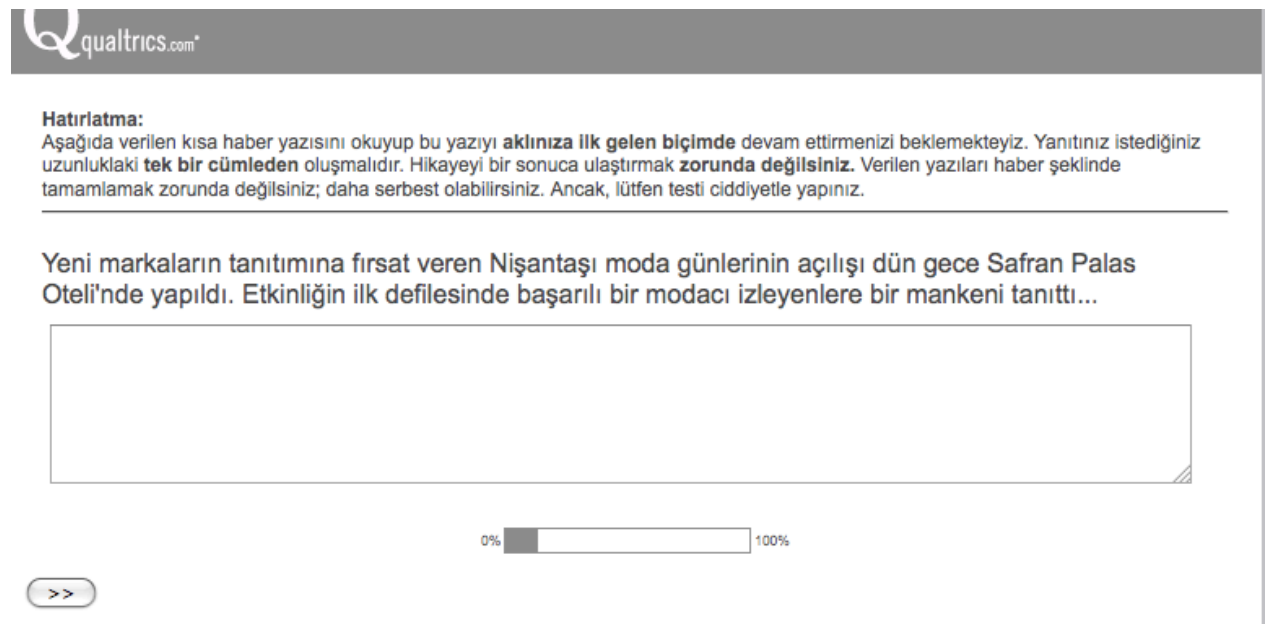
audience-Dat   a   model**(-Acc)**   introduced

'In the first show of the night, a successful designer introduced a model to the audience.'

We had 30 items composed of 12 critical items and 18 filler items. In constructing our critical items we aimed to make sure that: (i) the selected verb is grammatical both with an accusative-marked indefinite object and with a non-marked indefinite object; (ii) the absence versus the presence of the accusative marker does not interact with any nominal or intensional operators such as quantifiers, modals, tense etc.; (iii) the selected verb is structurally similar, namely ditransitive verbs with a dative argument in addition to the direct object. Filler items were structurally similar to the critical items and consisted of referents marked with various case markers other than Acc (e.g., ablative, genitive, dative).

5.2.3 Procedure

We conducted a web-based discourse continuation study, where participants were instructed to read the two-sentence news scripts and write a one-sentence continuation for them. Each script appeared individually on the computer screen and participants typed in their sentences into the empty box provided below each script. They were not able to go back to a previous response once they had entered a continuation (see Figure 1 for a sample screenshot).

Figure 1: A sample screenshot from the web-based discourse continuation study (see (10) above for the translation of the experimental item).[5]

5.3 Predictions

If Turkish DOM has a DSP effect on the indefinite NPs similar to the Romanian *pe* marker

---

[5] The instructions that appear on the top part of the screen translate as: 'We ask you to read the short news text below and write a continuation to it in the way that comes first to mind. Your response should be a single sentence without any length limit. You are not required to conclude the story. You do not have to continue the story in the news style; you are free to adopt any style you wish. However, please be serious in your response.'

