Interpretation and Production of Pronouns in Turkish

by

Nehir Aygül

Submitted to the Department of Linguistics in partial fulfilment of the requirements for the degree of

Master of Arts

at

Goethe University Frankfurt

6 April 2021

Supervisor: Prof. Dr. Markus Bader

Second reader: Dr. Yvonne Portele

Abstract

Many studies on pronoun resolution have investigated the interpretation of pronouns in implicit causality contexts and more recently in transfer-of-possession verb contexts and found that grammatical role alone do not suffice to account for the interpretation of pronouns. That is, syntactic functions are not the sole determinant of an antecedent of a pronoun. Instead, a combination of grammatical and semantically/pragmatically driven factors has been found to influence pronoun interpretation (Rohde & Kehler, 2014). Pronoun production, however, has been claimed to be insensitive to semantically/pragmatically driven factors. This thesis presents two passage completion experiments: Experiment 1 investigates the production of referential expressions in implicit causality contexts with subject-experiencer and object-experiencer verbs, and Experiment 2 analyses both interpretation and production of forms in transfer-of-possession contexts by manipulating prompt and verb-type. Results revealed that null pronouns in Turkish are used mainly for referring back to subject antecedents, but the interpretation of nulls is also modulated by semantic/pragmatic biases. Although overt pronouns may be interpreted as referring to either the object or the subject, they are not preferred in reference back to the object and instead proper names are favoured. This casts doubt on a division of labour between nulls and overts in Turkish. In addition, verb type had an effect across the board: in both experiments participants tended to re-mention the object antecedent more often when it was the stimulus or the goal referent. In order to reveal whether null and overt pronouns in Turkish pattern similar to those in other languages, the findings are discussed in light of previous cross-linguistic empirical work on pronouns.

Acknowledgement

I would like to take this opportunity to express my deepest gratitude to the individuals that have contributed to make this thesis and experimental work possible and to the people that helped me to stay sane during the difficult times of COVID-19 pandemic.

Firstly, I would like to thank my supervisor Markus Bader for his full-time guidance and support throughout the writing process and for helping me out with the analyses. I would also like to express my most sincere appreciation to my second reader, Yvonne Portele, whom I look forward to learning more and collaborating in projects for years to come.

Secondly, to my parents and my sister Deniz, for being my emotional anchors and my role models by having the courage to embark on new adventures. Had it not been for your constant and unconditional love and support, I would have never taken the decision to come to Germany to pursue my academic ambitions. Thank you for taking living seriously as the verses by Nazım says:

> "[i] mean, you must take living so seriously that even at seventy, for example, you'll plant olive trees — and not for your children, either, but because although you fear death you don't believe it, because living, I mean, weighs heavier…." (Nazım Hikmet Ran, *On living*)

Last but not least, I would like to thank my wonderful colleagues Nasim Bahmanian and Sebastian Buchczyk, and my friends in Turkey (Hazal, Ömer and Dilin) for believing in me sometimes more than I do myself. I was also very privileged to have received valuable remarks and suggestions from Geena Carlisle while she was thousands of miles away.

Contents

List of Tables

List of Figures

Abbreviations

ABL	ABLATIVE CASE
ACC	ACCUSATIVE CASE
CAUS	CAUSATIVE MORPHEME
CG	COMMON GROUND
DAT	DATIVE CASE
GEN	GENITIVE CASE
IC	IMPLICIT CAUSALITY
IMPF	IMPERFECTIVE
INS	INSTRUMENTAL CASE
NEG	NEGATION
NOM	NOMINATIVE CASE
NMLS	NOMINALISER
NP	NOUN PHRASE
PF	PERFECTIVE
PLU	PLURAL
POSS	POSSESSIVE
PRES	PRESENT TENSE
PROG	PROGRESSIVE ASPECT
PST	PAST TENSE
1SG	FIRST PERSON SINGULAR
2SG	SECOND PERSON SINGULAR
3SG	THIRD PERSON SINGULAR
2PLU	SECOND PERSON PLURAL

Chapter 1

Introduction

Languages provide speakers with the tools for creating sequences that are meaningful and coherent. Successful communication depends on the references established in the discourse. These references, be it referring back to events, entities or full propositions, need to be assigned correctly by the speaker. S/he must assume what the audience knows. As for the audience, the task is to draw inferences from what they have previously read or heard in order to interpret an appropriate reference. This thesis focuses on a particular type of reference, that is established by pronominal forms, and aims to investigate the factors that influence the production and interpretation of pronouns.

Studies on discourse anaphora have investigated how a pronoun in an ambiguous context, in which two antecedents are equally available, is resolved. The pronoun in such contexts is assumed to be generally resolved towards the most prominent referent in the discourse. However, what determines prominence is subject to much debate. Factors, such as grammatical role, thematic role, verb-type, information structure, coherence relations, and selected forms have been argued to influence pronoun interpretation. Over the last decades, a great deal of evidence that suggests pronoun interpretation is influenced by an interaction of multiple factors, rather than a single factor, has been accumulated. The empirical contribution of this thesis is to investigate these factors for Turkish, a typologically different language that has overt pronouns as well as null pronouns. To this end, two experiments are carried out in order to address the factors that influence production and interpretation of null and overt pronouns in Turkish.

The effects of the above-mentioned factors are subject to cross-linguistic variability. Therefore, by testing Turkish, I probe the universality of the mechanisms involved in reference production and interpretation with a focus on null and overt pronouns. A great majority of studies examining null pronouns have analysed nulls in several Romance languages, as well as in Korean and Japanese. Thus, to analyse whether nulls and overts in Turkish behave similarly to those of other previously studied languages will be one of the main goals of this thesis.

One commonly held conjecture in the literature on anaphora is that explicit forms are used for less prominent referents. Therefore, reduced forms such as null pronouns are argued to refer to prominent entities. Since the factors that determine prominence are a matter of debate, what guides speakers to choose a certain form over an alternative one remains a puzzle. In addition, recent studies have argued that prominence for pronoun interpretation is not identical to that of production (Kehler & Rohde, 2013). Consequently, these studies claim that there are two different senses of prominence, and that each one is affected by a unique set of factors (see Rohde, 2019: for an overview). According to such studies, both semantic-pragmatic and information-structural factors influence interpretation, but only the latter influences production. In this thesis, I empirically test this claim by examining whether production is sensitive to verb-type, coherence-relations and/or subjecthood.

The structure of the current paper is as follows. Chapter (2) gives an overview of the relevant literature on pronoun interpretation and production. It will specifically address the factors that influence these two processes, and presents models and hypotheses accounting for the distribution of referential forms. In the first part of Chapter 2, the factors that are argued to influence the choice of referential form are discussed. In particular, the idea that entity prominence closely correlates with the choice of form, and that this correlation can be captured by hierarchies of explicitness that relate expressions to hierarchies of cognitive prominence, is discussed.

Chapter (3) provides a brief outlook of the properties of Turkish pronominal forms. It presents the linguistic environments that allow for pro-drop, namely subject and non-subject drop. It provides background on theoretical work especially pertaining to the relationship between information structure and pronouns. Then, it discusses several related experimental studies on referential forms in Turkish.

Chapter 4, which consists of two parts, is devoted to empirically testing the reported patterns in Turkish. The first part introduces Experiment 1, which uses a sentence completion task to test implicit causality biases in production. Even though pronoun resolution in Turkish has been a theoretical and also experimental concern of several authors, production of referential forms in implicit causality contexts has not been extensively studied. Thus, in Section 4.1, I investigate these contexts and discuss the findings in light of previous literature from Turkish as well as from a broad range of other languages.

In the second part of Chapter 4, Experiment 2 is presented. This study employs the same paradigm as in Experiment 1, with the addition of different prompt types. Analysing prompts that involve pronouns, namely overt and null pronouns, and contrasting these conditions with the free prompt that lacks a pronoun, allows for an examination of not only the behaviour of but also the interpretation of pronouns. With this design, Experiment 2 puts the proposed asymmetries concerning pronoun production and interpretation to test (Rohde & Kehler, 2014).

Having provided the results from each experiment and discussed them in light of previous research presented, Chapter 5 concludes the present thesis by drawing general conclusions

and by stating the implications of the findings. Besides, directions for future endeavours are outlined by pointing out multiple issues that are yet to be resolved.

Chapter 2

An overview of previous literature

This chapter provides a summary of the factors that guide pronoun production and/or comprehension and presents empirical data for these factors. Moreover, several models that account for the role of these factors in interpretation and production are considered.

Section 2.1 is a general introduction to the first aspect of the current thesis, production. Following a discussion of what the process of reference production involves, the subsection 2.1.1 presents the models that capture the interaction between referential form and prominence. Section 2.2 examines the second aspect, interpretation, and offers a glimpse into how listeners interpret reference established by the speakers. Next, the factors at play in comprehending and producing anaphora are discussed, and empirical data for such factors are presented in Section 2.3. Before the interim summary of the chapter in Section 2.5, an overview of the models and theories that account for the phenomena observed in previous studies is given in Section 2.4.

2.1 What governs reference production?

In a conversational exchange, speakers establish links between sentences or propositions and create meaningful texts with ease and fluency. Nevertheless, the workings of this seemingly effortless process are not so simple to explain.

One of the integral parts of the task is to map a word or expression to a concrete object or concept, also called a referent. First of all, the speaker or writer[1] chooses the form with which s/he will introduce the referent into the discourse[2]. Languages provide speakers with a range of referential forms to choose from when referring to an entity. There is a decision process behind the production of every referential expression. For example, the expressions below in (1) have the same referent, the author of this thesis. Thus, when introducing a referent, the speaker has to make a selection between a number of potential forms.

---

[1]Henceforth, I use *speaker* interchangeably to refer to both *speakers and writers*, and *listeners* to refer to *listeners and readers*.

[2]Throughout the thesis, the terms *referential forms* or *referential expressions* are used to indicate the words or expressions that stand for things in the real or imaginary world, which include proper names (Nehir, Jim, Susan etc.) and pronouns (He, she, they etc.).

(1)   a. Nehir

     b. The person behind the monitor writing this example

     c. The person that is sitting on a yellow armchair typing this example (1c) in the midst of the corona pandemic

There is no language, to the best of my knowledge, that does not have any proper names to refer to people (e.g. Nehir). Forms that refer to entities are not universal, they differ from language to language. They can be constrained by morpho-syntactic information such as gender, person or number. For instance, it would be infelicitous for speakers to use a masculine personal pronoun such as *He* or a plural pronoun such as *They* (without introducing a second character to the discourse) for the referent in (1).

     Some languages provide different tool-sets for referring than that of English. For example, in pro-drop languages, such as Spanish, it is possible to establish a reference without overtly mentioning the referent because agreement on verbs (number and person) provide sufficient information. Below, (2) is an instance of pro-drop.

(2)   Lo veo
     it   see.1SG
     'I see it'

The pronoun "Yo" (*I*) is dropped entirely, and yet reference to the person uttering the sentence can, nevertheless, be processed and interpreted in the mind of the listeners. In other words, in languages that provide the tool for dropping arguments, it is possible, in most of the cases, to understand which entity the speaker refers to, irrespective of whether the entity is overtly mentioned or not.

     Once an entity is introduced to the discourse, speakers often need to refer back to them in subsequent sentences. This task, the task for re-mentioning, again requires selection among different forms. This time, it is to decide on an expression to maintain the previously introduced referent(s). Upon closer scrutiny, we can see that sometimes some forms are more fitting or appropriate than other variants for re-mentioning. For example, shorter forms, pronouns or dropped arguments as the one in (2), are preferred over others in maintaining reference. Consider the passage by Woolf (2000: 74) below:

(3)   *"Other people have faces; Susan and Jinny have faces; they are here. Their world is the real world. The things they lift are heavy. They say Yes, they say No; whereas I shift and change and am seen through in a second. If they meet a housemaid she looks at them without laughing. But she laughs at me. They know what to say if spoken to. They laugh really; they get angry really; while I have to look first and do what other people do when they have done it."*

In this paragraph, there are proper names *Susan* and *Jenny*, full noun phrases that denote a (kind of) entity (a housemaid) and many expressions establishing reference back to entities

such as *they, I* or *she.* Yet, there seems to be constraints on which form the speaker should use. For instance, if we were to exchange the third person plural pronoun (*they*) in the paragraph with *Susan and Jinny*, the text would be unnecessarily wordy and repetitive. This implies the existence of restrictions on what counts as an appropriate form, and how "explicit" a form should be (Arnold, 2010).

For a successful communicative exchange to take place, speakers need to shape their utterances in a certain way to be understood by the other listener. Accordingly, speakers' task is not only to choose forms that are concise enough to avoid any redundancies or verbosity but also to provide enough information for the addressee to identify the referent (Ariel, 1990; Chafe, 1976; Givón, 1983; Gundel et al., 1993). Some cases, especially when 3rd person pronouns (*he, she* and *it*) are used, result in ambiguity. Consider the sentence below (4) in which the pronoun (*he*) is ambiguous and can have co-reference with two potential antecedents, i.e. can stand for *John* or *Jim.*

(4)   John met Jim. He looked miserable.

In the same vein, pro-drop can lead to ambiguities, hence not only the issue of when speakers select zero forms but also which referent(s) the dropped argument typically refer to has been an interest of much work. So how exactly do speakers determine which form to utilise, and when do they decide to employ lexically richer forms such as full noun phrases (NPs) or reduced forms such as pronouns or null forms?

In an attempt to answer these questions, researchers turned towards the relationship between form and prominence. Many traditional approaches claim that speakers signal a referent to the listener by their choice of expressions, and this choice is dependent on the notion of *prominence.* In general, entities that stand-out in a discourse are prominent and are indicated by the form chosen by the speaker to help the listener to identify the referent correctly. Unsurprisingly, certain forms require a level of prominence in the discourse. For example, the pronoun in example (5) does not suffice to establish a reference without prior context[3] because a female entity has not been introduced to the discourse, thus no accessible antecedent is available.

(5)   John and Jim are cooking. She likes pancakes.

Throughout the decades, many scholars observed a correlation between referential form and prominence and proposed hierarchies of "explicitness" to capture this interaction (Arnold, 2010). These hierarchies arrange forms on a scale according to their levels of *explicitness* and map them to discourse properties of an entity (Ariel, 1990; Givón, 1983; Gundel et al., 1993: *inter alia*). Explicitness is constrained by discourse statuses of entities and non-linguistic constraints such as activation (Arnold & Griffin, 2007).

---

[3]We assume that the (5) occurs in isolation. The listener might, nevertheless, try to speculate a context in which there is an entity with +Female feature that *John* and *Jim* cook for. Under such speculative context the reference can be established. This shows us the need for coherence between structures which will be looked at in-depth in section 2.3.4

The phenomenon of prominence is operationalised under different notions such as *given-ness, accessibility, salience, topichood, center of attention* etc. Albeit not being identical terms, they relate and contribute to prominence. Numerous authors suggest that reduced forms (e.g. pronouns or zero forms) tend to be used for referents that are salient or prominent in discourse (Ariel, 1990; Givón, 1983; Gundel et al., 1993), and posited various factors such as thematic role, information structure semantics or world knowledge etc. to contribute to this notion of prominence. However, it remains an issue whether all of these factors influence both interpretation and production and whether determinants of prominence for anaphora interpretation and production are the same. Therefore, before delving into detail on what factors affect prominence, and whether they influence both production and interpretation (see section 2.1.1 for a discussion), we focus on the referential form and its relation to prominence and discuss the hierarchies for discourse prominence in the following subsection.

2.1.1 Cognitive prominence and referential form

Communication requires keeping "non-linguistic mental representations" of introduced characters and eventualities (Arnold, 2010: 188). This helps discourse participants, i.e. *inter-locutors*, to monitor what has been brought up to the discourse and the addressee to establish reference to the mental representations they have (Arnold, 2010). Yet not all information is equally prominent (or accessible in Arnold's (2010) terms[4]).

There have been theories to ground this notion of prominence on the degree of activation of an entity in the cognitive status of interlocutors (Ariel, 1990; Arnold, 2010; Grosz, Weinstein, & Joshi, 1995; Gundel et al., 1993). These theories take what is known both to the speaker and listener, that is the common ground, into consideration, and order linguistic expressions based on their information status.

One observed relation between linguistic expressions and information status is the one between pronouns/proper names and givenness/newness. Listeners need all the information they can get to create a representation for the newly introduced entity in their mental models. Thus, when a new entity is introduced to the discourse, descriptions or proper names are required (Chafe, 1976; Prince, 1981). In contrast, stating what already is in the common ground, *CG* (i.e. *given*) leads to uninformative expressions since the utterance does not provide any new information to the discourse. As a result, entities that are previously mentioned in the discourse and are salient, accessible, topical or highly activated, and therefore, they can be referred back to with pronouns or even be left out in an utterance (if the language allows dropping arguments) as in example (6). The first sentence in (6) introduces a new entity; hence an informative form, a name, is used. In the second sentence, the entity is re-mentioned by the use of a dropped pronoun. This is only allowed when the entity is salient in the discourse, and only with associated forms, existing referents are activated in mental

---

[4]According to Arnold (2010), accessibility does not stand for lexical accessibility but rather is a property of information.

representations of the listeners. In other words, the choice of form instructs the addressee on where to find the antecedent.

(6) Ali bir kitap satın al-dı.   Sonra $\emptyset_i$ evin-e   git-ti ve kitab-ı   okumaya
Ali$_i$ a   book buy get-PST. Then $\emptyset_i$ home-DAT go-PST and book-ACC reading
başla-dı.
begin-PST
'Ali bought a book. Then (He) went home and began to read the book.'

Initial work analysed the interaction between form and reference to interact with discourse status of givenness, that is given vs. new information (Chafe, 1976; Prince, 1981). Later in 1983, Givón proposed a theory for *topicality*[5] of a given entity. By departing from the traditional views of theme/rheme (given/focus) discussions carried out by the lead of Mathesius in Prague school, Givón (1983) maintains a cognitive approach to topicality in discourse. He claims that the choice of a referential expression is associated with this notion of topicality.

According to Givón (1983; 1992), each interlocutor seeks to create a mental model of the other's perspective. This, in turn, requires a need for the speaker to accommodate what s/he wants to say to what s/he thinks the hearer's perspective is. To do so, speakers make use of 'two quantifiable aspects of topicality' and their choice of form relies on prominence/topicality of entities. These aspects of topicality are *anaphoric* and *cataphoric*[6] one (Givón, 1992: 10). Anaphoric reference is defined as *referential accessibility* or *predictability* and cues the listeners to where to search for the referent in memory storage. On the other hand, cataphoric elements signal the importance of referents to the addressee for the addressee to activate the referent in the representation. In other words, speakers cue listeners to identify a reference towards the topical entity in discourse with the use of forms or 'referent-coding devices' (Givón, 1983).

In Givón's *Topic continuity theory*, choice of form depends on the degree of accessibility or continuity of an entity. This is measured with three factors: distance, decay and ambiguity. *Distance* is defined by how recent a topic NP is. That is measured by the number of clauses that appear in between an antecedent and referential form: the more recent an entity is, the more accessible it is. *Decay* is identified by the persistence of the presence of an entity in later discourse, i.e. whether it is mentioned at a later point in the discourse. Whereas *Ambiguity* is related to uniqueness of entities, that is whether there are any competing topics for the expression leading to multiple interpretations. Considering these factors, Givón (1983) subsequently provides the scale below for the link between referential form (referent-coding devices) and accessibility (referential continuity). He also shows that the scale holds for cross-linguistic data to a great extent (Givón, 1983; 1992).

---

[5]According to Féry and Krifka (2008: 126) *topic* describes an element that "identifies the entity or set of entities under which the information expressed in the comment constituent should be stored in the CG content". Hence, topics represent given or discourse-old information.

[6]In linguistics, anaphoric and cataphoric reference may be crudely defined as backward and forward reference respectively. Depending on the position of the entity they refer to, a pronoun can be cataphoric. That is, when the referent is introduced by a later expression, after the pronoun. E.g. 'If you are looking for *her*, *Professor X* is on the balcony.'

Figure 2.1: Referent-coding devices and referential continuity (simplified version of the figure by Givón 1983: 17)

Accessible/Continuous Topics ⟵		*Discontinuous/Inaccessible Topics* ⟶		
Zero anaphora	Unstressed Anaphoric pronoun	Stressed independent pronouns	Definite NPs	Indefinite NPs

According to the hierarchy *zero* signals the highest referential continuity which is an anaphoric relation concerning predictability. On the other hand, NPs in this hierarchy codes topics that are low in continuity, that are inaccessible. For the scale to make better predictions across languages, he additionally provides scales for specific syntactic properties of given languages, and takes word-order, morphology, intonation and phonological size into consideration. Besides, he stresses the role grammatical relations play. In particular, he maintains that subjects are more topical than direct objects which are more topical than the indirect object (Givón, 1983: 22), as indicated by the scale below in (7).

(7) SUBJECT > DIRECT OBJECT > OTHERS

He also states the need for a semantic scale that captures the hierarchy of case-roles, that correlates with the scale for grammatical roles[7] (See the scale in 8 adapted from Givón 1983: 22). We defer the discussion of the special status of the subject/agents for the time being and will discuss it in more detail in section (2.3.1).

(8) AGENTS > DATIVE/BENEFACTIVE > ACCUSATIVE > OTHERS

An independently motivated taxonomy, rather similar to Givón's (1983) was later developed by Ariel (1988). The difference between these two models is the term used to characterise the distribution of referential expressions. In contrast to Givón's (1983) topicality, Ariel (1988) argues for a model built around the notion of *Accessibility*. This theory computes the degree of activation of an entity in memory. The continuum of activation corresponds with how accessible an entity is in the mental representations of the discourse participants. In her theory, she uses the term 'accessibility markers' to refer to referential expressions, because she views anaphoric forms as markers that tell how accessible an entity is. Such markers guide the listeners in retrieving the correct mental representation (Ariel, 2001). Accessibility is assessed through four factors: topichood, inference, recency and role of frames (Ariel, 1988: 65)[8]. Depending on the level of accessibility, it is mapped to a referring expression on a scale.

---

[7]Case roles are susceptible to grammatical relations as there is a universal tendency for agents to be subjects and dative cased entities to be direct objects (Givón, 1983: 22)

[8]The first three factors correspond to Givón's (1983) *Decay*, *Ambiguity* and *Recency* respectively.

A simplified version of the scale, *The Accessibility Hierarchy*, proposed in Ariel (1990: 73) is provided below in Figure 2.2.