(Chiriacescu & von Heusinger 2010; Chiriacescu 2014), i.e. if the accusative case increases the DSP of the object referent, we predict an increase in the number of object rementions in cases where the object is Acc-marked compared to the cases where the object appears unmarked (or with zero case form). As for the salience structure, we also predict that Turkish speakers use more zero forms when referring back to an Acc-marked direct object than when they are referring back to an unmarked direct object. In general, we expect more references back to the subject and more null forms referring back to the subject antecedent.

5.4 Results

Responses, i.e. the one-sentence continuations provided by the informants, were coded as belonging to one of four categories: (i) the subject but not the object is rementioned (NP1-Subject), as in (11); (ii) the object but not the subject is rementioned (NP2-Object), as in (12); (iii) both the subject and object are rementioned (Both), as in (13); (iv) neither the subject nor the object is rementioned (Neither), as in (14). More than one occurrence of a referent type in the same response is counted as one. The examples (11) and (12) are sentence continuations of (9), and (13) and (14) are continuations of (10):

(11) *NP1-Subject*

Diyetin insan sağlığındaki öneminin altını çizen **diyetisyen**$_{NP1}$, sağlıklı bir diyetin mutlu ve sağlılı bir yaşamın temeli olduğunu söyledi.

'**The dietician**$_{NP1}$, who underlined the importance of diet in health, stated that a healthy diet is fundamental to a healthy and happy life.'

(12) *NP2-Object*

Diyabetli$_{NP2}$ yaşadığı sıkıntıları anlattı.

'**The diabetic**$_{NP2}$ told about the troubles he/she is experiencing.'

(13) *Both*

ve **pro**$_{NP1}$ **kendisinin**$_{NP2}$ podyumların vazgeçilmez ismi olacağını söyledi.

'and **she**$_{NP1}$ said **she**$_{NP2}$ will be an indispensible name on the catwalks.'

(14) *Neither*

Safran Palas Oteli'ndeki ortam görmeye değerdi.

'The atmosphere in Safran Palas Hotel was worth seeing.'

The results are summarized in Table 1 and Table 2. Table 1 provides the percentages and absolute numbers (in parentheses) of continuation sentences according to the four classes discussed above and according to the condition on case marking (case marked = Acc; unmarked = zero). Table 1 makes it clear that case marking does not affect the type of continuation sentence – we have exactly the same proportion across the four types with zero-marked direct objects (second line) or Acc-marked direct object (third line).

Case\ Referent	NP1-Subject	NP2-Object	Both	Neither	Total
Zero	9% (17)	43% (83)	18% (35)	31% (60)	100% (195)
Acc	8% (16)	39% (77)	23% (46)	29% (58)	100% (197)
Sum	8% (33)	41% (160)	21% (81)	30% (118)	100% (392)

Table 2 focuses on the number of anaphoric elements in the continuation sentences referring to NP1 and NP2. We calculated the sum of continuation sentences of "type NP1" and of "type Both" to yield the number of items referring back to NP1 and we did the same for NP2 responses.[6] Again we see that the relation between NP1-linked referents and NP2-linked referents is exactly the same with or without Acc-marking. Table 2 also shows that a continuation with an NP2-linked referent is much more probable than with an NP1-linked referent.

Case \ Referent	NP1-Linked	NP2-Linked	Total
Zero	31% (52)	69% (118)	100% (170)
Acc	34% (62)	66% (123)	100% (185)
Sum	32% (114)	68% (241)	100% (355)

Table 2: Number of anaphoric expressions linked to NP1-Subject and NP2-Object

The percentage of each referent type (NP1-Linked and NP2-Linked) for each condition (Acc/Zero) was subjected to the statistical analysis. We conducted an ANOVA with Case (Acc, Zero) X Referent-Type (NP1-Subject, NP2-Object). This revealed an effect of Referent-Type only, which was due to NP2-Object-bias *[F1(1,32) = 42.19; p = .000; F2(1,11) = 18.75, p = .001]*. Overall, the object was more likely to be rementioned (68% of the responses) compared to the subject (32% of the responses) (see Table 2). There were no other main effects or *interaction effects [p > .1]*.[7]

Concentrating on cases where participants rementioned only the subject or only the object, participants were more likely to add their sentences to the second discourse-initiating sentence (e.g. (13)) using the discourse connector *ve* ('and') when rementioning the subject (39% (13 out of 33)) compared to when rementioning the object (2% (3 out of 160)).

As for the type of referential expressions, there is no difference between case-marked and non-case-marked forms. Therefore, Table 3 has merged the figures for the two types; percentages indicate how often a null pronoun refers to the NP1 referent or the NP2 referent. As expected, pronouns primarily refer to NP1 referents (85% vs. 15%); there are nearly no full pronouns and full NPs primarily refer to NP2 referents (78% vs. 22%). The last column of

---

[6] Note that these numbers are different from the number of continuation sentences altogether as we have to count the two items NP1 and NP2 twice in the sentence type Both.

[7] We also conducted an alternative non-parametric analysis where we found similar results. For this, we calculated the number of each referent type (NP1-Subject, NP2-Object) for each condition. We derived a new dependent variable 'Object Preference' by subtracting the number NP1-Subject responses from the number of NP2-Object responses. Using this variable, we conducted a Wilcoxon Signed-Rank Test to test the effect of Case (Acc, Zero) on the Object Preference. The test revealed no significant effect of Case [*W* = 210.5, z = -.45, *p* = .65 r = 5.74].

Table 3 provides the percentages of the different anaphoric forms with respect to all anaphoric forms: 2% pronouns, 17% null pronouns and 81% full NPs. As we can see, there is a clear contrast: The majority of NP1 referents are represented by null pronouns, while the majority of NP2 referents are represented by full NPs (but there is no effect of case on the direct object).

Case \ Referent	NP1-Subject items	NP2-Object items	Total
Null pronoun	85% (51)	15% (9)	17% (60)
Pronoun	0% (0)	100% (6)	2% (6)
Full NP	22% (63)	78% (226)	81% (289)
			100% (355)

Table 3: Distribution of referential forms for each condition

6. Discussion

We set out to investigate whether the differential object marker on indefinite direct objects in Turkish (i) enhances the referential persistence of the object referent and (ii) the salience structure, which is mirrored in the form of the anaphoric expression linked to it. The results of our experiment showed that there is no effect of the case marker on either phenomenon.

In contrast to previous studies, we found in our sentence continuation study that Turkish speakers tend to remention the object in their upcoming discourse rather than the subject referent. That is, the object was referentially more persistent regardless of the case marking of the antecedent object referent. As for the salience structure, the referential expressions used by the participants indicated that they treated the subject as the more salient entity compared to the object. In what follows, we will discuss each of these findings in turn.

First, the overall object-mention-bias was rather unexpected given that most of the previous studies documented that speakers tend to start and continue their discourse with the subject referent (Turan 1996; Hoffman 1998). We will briefly mention some main perspectives on the discourse prominence of subject referents. In some models, as the subject constitutes the starting point or the perspective of the speaker, it is more likely that speakers will start their discourse with the subject. Whenever the perspective has to shift from subject to object during the course of the discourse, the processing load in the communication increases (MacWhinney 1977). This has been related to the fact that the subject is a generally more salient, active, or potent entity (Johnson-Laird 1968; Ertel 1977; Osgood & Bock 1977; MacWhinney 1977). In other models, speakers tend to start their utterance with the subject referent because in these models, the subject is easy to retrieve from the memory. That is, the subject is a conceptually more accessible referent than other referents in the discourse (Bock & Warren 1985). Centering Theory also puts the subject referent in the center of attention as the entity that remains prominent throughout the discourse via repeated rementions (Grosz, Joshi, & Weinstein 1995; Brennan 1995).[8] Indeed, previous studies have reported that subject arguments are more likely to be rementioned in the subsequent discourse compared to the object arguments (Arnold 1998). In Arnold's study, the subject

---

[8] These approaches do not make the distinction we make here between the referential persistence and salience, often treating the two notions as a unified concept.

referent was a discourse-old referent, which had already re-appeared once in the given discourse with a pronoun whereas the object was a discourse-new entity that was in the focus and did not re-appear as part of the given discourse. Even a focused object was not rementioned in the upcoming discourse as often as the subject. Many subsequent studies in anaphora resolution pointed to the overall subject-bias in various languages even though some features increased the selection of the object, such as using discourse-new subjects compared to discourse-old ones (Cowles, Walenski, & Kluender 2007), pronominalization of the object (Kaiser 2011), or ues of topic marking (Kim & Kaiser 2009) or Differential Object Marking on the object (Chiriacescu & von Heusinger 2010).