Figure 2.2: THE ACCESSIBILITY HIERARCHY (Ariel, 1990: 73)

Low Accessibility		$\longleftarrow$			$\longrightarrow$	High Accessibility	
Full name	Definite description	Distal DEM	Proximate DEM	Stressed pronoun	Unstressed pronoun	Cliticised pronoun	Zero pronoun

In this scale, referential expressions are arranged according to their activation levels, that is the values of accessibility in mental representations of the individuals who are partaking in a discourse. The higher the activation is, the easier it is to retrieve a representation. In accordance with the model by Givón (1983), speakers choose a particular form based on their predictions of the listener's cognitive state. In other words, speakers have to make inferences on the information available to the addressee and select a form accordingly. The categories to the left of the scale are used to refer to entities that are low in accessibility. On the other hand, speakers prefer reduced forms, such as cliticised pronouns or zeros, over full forms when establishing reference to more accessible discourse entities.

In a similar vein, Gundel et al. (1993) probe the distribution and interpretation of referential forms used in natural language discourse, and introduce a hierarchy. In comparison to the hierarchies by Ariel (1990) and Givón (1983), their hierarchy is one of mental states rather than "a hierarchy of degrees of accessibility" (Gundel, 2010: 148). They argue that different forms "[c]onventionally signal different (assumed) cognitive statuses" (Gundel et al., 1993: 274), and further classify six different cognitive statuses proposing the givenness hierarchy shown in Figure 2.3.

Figure 2.3: THE GIVENNESS HIERARCHY (Gundel et al., 1993: 275)

in focus	> activated	> familiar	uniquely > identifiable	> referential	type > identifiable
{*it*}	{*that, this*}	{*that N*} {*this N*}	{*the N*}	{*this$_{indef}$ N*}	{*a N*}

In this hierarchy, referents to the left rank higher in givenness and are at the center of attention (Gundel et al., 1993). These entities are commonly referred to by zero or unstressed pronoun forms. As opposed to the other two models, this is a true hierarchy. The categories that appear to the left in this hierarchy entail the ones to its right, hence there is no one-to-one mapping between form and cognitive status but rather a one-to-many mapping. For example, an entity that is in focus can be referred to by forms that are activated or familiar. As a result, the selection of forms is more flexible for given entities (entities that appear to left

of the hierarchy compared to the ones on the right). However, this relation is unidirectional, for example not all of the entities that are referential are per se familiar or activated.

The hierarchy they introduce interacts with other linguistic and non-linguistic principles that contribute to production and interpretation of referential expressions. One such principle is a Gricean Maxim: *Maxim of quantity* (Grice, 1975). *Maxim of quantity* is a conversational implicature that regulates informativeness. It states that contributions must be as informative as required but not more than necessary (Grice, 1975: 45). This maxim consequently places constraints on the appropriateness of an anaphoric form. When the communicative goal is to provide information about the familiarity of an entity to the hearer then the choice of form should follow from the hierarchy and the speaker chooses the left most (most restricting, most informative) option for an entity. What follows from this that the category uniquely identifiable, mapped to indefinite article *the*, implies unfamiliarity of the referent. This is due to the fact that anything that is familiar should be indicated by the left most option, and hence cannot be uniquely identifiable.

At the same time, the second part of Maxim of quantity (i.e. do not provide more or less information than needed) leads to a dissociation in the notions of cognitive state and familiarity (or accessibility). In other words, what is cognitively familiar does not necessarily need to be indicated in the form (by using *that/this N*) when an entity's familiarity is irrelevant for the conversation. This follows from the Maxim of quantity: speakers provide as much information as is needed in a communicative exchange, and in some contexts, giving information on how familiar an entity is might be more informative than needed. In her paper, Gundel (2010) presents the dialogues below by Gundel and Mulkern (1998) to illustrate this point.

(9)   Nancy: Where do you want to eat tonight?

     a. Silas: Let's go to the Italian restaurant on Poland Street.

     b. Silas: Let's go to that Italian restaurant on Poland Street.

In these dialogues, it may be irrelevant whether the restaurant is familiar (*that restaurant*) to Nancy or not since *Nancy* asks for where the addressee (*Silas*) wants to eat. Therefore, the choice of form does not necessarily entail how familiar the restaurant is to Nancy, and both (9a) and (9b) signal sufficient information for Nancy to pick up the referent. This provides evidence for the claim that the hierarchy itself cannot explain all instances of referential form, and other principles such as pragmatic principles interact with the hierarchy (Gundel, 2010).

To sum up, the different models all share the idea that the selection of forms is an addressee oriented process. Certain forms are chosen to aid the speaker in identifying the correct antecedent. Each model distinguishes referential forms according to their levels of prominence, albeit addressing the issue under different notions (e.g. activation, accessibility, topicality, cognitive or informational status). One important difference between these hierarchies is whether the relationship between cognitive prominence and the use of a given referential form is pre-defined or not. Although, in some hierarchies prominence of an entity

is assessed with respect to the prominence of other competing entities (Ariel, 1990; Givón, 1983; Grosz et al., 1995), in others, for instance the *Givenness hierarchy* (Gundel et al., 1993), the relationship is static and non-relational (von Heusinger & Schumacher, 2019). In such static hierarchies, there is no ranking of potential referents, and a particular referential form can only be used when the required cognitive status of the referent (on a fixed scale of statuses) is met. Therefore, as opposed to the Givenness Hierarchy (Gundel et al., 1993) which assigns referents to mutually exclusive cognitive statuses, both the Topic Continuity theory (Givón, 1983) and the Accessibility Hierarchy (Ariel, 1990) arrange referential forms according to their levels of prominence on a continuous hierarchy of statuses.

One shared characteristic of these theories is that they are argued to hold across languages while making different predictions for anaphoric devices available in a given language. In all languages, less informative anaphoric forms of are used for prominent referents in the preceding discourse. This claim has been supported by experimental work. For instance, Gordon, Grosz, and Gilliom (1993) observed a penalty for repetition of a full proper names or NPs in English. Their findings show that repetition of a name leads to longer reading times when the name retrieved a subject antecedent. Given that they compared identical passages and solely manipulated the referential expression (name vs. a pronoun), they argue that the processing difficulty could not have emerged from the form of the antecedent or from the content of the sentences, but rather emerges from the referential form itself (i.e. from the repeated name). Similar findings, supporting the claim that reduced forms are preferred for prominent entities, have been reported for Japanese (Shoji, Dubinsky, & Almor, 2017) and Spanish (Gelormini-Lezama, 2018).

Lastly, all of the models discussed acknowledge the role of other existing factors in determining the degree of prominence of an entity, such as grammatical role or information structure. In fact, there is supporting evidence for this argument, and subjects or topical referents are more likely to be pronominalised than non-subject/topic referents (Rohde & Kehler, 2014: *inter alia*). As for other languages, such as German and Finnish, similar observations have been made for the behaviour of personal pronouns as opposed to demonstrative pronouns and personal pronouns are reported to prefer subject antecedents whereas demonstratives prefer non-topics (Finnish: Kaiser and Trueswell 2008; German: Portele and Bader 2016; Bosch, Rozario, and Zhao 2003 for a corpus study).

Therefore, existing evidence suggests that there diverse factors determining prominence. However, before identifying some of the possible factors in section 2.3, we will first discuss what is involved in the interpretation of referential forms in the following section (2.2).

2.2 What governs reference interpretation?

To resolve a given referential form, listeners need to identify an appropriate antecedent (or postcedent[9]) to the forms they encounter. But what makes it possible to resolve or establish such pronominal dependencies?

First, pronouns, albeit carrying morpho-syntactic information such as person, gender and number[10] that helps to disambiguate referents in interpretation, do not provide much semantic information on their own. Therefore, listeners consequently make their inferences according to what they think the speaker established as prominent in the discourse. Yet, sometimes it is not easy to interpret pronouns and establish pronominal co-reference, especially when there are two plausible entities for reference. For instance, in example (10) below (adapted from Kehler 2002: 143), the pronoun *him* is ambiguous in both sentences and, hence, can refer to both *John* or *Bill*. However, when asking native speakers of English, the pronoun *him* in (10a) would be interpreted as referring to *John* (as indicated within square brackets). Traditionally, this has been associated with a pronoun resolution preference towards the subject of the preceding sentence (Crawley, Stevenson, & Kleinman, 1990; Walker, Iida, & Cote, 1996) or towards the first-mentioned noun phrase (NP1) (Gernsbacher & Hargreaves, 1988; Gernsbacher, Hargreaves, & Beeman, 1989). This preference in interpretation appears to be persistent even if we alter the structure of (10a) while keeping the propositional content constant. Consider the example (10b). It is structurally different than (10a): the position and the grammatical role of the two NPs (*John* and *Bill*) is reversed, yet the interpretation preference is again towards the subject or the first-mentioned NP. As a result, the pronoun *him* is interpreted as referring to *Bill* in (10b).

(10)   *Subjecthood/First-mention Preference* (adapted from Kameyama 1995: 13)

    a.   John kicked Bill. Mary told him to go home. [= John]

    b.   Bill was kicked by John. Mary told him to go home. [= Bill]

However, the strategy to assign pronouns to subjects or first-mentioned NPs does not account for all cases. For example, although the pronoun *he* in (11) is again ambiguous and can refer to both *John* or *Mike*, studies show that readers or listeners prefer to resolve the pronoun towards *Mike* in (11a) and towards *John* in (11b) (Garvey & Caramazza, 1974: among others). This, therefore, suggests that grammatical or linear structure is not the sole factor in interpretation, and sometimes pronouns are preferentially resolved towards the object of the preceding clause, as in (11a) in contrast to (11b).

(11)   a.   John admires Mike because he... [= Mike]

    b.   John amazes Mike because he.. [= John]

---

[9]Although our focus in the present thesis is on anaphora, similar mechanisms are utilised for cataphora. Besides, the thesis is mainly concerned on personal pronouns and, therefore, demonstratives are not analysed in-depth.

[10]All of these factors depend on how the language at issue marks nominals.

Such examples were studied extensively in the linguistic literature under the term *implicit causality* (IC), and this divergence in interpretation is claimed to be related to the verb type (See Brown and Fish 1983; Garvey and Caramazza 1974; Stevenson, Crawley, and Kleinman 1994 for initial works on the topic). These implicit causality verbs are attributed with pronoun biases. Depending on the verb type *admire* or *amaze*, they cause discrete preferences in interpretation of pronouns, either for the first noun (John) or the second (Mike), yet this bias is not constant and can be subsumed or reversed when the clause connector is changed. It is observed that sentence connectors can be congruent or incongruent to the biases and result in an inconsistency in interpretation. Compare the example (12) below with the sentence (11a).

(12)   John admires Mike and he.. [= John]

The theories presented in the previous section (2.1.1) posit that an identical notion of prominence underlies both production and interpretation. According to these views, speakers use reduced forms to refer to entities that are prominent, and, in a similar fashion, prominent entities are preferred as antecedents of referential expressions. Given the examples above, it is clear that prominence cannot only be a static notion that is inherently coded in the semantics of entities[11], and several factors seem to be at play in defining what is prominent in a discourse. Therefore, the task is to define which factors contribute to the notion of prominence and whether the same factors are equally at play in both production and interpretation of pronouns.

While many authors argue that a similar set of factors influence both production and interpretation, Kehler, Kertz, Rohde, and Elman (2008) claim that production is not affected by some of the factors that affect interpretation. In the rest of the chapter, the factors that are proposed to play a role in production and/or interpretation are considered, and empirical data from previous studies that either support or undermine these claims are provided. Given that an adequate model for discourse should be able to capture such biases as observed in (11) and (12), our focus will additionally be on the models that have been proposed to explain the empirical data discussed.

2.3   What contributes to prominence?

The notion of prominence has been examined as *salience, accessibility, center of attention* and *psychological focus* (see 2.1.1). Despite making similar claims concerning the interaction between anaphoric form and prominence of potential antecedents, the models based on such notions do not have a uniform list of factors that are assumed to influence entity accessibility.

The current section, therefore, focuses on the factors that influence prominence. It presents an overview of a large body of linguistic and psycholinguistic research and discusses

---

[11]Semantics of entities, inherent accessibility in (e.g. concreteness, animacy, gender) have also been shown to affect pronoun resolution.

the effect of grammatical and semantically/pragmatically driven factors on the interpretation and/or the production of referential expressions. Since there is no consensus within the literature on which set of factors actually exists, each factor will be presented under its own heading[12]. This thesis follows von Heusinger and Schumacher[13] in employing the term *prominence* as an overarching term and alluding to other concepts such as accessibility as 'prominence-lending cues' (von Heusinger & Schumacher, 2019: 119).

As previously mentioned, albeit analysed under a unified heading in the current thesis, not every factor is observed to influence both production and interpretation (Kehler et al., 2008). Thus data not only from interpretation but also from production studies are discussed.

2.3.1 Structural prominence

Language production is widely assumed to proceed in an incremental fashion (J. K. Bock, 1982; Levelt, 1989). Items or structures that are conceptually more accessible tend to be retrieved more easily from memory. According to the language production model by K. Bock and Levelt (1994), once the concept/message that is to be expressed has been selected, it undergoes grammatical encoding. Grammatical encoding comprises two stages: functional processing and positional processing. The former is responsible for identification and selection of lexical concepts, and assignment of syntactic/grammatical functions such as case (NOM, ACC, etc.) or grammatical role (subject, object etc.). The latter, positional processing, involves constituent assembly and inflectional processes in which the linear order of an utterance is created.

Since the ease of retrieval from memory determines the order that concepts undergo grammatical encoding, more accessible concepts or structures have priority in functional and/or positional processes. The twofold nature of the encoding process has created a divergence in the views as to what exactly accessibility has an impact on. There are two main viewpoints[14]: either it influences the linear order, and the most accessible concept occupies the sentence-initial position; or it affects grammatical role assignment, and the most accessible concept is assigned with the most prominent grammatical role, that is subjecthood. This would indicate two biases: a subject bias and a first-mention bias. With this in mind, we could thus expect a similar bias in production and interpretation of referential forms within a theory based on antecedent prominence.

With respect to interpretation, many researchers have investigated how ambiguous pronouns are interpreted. Preliminary work suggests a correlation between form and antecedent prominence, and maintains that pronouns refer to more prominent entities (Chafe, 1976:

---

[12]I have to note that this is not an exhaustive list, and other factors have been proposed, e.g. *animacy* (Bittner & Kuehnast, 2012); *agentivity* (Gernsbacher & Hargreaves, 1988), *recency* (Clark & Sengul, 1979).

[13]They diverge from several models discussed under this heading by claiming that interpretation and production does not depend on one-to-one mappings between items and mental representations but rather the relation is dependent on other competing items. They, therefore, suggest the relation between form and items to be dynamic.

[14]There is a third possibility that accessibility influences both but at different degrees.

inter alia). Additionally, it has been argued that the subject function is more prominent than other grammatical roles (Kieras, 1979). Experimental work has reported preferences for assigning the pronoun to the subject of the previous sentence (also called *subject-assignment strategy*). Crawley et al. (1990) investigated the interpretation of object pronouns in a self-paced reading task followed by an antecedent selection task. Participants read ambiguous and non-ambiguous object pronouns and selected an antecedent. An ambiguous experimental item by Crawley et al. (1990: 259) is given in (13).

(13) John and Sammy were playing in the garden. One of their classmates, Evelyn tried to join in their game. John pushed Sammy and Evelyn kicked him.

Although the pronoun *him* in this example can refer to either *John* or *Sammy*, Crawley et al. (1990) found a strong preference for the subject assignment strategy: *him* was more likely to be interpreted as referring back to *John*. On the other hand, when the pronoun was coded for a specific gender, which disambiguates the pronoun (i.e. when a female character was introduced instead of Sammy in 13), there was no effect of subject preference, and reading rates were lower. Thus, Crawley et al. (1990) suggest that morphological cues such as gender are stronger than other cues such as subject-assignment. Therefore, it appears that people rely on a structural cue, subjecthood, in determining the antecedent when they encounter an ambiguous pronoun. Subsequent experimental work on subject-assignment strategy proved to be consistent with these findings (Frederiksen, 1981; Stevenson et al., 1994).

In contrast to subject-assignment (albeit overlapping in many instances), it has been found that the first-mentioned entity in the preceding sentence is the preferred entity for pronoun resolution (Gernsbacher & Hargreaves, 1988; Gernsbacher et al., 1989). In many languages, it is not easy to tease apart these two biases. For example, in English, the subject is almost always the first-mentioned entity of a sentence, thus much work has clustered these two biases together. To untangle these two linguistic factors, Gernsbacher et al. (1989) examined sentences similar to the one below in (14).

(14) a. Because of Tina, Lisa was evicted from the apartment.
     b. Tina was evicted from the apartment because of Lisa.

In (14a), *Tina*, despite being the first-mentioned entity, is not the grammatical subject. Whereas in the second, it is both subject and first-mentioned. Gernsbacher et al. (1989) conducted a probe recognition task in which subjects are asked to judge whether a probe that occurs on the screen (either *Tina*, *Lisa*, or another name) was a part of the sentence they had just read e.g. (14a). Their findings showed that probe recognition of the first-mentioned entity *Tina* was faster than the second-mentioned entity. Thus, they conclude that first-mentioned entities are more accessible than other entities, and that this effect is independent of the subjecthood factor. Yet, this effect might have resulted from information structure properties, particularly due to the *because-clause* being moved to a focus position[15].

---

[15]There is evidence for focus constructions such as clefts in the literature to contribute to accessibility (See Arnold 1998 for a review). Section 2.3.2 provides a brief discussion on this argument.

Contrary to English, there are languages where it is possible to find subjects mentioned later in a sentence. For instance, in Finnish, a language that has flexible word order, Järvikivi, Van Gompel, Hyönä, and Bertram (2005) manipulated word order (SVO vs OVS) to test the two biases. Findings from their visual world eye-tracking experiment indicated a bias for both first-mention and subject antecedents. They further reported that these biases exist independently and do not interact with one another.

Opposed to the subject-assignment or first-mention preferences presented above, the object pronoun *him* in the example (15) below does not resolve preferentially towards the subject or the first-mentioned NP, but rather towards the object of the context sentence (i.e. *Bill*). So what could be mandating the intuitions in (15)?

(15)   John greeted Bill. Lily hugged him. [= Bill]

In order to account for such inferences, earlier work argued that this preference results from parallelisms in grammatical roles, in lieu of a subject assignment preference, (Sheldon 1974; Grober, Beardsley, and Caramazza 1978 for initial work supporting the claim). Grammatical parallelism refers to a strategy that connects pronouns in subject position to the entity that is the subject of the preceding sentence, and pronouns in the object position to the object of the preceding sentence. Although initial experiments for grammatical parallelism did not test this strategy as completely divorced from the preference for subject-assignment, the role of grammatical parallelisms is noted to exist. Yet, data from offline methods such as judgements or passage completion tasks have been inconsistent. For instance, while Smyth (1994) also found strong parallelism preferences both for subject and non-subject dependencies, Crawley et al. (1990) found a preference that originates from subject-preferences rather than grammatical role parallelisms (see example 13). Despite observing an effect of grammatical parallelism, Smyth (1994) reported an overall preference for subject assignment. As for online methods, a recent visual world eye-tracking study by Sauermann and Gagarina (2017) found an early and strong effect of grammatical parallelism in German, regardless of word-order. Their results showed more looks to the subject antecedent in subject anaphor condition (personal pronoun, *he* 'er'), and more looks to the object antecedent in object anaphor condition (personal pronoun *him* 'ihn') both in SOV and OVS word order.

The psycholinguistic literature has widely employed the completion paradigm to investigate anaphora production. This paradigm makes it possible to address questions regarding both production and interpretation. While continuations to (16a) merely show what the preferred antecedent for a pronoun is (i.e. pronoun interpretation), participants are free to select any referring expression in (16b), and hence such sentences provide data to investigate production preferences.

(16)   a. John scolded Jim. He . . .

       b. John scolded Jim. . . .

As mentioned previously, *he* in (16a) generally refers to the subject of the previous sentence, *John.* With regard to sentences similar to (16b), studies reported a preference to refer to subjects with pronouns (Kehler et al., 2008; Rohde & Kehler, 2014; Stevenson et al., 1994). In other words, there appears to be a bias for pronominalising subjects. Given that subject is a prominent position, these results from interpretation and production fit the predictions made by the models observed in 2.1.1. In languages that allow for null pronouns, comparable findings have been reported. Recall that nulls are uninformative expressions and refer to salient antecedents (Givón, 1983: *inter alia*). Having seen the evidence supporting the special status of subjects (or first-mentioned entities), one could anticipate that nulls would be used to refer to syntactic subjects. This is exactly what various studies show: nulls prefer subject antecedents, whereas overts prefer non-subject antecedents (See Cameron 1992; Silva-Corvalán 1977 for evidence from corpus studies; Carminati 2002; Keating, Jegerski, and VanPatten 2016; Mayol 2018; Mayol and Clark 2010 for experimental work). Carminati (2002) observes this disparity in Italian intra-sentential anaphora and identifies it by the hypothesis given below in (17).

(17) <u>*Position of Antecedent Hypothesis* (PAH)</u>   (Adapted from Carminati 2002: 57)
      The null pronoun prefers an antecedent which is in the Spec IP position, while the overt pronoun prefers an antecedent which is not in the Spec IP position

In generative grammar, Spec IP traditionally stands for the specifier position of the inflectional phrase (IP) position on a syntax tree, that is the subject position. Therefore, the hypothesis associates nulls with subject antecedents and overt pronouns with object antecedents. The asymmetry between forms has been referred to as 'division of labor', the complementary distribution of the two forms: when one occurs it signals a property that is lacking in the other. Testing sentences with a potentially ambiguous null ($\emptyset$) or an overt pronoun (*lui* 'he'), as shown in (18), Carminati (2002) reported a subject preference for nulls and object preference for overts, which is predicted by PAH. She concludes that the preference is motivated by accessibility of referents as suggested by Ariel (1990).

(18) Quando Mario$_i$ ha   telefonato a Giovanni$_j$, $\emptyset_i$/lui$_j$ aveva   apppena
      When   Mario have-3SG telephoned to Giovanni$_j$, $\emptyset_i$/he$_j$ haveSG-PST just
      finite   di mangiare.
      finished of eat-INF
      'When Mario$_i$ telephoned Giovanni$_j$, $\emptyset_i$/he$_j$ had just finished eating'
      (Adapted from Carminati 2002: 58)

Although Carminati (2002) only provides evidence for intra-sentential pronominal dependencies (see 18), subsequent studies replicated the results in both offline and online studies and showed that the PAH holds to be valid for inter-sentential dependencies (Alonso-Ovalle, Fernández-Solera, Frazier, & Clifton, 2002; Filiaci, 2010; Keating, Vanpatten, & Jegerski, 2011).

2.3.2 Information structure

In addition to the biases in linear and syntactic structure, information structure has been shown to have implications on antecedent accessibility. In fact, the hierarchies for 'explicitness' presented in 2.1.1 suggest that information-structural properties, topichood in particular, influence pronoun interpretation and production. Traditionally, there are two main information structure categories: *topic* and *focus*. Topic is defined as 'what is being talked about' in a discourse, and thus, related to 'aboutness' (Reinhart, 1981). On the other hand, focus refers to information that is new and unpredictable. Throughout the literature, observing a relation between form and prominence, many researchers suggested that pronouns prefer topic antecedents, while full-names prefer focus antecedents (Ariel, 1988; Arnold, 1998; Grosz et al., 1995; Gundel et al., 1993).

However, since subject position is considered to be a topical position (Chafe, 1976), there is a need to examine the effect of topicality independent of subjecthood. Analysing the two effects is possible when object constituents of canonical active sentences are placed in particular positions. In English, it is not possible to scramble constituents and have OVS sentences. Thus to examine topichood and subjecthood preferences, voice was manipulated in several studies. In a study by Rohde and Kehler (2014), active and passive sentences such as the one below in (19) were analysed in two conditions namely no-pronoun and with an overt pronoun. The underlying idea is that passive constructions promote a constituent to the subject position in order to mark topicality, and thus are stronger signals for indicating that the subject is the topic (see Rohde and Kehler 2014 for a discussion on *Topichood Hypothesis*). Within this assumption, a higher rate of pronominalisation is predicted for passives.

(19)   a.   Amanda amazed Brittany. (She) . . .   [active]

        b.   Brittany was amazed by Amanda. (She) . . .   [passive]

The results from two experiments, showed more re-mentions for the subject in (19a) compared to (19b) in no-pronoun prompts (i.e. sentences without *she*). Consistent with the predictions, a higher rate of pronominalisation for passives in comparison to actives was observed. In sentences with pronoun prompt (she) a high rate to refer back to the subject (to *Amanda* in (19a); to *Brittany* in (19b)), was found. They conclude that passive voice is a stronger indicator of topicality than active voice, and both grammatical and information-structural properties influence pronoun production. In addition, their results revealed a lower effect for subject antecedent preferences when the verb type was changed. This effect was only present with pronoun prompts, therefore suggesting an asymmetry in production and comprehension; this will be discussed in more detail in the following sections.

Moreover, other structures including clefts, left-dislocation and relative clauses also allow researchers to disentangle subjecthood and topichood effects, as they explicitly mark new information/focus (Chafe, 1976; Colonna, Schimke, & Hemforth, 2014; Cowles, Matthew, & Kluender, 2007; Foraker & McElree, 2007; Prince, 1981)[16]. Such explicit markers are em-

---

[16]There are also languages that overtly mark topics morphologically such as Japanese.

ployed in investigating the effect of topicality, and a general preference for focus antecedents is found inter-sententially in English and Romance languages[17] (Arnold, 1998; Cowles et al., 2007; de la Fuente, 2015).

Topicality is shown to be extremely pertinent to the type of pronoun. Demonstrative pronouns, for instance, have been argued to especially prefer non-topic antecedents, while personal pronouns prefer antecedents that are topics (Comrie, 1997). Despite not making a distinction between the effect of subjecthood and topichood, Bosch et al. (2003) argued that findings from their corpus study support this claim: topicality/subjecthood appeared a decisive factor in the production of these two types of pronouns. Given this observation, Bosch et al. conjectured a hypothesis to argue that p-pronouns and d-pronouns show complementary preferences (in a similar manner to the distribution of nulls and overts in Carminati's (2002) *PAH*). The hypothesis is as follows:

(20) Complementary Hypothesis (Bosch et al., 2003: 4)
      Anaphoric personal pronouns prefer referents that are established as discourse topics while demonstratives prefer non-topical referents

Findings from various pronoun resolution studies conform to the hypothesis (Colonna, Schimke, & Hemforth, 2012; Schumacher, Dangl, & Uzun, 2016). Yet, other studies suggest that topichood is not the sole factor determining pronoun production, but rather there are several factors at play. These studies show that production of p-pronouns is influenced by both topichood and subjecthood (Portele & Bader, 2020). In a similar manner, when it comes to interpretation, p-pronouns are reported to prefer topic antecedents that are subjects while d-pronouns prefer non-topic antecedents that occur in clause positions (Bader & Portele, 2019; Kaiser, 2010; 2011; Kaiser & Trueswell, 2008; Portele & Bader, 2020). Therefore, the preferences cannot be cast solely in terms of topicality. Yet, both forms are shown to be sensitive to semantic bias created by verbs (Kaiser & Trueswell, 2008; Portele & Bader, 2020).

As for null pronouns, mixed findings of topicality effects have been found. While Kwon and Sturt (2013) reported an influence of topicality in null pronoun resolution in Korean, recent studies in Japanese (Ueno & Kehler, 2016) and Korean (Song & Kaiser, 2020) did not find any effects[18].

2.3.3   Thematic prominence and Verb Semantics

As previously mentioned, biases associated with verb semantics have been shown to intersect with structural biases. Garvey and Caramazza (1974) observed that some verbs do not abide by structural biases, but rather induce their own. The sentences in (21) are minimal pairs, varying only in the verb, yet when asked, a native speaker of English would

---

[17]In clefts with intra-sentential pronominal dependencies such as "It was Peter who slapped John when he was a student." the preference appears to be reversed, and a general dis-preference for focused antecedents are noted (See de la Fuente 2015 for a review on the studies).

[18]These studies argue that the absence of such effects might have resulted from insufficient discourse context.

interpret the pronoun differently: they would resolve it as referring to *Jim* in (21a) and to *Tim* in (21b).

(21) a. Jim fascinated Tim because he . . . [he = Jim]

  b. Jim liked Tim because he . . . [he = Tim]

Garvey and Caramazza (1974) relate this interpretation preference to what the verb (implicitly) encodes as the cause of an event. They claim that some verbs such as *fascinate* mark the first-mentioned referent as cause (21a); while others, such as *like*, mark the second-mentioned (21b). Upon encountering a pronoun in sentences similar to the ones in (21), speakers preferentially choose the causally implicated referent as the antecedent. These two types of verbs are characterised in the literature as *object-experiencer* (OE) verbs and *subject-experiencer* (SE), respectively[19]. To explain the preference in interpretation many researchers allude to thematic prominence.

Thematic prominence depends on the interaction between a predicate (the verb) and its arguments. Each predicate assigns roles to its arguments. The roles are lexically specified. For this reason, thematic prominence cuts across syntactic categories such as grammatical roles[20]. A list of thematic roles relevant to the focus of the current thesis and example sentences are given in Table 2.1[21].

Table 2.1: A list of thematic roles

Roles	Examples
AGENT – PATIENT	Sally$_{agent}$ ate an orange$_{patient}$.
SOURCE – GOAL	Ben$_{source}$ bought Jane$_{goal}$ a pen.
STIMULUS – EXPERIENCER	Mary$_{stimulus}$ amazed Jim$_{experiencer}$.
AGENT – THEME	Louis$_{agent}$ put the book$_{theme}$ on the table.

IC verbs encode a stimulus (i.e. causer), and an experiencer (i.e. causee). The experiencer is the entity that undergoes the emotional state and is by definition animate (Verhoeven, 2014). On the other hand, the stimulus is the instigator of the state and can either be animate or inanimate. The observed preference in (21) is towards the stimulus argument of a clause and has been replicated in a variety of typologically different languages such as Mandarin (Brown & Fish, 1983; Hartshorne, Sudo, & Uruwashi, 2012), Catalan (Mayol, 2018), Norwegian (Bott & Solstad, 2014), German (Portele & Bader, 2020), Dutch (Cozijn, Commandeur, Vonk, & Noordman, 2011; Koornneef & Van Berkum, 2006), Finnish (Järvikivi et al., 2005), Spanish (Goikoetxea, Pascual, & Acha, 2008), Japanese and Russian (Hartshorne et al., 2012).

---

[19]These verb types have been categorised under other labels such as *stimulus-experiencer* (SE) and *experiencer-stimulus* (ES).

[20]Although thematic role labels are assigned independently of syntactic position, there are strong connections between grammatical positions and thematic roles. For instance, the agent is usually (but not necessarily) the subject.

[21]In the literature other distinctions, such as *benefactive, location, recipient,* and *instrument*, exists.

Over the years, there have been different approaches as to how the bias surfaces itself. Garvey and Caramazza (1974), for instance, maintain that the semantics of the verb encodes for causality, and that the direction of the cause is indicated in the lexicon. Subsequent theoretical work argues that a verb encodes causality as a part of its thematic structure (Bott & Solstad, 2014). Following such approaches, assuming IC verbs universally code a particular entity as more prominent than the other, we should expect to find the similar preferences in both interpretation and production. Thus, IC bias should be expected in similar degree in (22a-22c) towards referring to *Sally* and (22b-22d) in referring to *Jane*.

(22) Story continuation results of Stevenson et al. (1994) (Adapted from Rohde 2019)

    a. Sally infuriated Jane. She . . . [63% Sally]

    b. Sally scolded Jane. She . . . [72% Jane]

    c. Sally infuriated Jane. . . . [78% Sally]

    d. Sally scolded Jane. . . . [66% Jane]

However, upon examining pronominal dependencies similar to (22), an interesting observation was made by Stevenson et al. (1994). When the pronoun was given (22a-22b), participants preferred to mention the stimulus next (albeit not being equally strong in both verb types), thus conforming to the IC bias. On the other hand, in no-pronoun prompts, (22c-22d), the IC effect was stronger for (22c) and weaker for (22d).

Thematic role effects were also found in different verb contexts. Stevenson et al. (1994) analysed transitive verbs, which encode an agent and a patient, and found patients to be preferred as antecedents over agents. On the other hand, when they analysed transfer-of-possession verbs, which encode a source and a goal argument, they showed a clear preference for goals over sources in interpretation. However, Stevenson et al. (1994) also noted that biasing effects can be strengthened, weakened or even reversed with respect to the type of prompts used. For instance, when an IC verb is used with the sentence connector *so*, the IC bias seems to be overridden, and continuations favour the Experiencer (Kehler et al., 2008; Rohde, Kehler, & Elman, 2006; Stevenson et al., 1994). This further indicates that the picture regarding interpretation and production of pronouns may rely on a more complex process than just the semantics of verbs. Therefore, there seems to be a need for further analysis of the relationship between contexts and continuations (see 2.3.4).

As for online measures, there have been several studies employing eye-tracking and self-paced reading methodologies in order to study the effects of thematic role biases in anaphora processing in recent years. These studies were mainly interested in the time course of the causality bias and tested pronouns congruent or incongruent with the biases. Consider the following example in (23) (adapted from Goikoetxea et al. 2008: 765). The IC bias, in general, prefers stimulus antecedents, yet in (23b) the continuation (already knew the story) requires pronoun resolution towards the experiencer argument, *Gabriel*.

(23) a. Ana bored Gabriel because s/he talked slowly. (congruent)

b. Ana bored Gabriel because s/he already knew the story.(incongruent)

Testing such sentences, a number of experiments reported a congruency effect for implicit causality right at the pronoun region (Featherstone & Sturt, 2010; Koornneef & Van Berkum, 2006). In a number of experiments employing visual world paradigm, the expectation effect was observed even before the region of "because" (Pyykkönen-Klauck and Järvikivi 2010; Cozijn et al. 2011). Similar effects of implicit causality bias were observed in event-related potentials (ERPs), and Van Berkum, Koornneef, Otten, and Nieuwland (2007) found a P600 effect for a pronoun which was incongruent with the implicit causality bias. These results suggest that comprehenders have predictions or anticipations as to how the discourse continues before the second clause is unveiled. This was taken as evidence for approaches that argue that verb semantics trigger this bias (Bott & Solstad, 2014). Another interesting observation is that, continuations such as the one in (23a) appear to be more anticipated in contrast to (23b). This suggests that there are restrictions on what kind of relations can hold between clauses. This brings us to the next factors: world-knowledge and coherence relations.

2.3.4 World-knowledge and Coherence relations

Hobbs (1979) states that simply referring to the same entity is not sufficient for successive sentences to create a coherent discourse. He provides the example below in (24).

(24) John took a train from Paris to Istanbul. He likes spinach. (Hobbs, 1979: 67)

Although readers can understand that *he* refers to *John* in this example, the second sentence is an odd sequence. As a matter of fact, one may even ask what does liking spinach have to do with travelling to Istanbul. Employing this example, Hobbs argues that something else comes into play "which is deeper than the notion of a discourse just being 'about' some set of entities" (1979: 67). According to him, coherence problems, such as the one in (24), arise as a by-product of relations that are required to hold between utterances. When readers or listeners encounter a pronoun that is ambiguous in a discourse, they try to infer what relations may have been established to hold between clauses by the speaker (Rohde, 2019). Hobbs (1979) characterises several relations that link clauses, which he terms "coherence relations", and provides formalisms for each relation with the aim to adequately model communication[22].

- OCCASION describes a "mere succession" of events, temporally following one another (Hobbs, 1990: 86).

- EVALUATION is a "metacomment" in which S_1 or S_0 indicates what the conversational goal is[23].

- GROUND-FIGURE and EXPLANATION relate a part of the discourse to what the speaker knows.

---

[22]Hobbs (1990) further sub-categorises OCCASION under two relations: *cause* and *enablement*.

[23]While S_1 stands for the current clause or segment, S_0 refers to the segment immediately preceding it.

- EXPANSION RELATIONS

    - PARALLEL describes a relation in which similar eventualities can be inferred from S_1 and S_0.

    - ELABORATION is when an identical proposition can be inferred from two sentences.

    - EXEMPLIFICATION is when the second sentence expresses an example to the first.

    - CONTRAST describes a relation in which the inferred proposition of the second sentence is a negation of the first (or it negates a property of the first).

    - VIOLATED EXPECTATION describes an unexpected result of the eventuality.

According to Hobbs, each relation addresses a communicative need: the speaker establishes a relation to convey a message which pertains to a communicative goal. Correspondingly, comprehension depends on "[d]educing the speaker's goals" (Hobbs, 1979: 87). Thus, he suggests that pronoun interpretation revolves around an inferential process, that is the coherence relation identified by the listener, and that this relation determines the selected antecedent for reference. However, merely identifying the coherence relation is not sufficient to interpret anaphora and people must also rely on their world-knowledge. Consider the example (25) below adapted from Winograd (1972).

(25)   The city council denied the demonstrators a permit because they...

    a. feared violence

    b. advocated violence

The two sequences in (25) have different interpretations of the pronoun *they*: in the former it refers to the city council and to the demonstrators in the latter. Although both clauses carry an Explanation relation with the main clause, different reasoning processes underlie the two continuations. Since our world-knowledge dictates what is true or more likely for our world, we can infer that the reason behind the city council's denial of the permit could neither be due to the city council advocating for violence nor demonstrators fearing violence. As such, world knowledge mandates the coherence relation identified by the listeners. In a similar vein, without world-knowledge identifying a coherence relation would not be possible.

Following an approach that is framed around coherence relations appears to provide an explanation for seemingly puzzling data. Recall the experiments by Stevenson et al. (1994) who reported that the preferred antecedents for pronouns is influenced by the connectives used in prompts (*because, so, and, full stop*). For fragments with *because* and *so*, they found a strengthened preference for Goal (in ToP) and Patient (in action verbs). On the other hand, in implicit causality contexts, *because* reinforced the IC bias towards the stimulus/subject (26d), while *so* weakened the bias and indicated a preference for the experiencer/object (26b). See the percentages of prompt type from implicit causality context below in (26) (adapted from Stevenson, Knott, Oberlander, and McDonald 2000: 228).

(26)   Ken impressed Geoff

    a. and he ... [Geoff 76%]

    b. so he ... [Geoff 89%]

    c. (full stop). He ... [Ken 73%]

    d. because he [Ken 87%]

As previously mentioned, the verb in (26) is an IC verb, in particular it is a OE verb. The OE verbs are well-known for being subject biased, which is apparent in (26c) and (26d). However, the continuations (26a) and (26b) are not subject biased which suggests that sentence connectives might have an influence on the pronoun interpretation. In fact, Stevenson et al. (1994) maintained that the conjunction *because* differs from *so* as it represents a contrasting coherence relation: the former conjunction bears a causal relation, and the latter bears a consequence relation (also called "implicit consequentiality" by Pickering and Majid 2007). Observing such patterns, it is argued that connectives carry their own biases which can modify the semantic biases (Stevenson et al., 1994; 2000). Under a coherence-driven approach, one could view sentence connectors as explicit signals of specific coherence relations (Hobbs, 1979). For instance, *because* in (26d) signals an Explanation relation. This relation in turn is biased toward stimulus. As IC verbs are demonstrated to give way to Explanation continuations, an explicit *because* appears to reinforce the stimulus bias which is indicated by the percentages for continuations (26d) compared to other connectors (Kehler et al., 2008).

    However, one particular observation by Stevenson et al. (1994) regarding ToP sentences without any explicit sentence connectors remains puzzling. Although they found a Goal/Subject bias for (Goal-Source) transfer-of-possession verbs in the full stop condition (85%, see 27b), the preference was 50/50 for the two antecedents when the order of potential thematic roles was changed, i.e. in Source-Goal sentences (27b). One question that arises, then, is why the subject preference is absent in (27b).

(27)   Transfer-of-possession contexts from Stevenson et al. (1994)

    a. John seized the comic from Bill. He ... [John 85%]   (Goal-Source)

    b. John passed the comic to Bill. He ... [John 51%]   (Source-Goal)

Stevenson et al. (1994) conclude that there are two different biases at play in such examples, namely a bias favouring goals and another favouring subjects. The first bias, the goal-bias, is explained on the basis of an end-state bias encoded by the ToP verbs[24]. They hypothesise that referents associated with the end-state of events make more prominent antecedents, and hence pronouns are interpreted towards the end-state antecedents. To test this hypothesis, Rohde et al. (2006) manipulate verbal aspect by comparing pronoun resolution in an experiment with imperfective and perfective context sentences. In completed events (perfective), the entity associated with the end-state is the goal and is more salient. When the continuation

---

[24]One can think of an end-state as a completion of an event. In imperfective events, one cannot know for sure if the transfer is completed. For example, we do not know in (29) if Mary has the glass of water or not.

describes what happens after the event, as in (31a), the bias, unsurprisingly, favours the goal. Given that salient entities are preferred to be assigned as antecedents, the sentence in (28) is expected to yield more goal (*Mary*) interpretations than (29). On the other hand, the end state has not been reached in imperfective events and therefore continuations are not necessarily about what happens after the event.

(28) Sue$_{Source}$ <u>handed</u> a glass of water to Mary$_{Goal}$. She . . . (PERFECTIVE EVENT)

(29) Sue$_{Source}$ <u>was handing</u> a glass of water to Mary$_{Goal}$. She . . . (IMPERFECTIVE EVENT)

In fact, early reports of the end-state bias examining imperfective events which lack an end-state, showed that while people referred to the goal referent more than the source in perfective (28), in imperfective (29) they referred to the goal as often as the source (Rohde et al., 2006). This suggests that pronouns in perfective conditions tend to refer to the goal, which is affiliated with the end-point of the verb. However, as the imperfective lacks an end-state, participants show a significantly higher preference for source linked interpretations. In light of the data, Rohde et al. (2006) conclude that the interpretations of pronouns reflect event representations, rather than thematic role biases, such as the goal bias.

Taken together with the widely recognised subject bias, the bias for end-states provides an explanation for the seemingly puzzling preferences observed by Stevenson et al. (1994). In their study, when the context sentence consisted of a Goal-Source sequence, participants were more likely to resolve the pronoun (*He*) towards the goal antecedent, which was also the subject. Consider the example in (27a), repeated below in (30).

(30) John$_{Goal}$ seized the comic from Bill$_{Source}$. He . . . [John 85%]

The two biases (end-state and subject bias) reported in the literature are uniform in Goal-Source sentences (30), the entity associated with the end-state (i.e. goal) is also the subject of the context sentence. Yet, in Source-Goal sentences (31) these two biases are incompatible and the subject is not the entity affiliated with the end-state but the indirect object is. Therefore, unlike (30), participants have to decide whether to resolve the pronoun towards the end-state of the event, *Bill* as in (31a) or towards the subject, *John* as in (31b), and write continuations accordingly. This clash of biases is reflected in the 50/50 continuations when the biases are discrepant (31), while a high percentage of goal/subject preference emerges when the two biases are congruous as in (30).

(31) John$_{Source}$ handed a book to Bill$_{Goal}$. He . . . (adapted from Rohde 2019: 457)
    a. opened it to the first page. [He=Bill$_{goal}$]
    b. handed him War and Peace. [He=John$_{source}$]

Analogous findings were found in a study by Kehler et al. (2008), and the pronoun did not show any goal preference in Source-Goal sequences in a full stop prompt. Though an explanation could be provided via the two biases, when the analysis was broken down to

the coherence relations, following Hobbs (1990), an interesting picture emerged: OCCASION and RESULT were goal biased, while ELABORATION and EXPLANATION were source biased (Kehler, 2002; Kehler et al., 2008; Rohde, Kehler, & Elman, 2007). (31a) is an example of occasion: he opens the book after getting it. As for example (31b), the coherence relation is elaboration: it provides description of the same event, i.e. handing a book. For this reason, (31a) is goal/object biased, and (31b) is source/subject biased, respectively. Based on these observations, Kehler et al. (2008) subsequently argue that the thematic role biases stem from coherence relations[25]. In other words, the goal-bias, which is a symptom of a bias for end-states, is argued to be an epiphenomenon of a more general mechanism for establishing coherence between clauses (Kehler et al., 2008).

Even though an analysis based on coherence relations cannot account for interpretation on its own, it has proven to be valuable tool in accounting for the cases that cannot be explained by either the grammatical or the semantic role. For example, the continuation in (32b) cannot be explained by subject or stimulus bias, whereas the coherence relations can readily account for the data: the stimulus is preferred for explanation relations (32a), and the object/experiencer is preferred in result relations.