In our study, both entities (subject always before the object) were introduced as full NPs within the same sentence, so both were discourse-new entities with no pronominalization and the case marking on the object varied between the zero and the accusative. Thus, one would have expected a subject-bias in line with the previous findings. Why this was not the case in our study is a question that we can only partially address here, given the limitations of the data at hand. As stated above, when they rementioned the subject referent, the participants tended to continue the discourse using the discourse connector *and*, as in (15), more often than when they rementioned the object, as in (16), both as continuations to (10).

(15)   …ve onunla bir sonraki defilelerde de calisacagini belirtti.
         '…and she (*pro*) stated that she (*pro*) will work with her in the upcoming shows as well.'
(16)   Manken doğal güzelliği ve samimi davranışlarıyla dikkatleri üzerinde topladı.
         'The model attracted attention with her natural beauty and warm demeanor.'

This tendency might reveal an interesting pattern about the discourse structure in Turkish. Since Turkish allows argument drop, it may be more convenient (as it decreases ambiguity as to which referent is dropped) to link the new sentence to the already available discourse when speakers are continuing their utterances with the subject of the available sentence. Similarly, starting a sentence that is not connected to the available sentence by an overt discourse connector or prosodic means might be enabling the speaker to indicate that the focus of attention has shifted from the subject to the object.[9] The remention of the subject entity as a null pronoun has been observed previously (Turan 1996; Hoffman 1998), but the observation that the remention of the subject entity is linked to the available sentence via connectors (especially *and*) is a novel observation. Thus, the hypothesis is that the use of *and* following the discourse-introducing sentence in a similar task to the one we described above should lead the participants to remention the subject more often than the object in their subsequent discourse (see also Holler & Suckow 2016 for discussion). We are currently adapting the material here to test this possibility.

Also, several other factors have been observed to weaken the subject-bias in previous studies. It has been observed that not all verb types lead to the referential persistence of the subjects in subsequent discourse. For instance, verbs of transfer in English were reported to lead to an unexpected object-bias in speakers' discourse continuations (Stevenson et al.

---

[9] In Turkish, declarative intonation has a sentence-final falling boundary. Discourse subordination is signaled by a rising intonation at the end of a sentence.

1994; Arnold 2001). We used ditransitive verbs in our study, which might have acted in the same way. Using ditransitive verbs increases the number of arguments, which might in turn have increased the referential persistence of the object referents. In our test items, the manipulated referent appeared as the direct object in a ditransitive construction where the order of referents was subject, indirect object, and direct object (e.g., *modacı izleyenler-e manken(-i) tanıttı / designer audience-Dat model(-Acc) introduced*). The fact that there is another intervening referent with a full NP (i.e. *the audience*) between the subject and the direct object might have reduced the referential persistence of the subject. This is in line with studies reporting a recency effect in anaphora resolution (e.g., Kim & Kaiser 2013; c.f., Cowles et al. 2007). We think that the intervening indirect object (i.e. the third referential NP between the subject and the direct object) might have increased the recency effect (as in the *Dependency Locality Theory*, which underlines the effect of an intervening referential NP on the ease of integration of these NPs into the structure, Gibson 1998).

Regarding the DSP effect of the accusative case, we did not find any significant effects of the accusative case on the object rementions. This by no means suggests anything conclusive about the DSP account. One possibility is that overall object-bias might have prevented us from seeing any possible discourse-structuring effects of the accusative case. In the case of Romanian, the *-pe* marking on direct objects diminished the likelihood of the remention of the subject argument in speakers' upcoming discourse (Chiriacescu & von Heusinger 2011, von Heusinger & Chiriacescu 2013, for similar findings for the German indefinite *dies* and for the English indefinite *this*, see Deichsel & von Heusinger 2010 and Chiriacescu 2014, respectively). That is, there was an overall subject-bias in these previous studies and the related markers increased the DSP of the object relative to cases where the object NP appeared with no markers. In our case, there is already a predominant selection of the object referent, which might have masked any possible effects of the case marking. In this connection, future studies should ensure cases where the subject is the more dominantly rementioned entity rather than the object. We suggest that presenting the same material with discourse connectors (*and, but, because*) might increase the overall subject-bias in the responses, thereby enabling one to test the effect of the case marking in contexts more comparable to the previous studies in terms of an overall subject-bias.