(32) Implicit causality contexts (adapted from Rohde 2019: 457)

    a. John amazed Bill. He could walk on his hands. [He=John$_{stimulus}$]

    b. John amazed Bill. He clapped his hands in awe. [He=Bill$_{experiencer}$]

The conclusion arrived at here is that coherence relations interact with semantic factors such as verb type and world knowledge, and that these relations, at the same time, carry their own biases. Together with structural factors, coherence-relations and semantic biases (henceforth semantic-pragmatic factors) determine pronoun interpretation. Yet, the two structural and semantic-pragmatic factors influence the interpretation in a different manner (Stevenson et al., 1994). The semantic-pragmatic factors come into being through mechanisms that establish prominence and predict the next mentioned referent (even before encountering the pronoun), and thus this mechanism is a "top-down" mechanism. The structural factors, however, constitute a bottom-up mechanism: upon encountering a pronoun, comprehenders assume that it refers back to the subject (Stevenson et al., 1994; 2000). Evidence for assuming such an architecture of mechanisms is found when different prompts are compared with one another. As previously discussed in (22), upon testing no-pronoun and pronoun prompts, Stevenson et al. (1994) found a discrepancy in the IC bias, see (33) for percentages.

(33) a. Sally infuriated Jane. She [78% Sally] vs. No-pronoun prompt [63% Sally]

    b. Sally scolded Jane. She [72% Jane] vs No-pronoun prompt [66% Jane]

---

[25]Though following Hobbs's (1990) coherence relations, Kehler departs from Hobbs by claiming that "the categorisation is based on the types of coherence that we extract from the external world" (2002: 31), rather than arguing in favour of a categorisation of relations based on communicative needs.

The percentages shows that the presence of a pronoun (she) increases the continuations for the subject, regardless of whether the subject is causally implicated (33a) or not (33b). Later, in a study by Rohde and Kehler (2014) analogous observations were made. They found more explanation continuations for both types of IC which resulted in more subject related continuations in OE verbs and more object related ones in SE. So, the semantic-pragmatic factors are shown to affect the probability of who to mention next (re-mention). On the other hand, how the entity will be mentioned is argued not to be influenced by these factors (Fukumura & Van Gompel, 2010; Kehler et al., 2008; Rohde & Kehler, 2014). Even though SE verbs prefer object antecedents, pronominalisation of the object was not favoured, and the rate of pronominalisation was higher when the antecedent was the subject. Following these observations a proposal for capturing the relation between interpretation and production of pronouns via Bayesian principles was put forward, which will be discussed in section 2.4.3.

Several studies, not only for English (Fukumura & Van Gompel, 2010; Rohde & Kehler, 2014) but also for other languages (German: Portele and Bader 2020; Catalan: Mayol 2018) obtained results confirming the claim that likelihood of re-mention (which is also known as predictability) does not affect pronoun production. However, conflicting data was also found. A corpus study by Arnold (2001) showed that Goals, in comparison to Sources, are more likely to be mentioned again, later in the discourse. Empirically testing this finding in a story completion experiment consisting of three sentence contexts, Arnold (2001) revealed that goals were pronominalised at a greater rate than sources when referring back to the second entity. This was replicated in a recent study by Rosa and Arnold (2017) who investigated whether predictability affects reference form production. They conducted three experiments: two story continuation tasks, and one picture description task that involved longer contexts (compared to the widely utilised one sentence contexts). Given that the widely employed continuation paradigm is generally more de-contextualised than natural language, as it consists of single-sentence contexts, Rosa and Arnold initially suggested that richness of context might be responsible for the conflicting results by Arnold (2001). However, they provided evidence against such a hypothesis and observed an effect of thematic roles in production of referential form, even in the tasks that do not involve a story like context, and they found a significant effect for thematic roles and pronominalisation: pronouns were used more often for goal entities than source entities. Importantly, the interaction between form and thematic role was smaller for the tasks with shorter story contexts. Thus, it remains unclear whether semantic-pragmatic factors have an influence on referential form, and whether the absence of such interaction found in many studies (Rohde & Kehler, 2014: *inter alia*) results from paradigm related mechanisms.

What is clear, however, is that purely syntactic or semantic proposals cannot adequately capture all of the observed facts as suggested by the majority of the literature. Therefore, there is a need for a model that takes not only syntactic and information-structural factors into consideration but also semantic factors. Additionally, this model should be able to correctly predict which entity will be mentioned next and in which form it will be mentioned.

It should also predict which entity will be mentioned in the case of a specific referential form. In the following sections, approaches that attempt to predict discourse coherence are presented.

2.4   Approaches to Anaphora and Discourse Coherence

Over the years there have been several approaches to how coherence is established in a discourse, many of which were concerned with the interpretation and production of anaphora. While some authors proposed models by analysing the distribution of referring expressions and correlating them with entity prominence (Ariel, 1988; Gundel, 2010), others emphasised the importance of competition between potential topics (Gordon et al., 1993; Grosz et al., 1995). Although these approaches entail a similar process behind production and interpretation of anaphora, some suggested that these two processes are sensitive to different factors. In this section, we analyse the theories that attempt to model/predict coherence and account for the data observed in the literature. The accounts are provided under three main headings: Mirror models, Expectancy hypothesis and Bayesian account.

2.4.1   Mirror models

The heuristic based models that were discussed in 2.1.1 (Ariel, 1990; Arnold, 1998; Gundel et al., 1993) and the Centering Theory (Grosz et al., 1995) are widely referred to as Mirror models by Rohde and Kehler (2014). They employ the term "mirror" to "[c]apture the idea that pronoun interpretation biases mirror the biases that underlie the speaker's decision to produce a pronoun." (Kehler & Rohde, 2019: 3). In fact, these approaches argue that the speaker uses a referential form to refer back to a referent that they believe to be prominent in the discourse, and, in turn, the listener interprets the form by assuming that the speaker would use it to denote the most prominent referent. As a result, interpretation and production are mirror images of each other (Rohde & Kehler, 2014: 1). The same factors contribute to what speakers and hearers consider as prominent in a discourse. Therefore, much experimental research has focused on identifying the factors that determine entity prominence (many of which were discussed in section 2.3). The formalism for such mirror models is as follows:

(34)   P(referent | pronoun) ∝ P(pronoun | referent)

On the left side *P(referent | pronoun)* expresses the interpretation bias. It is the probability that a pronoun that has encountered (| pronoun) refers to a particular referent. This represents the process in which the comprehender assigns a referent to a pronoun s/he heard. Conversely, *P(pronoun | referent)* expresses the production bias. It is the probability of using a pronoun, given that a specific referent is to be mentioned (| referent). In this equation, the operator ∝ indicates that interpretation biases are predicted to be directly proportional to the ones of production. Provided that production and interpretation are mirror images of one another, this account predicts interpretation and production preferences in continuation experiments to be comparable for the free prompt and the pronoun condition.

Although the Mirror models account for the differences between subject vs non-subject referents' distribution in production, it is shown to underestimate the semantic and pragmatic effects that result in next-mention biases (Kehler & Rohde, 2019; Portele & Bader, 2020; Rohde, 2019; Rohde & Kehler, 2014). The problem of the model is that it does not make predictions on the likelihood of mention, and hence ignores the expectations of non-subject entities that are created by the end-state bias seen in sentences such as "Mary passed the book to Jenny. She ..." (Rohde, 2019). Rather, it predicts the pronoun to be resolved towards the subject. In other words, the model does not distinguish between which form is more likely to be used and which entity is more likely to be mentioned.

Centering Theory Centering theory is a computational approach to discourse anaphora proposed by Grosz et al. (1995). This theory focuses on three interacting components: attentional limits of interlocutors, intention and language structure. It gets its name from *centers* of an utterance, which are *centers of attention*. Grosz et al. (1995) argues that certain entities are more central than others and this notion of centrality consequently sets restrictions on speakers' use of referential expressions. According to their theory, there are two categories of centers: the *backward looking center* (C_b) and *forward looking centers* ($C_{f...f_n}$). Each sentence has one backward looking center and a list of forward looking centers. This distinction leads to a dissociation of the processes involved in looking forward and looking backward (Grosz et al., 1995). Forward looking centers are ranked for prominence, and the most prominent center becomes the preferred subsequent center (C_p). The backward looking center is the most topical center that is the highest ranked element in the list. It denotes the C_p and selects it as its antecedent. Although how a pronoun is comprehended depends on a variety of factors, the algorithm for ranking topics (forward looking centers) is determined by a grammatical roles hierarchy. The subject is ranked highest, then comes the object, and then other elements (Brennan, Friedman, & Pollard, 1987). As C_b is the most prominent center in the ordered list, only the backward looking center can be pronominalised irrespective of the forward looking centers. In contrast, C_f cannot be realised as a pronoun if C_b is not pronominalised. Recency of topics also plays a crucial role in the theory, when entities are repeatedly mentioned in subject position they become more topical, and hence are ranked higher in prominence. Further, Grosz et al. (1995) posit three types of transition relations:

1. CENTER CONTINUATION: A backward looking center remains to be the most prominent entity as a forward looking center

2. RETAIN: Despite having been ranked lower on the C_f list a previous entity continues to be C_b

3. SHIFT: The antecedent of C_b changes

Below in (35), an example for transitions by Grosz et al. (1995: 217) is provided. In (35b), *John* is the C_b and therefore the pronoun is resolved towards it. It is, at the same time, the highest ranked C_f. In (35c), *John* continues to be the C_b, yet another character (Mike) is

introduced which enters the ranking for C_f. Since *John* ranks higher than *Mike* in (35c), it would not be felicitous to pronominalise *Mike* and refer to *John* with the name in (35d). Although ranking higher, *John* is retained in (35d) and *Mike* becomes the highest ranked entity. For this reason, in (35e), *Mike* is the antecedent of *he* and a shift of C_b is observed.

(35)   a. John has been having a lot of trouble arranging his vacation.

        b. He cannot find anyone to take over his responsibilities. (he = John) C_b = John; C_f = John

        c. He called up Mike yesterday to work out a plan. (he = John) C_b – John; C_f = John, Mike (CONTINUE)

        d. Mike has annoyed him a lot recently. C_b – John; C_f = Mike, John (RETAIN)

        e. He called John at 5 AM on Friday last week. (he = Mike) C_b = Mike; C_f = Mike, John (SHIFT)

This is a "production-oriented claim" that focuses on the influence of grammatical and information-structural constraints on reference production (Kehler & Rohde, 2013: 5). It specifically argues that the speaker chooses a referential form to signal to the hearer how to correctly interpret a pronoun. It does not acknowledge any influence of semantics or world-knowledge in pronoun production/resolution. It is shown to widely predict the subject antecedent preference and the restriction on using names (e.g. *Repeated Name Penalty*). The theory revolves around three interacting components: attentional limits of interlocutors, intention and language structure. It claims that if an entity is pronominalised, it will be the topic. It provides explanation for the data reporting a preference for subject or topic assignment of pronouns and also for the repeated name penalty, despite the theory lacking any distinction between information structural entities from the grammatical ones. This theory is shown to be consistent with the recent observations on the distribution of demonstrative and personal pronouns in German. Portele and Bader (2020) reported an effect of topichood on participants' selection of the referent to re-mention. However, this theory does not explain why a 50/50 reference for subjects in Source-Goal verbs is found in the studies by Stevenson et al. (1994) since it does not acknowledge the semantic pragmatic biases identified in previous experimental studies. Only prominence determined by topicality and subjecthood is relevant for interpretation.

2.4.2   Expectancy Hypothesis

In her proposal put forward in 1998, Arnold claims that "discourse features influence accessibility by providing information about the predictability of upcoming reference" (Arnold, 2010: 192). That is, listeners interpret a pronoun based on their expectations on what the speaker will most likely mention next. So expectancy is calculated in terms of likelihood of re-mention, and different syntactic and semantic factors (e.g. recency, subjecthood, grammatical role and information structure) are posited to influence a referent's expectancy (i.e. predictability). The model suggests that subjects are more likely to be mentioned again than

non-subjects, and recently mentioned entities are more likely to be mentioned again than less recent entities. Accessible entities are high in expectancy and tend to be included in a subsequent clause. Expectancy model equates the likelihood of a referent being selected as an antecedent when given a pronoun, with the probability that the referent will be named again, as indicated in the formalism below.

(36)   P(referent | pronoun) $\propto$ P(referent)

This model differs from the others in how pivotal the role of predictability is for determining referential form. It is a 'top-down' model that makes predictions about the hearer's expectations of which entity will be re-mentioned, via various semantic and structural cues. The hearer interprets the pronoun according to the referent s/he expects the speaker to mention next. According to Rohde (2019) a shortcoming of this model is that it does not account for the subject preference of pronouns. For this reason, she states that the model overestimates object interpretations of object biased verbs (Experiencer-Stimulus verbs or Source-Goal verbs), and at the same time underestimates subject interpretations of subject biased verbs (Rohde, 2019).

2.4.3   Hybrid Models: The Bayesian account

The Bayesian account is not only cast in terms of entity prominence but rather follows from a more general mechanism responsible for discourse coherence. It is a hybrid model as it "combines the comprehender's prior coherence-driven expectations about what entities will be referred to next and centering-driven likelihoods that govern the speaker's choice of referential form."(Kehler & Rohde, 2013: 1). Therefore, it could be regarded as a conciliation of the Centering Theory and the Hobbsian coherence-driven model. The Bayesian model departs from the idea that interpretation and production are mirror images of each other, positing that the interaction between the two notions is more complex (Kehler et al., 2008). To predict the probability of the speaker using a specific referential form when mentioning a referent, Kehler et al. (2008) makes use of the Bayes' rule. It computes this probability on the grounds of two probabilities: the probability of a referent occurring in the subsequent discourse and the probability of it being conveyed in a pronominal form. The formulation is as follows:

(37)   BAYES' RULE by Kehler and Rohde (2013: 15):

$$P(\textit{referent} \mid \textit{pronoun}) = \frac{P(\textit{pronoun} \mid \textit{referent}) \ P(\textit{referent})}{\sum\limits_{\text{referent} \in \text{referents}} P(\textit{pronoun} \mid \textit{referent}) \ P(\textit{referent})}$$

The probability on left side of the equation, *P(referent | pronoun)*, indicates the interpretation bias, and *P(pronoun | referent)* expresses the production bias. The production bias is the hearer's estimation of how probable it is that a pronoun will be used to refer back to a particular referent. This is dependent on information-structural and grammatical biases such as topichood or subjecthood (also referred to as the *likelihood* term). This likelihood term interacts with the belief of a referent being mentioned in subsequent discourse, *P(referent)*, regardless

of the referential form. Furthermore, the denominator can be simplified as *P(pronoun)*. It represents the likelihood of a pronoun being used which is also called *prior*. As stated by Kehler and Rohde (2019: 8) in a footnote, this probability can be estimated by calculating the value in the numerator (*P(pronoun | referent) P(referent)*) for each possible referent and summing the result. In other words, the model computes the probability of a pronoun occurring by considering other possible forms that might be produced by the speaker. As Rohde (2019: 466) posits, the rule itself is not a model for anaphora but is "simply a statement of mathematical truth" which is useful for predicting pronoun production and interpretation. This rule also makes it possible to tell apart who will be mentioned next and how they will be mentioned. These two notions are proposed to be influenced by different factors. The probability of mention is argued to be affected by semantic-pragmatic factors. While the probability of particular forms being used is insensitive to such factors, it is influenced by structural factors. This relation is captured with the division of labour proposal:

(38) Proposal for a division of labor in a Bayesian model of pronouns (Kehler & Rohde, 2013)
      Different factors influence the probability of mention (the prior) and the probability of pronominalization (the likelihood). The prior probability of mention is influenced primarily by semantic and coherence-driven cues. The likelihood of forms is influenced primarily by structural properties of the preceding discourse.

The Bayesian account has been shown to outperform other models in several studies (Rohde and Kehler 2014; Bader and Portele 2019; Portele and Bader 2020; Godoy, Weissheimer, and Mafra 2018). For instance, after conducting two discourse continuation tasks to investigate production and interpretation of Catalan pronouns, Mayol (2018) computed the predicted probability of nulls referring to the subject and compares it with the observed probability. She reported a significant correlation between the measured and predicted probabilities in Bayesian account for both experiments. Yet, in the case of the Mirror model, the correlation was not significant. The correlation was significant for the Expectancy hypothesis but is found to be lower than that of the Bayesian Model[26]. Moreover, findings from her study support the division of labour introduced in (38) for pronoun production and interpretation. She reports that, in production, nulls are subject biased while overts are non-subject biased and insensitive to semantic-pragmatic biases, yet interpretation of these forms is affected by both factors.

2.5 Interim Summary

As is presented in the chapter, throughout the years many authors observed a correlation of referential form and (some notion of entity) prominence and proposed hierarchies

---

[26]Following the formalisms provided in the section 2.4, the Mirror model equates the likelihood of referring to a subject when given a null, i.e. P(subject | null), with the probability of a null is used given that the preferred referent is the subject , i.e. P(null | subject). As for the Expectancy model, P(subject | null) is equated with the probability that the subject will be mentioned again.

to capture this relation (see section 2.1.1). Both theoretical and empirical work has supported the idea that referential forms follow the hierarchies, and that reduced forms, such as pronouns or nulls, are preferably used for more accessible entities rather than for the less prominent/informative referents such as full noun phrases. Yet, this is argued to not be a case of one-to-one correspondence of form and activation but rather the selection of forms is dependent on the competition of two forms and their activation levels. Therefore, referential form is contingent on the referential form chosen in referring to other entities.

Moreover, as discussed in section 2.3, there is a growing body of research investigating the potential factors that induce the prominence of antecedents. Many have provided evidence that syntax-based factors play a role both in production and interpretation (Rohde & Kehler, 2014: *inter alia*). This assumption appears to be line with the proposed hierarchies in section 2.1.1 that also observe a correlation in referential form and structural position. Pronouns have generally been observed to favour subjects as antecedents. The relation between different forms is claimed to be captured in hypotheses for division of labour. These hypotheses generally argue for a single factor, namely the subjecthood, to be determining for interpretive preferences of forms and that these forms are complementary to one another. However, this appears to depend on the inventory of referential forms of a language and some forms, such as demonstrative pronouns or overt pronouns in languages that allow for null subjects, are shown to behave differently than the pronoun in English and carry their own biases. Thus, these hypotheses were not borne out for pronoun resolution in every language. Recent studies on nulls in Korean (Kwon & Sturt, 2013; Song & Kaiser, 2020), Japanese (Ueno & Kehler, 2016) and Italian (Fedele, 2016) show that the distribution of forms are overlapping rather than complementary in terms of grammatical biases, and that they are, in fact, sensitive to semantic and discourse factors.

In addition, there is considerable overlap between different factors, and when closely analysed, several factors appear to interact with one another. For instance, syntactic and information structural properties are found to influence both production and interpretation of pronouns (See section 2.3.2 for details). Yet, in many studies no differentiation is made between the two. Thus, there is an ongoing debate as to whether the subject bias results from a preference for topic antecedents. As summarised in section 2.4, some researchers assume syntax-driven factors to encode entity prominence. Others, on the other hand, claim that this effect is a by-product of other preferences, for instance coherence relations. Rohde (2019: 455) states that "[o]ne can see the appearance of syntax based references might depend on the way that the sentences combine to form a coherent discourse".

With regard to pronoun interpretation, verb meaning is observed to have an influence on the interpretation of pronouns (Stevenson et al., 1994: among others). Verb induced biases are argued to be symptoms of our world-knowledge that guide us when determining the relationship between two sentences. This is subsequently claimed to impact how people interpret ambiguous pronouns (Kehler & Rohde, 2013). Some authors have even claimed

that world-knowledge influences production of referential expressions (Arnold, 2001; Rosa & Arnold, 2017).

Previous literature provides evidence to suggest that production and interpretation are susceptible to different types of biases (see Rohde 2019 for a review). While production is observed to be more prone to structural biases, interpretation is affected by both semantic/pragmatic and structural biases. Yet, there are studies that challenge this claim by positing that coherence-driven effects do exist in production when a richer context is provided, therefore suggesting that the lack of effect may have resulted from the sentence continuation paradigm that is widely used (Arnold, 2001; Rosa & Arnold, 2017). Though the effects reported are marginal, further work is required to clarify the relationship between production and interpretation.

Chapter 3

Referential Expressions in Turkish

This chapter reviews the theoretical and experimental work in the literature of Turkish linguistics. In the first part of the chapter, section 3.1, properties of Turkish pronominal forms are presented. Then, the distribution and function of forms, as well as the role that information structure plays in establishing reference are discussed. The second part of the chapter, section 3.2, is concerned with the recent experimental work regarding the interpretation and production of referring expressions. Lastly, section 3.2.1 discusses findings of the experimental work thus far for Turkish linguistics in light of the previous body of knowledge on anaphora. This paves the way for several predictions to be made for the experiments in Chapter 4.

3.1   General characteristics of Turkish

Turkish is a head-final language with the standard word-order of SOV. Pronominal forms, like other noun phrases, inflect for case. They receive nominative case marking in subject position[1]. The personal pronouns in Turkish are as follows:

- *Ben* 'I'

- *Sen* 'You' (singular)

- *O* 'He/she/it'

- *Biz* 'We'

- *Siz* 'You' (plural, formal)

- *Onlar* 'They'

As indicated in the list above, there is no gender distinction in Turkish pronouns, and the third person singular pronoun, *o*, is equivalent to *he*, *she* and *it* in English. The third-person singular pronoun, which is relevant for the experiments provided in Chapter 4, is

---

[1]The nominative in Turkish is a zero morpheme. Albeit being phonetically empty, it has the function of signalling nominative case.

homophonous to the distal demonstrative pronoun, *that* in English (Göksel & Kerslake, 2005). The distal demonstrative pronoun does not inflect for number, and it encodes neither animacy nor gender. Therefore, in many cases, it is not possible to make a clear-cut distinction between the two elements. For instance, below in (39), *o* is ambiguous. It can be interpreted either as a demonstrative with a dropped subject (39a) or as an overt pronoun (39b).

(39) O   kitab-ı   oku-du.
      that/he book-ACC read-PST.3SG

    a. $\emptyset_{s/he}$ read that book

    b. S/he read the book.

Turkish lacks an article system to mark definiteness, instead word-order, case and the indefinite determiner *bir* 'one' are used to distinguish definite expressions from indefinite ones (Aksu-Koç & Nicolopoulou, 2015). In example (40), (in)definiteness marking of direct objects with the accusative case (40a) or the indefinite determiner (40b) is given. The indefinite determiner[2] is homophonous with the numeral 'one'. The position of *bir* relative to the adjective allows us to differentiate between the numeral and indefinite determiner. If *bir* precedes the adjective it is interpreted as a numeral. If it follows the adjective it receives the indefinite determiner interpretation. In addition, when the direct object in pre-verbal position is unmarked, it receives an indefinite and non-specific interpretation, as in (40c).

(40) a. kitab-ı   oku-du-m
         book-ACC read-PST.1SG
         'I read the book'

    b. kalın bir kitap oku-du-m
         thick a   book read-PST.1SG
         'I read a thick book'

    c. kitap oku-du-m
         book read-PST.1SG
         'I read a/any book'

Verbs in Turkish agree with the subject in terms of number and person morphology. This allows for dropping subjects, and hence the occurrence of the subject is optional. In other words, Turkish has overt and covert pronominals and, therefore, is classified as a pro-drop language. In languages such as Turkish, subject omission is generally licensed by the verbal morphology[3]. For instance, in (41), the subject could be dropped even when the sentence is not uttered as an answer to the question in (41a). The listener can unambiguously interpret the null subject since the first person singular inflection on the verb provides enough information as to who the subject is. Besides subject pro-drop, Turkish also allows for object

---

[2]See Kornfilt (2018) for arguments regarding the determiner status of *bir*.

[3]Romance languages are similar to Turkish in terms of their rich verbal morphology. Yet, there are languages like Chinese that lack inflectional morphology, but nevertheless, allow for pro-drop.

pro-drop, as indicated with parentheses around *o-nu* in (41b). Compared to subject-drop, object-drop is more restricted. Objects can be dropped only when the referent is recoverable from the discourse. This is because the verb does not carry any morphological cues that help to interpret the null object. For this reason, listeners rely on semantic/pragmatic cues. Once a referent is introduced to the discourse, a null form may be used to refer back to it in object position. For this reason, an answer with a dropped object would not be felicitous, if it were to occur in isolation, in the absence of the question (41a). When used to answer the question, both subject and object may be omitted at the same time (41b).

(41)   a. Have you seen James today?

      b. (Ben) (o-nu)   market-te gör-dü-m.
         (I)   (he-ACC) shop-DAT saw-PAST-1SG
        'I saw him at the shop.'

Analysing the distribution of pronouns and comparing it to the one of English, Kerslake (1987) states that null subjects in Turkish correspond to unstressed pronouns in English, whereas overt pronouns correspond to stressed pronouns. As for their function, Enç (1986) maintains that full forms are generally used in situations in which the speaker wants to emphasise a referent in order to convey new information. Therefore, when the question asks for a specific person, pro-drop gives rise to infelicity, as in (42b).

(42) Kim-in   ev-i   satılık?
     who-GEN house-3SG.POSS for.sale

     'Whose house is for sale?'

      a. Ben-im ev-im   satılık.
        I-GEN   house-1SG.POSS for.sale
        'My house is for sale.'

      b. *∅ ev-im satılık.   (Examples adapted from Enç, 1986: 212)

Preliminary theoretical work on anaphora in Turkish focused on the behaviour of null and overt pronouns. Null pronouns in pro-drop languages adhere to the hierarchies discussed in 2.1.1. That is, for a referent to be realised with a phonologically empty form, the referent is required to be highly accessible (prominent) in a given context (Enç, 1986). In one of the first studies on anaphora in Turkish, Turan (1996) systematically analysed the behaviour of overts and nulls within the framework of Centering Theory. She proposed that while covert subjects are commonly used in maintaining topics, overts signal topic-shift. Given that subjects rank higher than other potential topics in Centering Theory, and that the highest ranking topic is selected in maintaining topics, Turan (1996) rephrased this observation, maintaining that the subject referent tends to be referred to with null pronouns.

      Furthermore, Turan states that the subject must be more reduced in form than the object. This squares well with the models that claim subjects are generally the most prominent element in a discourse, and that the most prominent entity is generally realised with

the most reduced form (Givón, 1983: *inter alia*). By providing the examples below, Turan (1996: 87) shows that, when the object occurs in a less explicit form than the subject, the continuation may lead to infelicity (indicated by the hash symbol, #). This follows from the sentences in (43b) and (43c), since employing an overt pronoun for the object antecedent and realising the subject antecedent in pronoun or NP form appears to be infelicitous[4].

(43)   Ali$_i$ Murat-ı$_k$ davet   etti.
       Ali   Murat-ACC invited

     'Ali invited Murat.'

     a. $\emptyset_i$ o-na$_k$ içki   ikram etti.
         $\emptyset_i$ him-DAT$_k$ drink served.

         '(He$_i$) served him$_k$ a drink.'

     b.   # '$\emptyset_k$ ona$_i$ içki ikram etti.'

     c.   # $\emptyset_k$ Ali-ye   içki   ikram etti.
         $\emptyset_k$ Ali-DAT drink served.

         # '(He$_k$) served Ali a drink.'

     d.   Murat ona$_i$ içki   ikram etti.[5]
         Murat him-DAT$_k$ drink served.

         'Murat served him$_k$ a drink.'

Nulls in Turkish appear to prefer subject entities regardless of the subject's position in a sentence. In fact, U. Özge, Özge, and von Heusinger (2016) show in (44) that the same observation holds even when the order of mention is changed so that the subject occurs after the object:

(44)   Ali-yi$_k$ Murat$_i$ davet etti.
       Ali-ACC$_k$ Murat$_i$ invited

     'Murat invited Ali.' (Lit. 'Ali, Murat invited.')

     a. $\emptyset_i$ on-a$_k$ çiçek   getirdi.
         $\emptyset_i$ he-DAT$_k$ flower brought

         '(He$_i$) brought him$_k$ flowers.'

     b.   # $\emptyset_k$ ona$_i$ çiçek   getirdi.
         $\emptyset_k$ he-DAT$_i$ flower brought

         # '(He$_k$) brought him$_i$ flowers.'

In terms of their classification, some scholars argue that null objects and subjects in Turkish are instances of the category *pro* (Öztürk, 2005). The small *pro* in theories of syntax, such as the Government and Binding theory, is an empty category that stands for null pronominals.

---

[4]Even though it is felicitous to use two proper names as in 'Ali served Murat a drink', this leads to redundancy and is generally dispreferred.

[5]Object drop in this sentence would lead to a generic reading (Murat served drinks). The verb *serve* requires a dative marked argument, and thus, only when the previous occurrence of the referent carries the same case marking, the same reading in (43d) is available for a sentence with null object.

Yet, Kornfilt (2018) states that there is considerable evidence against the *pro* analysis of silent objects in Turkish. She argues that the empty category is merely an instance of an 'A′-bound empty variable' (Kornfilt, 2018: 152). In other words, null-objects in Turkish are discourse elided categories bound by a topic operator. Therefore, she suggests that silent objects in Turkish are not an instance of a Spanish pro-drop but of a mechanism similar to topic-drop.

Given that Turkish allows for pro-drop, the use of overt forms are more restricted when compared to the overts in languages that prohibit omission of subjects (Gračanin-Yuksek, Lago, Şafak, Demir, & Kırkıcı, 2017). Overt pronouns are conditioned by different pragmatic and information-structural constraints. They can, therefore, function as signals of pragmatic status such as topic shift, similarity or contrast (Enç, 1986). In fact, the overt pronouns are argued to be markers of similarity or contrast between two referents and/or actions (Azar, Backus, & Özyürek, 2019). In (46c) and (45), these pragmatic functions of overt pronouns are exemplified. Azar et al. (2019) state that the speaker of (45d), employs the overt pronoun to mark that the eventualities carried by the referents (Pelin and the speaker) is similar: they both ordered coffee[6]. Contrastingly, the overt pronoun in the example (46c) does not mark similarity. Instead, it marks the contrast of actions: one fails the exam while the other passes.

(45) a. Dün $\emptyset_j$ Boğaz-da   kahve içi-yor-du-m.
          Yesterday $\emptyset_j$ Bosphorus-DAT coffee drink-PROG-PST-1SG

          'Yesterday (I) was having coffee by the Bosphorus.'

      b. *Pelin$_k$* beni gör-müş.
          Pelin   me   see-PF

          '*Pelin$_k$* saw me.'

      c. $\emptyset_k$ hemen yanıma   gel-di.
          $\emptyset_k$ quickly next.to.me come-PST

          '(She) came by at once.'

      d. *O$_k$* da bir kahve söyle-di.
          She also one coffee order-PST

          'She, too, ordered coffee.'

(46) a. *Ahmet$_i$* matematik sınavın-dan   kal-mış.
          Ahmet math   exam-POSS.3SG-ABL fail-PF

          '*Ahmet$_i$* failed the math exam.'

      b. *Selin$_m$* de aynı sınav-a   gir-miş.
          Selin   too same exam-DAT enter-PF
          '*Selin$_m$* too took the same exam.'

      c. Ama *o$_m$* geç-miş.
          But she pass-PF

---

[6]Azar, Backus, and Özyürek (2016) note that the use of the clitic *-dA* 'also' is widely used with overt pronouns, as in 45d (82% percent of all overt pronouns in their data). They suggest that this clitic reinforces the emphatic function of the overt pronoun in Turkish.

'But she_m passed.'   (Examples adapted from Azar et al., 2019: 557)

In addition to the pronouns discussed so far, there is a reflexive pronoun expressed with *kendi* 'self' in Turkish (Göksel & Kerslake, 2005). The reflexive gets a possessive suffix and carries person and number agreement marking of its possessor (Dinçtopal-Deniz, 2009). Similar to the other nominals, the reflexive pronoun lacks gender marking. Moreover, an inflected form of the reflexive *kendi* exists in third person singular: *kendisi*. This form behaves similarly to the 3rd person pronoun *o* 'he/she/it' (Göksel & Kerslake, 2005). Although this is beyond the scope of the present thesis, I have to note that *kendisi* 'self' deviates from the Binding Principles (See Dinçtopal-Deniz 2009 for more information about this particular form). Unlike the bare reflexive form *kendi* which only allows for local attachment (47b), *kendisi* in (47a) can be co-referential with referents that are within (Mehmet) or outside the clause boundaries (Ahmet), or with extra-sentential referents, as indicated by the index *k*.

(47)   a. Ahmet$_i$ [Mehmet'in$_j$ kendisi-ni$_{i/j/k}$ bugün gör-düğ-ünü]
           Ahmet [Mehmet-GEN kendi/kendisi-ACC today see-NMLS-3SG.POSS.ACC]
           duydu.
           hear-PST
           'Ahmet heard that Mehmet saw himself/him/her$_{i/j/k}$'

       b. Ahmet$_i$ [Mehmet'in$_j$ kendi-ni$_{*i/j/*k}$ bugün gör-düğ-ünü] duydu.
           'Ahmet heard that Mehmet saw himself$_{*i/j/*k}$'

3.2   Experimental Work

Many studies on anaphora in Turkish were framed around intuitions of the researchers who attempted to account for the distribution of different pronouns through various approaches or models (Enç, 1986; Erguvanlı-Taylan, 1986; Turan, 1996). Only a few studies have tested anaphora in Turkish experimentally, and existing studies have been mostly concerned with pronoun resolution.

The interpretation of the tripartite anaphoric system in Turkish comprising two reflexive forms (*kendi/kendisi*) and an overt pronoun *o* 'he/she/it' was empirically tested in two experiments by Gračanin-Yuksek et al. (2017). In their Experiment 1, they first analysed acceptability judgements of these forms when they occurred as embedded direct objects. Participants were asked to answer a question with three choices:

(48)   Emre realized right away that Cem is blaming *kendi/kendisi/o*.
       Who was being blamed?
       A)Emre   B)Cem   C)Someone else

                                        (Gračanin-Yuksek et al., 2017: 1404)

The results demonstrated that *kendi* and *kendisi* in general favoured both local and long-distance reference (choices A and B in 48), and *o* preferred non-local reference (both long-distance and extra-sentential, A and C in 48). This was mostly in line with the observations

of the behaviour of the anaphoric forms, yet *kendi* also preferred long-distance antecedents which was unexpected for several syntactic accounts (Enç, 1986; Göksel & Kerslake, 2005). They also report that, albeit very rarely, local attachment is preferred for *o* (11% of the trials). In their Experiment 2, they conducted a self-paced reading task and manipulated contextual information. They tested configurations that were either compatible or incompatible with the preferred referent for the interpretation identified in Experiment 1. The manipulation involved changing the referring expression that occurred in the two-sentence dialogue preceding the critical item. For example, the biasing conditions for the sentence (48) included sentences such as "Cem: I am so rude! I wish I hadn't hurt my mother." for local antecedent. The pronoun 'I' was changed to 'you' for long-distance antecedent and to another name that is not included in the dialogue like 'Ali' for extra-sentential antecedent (Gračanin-Yuksek et al., 2017: 1410). Results showed that context has an effect on the interpretations. The observed distribution is changed when context favours a referent that is not normally preferred, yet this effect in not strong enough to override the general interpretive preferences especially in those forms that are syntactically constrained, such as *kendi* and *o*. Thus, syntactically more constrained forms that either prefer more local binding, like *kendi*, or non-local binding, such as the overt pronoun, may receive interpretations incompatible to these constraints when context biases those antecedents (non-local for *kendi* and local for *o*, respectively). Yet, contextual information exerts a greater impact when a form is not syntactically constrained such as the *kendisi*.

Rather than focusing on the interpretation of reflexive pronouns, several other studies focused on independent clauses and investigated the behaviour of nulls in comparison to overt pronouns in inter-clausal referential dependencies. In a study employing the design by Hartshorne (2013), D. Özge, Hartshorne, and Snedeker (2017) investigated the question of how an ambiguous pronoun is interpreted when the context contains two potential antecedents. They tested overt and null pronouns in two different experiments with two types of IC verbs: OE and SE, and conducted two rating studies. In this design, two clauses are conjoined, and the verb in the second clause is a nonce, meaningless, adverbial predicate (e.g. *daxed*). Participants are asked to select an antecedent for the ambiguous pronoun, see (49) for an example experimental item.

(49) Bahar Ceren'i   çok büyülü-yor/arzulu-yor   çünkü (o)   son derece dakmuk.
      Bahar Ceren-ACC a.lot dazzle-PROG/desire-PROG because (she) extremely dakmuk
      'Bahar dazzles$_{OE}$/desires$_{SE}$ Ceren a lot because she is extremely dax.'
      - 'Who is dax?'

Their findings showed a stimulus/subject bias both for overt and null pronouns in stimulus-experiencer verb condition. Object-experiencer verbs did not have a uniform preference for subject or object but showed a significant effect for prompt manipulation: the null pronoun increased the likelihood of subject/experiencer antecedent interpretations, while overts in this condition exerted a bias for stimulus/object. So, D. Özge et al. (2017) found a bias for

stimulus antecedents, which they propose to be due to the conjunct *because*. The observed pattern appears to be consistent with that of Japanese (Ueno & Kehler, 2016). In Japanese, nulls, in comparison to overt pronouns, are shown to be less sensitive to contextual biases (Ueno & Kehler, 2016). Yet, unlike Japanese which demonstrates a subject bias for both null and overt pronouns, overt pronouns in Turkish did not reveal a bias towards subject but rather showed a preference for stimulus entities. The distribution found by D. Özge et al. (2017) suggests a complex interplay between different factors. Therefore, their results are in contrast to the observations on null pronouns in Korean (Kwon & Sturt, 2013) which do not show any effect for pronoun form but only thematic roles. Instead, results from D. Özge et al. (2017) indicate that subject-experiencer verbs are sensitive to referential form. Following this, null forms are found to increase subject/experiencer interpretations (D. Özge et al., 2017).

In another study, D. Özge, Evcen, Kesici, and Köse (2018) analysed ToP contexts and tested three different types of verbs: Source-Goal, Agent-Patient, and Agent-Benefactor. Again using the model of Hartshorne (2013), they manipulated sentence connectors and examined configurations with *and, so* and *because*. In addition to these factors of verb type and conjunction type, which varied within subjects, they included a between-subject factor of pronoun type and contrasted nulls with overts. The results indicated that nulls, in general, prefer the subject as antecedent in continuations that trigger Elaboration and Occasions (i.e. *and, so*). This preference of nulls prevailed, regardless of the thematic role of the subject referent (agent or source). As for the overt pronoun, despite a differences in rates, it demonstrated a non-subject bias and acted as a topic-shifter. This was in accordance with the observations of previous theoretical work (Enç, 1986: inter alia). However, continuations with *because* preferred both the null and the overt pronoun to be interpreted as referring to the entity associated with the non-subject, which was either the goal, benefactor or patient. They also report that in Occasion (*because* prompt) and Result (*and* prompt), the distribution of forms confirms Carminati's (2002) division of labour hypothesis for Turkish: null pronouns favour subject antecedents and overt pronouns non-subject. This observation is inconsistent with the distribution of overts in Japanese (Ueno & Kehler, 2016) and Korean (Kim, Theres, & Schafer, 2013), both of which show a subject-bias. Instead, pronouns in Turkish appear to be influenced by both verb-type and coherence relations. The interpretation also depends on the grammatical factors: while nulls favoured subject antecedents, overts favour non-subjects. They concluded that nulls in Turkish behave like pronouns in English and not like nulls in Korean or Japanese. Both nulls and overt pronouns in Turkish are influenced by grammatical and semantic-pragmatic factors but overt pronouns are more sensitive to grammatical factors, and they almost always refer to the object.

In addition to the effect of verb type and coherence relations, the effect of information-structure on pronoun resolution was tested in a study by Özge and Evcen (2020) who manipulated word order and tested SOV and OSV sentences. They found a greater subject bias for OSV sentences than SOV sentences. In other words, when the subject was placed in

the focus position, which is immediately pre-verbal in Turkish (Erguvanlı-Taylan, 1986), the pronoun is more likely to be attributed to the subject antecedent. In light of the hierarchies discussed in section 2.1.1, they suggest that both focus status and subjecthood add to a referent's prominence. In addition to this, they analysed the effect of verb valence and found that nulls in the canonical SOV order are interpreted as referring to the subject irrespective of the verb valence. Conversely, interpretation of overt pronouns in SOV sentences are affected by verb valence: positive valence yields object interpretation tendencies and negative valence yields subject interpretations. This was taken as evidence that overts are topic-shifters as previously claimed by Enç (1986).

Studies concerning the production of referential forms have either focused on the contribution of specificity in establishing reference, or on general reference tracking strategies in narratives. For the former, a study by U. Özge et al. (2016) investigated the influence of differential object marking (DOM), which is the optional accusative marking on direct objects that denote specificity, on the selection of referential form and also likelihood of re-mention. They conducted a discourse completion study to examine continuations of sentences in which two referents are introduced with noun phrases. Below in (50), an example item from their experiment is provided. The stimuli consisted of a context sentence and a critical sentence in which the DOM marking on the object (*-yi*) was manipulated.

(50)   Diabetes, which is currently in the focus of health media, has been analysed on the Health TV channel last week.

Serinin   ilk program-ın-da   tanınmış   bir diyetisyen program-a   bir
Series-GEN first program-GEN-LOC well-known a   dietician   program-DAT a
diyabetli(-yi)   konuk etti. . . .
diabetic(-ACC) hosted

'In the first program of the series, a well-known dietitian hosted a diabetic in the show.' . . .   (U. Özge et al., 2016: 9)

Results indicated more re-mentions to objects antecedents in conditions with or without DOM. This does not appear to corroborate the argument of Turan (1996) that speakers generally tend to mention the subject in subsequent discourse. U. Özge and colleagues suggest that the distribution observed in re-mentions that favour the object over the subject might be a symptom of the verb meaning, in particular the ditransitive verb used. Yet, when mentioning the object entity participants used more overt forms. When mentioning subject entities, participants preferred a null pronoun rather than an overt one. They conclude that these results suggest a partition of referential persistence (which is indicated by re-mention preferences) and entity prominence (which is reflected by the referential form chosen for re-mention), as argued by Kehler and Rohde (2013). That is, nulls and overts carry their own

biases despite the fact that objects were more likely to be mentioned next in the upcoming discourse[7].

Other studies on pronoun production in Turkish have examined the function of different referring expressions in reference tracking. Azar et al. (2016) analysed the function of null and overt pronouns in Turkish and Dutch, particularly in marking maintained and re-introduced referents. In their experiment, participants were instructed to provide narratives to silent videos. Their answers were later coded according to the referential form used. The results demonstrated that both in Dutch and in Turkish participants used full NPs when they are (re)introducing entities to the discourse and reduced forms, such as null pronouns, when maintaining referents. Overt pronouns were widely used in Dutch but they were used very infrequently in Turkish. They further report that overt pronouns in Turkish are predominantly used as a strategy to mark discourse status, particularly in emphasising a referent. Taking the differences in the reference tracking strategies in Turkish and Dutch into account, Azar et al. (2016) argue that referential systems of languages do not reveal a uniform set of universally available principles that guide speakers when establishing/maintaining reference.

In a later study, again employing an elicited narrative task, Azar et al. (2019) investigated the interaction between speech and gesture during reference tracking. Their results from speech data replicated that of Azar et al. (2016). That is, speakers used reduced forms to maintain referents and more explicit forms to re-introduce referents. Additionally, they maintained that overt pronouns are strongly influenced by pragmatic context, and that they are used to mark similarity or contrast between eventualities.

Recently, a study by Konuk and von Heusinger (To appear) investigated interpretation and production of null and overt pronouns in IC contexts. For interpretation they provided sentences with ambiguous pronouns, and asked participants to select an antecedent. In order to control for IC bias, they included a causal adjunct in their experimental items. Because of the adjunct, participants had to provide continuations other than explanations in the production task. Consequently, they predicted a bias for experiencer referents as it is more agentive than stimulus referents. (51) presents an experimental item with a subject-experiencer verb.

(51) Gökhan dünkü   kahvaltı daveti   sonrasında Naz-ı   [büyüleyici
     Gökhan yesterday's breakfast invitation after   Naz-ACC charming
     gülüşünden   dolayı] gün boyunca düşledi.
     smile-NOM-3SG-ABL because.of day long   dream-PST
     'After yesterday's breakfast invitation, Gökhan dreamed of Naz all the time because of her charming smile.'

   a. Ø sekiz-de   mail   at-tı.
      Ø eight-LOC e-mail send-PST-3SG

---

[7]As discussed in section 2.4.3, this follows from the disassociation of *P(pronoun | referent)* from *P(referent)*.

b. O   sekiz-de   mail   at-tı.
   she/he eight-LOC e-mail send-PST-3SG
   'She/He sent an e-mail at 8 p.m.'   (Konuk & von Heusinger, To appear)

Konuk and von Heusinger (To appear) found that both verb type and grammatical function affect antecedent selection: participants tended to interpret both overt and null pronouns as referring to the subject in subject-experiencer contexts, while they did not show any preference for either null or overt pronouns in object-experiencer contexts. In other words, participants were biased to select the subject as the antecedent (regardless of the pronoun used) only when the subject was also the experiencer referent. This provides evidence against a contrast between null and overt pronouns. For production, their results demonstrated a re-mention bias towards experiencer: in SE contexts the subject was more likely to be re-mentioned, and in OE contexts the object was more likely to be re-mentioned. Participants used a null pronoun to refer back to the subject with SE verbs. Yet, they did not show a preference for OE verbs and employed other referential forms. Overall, they did not find many cases of overt pronoun. Instead, proper names were preferred for less prominent entities.