As for the salience structure, the participants referred to the subject with null pronoun rather than overt pronoun. They referred to the object entity using overt lexical items by either repeating the initial lexical item or adding a modifying expression that would increase the salience of this rementioned entity. This is in line with previous studies (Turan 1996; Hoffman 1998) and with our expectations. Centering Theory suggests that the means of referring to the less salient entities in the discourse serve the function of increasing the salience of these entities for the audience. Thus, one reason why our participants used more overt referential forms while referring to the object might be to signal that they have shifted the center of attention to this less salient entity (i.e. one that would not normally be expected in the center of attention) in a manner disambiguating the rementioned referent from the subject that would normally be expected in the center of attention.

Finally, in our experiment, objects were more frequently rementioned while being treated as a more salient entity via reduced referential forms. In other words, the object entity was

referentially more persistent in the upcoming discourse despite being less salient. This leads us to conjecture that the forward-looking discourse properties of nominal expressions come in two distinct categories: (i) the discourse structuring potential of a nominal, which determines the referential persistence of the upcoming discourse regarding which entities will reoccur in the discourse and (ii) the level of salience, which determines the form of anaphoric expressions that target a nominal. Interestingly, while objects were higher on the first category described here, subjects took the lead in the second one in our study.

7. Conclusion

To sum up, we questioned whether the differential object marker on Turkish indefinite direct objects increases the referential persistence of the object referent and the salience structure of the discourse. We did not find any effect of the DOM marker on the likelihood of object rementions in participants' continuations in our story completion task. In previous studies, the DOM marker counteracted the overall subject-bias, increasing the likelihood of rementioning the object. In our study, the participants predominantly tended to remention the object, which we think might have masked the overall effect of the case marking. Subject rementions tended to appear as a connected sentence to the referent-introducing discourse in comparison to object rementions. This led us to speculate that using discourse connectors such as *and* following the referent-introducing sentence might increase the subject-bias, hence enabling us to test the effect of the DOM marker in a context comparable to that of previous studies. As for the salience structure, we observed that the participants used more overt referential forms to refer to the object while using null pronouns to refer to the subject, which confirms that the referential persistence of a discourse entity and the level of its saliency are two distinct properties, as recently discussed by Kehler & Rohde (2013).

References

Arnold, Jennifer E. 1998. *Reference form and discourse patterns*. Stanford, CA: Stanford University dissertation.

Arnold, Jennifer E. 2001. The effect of thematic roles on pronoun use and frequency of reference continuation. *Discourse Processes* 31(2). 137–162.

Arnold, Jennifer E. 2003. Multiple constraints on reference form: Null, pronominal, and full reference in Mapudungun. In John W. Du Bois, Lorraine E. Kumpf, & William J. Ashby (eds.), *Preferred Argument Structure: Grammar as Architecture for Function*, 225–245. Amsterdam: Benjamins.

Bittner, Maria. 2012. Perspectival discourse referents for indexicals. In Hannah Green (ed.), *Proceedings of SULA 7*, 1-22. Amherst, MA: GLSA.

Bock, J. Kathryn, & Richard K. Warren. 1985. Conceptual accessibility and syntactic structure in sentence formulation. *Cognition* 21. 47–67.

Brennan, Susan. 1995. Centering attention in discourse. *Language and Cognitive Processes* 102. 137–167.

Chiriacescu, Sofiana. 2014. *The discourse structuring potential of indefinite noun phrases. Special markers in Romanian, German and English*. Stuttgart: Universität Stuttgart dissertation.

Chiriacescu, Sofiana, & Klaus von Heusinger. 2010. Discourse prominence and *pe*-marking in Romanian. *International Review of Pragmatics* 2(2). 298–332.

Cowles, H. Wind, Matthew Walenski, & Robert Kluender. 2007. Linguistic and cognitive prominence on anaphor resolution: Topic, contrastive focus and pronouns. *Topoi* 26. 3–18.

Dede, Müserref. 1986. Definiteness and referentiality in Turkish verbal sentences. In Dan I. Slobin, & Karl Zimmer (eds.), *Studies in Turkish Linguistics*, 147–163. Amsterdam: Benjamins.

Deichsel, Annika, & Klaus von Heusinger. 2011. The cataphoric potential of indefinites in German. In

Iris Hendrickx, Sobha Lalitha Devi, Aantónio Branco, & Ruslan Mitkov (eds.), *Anaphora and Reference Resolution.* Lecture Notes in Artificial Intelligence, Vol. 7099, 144–156. Heidelberg: Springer.