3.2.1 Interim Summary

In Turkish, a language that allows for subject (and also object) omission, null pronouns are theoretically and empirically observed to prefer subject antecedents. Yet different factors are observed to affect this interpretive preference. First of all, coherence markers like clausal connectives are reported to be such a factor. D. Özge et al. (2018) showed a gradient decline for subject preference in Source-Goal ToP contexts (and>so>because). This was also observed by studies in other languages like English (Hartshorne, 2013). Second, verb type appears to affect how nulls are resolved: it influences which referent is interpreted as the antecedent as well as the rate at which this happens (D. Özge et al., 2018; 2017). In line with the preliminary findings by Garvey and Caramazza (1974) and Stevenson et al. (1994); when the subject is also the thematically prominent entity, the subject bias of nulls appears to be stronger in Turkish (Konuk & von Heusinger, To appear; D. Özge et al., 2018; 2017). Lastly, information structure, focushood in particular, is also observed to have a significant effect on the interpretation of null pronouns (Özge & Evcen, 2020). Although referential forms appear to carry their own biases, through the influence of context, these biases may be reduced (Gračanin-Yuksek et al., 2017).

With regards to the overt pronouns in Turkish, they resemble those of other pro-drop languages (e.g. overts in Catalan as reported by Mayol, 2018). They do not appear to prefer subject antecedents, instead showing a bias for non-subject antecedents, and acting as topic-shifters (D. Özge et al., 2018). It is, therefore, argued that they do not pattern parallel to the overts in Japanese or Korean which show a general preference for subjects. However, since the studies contrast overt pronouns with nulls, they do not differentiate the bias of form itself from general interpretation biases. Following the argumentation of Kehler and Rohde (2013), to test how the appearance of a null or overt pronoun in a prompt biases re-mention,

re-mention should be analysed independently of the behaviour of pronominal reference. That is, only comparing a condition in which participants are free in selecting the referential form (i.e. a free prompt) to a condition with a given pronoun would show the bias created by the pronoun.

Following from the models discussed in Section 2.1.1, it is reported that, in Turkish, fuller forms are used for reference to the less prominent referents or in introducing entities (Azar et al., 2016; Konuk & von Heusinger, To appear). Null forms are used in reference to prominent entities, and they are observed to be employed in maintaining reference. The use of overt pronouns, on the other hand, is not a common strategy in reference tracking and they predominantly function as markers of contrast or emphasis.

Table 3.1: Overview of experiments testing the interpretation and production of of null and overt pronouns in Turkish. The table shows whether subject or object is reported as the preferred antecedent of the pronoun. Interpretation experiments are marked by 'Intpn' and production experiments by 'Prod', and 'n.s' stands for non-significant effects.

	Focus	Context	Null Pronoun	Overt Pronoun
D. Özge et al. (2017)	Intpn	OE	Subject	Subject
		SE	Subject n.s.	Object
D. Özge et al. (2018) source-goal ToP contexts	Intpn	and	Subject	Object
		so	No-Preference	No-Preference
		because	Object	Object
Konuk and von Heusinger (To appear)	Intpn	OE	No-Preference	No-Preference
		SE	Subject	Subject
	Prod	OE	No-Preference	Barely used
		SE	Subject	Barely used
Özge and Evcen (2020) physical contact verbs	Intpn	SOV	Subject	Mixed[7]
		OSV	Mixed[8]	Subject
Azar et al. (2016)	Prod	-[9]	Subject (Maintenance)	Barely used (Emphasis marking)

Although pronoun interpretation and production has been demonstrated to be sensitive to different sets of factors, studies in Turkish so far have focused on either production or comprehension. No study exists, to the best of my knowledge, that contrasts the behaviour of interpretation and production of forms. Though grammatical factors have been reported to have an effect on anaphora resolution and also in establishing reference in Turkish, whether

---

[7]Özge and Evcen (2020) investigated verb valence together with information structure. Verb valence had a significant effect for overt pronouns: they were interpreted as coreferential with the object in the positive valence verbs and with the subject in the negative valence verbs.

[8]Positive verb valence yield significantly more subject responses for OSV sentences.

[9]Context type in Azar et al. (2016) included reference tracking strategies, namely maintenance and re-introduction of referents.

semantically or pragmatically driven factors affect the two in similar ways is, therefore, yet to be investigated. Moreover, it is unclear if overt and null pronouns in Turkish are in complementary distribution as is argued for Romance languages with pro-drop (Carminati, 2002). That is, whether overts prefer object antecedents and nulls subject antecedents, or whether they pattern similarly to the overts and nulls in Japanese and Korean in terms of showing a general grammatical bias towards the subject. Reported in the following chapter, Chapter 4, are the results of two studies on anaphora in Turkish that attempt to find answers to these questions.

Chapter 4

Experiments

In this chapter, two passage completion experiments investigating the biases created by two types of verbs, namely implicit causality and transfer of possession verbs, are presented. The subsections contain a detailed presentation of the experimental design, predictions, stimuli and results. Section 4.1 presents Experiment 1, a sentence continuation experiment analysing the biasing effect of IC verbs in the production of referential expressions. This experiment examines whether verb type and subjecthood has an effect on the referent to re-mention and on the referential form chosen. Section 4.2 outlines Experiment 2 which uses a passage completion task. With the use of different prompts, this experiment investigates the interpretation and production of null and overt pronouns in transfer-of-possession contexts. Verb type is also manipulated, and Goal-Source and Source-Goal verbs are contrasted to analyse the influence of semantically/pragmatically driven factors in interpretation and/or production of referential forms. After presenting and discussing the results from each experiment, a general discussion of the findings is given in Section 4.3. Since these verb types have been extensively studied in the field, the results are compared to previous findings discussed in Chapter 2 and 3.

4.1   Experiment 1: IC bias in Turkish

Pronoun interpretation has been shown to be sensitive to grammatical and semantic-pragmatic factors in many languages. On the other hand, production of such forms seems to be sensitive solely to grammatical factors and not to the semantically/pragmatically driven ones. In this section, I present a sentence completion study similar to the design of Stevenson et al.'s (1994), investigating the production of referential expressions in Turkish. In particular, implicit causality verbs (IC) are under scrutiny. As previously mentioned, different types of IC verbs are typically observed to trigger different continuations, and a general bias to mention stimulus entities in the upcoming discourse has been reported. Subject-experiencer verbs are known to show a bias towards object re-mentions while object-experiencer verbs show a bias towards subject. Although IC contexts in Turkish have been investigated for pronoun interpretation by D. Özge et al. (2017), there are not any studies that analyse the production

of forms in these contexts in Turkish. Therefore, I investigate reference production and discuss the results in light of the observations for anaphora interpretation in IC contexts.

4.1.1 Research Questions and Predictions

This experiment attempts to answer four research questions:

(i) Which referent is more likely to be mentioned next?

(ii) Is there a correlation between the referent and verb-type?

(iii) Which form of expression is used in referring to the subject/object?

(iv) Is there a correlation between the form of expression used and verb-type?

In this experiment, only prompts that do not include a pronoun are analysed, and hence, participants are free in choosing a referential form. All prompts ended with the connective *because*, and verb type was manipulated. An example item for each type of verbs is given one below in (52).

(52) a. Ressam Ece'yi   kork-ut-tu   çünkü ... [Object-Experiencer]
          Painter Ece-ACC scare-CAUS-PST because
          'The painter scared Ece because ...'

      b. Ece ressamdan kork-tu   çünkü.... [Subject-Experiencer]
          Ece painter-ABL scare-PST because
          'Ece was scared of the painter because ...'

Analysing the form would allow us to observe which referent is preferred in subsequent mentions and which factors determine the choice of referential forms. In the research questions (i) and (ii), I address the re-mention bias and the contributing factors to this bias. As noted in Section 2.3.3, previous work on implicit causality contexts suggest that the verb type influences which referent is mentioned next (Rohde et al., 2006; Stevenson et al., 1994). In particular, many note a bias for re-mentioning the stimulus entities. Several authors attribute this bias to the expectations on coherence-relations created by IC verbs (Kehler et al., 2008). According to this view, sentences with IC verbs are generally predicted to be followed with an explanation, this in turn creates next-mention biases which differ for the two type of IC verbs (see Kehler & Rohde, 2019: for recent work). As a consequence, object-experiencer verbs are strongly biased to the subject. If Turkish behaves similarly to English in this sense, the likelihood of re-mentions should reflect a subject bias for object-experiencer verbs and an object bias for subject-experiencer verbs.

As for the third question, (iii), I examine the referential forms used. Since Turkish is a pro-drop language, I investigate overt pronouns, null pronouns, full DPs (proper names) as well as reflexive forms (both bare and the inflected form). Taking the literature on the distribution of forms in Turkish (Enç, 1986) and the proposed hierarchy for prominence (Section 2.1.1) into account, nulls are expected to refer back to subjects more often than to

non-subjects. Fuller forms such as overt pronouns or DPs, on the other hand, are predicted to encode objects rather than subjects. Therefore, I expect to find nulls to be used for subject reference and overts for object reference. If this relation holds, it would argue in favour of the division of labour hypothesis for nulls and overts in pro-drop languages (Carminati, 2002). However, it is also possible that overts occur very rarely, and when they do occur, they may do so to mark similarity or contrast (as observed by Azar et al., 2016).

With regard to research question (iv), if thematic roles influence referential form as argued by Rosa and Arnold (2017), the non-subject stimulus entity should be referred to by a reduced form such as the null pronoun. However, following Rohde and Kehler (2014), if thematic roles or coherence driven factors do not affect production, the rate of pronominalisation or null form usage is expected to be similar in both SE and OE verbs.

As previously mentioned in Chapter 3, D. Özge et al. (2017) tested IC verbs for pronoun interpretation in Turkish and found that both null and overt pronouns are resolved towards the subject antecedent in object-experiencer contexts. In subject-experiencer contexts, however, participants tend to interpret the overt as referring to the object more than the subject. Nulls, in this context, were still interpreted as referring to the subject but the rate of this was much lower when compared to the object-experiencer contexts. In short, pronouns in OE contexts appear to prefer subject/stimulus antecedents regardless of the referential form. On the other hand, form remains a determining factor for interpretation, especially in SE contexts, and there is a tendency for having reference to stimulus antecedents. If verb type affects production and interpretation similarly, there should be more object continuations with both overts and nulls in SE contexts. Yet, if forms carry their own bias, null pronouns should be expected in subject reference for both OE and SE contexts.

4.1.2 Method

Participants In total 23 volunteers participated in the questionnaire. Two of them were excluded from the analysis as they completed the questionnaire only partially, and left most of the continuations empty. The majority of participants were university students or recent graduates (Mean age: 28.3).

Material Twenty experimental items were constructed. Each sentence introduced two individuals, one with a proper name (e.g. *Ahmet*) and the other with a noun e.g. *oyuncu* ('actor/player'). All experimental sentences were constructed with two varying types of psych-verbs. These were Subject-Experiencer (SE) and Object-Experiencer (OE) verbs. The verb-type variation was created with the use of either the passive suffixes *-Il* and *-n* or causative morphemes *-DIr* and *-It* [1]. Therefore, the SE and OE verbs were minimal pairs. The complete verb list is provided in Table 4.1 below.

---

[1]The capital letters in the participles indicate the positions that phonological processes occur. These operations take place due to vowel harmony or consonant assimilation and lead to different realisations of the same participle. As for the capitalised consonants they refer to the voicing alterations.

Table 4.1: Verb list of Experiment 1

üzmek–üzülmek	'sadden–be sad'
şaşırtmak–şaşırmak	'surprise–be surprised'
kandırmak–kanmak	'deceive–be deceived'
korkutmak–korkmak	'scare–be scared'
kızdırmak-kızmak	'anger–be angry'
sevindirmek–sevinmek	'delight–be delighted'
tiksindirmek–tiksinmek	'disgust-be disgusted'
utandırmak–utanmak	'ashame–be ashamed'
etkilemek–etkilenmek	'influence–be influenced'
büyülemek–büyülenmek	'charm–be charmed'

There was a sentence connector, namely *çünkü* ('because'), after each sentence to elicit sentence continuations. This conjunction is used to express a causal relation (see Solstad & Bott, 2017: for a fine-grained semantic account). Moreover, all of the chosen Object-Experiencer verbs select a nominative subject and an accusative object, while the Subject-Experiencer verbs either select a nominative subject and a dative object (n=5) or a nominative subject and an ablative object (n=5), as shown in Table 4.2 below. In addition, the position of the individuals were pseudo-randomised in sentences, either occurring as the subject or the object across different verb types. There were also twenty filler items created in the same manner but with agentive verbs. The task was the same for filler items, and participants again wrote sentence continuations. The complete list of critical items and their English translations can be seen in Appendix.

Table 4.2: Overview of Experimental Items

TYPE	CASE	EXAMPLE
OE	ACC	Adam Melih-i kandı-r-dı çünkü ... Man Melih-ACC deceive-CAUS-PST because ...
SE	DAT	Melih adam-a kan-dı çünkü ... Melih man-DAT deceive-PST because ...
OE	ACC	Ressam Ece-yi kork-ut-tu çünkü ... Painter Ece-ACC frighten-CAUS-PST because ...
SE	ABL	Ece ressam-dan kork-tu çünkü ... Ece painter-ABL scare-PST because ...

World knowledge came into play, especially when one of the NPs was a noun for an occupation. For example, upon seeing the main clause (53a) below, one might have an opinion as to what a plausible reason might be for a teacher to upset someone.

(53)   a.   Öğretmen   Ahmet-i   üz-dü   çünkü   ...
              teacher.(NOM) Ahmet-ACC upset-PST because ...
              'The teacher upset Ahmet because ...'

  b. Ø düşük not   ver-di
      Ø low   grade give-PST.
      'The teacher upset Ahmet because s/he gave a low grade.'

To balance this, some sentences were created against the stereotypical roles of characters that partake in the eventuality. An example can be seen below in (54a). In this example, the fisherman is the one being disgusted. Two of the participants noted that it was more difficult to find continuations to such cases.

(54) a. Bekir   balıkçı-yı   tiksin-dir-di   çünkü …
          Bekir.(NOM) fisherman-ACC disgust-PST-CAUS because …
          'Bekir disgusted the fisherman because …'

      b. Ø Kirli ellerle balığ-a dokun-muş-tu.
          Ø dirty hands fishACC touch-IMPF-PST.
          '(He) touched the fish with dirty hands.'

Procedure Items were presented in an online survey through *Google Forms*. Participants accessed the survey via a link. Each participant saw the items in a randomised order, and they were instructed to write coherent continuations, avoiding humour, in the blank space after *because*. The task was not timed but participants were asked to write the first continuation that came to mind and to not spend too long on each item. At the beginning of the task, participants were ensured anonymity and were required to provide a verbal consent.

Analysis A total of 420 completions (20 items × 21 participants) were collected. Each completion provided by the participants was annotated for the referential form used. Classifications of forms were made under four labels: proper name, null pronoun, overt pronoun, reflexive. Since both *kendi* and *kendisi* can have non-local antecedents (Gračanin-Yuksek et al., 2017), both forms were examined. Completions were also coded for antecedent. The base of the coding was the subject of the matrix clause, and subjects of embedded clauses are not coded. Antecedents were analysed in three categories: (i) referential expression referring back to the subject (the first NP) or the (ii) object (the second NP) of the context sentence or (iii) a joint reference (plural reference). No cases of joint reference, in which the expression refers back to both the subject and the object, were observed. When the referents were unclear, completions were judged as ambiguous. Ambiguous completions (n=12) were not included in the analysis. An example of this sort is provided below in (55).

(55) Öğretmen$_i$ Ahmet-i$_j$ üz-dü   çünkü …
      teacher.(NOM) Ahmet-ACC upset-PST because …
      'The teacher upset Ahmet because …'

      a. Ø şapka-sı-nı   al-mıştı.
          Ø hat-3SG.POSS-ACC take-PF
          '(He/she) took his/her hat.'

As the experimental material incorporates a *because*-clause, the data were not analysed for coherence relations. This was due to the connective *because* being an overt indicator of a causal relation and, for this reason, completions for the *Explanation* were expected. In fact, all continuations in the data fell under the Explanation relation. This is consistent with previous observations that IC verbs are widely biased to induce Explanation relations, and that a causal connective used in prompts can reinforce such biases (Stevenson et al., 1994). In a study, Kehler et al. (2008) found that more than half of the continuations (60%) had an Explanation relation even when the prompt is a full stop and no sentence connector is used.

4.1.3 Results

Let us start by discussing the results of re-mention. Figure 4.1 shows the rate at which participants wrote continuations mentioning the subject or the object grouped by the two verb types, namely OE and SE verbs. As expected, reference was influenced by verb type: participants wrote more continuations about the causally implicated entity (i.e. the Stimulus). OE verbs were biased towards re-mentioning the subject whereas the SE verbs created a bias towards the object. This indicates that likelihood of re-mention was consistently higher for the stimulus entity, which was the subject in OE verbs (84%) and object in SE verbs (73%).

Figure 4.1: Proportion of continuations about the subject or the object by verb type
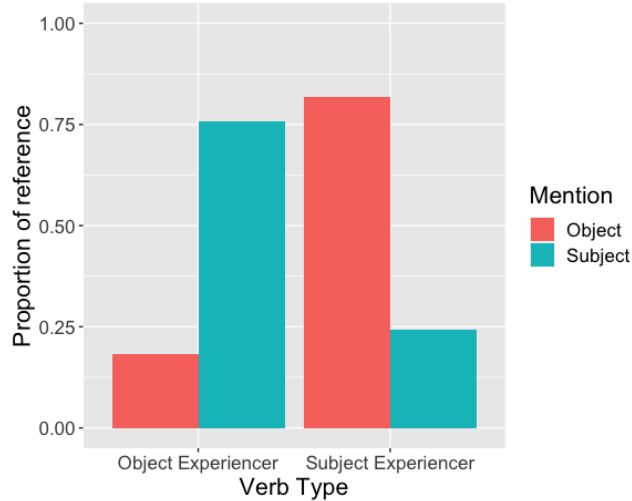
Let us next take a look at the referential forms used by the participants. Table 4.3 breaks down the data, presenting the referential form that was selected to refer back to the subject or object. Using a null is a common strategy when referring back to the subject antecedent. When referring back to the object, nulls appear to be less preferred and, in many cases, proper names (i.e. full DPs) are used. In general, overt pronouns were not used often. As for the reflexive, there were five occurrences of the bare reflexive form *kendi* (subject reference n=4) and one instance of the inflected reflexive *kendisi*, which was used in subject reference.

Table 4.3: Percentages of referential form by re-mention

	Reference	
Form	**Subject**	**Object**
Null Pronoun	85.33 (n= 192)	56.6 (n= 103)
Proper Name	11.11 (n= 25)	41.8 (n= 76)
Overt Pronoun	1.33 (n= 3)	1.1 (n= 2)
Reflexive	2.22 (n= 5)	0.55 (n= 1)

Table 4.3 conflates the verb-type. Therefore, to see whether verb type has an influence on the referring expression that participants use to refer to the subject or the object antecedent, the OE and SE prompts must be examined separately. Figure 4.2 provides the proportions of continuations that mention the subject antecedent by verb type. As shown in Table 4.3, null forms are preferred when mentioning the subject. This pattern is observed regardless of verb type (OE= 87%, SE= 81%). In other words, subjects are dropped when referring back to the subject of the context sentence both in OE and SE verbs. This suggests that the thematic role of the subject antecedent does not have a deciding impact on the use of null pronouns, and participants tend to drop the pronoun even when the subject antecedent is not the causally implicated referent.

Figure 4.2: Proportion of continuations about the subject by form and verb type
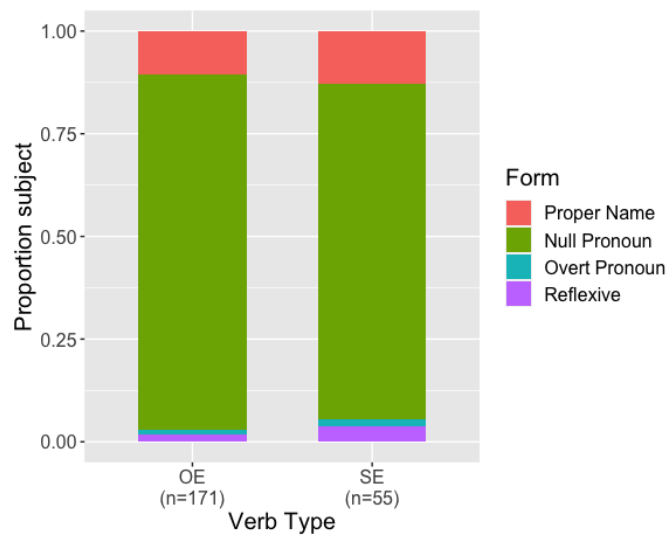
For the cases when the object of the context sentence is referred back to, Figure 4.3 shows the distribution of referring expressions used by verb type. When the object is the experiencer, there is an increase of fuller DPs. In particular, proper names (70%) are preferred over null forms (27%) in OE context. Yet in Subject-Experiencer verbs, null pronouns appear to be used frequently when mentioning the causally implicated object referent. Although null pronouns are generally argued to be subject biased, they are preferred when the object is

also the stimulus, as with SE verbs (63%), but not so often with OE verbs. Therefore, verb type seems to make a difference to references back to the Object.

Figure 4.3: Proportion of continuations about the object by form and verb type
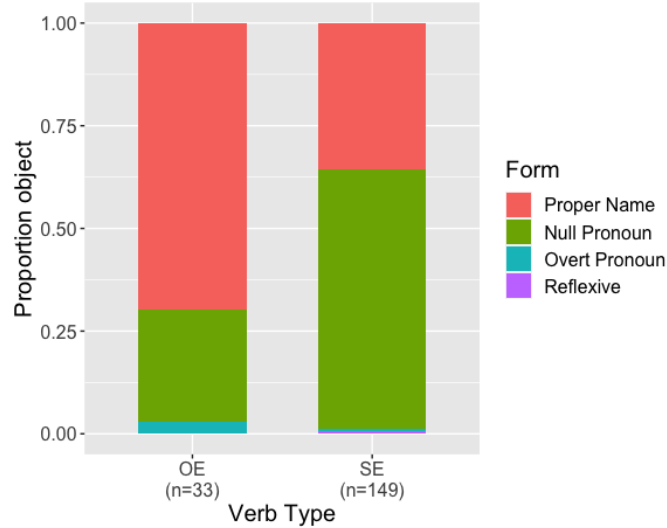
4.1.4 Experiment 1: Summary of Results

Having analysed the results from the production of referential forms in IC contexts in Turkish, preliminary answers to the questions raised in 4.1.1 can now be provided. For question (i), concerning the preferred referent to re-mention, a stimulus bias was found. This bias was induced by verb semantics, and yielded a preference for subsequent mention of the causally implicated stimulus entity. This additionally provides an answer to question (ii) pertaining to whether there is a correlation between the selected referent and verb-type. Clearly the answer is yes. Therefore, the hypothesis that re-mention is sensitive to semantic-pragmatic factors was borne out. However, since only prompts with the connective *because* were tested, further analyses are needed to investigate whether this distribution holds for full stop prompts, whereby the causal relations is not reinforced with a connective.

In question (iii), I addressed which forms are used for subject/object reference. As shown in Figure 4.2, participants overwhelmingly used a null pronoun for subject reference both in OE and SE contexts. In reference back to the object, the picture was more complex. There were more cases of proper names for both verb types when compared to the distribution of proper names to refer to the subject. However, the rate at which proper names are used differed for verb type, and participants preferred the null pronoun to refer to the object when it was also the stimulus. This brings us to the research question (iv) concerning the potential correlation between referential form and verb-type. It seems that there is a correlation, but only in reference back to the object antecedents: participants prefer to use nulls to denote the stimulus referent and proper names to denote the non-stimulus referent. Similar findings were found for the interpretation of nulls in a study by D. Özge et al. (2017). They report that

nulls in Object-Experiencer contexts are interpreted as referring to the subject but the rate at which this occurred was much lower when compared to the Subject-Experiencer contexts. So nulls were used in reference to the non-subject antecedent only in prompts that contained Subject-Experiencer verbs.

In addition, the use of overt pronouns was extremely scarce. It is a possible that this originates from the design of the experiment, specifically the prompt used in the experiment. In fact, the causal structure of the prompt and the connective used may have led to overts, in general, being disfavoured. This is not so surprising if the account that overt pronouns in Turkish act as topic shifters is taken into consideration (Turan, 1996). The existence of the conjunction (because) signals an Explanation relation explicitly and does not leave room for other coherence relations. In fact, if overt pronouns do act as topic shifters, one should expect to find this form with unexpected continuations such as Results or Violated Expectations. Similar distribution of overt pronouns was also observed in Catalan by Mayol (2018). Both of her experiments revealed a low percentage of overt pronouns used in production, and she found that these pronouns occurred in relations that are unexpected. Based on these observations, she concludes that their use is "favoured by discourse contrastivity" (Mayol, 2018: 18).

4.2 Experiment 2: ToP Bias in Turkish

Pronoun interpretation in many languages has proven to be influenced by both grammatical and semantic-pragmatic factors (See Rohde, 2019: for an overview). However, the production of such forms does not appear to be sensitive to semantically-pragmatically driven factors. This pattern has been taken as evidence for an asymmetry in pronoun production and interpretation (Kehler & Rohde, 2013). The current experiment asks whether the same asymmetry holds for Turkish. It reports the results from a passage completion study in which the context sentence introduces two entities and an object with a transfer-of-possession verb. Moreover, two types of verbs, Source-Goal (e.g. give) and Goal-Source verbs (e.g. get), were tested to analyse how the direction of transfer affects interpretation and/or production of pronouns. To examine interpretation and production independently, prompt type was manipulated and conditions without a pronoun (free prompt) and containing an overt pronoun (overt prompt) or a null pronoun (null prompt) were tested.

4.2.1 Research Questions and Predictions

Through a discourse continuation task, this experiment investigates how Turkish, an empirically understudied language, patterns in light of the previous body of knowledge on typologically different languages. In particular, the study attempts to provide answers to the research questions below:

(i) Which factors affect the interpretation of overt and null pronouns in Turkish?

- Do they affect different forms in a similar way?

- How do different referential forms interact, and do they pattern similarly to those of Japanese or English?

- Is there a division of labour between Overts and Nulls in Turkish?

(ii) Do the same factors influence referential choice?

- What determines entity prominence?

- Is the distribution of form accounted for by the hierarchies?

(iii) What contributes in selecting a referent to re-mention?

- Do the syntactically driven factors and/or end-state bias affect how likely a referent is mentioned?

(iv) Does the distribution of coherence relations accompany the expectations of pronominal reference?

- What biases do coherence relations have in ToP events?

- Do different pronouns favour different coherence relations?

The first major question (i) concerns the factors that affect interpretation of the two referential forms in Turkish, namely the overt and null pronoun. Although the literature on pronoun interpretation has mainly focused on languages that only have an overt pronoun like in English, recent studies showed that cross-linguistic differences in pronominal systems give rise to different patterns of pronoun interpretation. Especially in pro-drop languages, overts are argued to behave differently than those of English (Carminati, 2002). While some studies reported a strong object bias for overts (Mayol, 2018), other studies found that this is not the case for every language (Kim et al., 2013; Ueno & Kehler, 2016). Referential forms across languages are also shown to display biases towards referents at different rates. Therefore, the experiment addresses whether different forms are influenced by the same set of factors analogously.

If the null pronoun in Turkish is assumed to denote the most prominent antecedent in line with the accounts in Section 2.1.1, nulls should pattern similar to the overts in English. They should be influenced by thematic role, and more subject biased continuations in the null prompt, in comparison to other prompts, should be observed. As a matter of fact, theoretical work in Turkish linguistics reports that nulls prefer subject antecedents. Yet, as noted in the previous chapter, overt pronouns prefer object antecedents rather than subject antecedents (Turan, 1996). Given this argument, one could expect to observe a greater object bias for the overt prompt, possibly for both verb types (Source-Goal and Goal-Source). This would additionally suggest a division of labour effect, whereby the use of an overt indicates that the antecedent is not the one that is preferred by the null pronoun. In fact, such an effect is found in Catalan despite surfacing only in interpretation and not in production (Mayol, 2018).

However, if it is the case that both null and overt pronouns in Turkish pattern similar to the overt pronouns in English[2], one would expect more subject-biased continuations for both overt and null pronouns. Consequently, this would provide counter-evidence for a division of labour between the two forms.

Recent cross-linguistic work demonstrated that corresponding forms in typologically related or unrelated languages may show dissimilar behaviour in their sensitivity or resilience to different factors (Ueno & Kehler, 2016: *inter alia*). Thus, it is also possible that nulls and/or overts in Turkish behave differently than the overt pronouns in English or overts and nulls in other languages. Such findings would then support a form-specific approach like the one of Kaiser and Trueswell (2008), suggesting a cross-linguistic variation of the weight assigned to each factor that influences pronoun interpretation.

The second question (ii) addresses the asymmetry between production and interpretation, as predicted by the Bayesian model (Kehler & Rohde, 2013). The Bayesian model argues that not all factors influencing interpretation of pronouns influence production. More importantly, it suggests that production of referential form is insensitive to semantic-pragmatic factors. Thus, according to this model, the rate of the pro-drop strategy (and/or the distribution of the overt pronoun) found for subject antecedents in Goal-Source contexts should be commensurate with that of Source-Goal contexts. By contrast, Rosa and Arnold (2017) claim that thematic role influences the rate of pronominalisation. They argue that a referent's predictability (likelihood of re-mention) is determined by several factors, such as the referents' recency and grammatical or thematic role, and that predictability in turn affects referential form. Thus, according to Rosa and Arnold (2017), selection of form, in fact, depends on semantically-pragmatically driven factors. This being said, a preference for using reduced expressions for goal characters more than source characters could be anticipated.

As for the third major question (iii), in line with the next-mention biases reported by Stevenson et al. (1994) and Arnold (2001), semantically and pragmatically driven factors are expected to influence how the discourse is continued. In ToP contexts, participants should write continuations mentioning the Goal referent of the context sentence more often than the Source referent, as found in several studies (Arnold, 2001; Rohde & Kehler, 2014; Stevenson et al., 1994). Prompt type should also reinforce or undermine this Goal-bias as pronouns have their own biases. For English, Stevenson et al. (1994) showed that prompt type reinforces the Goal bias in Goal-Source contexts. They reported that the pronoun prompt leads to a stronger bias for the Goal antecedent in comparison to the free prompt. This effect was stronger for Goal-Source verbs than for the Source-Goal verbs. Following this, I expect to find interpretive preferences for subjects in Goal-Source verbs and to a lesser degree a subject preference for Source-Goal verbs. This is because both subject bias and goal bias are congruent in the former type of verbs. Findings of this sort would additionally confirm the results of Experiment 1 that re-mention is influenced by the verb type. Yet, if the re-mention

---

[2]In fact, nulls and overts in Japanese (Ueno & Kehler, 2016) and Korean (Kim et al., 2013) are found to behave similarly to overts in English, as they show a subject/source bias in Source-Goal.

bias observed in Experiment 1 was a symptom of the connective (because) used, ToP verbs might not exert such an effect in Experiment 2 as only prompts with full stop will be tested. Re-mention has been shown to be extremely pertinent to coherence relations. As for overt prompts, we may observe a reinforced bias for object antecedents, if overt pronouns prefer non-subjects. In a study similar to the present one, Mayol (2018) tested Source-Goal contexts in Catalan. Findings indicated an object bias for all conditions, with an especially strong effect for overt conditions, and 90% of the pronouns in completions had the object of the preceding sentence as its antecedent. The effect was weaker for the Free prompt (76%) and lowest for the Null condition (62%). Main effect of prompt type was significant.

This brings us to the final research question, (iv), on the role of coherence relations. To address this question, I analyse coherence relations and whether they reflect the pattern observed for reference. Crucial to this question is the role of event structure. I address whether the aforementioned argument, that the most prominent antecedent selected for re-mention differs for different coherence relations, holds for Turkish. As such, the antecedent that is associated with the end-state of the context sentence should be chosen as antecedent in an Occasion or Result relation (Rohde et al., 2006). Other relations, such as Elaboration and Explanation, should not focus on the end-state of the sentence, but rather on the initial state and are reported to be source biased (Rohde et al., 2006). The source bias for these two relations have been attested not only in English (Kehler et al., 2008; Rohde, 2008) but also in Korean (Kim et al., 2013) and Japanese (Ueno & Kehler, 2016). With that being said, an increase in the number of Explanation and Elaboration relations is predicted to lead to an increase in Source-biased continuations. Observing such effects would additionally provide support for Kehler's (2002) claims that pronoun interpretation is a by-product of the process in which coherent discourse is created.

Moreover, referential biases (*p(pronoun/referent)*) are reported to influence the distribution of coherence relations (Kehler et al., 2008). This suggests that different prompts affect the likelihood of coherence relations. Therefore, the null prompt may yield more Source-biased relations in Source-Goal verbs than the free prompt. In contrast, if Turkish overt pronouns are interpreted towards the object, we should find an increase in relations attributed to the object, like Occasion and Result relations.

4.2.2 Experiment 2: Method

Participants Thirty participants, the majority of which were university students or recent graduates (n=24), participated in Experiment 2. Mean age was 24.4 (range= 19-45; 22 female). Each participant wrote continuations to 60 one-sentence contexts (30 critical and 30 filler items). They entered a raffle to win a gift coupon. All participants were native speakers of Turkish and were naive with respect to the goal of the experiment.

Material For Experiment 2, a total of sixty sentences were created, thirty of which were filler items. Experimental items introduced either two male or two female individuals and a

direct object[3]. Six different verb pairs were used. Three of them occurred twice (*ödünç almak/ ödünç vermek* 'lend/borrow', *devralmak/devretmek* 'take over/transfer') and (*satmak/satın almak* 'sell/buy', and the other three were employed thrice *vermek/almak* 'give/receive', *uzatmak/kapmak* 'pass/snatch', *teslim etmek/teslim almak* 'hand over/receive')[4]. Verb type (Source-Goal vs. Goal-Source) was a within-subject factor. As for filler items, they were constructed in a similar manner but were used with different types of verbs such as transitive, intransitive or reciprocal.

The context sentence was followed by three different types of prompts. This was a within-subject factor in a Latin-Square design so that no participant saw the same sentence more than once with another prompt. The prompt started with an ellipsis (...), an overt pronoun (*O* 'he/she') or a phrase indicating 'no subject' ([ÖZNE YOK]). From here on these prompts will be referred as free prompt, overt prompt and null prompt respectively. The overt pronoun 'o' in Turkish is ambiguous with regards to gender, and it occurred in nominative form as a prompt. The nominative is a zero morpheme.

This design replicates the discourse continuation paradigm used by Stevenson et al. (1994) (and later Kehler et al., 2008) with the inclusion of the null prompt used in the experiments by Mayol (2018). Given that participants were free to write whatever they wanted as a continuation in the free prompts, they had to select which form to use to refer back to an antecedent. Therefore, data from this prompt will be analysed as production data. While free prompts provide us a way to analyse production of referential forms, other prompts, null and overt prompt, reflect interpretive preferences of null and overt pronouns. In these prompts, production is conditioned by the type of pronoun used in a manner similar to forced production tasks. Participants must interpret the pronoun before providing coherent continuations. Therefore, Overt and Null pronoun conditions enable us to analyse how these forms are interpreted in potentially ambiguous contexts with two potential antecedents. An exemplary item list is shown below in (56). The complete list of critical items of Experiment 2 and their English translations is provided in the Appendix.

(56)   a. Condition 1: Null Prompt
        Nehir Aylin-e   tuz-u   uzat-tı.   **[ÖZNE YOK]** ...
        Nehir Aylin-DAT salt-ACC pass-PST. **[NO SUBJECT]** ...

        'Nehir passed the salt to Aylin. **[NO SUBJECT]** ...'

    b. Condition 2: Overt Prompt
        Nehir Aylin-e tuz-u uzat-tı. **O** ...
        'Nehir passed the salt to Aylin. **She** ...'

---

[3]As the direct object is not relevant for the aim of this experiment, I do not discuss this form further. For this reason, the term 'object' (when not specified as indirect or direct) is used to denote the indirect object henceforth.

[4]The same verbs were needed to be used more than once due to the low number of ditransitive Goal-Source verbs available in Turkish.

c. Condition 3: Free Prompt

Nehir Aylin-e tuz-u uzat-tı. . . .

'Nehir passed the salt to Aylin. . . .'

Even though using a sentence coordinator such as *and* ('ve') in the prompt could make it less challenging for participants to find continuations for the null condition, no sentence connector (because, and, so etc.) was used. The reason behind this decision was to avoid any biases created by such conjunctions. As previously demonstrated, different connectors interact with verb semantics in creating specific biases (Kehler et al., 2008). Since this thesis attempts to analyse coherence relations associated with the ToP verbs, only context sentences ending with a full stop are tested.

Procedure Stimuli were presented using Ibex Farm (Drummond, 2013). An introduction familiarised participants to the prompt type. This process included several example sentences and continuations. For each sentence, at least two potential continuations, which either had reference back to the object or to the subject of the context sentence, were provided. The example sentences involved varying coherence relations so as to not prime the participants to a specific relation. Two practice items appeared after the instructions. The function of the practice items was to acquaint the participants to the design of the continuation task. Sentences were presented one by one; prompts appeared in bold font and each sentence was followed by a blank space into which participants could write their continuations. There was a progress bar on top of the page, and only after writing a continuation on the blank space, were participants allowed to proceed to the next sentence by clicking the button *Devam etmek için lütfen tıklayınız.* ('Please press here to continue.'). This ensured participants provided a continuation for each item.

The condition presented was determined by a Latin square design. Three different lists for prompt types were created. This ensured each participant only saw one version of each sentence. In addition, all items were randomised within different lists and participants read the contexts in different sequences. Given below is a screenshot from the task.

Figure 4.4: Sample trial stimuli
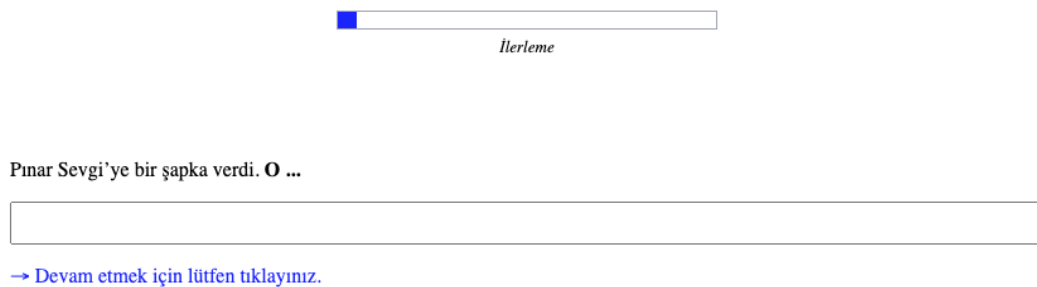
Analysis   A total of 900 continuations (30 items * 30 participants) were collected. Once the continuations coded under the label 'other' (e.g. ambiguous or wrong continuations) were excluded, the remaining data-set contained 725 continuations (81% of the total data). Each sentence was scored for reference (subject/object) and coherence relation used for linking the successive sentences. The coding of coherence relations comprised Occasion, Elaboration, Explanation, Result, Violated Expectation and Parallel, which were introduced in Section 2.3.4, with an addition of two other relations: Background and Denial. The term Background stands for the continuations that provide background to the event described in the prior sentence (Ueno & Kehler, 2016). Denial provides an explanation as to why an event in the prior sentence is unexpected. In addition, the referential form used by the participants was annotated for the free prompt condition to examine the production of forms. As in Experiment 1, production data were categorised under four labels: proper name, null pronoun, overt pronoun, reflexive.

Exclusions from the analysis   In total 185 continuations were excluded from the analysis because they were coded as joint (n=13), other[5] (n=143) or ambiguous (n=29). Out of the discarded categories, there were 28 cases in which participants interpreted the overt pronoun prompt as demonstrative pronoun. This was due to the pronoun being homophonous with the demonstrative 'o'. See (57) for an example.

(57)   Kemal Engin'den hediye çek-i-ni   teslim   al-dı.   O   çek
       Kemal Engin-ABL gift   certificate-POSS-ACC delivery get-PST. That certificate
       hayat-ı-nı   değiş-tir-di.
       life-POSS-ACC change-CAUS-PST

       'Kemal received his gift certificate from Engin. That certificate changed his life.'

In this example, the pronoun 'o' is interpreted as the demonstrative 'o', and the participant used it as a modifier for the *gift certificate*. Moreover, there were cases in which the two forms were ambiguous. Consider the following:

(58)   Alara Berfin'den kulaklık satın aldı. 'Alara bought headphones from Berfin.'

   a. O   kulaklık-lar-ı   uzun zamandır isti-yor-du
      she headphone-PLU-ACC long time   want-PROG-PST

      'She wanted headphones for a long time.'

   b. ∅   O   kulaklık-lar-ı   uzun zamandır isti-yor-du
      (she) those headphone-PLU-ACC long time   want-PROG-PST

      '(She) wanted those headphones for a long time'

This sentence in (58) has two different readings. It is possible to interpret the 'O' as a personal pronoun as in the gloss (58a) or as a demonstrative (58b), and only intonation or a comma in orthography disambiguates the two readings. Such ambiguous continuations, together

---

[5]This category consisted of generic (n=96), demonstrative (n=28) and incomplete continuations (n=19).

with the continuations that involved unambiguous demonstrative use of 'o', were excluded from the analysis. This was a drawback of the design since participants were not instructed to avoid the use of demonstrative pronouns. Yet, familiarising the participants with the difference between personal pronoun and the demonstrative 'O' might have created additional problems. Especially for cases in which familiarity of the direct object is required for the demonstrative reading. Below in (59), the pronoun prompt can potentially be interpreted as a demonstrative modifying the noun *friend*. Yet, this requires familiarity of the noun *friend*, which does not appear to be available in the context given[6]. Therefore, familiarising the participants with the difference between demonstrative and pronoun use of *o* might have hindered continuations similar to the one in (59).

(59) Kaya Can-dan arac-ı   teslim al-dı. O arkadaş-ı-nın   malı-na   zarar
     Kaya Can-ABL car-ACC receive-PST. He friend-ACC-GEN property-DAT damage
     ver-me-mek için çok   dikkatli kullan-dı.
     do-NEG-VN for   very careful drive-PST
     'Kaya$_i$ received the car from Can$_j$. He$_i$ drove very carefully to not damage his friend's$_j$
     property.'

For the null prompt, there were three continuations in which participants did not omit the subject and seven in which participants did not write a full sentence. This verifies that participants were, in general, capable of understanding what a null prompt was and writing an appropriate continuation. There were also some instances of joint reference. For instance, in (60) the participant used null but referred back to both of the characters.

(60) Esra Handan'a   raporu   teslim etti.   ÖZNE YOK Sonra yemeğe
     Esra Handan-DAT report-ACC hand.over-PST. NO SUBJECT later   dinner-DAT
     çıktılar
     went.out-PLU
     Lit. 'Esra handed in the report to Handan. (They) later went out for dinner.'

In addition, there were ambiguities resulting from ambiguous continuations. For example, in (61) the continuation can either be EXPLANATION or VIOLATED EXPECTATION. In the former, the reason behind Nehir's passing the salt is that she doesn't like salty food; whereas in the latter, Nehir might be passing the salt, and yet Aylin might reject the offer and say she doesn't like it.

(61) Nehir Aylin-e   tuz-u   uzat-tı. Ø Yemekleri   az   tuzlu
     Nehir Aylin-DAT salt-ACC pass-PST. Ø food-PLU-ACC little salty
     sev-diğ-ini   ve kaldır-ma-sı   gerek-tiğ-ini   söyle-di.
     like-NMLS-3SG.POSS and get.rid-NMLS-3SG.POSS requireNMLS-3SG.POSS tell-PST
     'Nehir$_j$ passed the salt to Aylin$_j$. (She$_{i/j}$) said she liked less salty food and that she$_{j/i}$
     should get rid of it.

---

[6]This observation appear to square well with Gundel et al.'s (1993) Givenness Hierarchy. As such, the speaker would not use to refer to the *friend* with a demonstrative *that* since it does not appear to be familiar in the context.

Lastly, cases in where the continuations contained elaborations on the direct object or on the event of the context sentence were also excluded from the analysis. These cases were coded as generic. An example of elaboration about the direct object is given below in (62).