Enç, Mürvet. 1991. The semantics of specificity. *Linguistic Inquiry* 22. 1–25.

Erguvanlı, Eser Emine, & Karl Zimmer. 1994. Case marking in Turkish indefinite object constructions. In *Proceedings of the Twentieth Annual Meeting of the Berkeley Linguistics Society: General Session dedicated to the Contributions of Charles J. Fillmore*, 547–552. Berkeley, CA: Berkeley Linguistics Society.

Ertel, Suitbert. 1977. Where do the subjects of sentences come from? In Sheldon Rosenberg (ed.), *Sentence Production: Developments in Research and Theory*, 141–167. Hillsdale, NJ: Lawrence Erlbaum Associates.

Farkas, Donka. 1994. Specificity and scope. In Lea Nash, & Georges Tsoulas (eds.), *Actes du Premier Colloque Langues & Grammaire*, vol. 1, 119–137. Paris: Université Paris-8.

Gernsbacher, Morton Ann, & Suzanne Shroyer. 1989. The cataphoric use of the indefinite *this* in spoken narratives. *Memory & Cognition* 17(5). 536–540.

Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68. 1–76.

Givón, Talmy. 1983. Topic continuity in discourse: An introduction. In Talmy Givón (ed.), *Topic Continuity in Discourse: A Quantitative Cross-Language Study*, 1–42. Amsterdam: Benjamins.

Grosz, Barbara, Aravind Joshi, & Scott Weinstein. 1995. Centering: A framework for modelling the local coherence of discourse. *Computational Linguistics* 21. 203–226.

von Heusinger, Klaus. 2007. Accessibility and definite noun phrases. In Monika Schwarz-Friesel, Manfred Consten, & Mareile Knees (eds.), *Anaphors in Text: Cognitive, Formal and Applied Approaches to Anaphoric Reference*, 123–144. Amsterdam: Benjamins.

von Heusinger, Klaus, & Sofiana Chiriacescu. 2013. The Discourse Structuring Potential of Differential Object Marking. The case of indefinite and definite direct objects in Romanian. *Revue Roumaine de Linguistique* LVIII. 437–453.

von Heusinger, Klaus, & Jaklin Kornfilt. 2005. The case of the direct object in Turkish: Semantics, syntax and morphology. *Turkic Languages* 9. 3–44.

Hoffman, Beryl. 1998. Word order, information structure, and centering in Turkish. In Marilyn A. Walker, Aravind K. Joshi, & Ellen F. Prince (eds.), *Centering in Discourse*, 253–271. Oxford: Oxford University Press.

Holler, Anke, & Katja Suckow. 2016. How clausal linking affects noun phrase salience in pronoun resolution. In Anke Holler, Christine Göb, & Katja Suckow (eds.), *Empirical Perspectives on Anaphora Resolution*. Berlin: de Gruyter.

Ionin, Tania. 2006. *This* is definitely specific: Specificity and definiteness in article systems. *Natural Language Semantics* 14. 175–234.

Johnson-Laird, Philip. 1968. The choice of passive voice in a communicative task. *British Journal of Psychology* 59. 7–15.

Kaiser, Elsi. 2011. Focusing on pronouns: Consequences of subjecthood, pronominalisation, and contrastive focus. *Language and Cognitive Processes* 26(10). 1625–1666.

Kamp, Hans. 2014. Dividing the province of indefinite noun phrase uses into three parts. Ms. Universität Stuttgart / University of Texas at Austin.

Karttunen, Lauri. 1969/1976. Discourse referents. In *International Conference on Computational Linguistics COLING* 70, 1–37. Stockholm: Research Group For Quantitative Linguistics. Reprinted in: James D. McCawley (ed.). 1976. *Syntax and Semantics*, 363–385. New York: Academic Press.

Kehler, Andrew, Laura Kertz, Hannah Rohde, & Jeffrey L. Elman. 2008. Coherence and coreference revisited. *Journal of Semantics* 25. 1–44.

Kehler, Andrew, & Hannah Rohde. 2013. A probabilistic reconciliation of coherence-driven and centering-driven theories of pronoun interpretation. *Theoretical Linguistics* 39. 1–37.

Kelepir, Meltem. 2001. *Topics in Turkish syntax: Clausal structure and scope*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Kılıçaslan, Yilmaz. 2006. A situation-theoretic approach to case marking semantics in Turkish. *Lingua* 116. 112–144.

Kim, Lucy Kyoungsook & Elsi Kaiser. 2009. Effects of honorific agreement on null subject interpretation in Korean. In *Proceedings of 2009 Seoul International Conference on Linguistic Interfaces*, Seoul: Yonsei University.

Kim, Lucy Kyoungsook, & Elsi Kaiser. 2013. Asymmetrical effects of topic-marking on discourse processing and memory retention. In Umut Özge (ed.), *Proceedings of the 8th Workshop on Formal Altaic Linguistics, WAFL 8*, 199–211. Cambridge, MA: MIT Working Papers in Linguistics.

Kornfilt, Jaklin. 1997. *Turkish*. London: Routledge.

MacWhinney, Brian. 1977. Starting points. *Language* 53(1). 152–168.

Milsark, Gary. 1974. *Existential sentences in English*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Nilsson, Birgit. 1985. *Case marking semantics in Turkish*. Stockholm: University of Stockholm dissertation.

Osgood, Charles Egerton, & J. Kathryn Bock. 1977. Salience and sentencing: Some production principles. In Sheldon Rosenberg (ed.), *Sentence Production: Developments in Research and Theory*, 89–140. Hillsdale, NJ: Lawrence Erlbaum Associates.

Özge, Umut. 2011. Turkish indefinites and accusative marking. In Andrew Simpson (ed.), *Proceedings of the 7th Workshop on Altaic Formal Linguistics, WAFL 8*, 253–267. Cambridge, MA: MIT Working Papers in Linguistics.

Özge, Umut. 2013a. "Strong" indefiniteness and topicality. In Emmanuel Chemla, Vincent Homer, & Grégoire Winterstein (eds.), *Proceedings of Sinn und Bedeutung 17*, 399–408. Paris. http://semanticsarchive.net/sub2012/.

Özge, Umut. 2013b. What does it mean for an indefinite to be presuppositional? In Guram Bezhanishvili, Sebastian Löbner, Vincenzo Marra, & Frank Richter (eds.), *Logic, Language, and Computation - 9th International Tbilisi Symposium on Logic, Language, and Computation, TbiLLC 2011, Kutaisi, Georgia, September 26-30, 2011, Revised Selected Papers*, 138–154. Berlin: Springer.

Pesetsky, David. 1987. Wh-in-situ: Movement and unselective binding. In Eric J. Reuland, & Alice G. B. ter Meulen (eds.), *The Representation of (In)definiteness*, 98–129. Cambridge, MA: MIT Press.

Prince, Ellen F. 1981. On the inferencing of indefinite-*this* NPs. In Bonnie Webber, Aravind K. Joshi, & Ivan Sag (eds.), *Elements of Discourse Understanding*, 231–250. Cambridge: Cambridge University Press.

Runner, Jeffrey, & Alyssa Ibarra. 2016. Information structure effects on null and overt subject comprehension in Spanish. In Anke Holler, Christine Göb, & Katja Suckow (eds), *Empirical Perspectives on Anaphora Resolution*. Berlin: de Gruyter.

Schumacher, Petra B., Manuel Dangl, & Elyesa Uzun. 2016. Thematic role as prominence cue during pronoun resolution in German. In Anke Holler, Christine Göb, & Katja Suckow (eds.), *Empirical Perspectives on Anaphora Resolution*. Berlin: de Gruyter.

Stevenson, Rosemary, Rosalind Crawley, & David Kleinman. 1994. Thematic roles, focus, and the representation of events. *Language and Cognitive Processes* 9(4). 519–548.

Turan, Ümit Deniz. 1996. *Null vs. overt subjects in Turkish discourse: A centering analysis*. Philadelphia, PA: University of Pennsylvania dissertation.

Turan, Ümit Deniz. 1998. Ranking forward looking centers in Turkish: Universal and language specific properties. In Marilyn A. Walker, Aravind K. Joshi, & Ellen F. Prince (eds.), *Centering in Discourse*, 139–160. Oxford: Oxford University Press.

Zidani-Eroğlu, Leyla. 1997. *Indefinite noun phrases in Turkish*. Madison, WI: University of Wisconsin-Madison dissertation.