(62) Ecem Gülnur'a   dükkan-ı devretti.   Dükkan-ın çevresin-de bir park ve
      Ecem Gülnur-DAT shop-ACC hand.over-PAST. shop-GEN   around-DAT one park and
      bir okul   var-dı.
      one school exist-PST
      'Ecem handed over the shop to Gülnur. Around the shop was a park and a school.'

4.2.3   Experiment 2: Results

A total of 900 sentences were coded for antecedent and coherence relation. Once the continuations that fell under a label other than subject or object were excluded, 715 continuations were analysed. In some continuations scrambling was observed (n=5); such cases were not excluded and were coded for reference. All cases with non-canonical word order re-mentioned the subject of the context sentence and full DP was used. This could be due to information-structure, namely focus position. In canonical SOV order, sentence initial position is reserved for given information while immediate preverbal position is for new information (Erguvanlı-Taylan, 1986). In scrambled word order, the marked postverbal slots is for presupposed information (Erguvanlı-Taylan, 1986). Postposing is argued to be pragmatically motivated and is the position for back-grounded constituents (Kornfilt, 2008). For example in (63) the subject *Esra* occurs after the verb. Although the canonical word-order is SOV in Turkish, here it is OVS. Considering that the immediately preverbal position is the unmarked focus position, participants might have chosen to scramble the word order to demote the proper name and possibly avoid the penalty for repeating names.

(63) Esra Handan'a   rapor-u   teslim et-ti.   Handan'a   güvenmeye
      Esra Handan-DAT report-ACC hand.over-PST. Handan-ACC trust
      başla-mış-tı   Esra.
      start-PERF-PST **Esra**
      'Esra handed over the report to Handan. Esra has begun to trust Handan.'

There were also cases in which participants explicitly provided sentence connectors. As the prompt included a full stop, and sentence connectors can link two independent clauses, *ve* 'and' (n=5), *çünkü* 'because' (n=7) and *ama/fakat* 'but' (n=14) were occasionally written in the continuations. In total, the analysed continuations consisted of 536 continuations referring back to the subject and 179 to the object. Below in Table 4.4, tokens for the distribution of reference to the subject or the object by verb and prompt type is provided. There were more continuations about the object for Source-Goal verbs (n=147) than Goal-Source verbs (n=32).

Table 4.4: Raw data for subject or object reference by prompt and verb type

| Verb	Prompt	Subject	Object
	Free	107	13
Goal-Source	Overt	105	15
	Null	111	4
	Free	77	48
Source-Goal	Overt	60	64
	Null	76	35

Let us first look at the results for Goal-Source verbs, in which the subject is also the end-state related entity (i.e. Goal). Figure 4.5, shows the distributions of subject and object continuations by prompt type. In all three conditions, a strong subject bias is found. The bias is strongest in the null condition (Free= 89%, Overt= 87%, Null= 97%).

Figure 4.5: Subject and Object continuations in Goal-Source verbs by Prompt
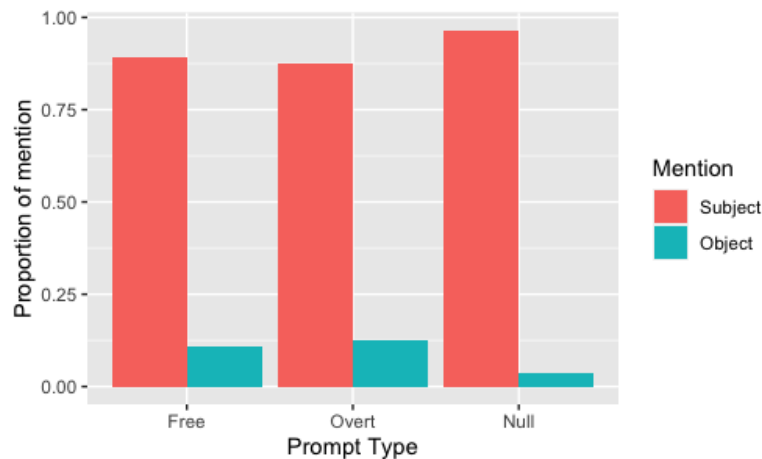
As for Goal-Source verbs, there were more mentions to the object of the context sentence than for the Source-Goal verbs. See Figure 4.6 for distributions. When participants had to select the referential form to continue with, which is the free prompt, they tended to write continuations about the subject more often than the object (Subject= 62%, Object= 38%). The distribution of object and subject continuations was almost 50/50 in the overt prompt (Subject= 48%, Object= 51%). The null prompt, on the other hand, was more likely have subject antecedents (Subject= 68%, Object= 32%). The percentages suggest that null prompts increase the likelihood of subject bias and overt prompts the object bias (the bias for end-state). Yet, to evaluate the statistical significance of this pattern, further analyses are carried out.

Figure 4.6: Subject and Object continuations in Source-Goal verbs by Prompt
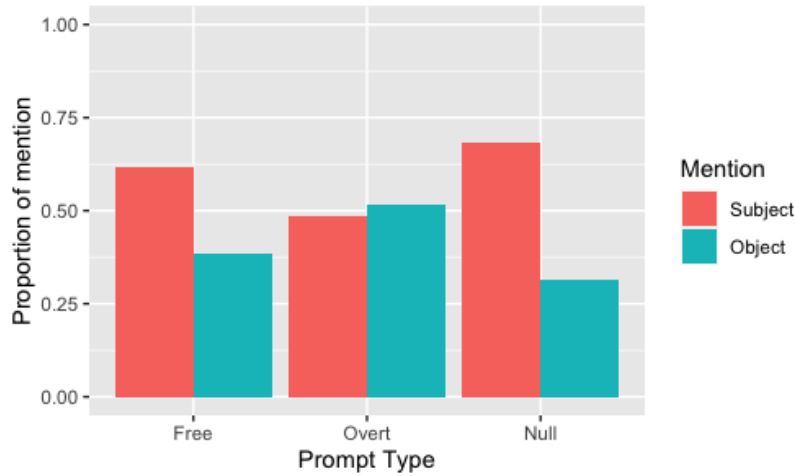
To test for a main effect of prompt and verb type, a mixed-effect logistic regression was performed using R (R Core Team, 2017) with *lme4* (Bates et al., 2015). Reference to subject was the dependent variable. To assess the role of prompt type in inducing biases, free prompt was taken as the baseline, and contrasts were coded to compare the null prompt with the free prompt, and overt prompt to be compared to the free prompt. According to Kehler and Rohde (2013), treating free prompt as the baseline provides the means to appropriately account for the referential biases. They argue that analysing interpretation biases of a form by comparing the distribution of subject and non-subject references would not differentiate between the re-mention bias (interpretive bias) and the bias contributed by the form itself, P(referent)[7]. Instead comparing the no-prompt to the prompt condition would adequately capture the referential biases independent of the bias that the form carries, for instance the subject bias of English overt pronouns. The model included participants and items as random effects.

Table 4.5 shows the estimate, standard errors, z-values and p-values for the intercept (the unweighted grand mean), the effect of verb, the contrasts for prompts and their interaction with verb type. In general, there was an overall preference for subject interpretations, as indicated by the intercept. There was a main effect of verb, with subject interpretations being more common for Goal-Source verbs, as shown in Figure 4.7. There was also a main effect of prompt type but only for one of the contrasts: when null prompts and free prompts were compared (null vs. free), the rate of subject reference was significantly higher for nulls. The interaction was not significant when the overt prompt was contrasted with the free prompt (overt vs. free). The analysis showed no interaction between prompt contrasts and verb type.

---

[7]For the heuristic based approaches, this does not constitute a problem since no differentiation between re-mention and pronoun biases is made, and only P(pronoun | referent) is relevant for interpretation Kehler and Rohde (2013).

Table 4.5: Mixed-effects model for reference to subject of the context sentence

Fixed Effects	Estimate	Std. Error	z value	Pr($>$|z|)	
(Intercept)	1.6294	0.2213	7.36	$<$.0001	***
Verb	-2.3313	0.3260	-7.15	$<$.0001	***
Null vs. Free	0.8760	0.3314	2.64	0.0082	**
Overt vs. Free	0.3590	0.2507	1.43	0.1521	
Verb $\times$ Null vs. Free	-0.9183	0.6585	-1.39	0.1631	
Verb $\times$ Overt vs. Free	0.3645	0.4987	0.73	0.4648	

subreference $\sim$ Verb*Prompt $+(1 \mid$ subject$)+(1 \mid$ sentence$)$

Figure 4.7: Proportion of continuations about the subject by verb type
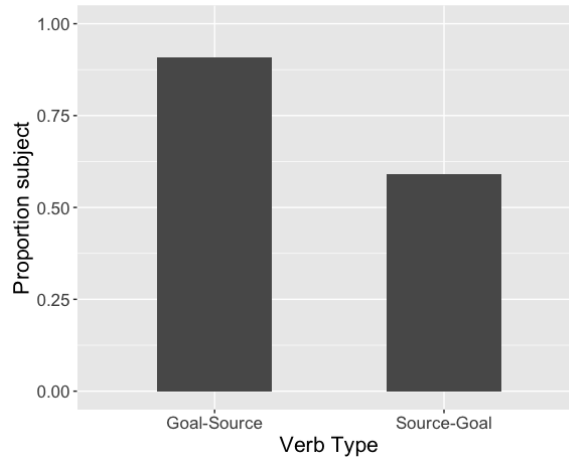
Let us now have a look at the production data. Although participants wrote continuations about the subject in 89% of the free prompts, as shown in Figure 4.5, one needs to observe which referential form participants select in referring to the antecedent of their choice. Such an analysis will indicate participants' preferences for referential choice for subject or object antecedents. In Figure 4.8, distribution of forms in Goal-Source contexts by reference type is presented. When the subject entity was re-mentioned, nulls were preferred (87%), followed by proper names (12%). While Null pronouns are observed to be subject biased, proper names (full DPs) were preferred for mentioning non-subject source entity (85%). In general, Overt pronouns were very uncommon and consisted of 8% of the object reference (n=3) and 2% of subject reference. In Figure 4.9, data form the free prompt in Source-Goal contexts are summarised. Similar to the Goal-Source contexts the Null (80%), and then a proper name (18%) was preferred for the subject. Referring back to the object of the context sentence, participants used a proper name preferably (79%), followed by the null form (13%). While the overt pronoun was used occasionally in object reference (8%), it was not used at all in

subject reference. The reflexive occurred once in referring back to the subject[8]. It was the inflected form, *kendisi*.

Figure 4.8: Distribution of referential form by Subject and Object reference in Free prompt with Goal-Source verbs
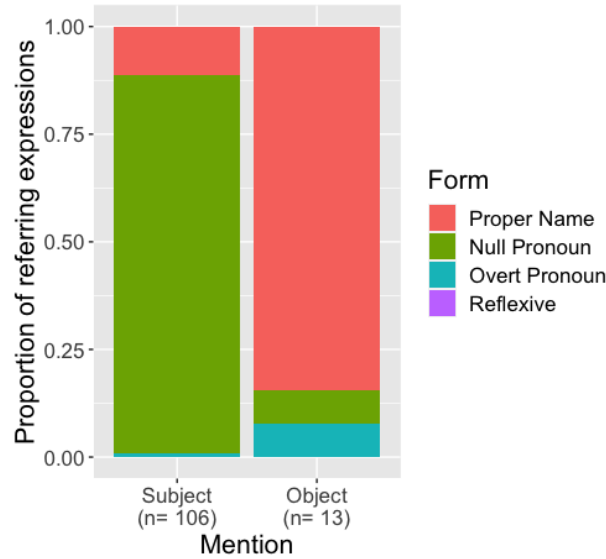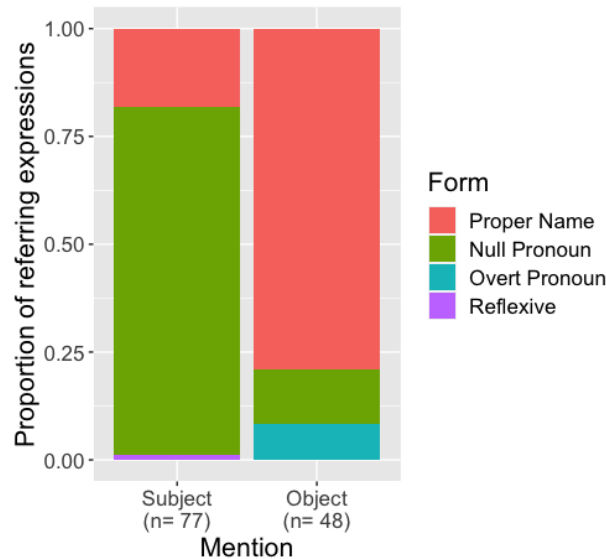

Figure 4.9: Distribution of referential form by Subject and Object reference in Free prompt with Source-Goal verbs

Here a similar pattern between Figure 4.9 and 4.8 can be observed. Reduced forms, such as nulls, are used in referring back to the subject antecedent and fuller forms in reference to the object antecedent. This observation is consistent with the previous literature in Turkish that

---

[8]There were other instances of bare and inflected reflexive forms; however, all occurred in embedded clauses. Therefore, those cases were not included in the analysis.

claims subjects are referred to with null forms (U. Özge et al., 2016; Turan, 1996) as well as with the hierarchies of prominence (see § 2.1.1) that suggest a reduced expression is used to refer back to the most prominent entity, with subjects usually being more prominent than non-subject entities. Thus, notwithstanding the context type (Source-Goal or Goal-Source), null pronouns in Turkish appear to be subject biased. However, since overt pronouns were used very rarely (n=7 in total), it is not possible to decide whether they have a bias towards an antecedent.

Let us now examine the data concerning the coherence relations. The analysis included eight relations taken from Hobbs (1990) and Kehler (2002). Below is a table of the relations relevant for the analysis. All examples are taken from the data and translated to English. Parentheses are used to indicate null subjects, and pronouns stand for overt pronouns in Turkish. There were also cases in which two coherence relations were equally felicitous. These cases were coded as ambiguous and excluded from the analysis.

Table 4.6: Examples of coherence relations from the data

Explanation	Ümit$_i$ sold his phone to Nejat$_j$.
	(He)$_i$ wanted to buy a new phone.
Elaboration	Neslihan$_i$ sold her fridge to İlayda$_j$.
	(She$_i$) did it for a reasonable price.
Occasion	Onur$_i$ borrowed an eraser from Emir$_j$.
	(He)$_i$ gave it back after deleting what he had written.
Background	Koray$_i$ gave a scarf to Canberk$_j$.
	He$_i$ learned how to sew for Canberk.
Violated Expectation	Pınar$_i$ gave a hat to Sevgi$_j$.
	(She)$_j$ did not like the hat (she)$_i$ gave at all.[9]
Result	Simay$_i$ lent her phone charger to Ipek$_j$.
	She$_j$ was very pleased with this.
Denial	Simay$_i$ lent her phone charger to Ipek$_j$.
	She$_i$ usually does not like to share her charger.
Parallel	Koray$_i$ gave Canberk$_j$ a scarf.
	He$_j$, therefore, gave him$_i$ a hat.

Considering the distribution of coherence relations, there was a fair share of Explanation (n=204) and Occasion (n=199) relations, followed by Result (n=125), then Elaboration

---

[9] Although the English translation appears to be ambiguous, the original continuation is disambiguated by the simple past tense of the verb in the matrix clause. In Turkish, the first pronoun, (She)$_j$, can only be coindexed with Pınar$_i$ when the verb *like* occurs in imperfective aspect.

(n=99). There were also some continuations with Background (n=38) and Violated Expectation (n=37). Parallel (n=9) and Denial (n=4) occurred in very few cases, and hence, will not be discussed further. Below, in Table 4.7, percentages of rhetorical relations for each verb type is shown.

Table 4.7: Percentages of coherence relations by Verb Type

Relation	Goal-Source (get)	Source-Goal (give)
BACKGROUND	4.2	6.4
DENIAL	0.3	0.8
ELABORATION	11.3	16.4
EXPLANATION	29.6	27.5
OCCASION	36.6	19.2
PARALLEL	1.1	1.4
RESULT	14.4	20.6
VIOLATED EXPECTATION	2.5	7.8
Total	100	100

As expected, ToP contexts involved many Occasion relations, and participants completed the story by providing what happened next. Many Explanation and Elaboration relations were also observed. As previously mentioned, studies reported these relations to be subject-biased even when the subject is not the end-state related entity (Kehler, 2002; Kehler et al., 2008). Thus, this observation is not surprising given the high number of subject related continuations reported in Table 4.4.

To analyse the effect of prompts, Figure 4.10 breaks down the coherence data by prompt and verb type. In Goal-Source verbs, more than 60% of all continuations in each prompt was either Explanation or Occasion. In Source-Goal verbs, continuations were more evenly distributed across relations, and there were more instances of relations other than Explanation and Occasion. Free prompts show which relation ToP verbs prefer by default. In a great deal of continuations, participants tend to link clauses with an Occasion relation in Goal-Source context. When the free prompt is taken as the baseline, the null prompt shows a near uniform distribution of relations for both verb types. As for overt prompts, a slightly different pattern is found. There seems to be a shift to Explanation relations, and less Occasion relations are observed. To get a full understanding of coherence relations, let us now analyse which antecedent is preferred for each relation. Violated Expectations and Background relations were few in number, and therefore not analysed further. Yet, it must be noted that Background relations were in general subject biased for both verb types (Subject n= 29, Object n= 9), and Violated Expectations were found to be object biased in Source-Goal verbs (Subject n= 19, Object n=9).

When the four most frequently used coherence relations are analysed, Figure 4.11 showed that there are more object related continuations in Source-Goal (i.e. *give* type) compared to Goal-Source (i.e. *get* type) verbs.

Figure 4.10: Proportion of coherence relations for all conditions

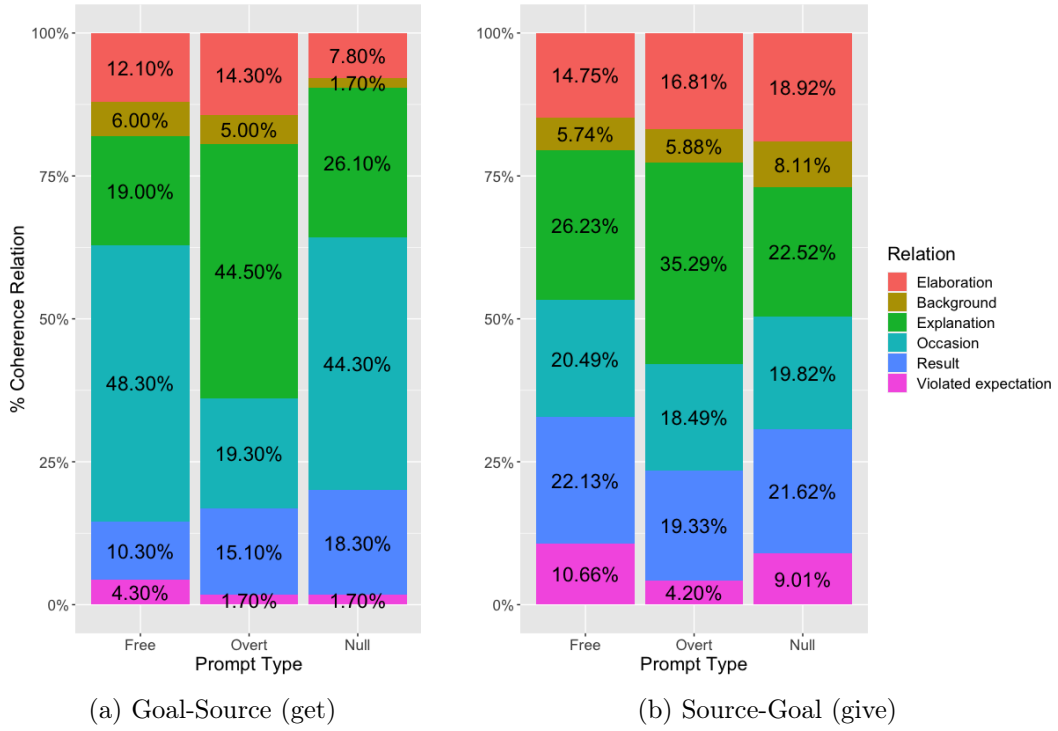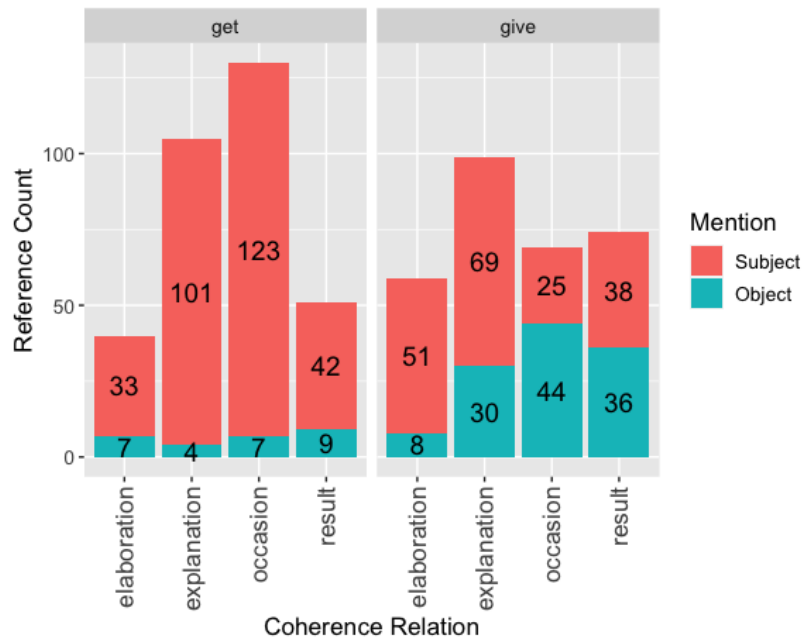(a) Goal-Source (get)   (b) Source-Goal (give)

Figure 4.11: Tokens per reference to subject or object by coherence relation collapsed across different prompts

The major share of relations was Occasion and Explanation for Goal-Source verbs. The relations were more equally distributed for Source-Goal verbs. For this reason, this thesis will focus on these particular verb types. The results indicate that Elaboration is highly

subject biased, even when the subject is not the goal referent, and there were only a handful of object related Elaborations. This is argued to be the case for English (Rohde, 2008; Rohde et al., 2006), Japanese (Ueno & Kehler, 2016), Korean (Kim et al., 2013) as well as Catalan (Mayol, 2018). Explanation is also reported to be source biased for these languages except Catalan. Although the data appear to support this claim, quite a few Explanation continuations about the object in Source-Goal contexts are still found. In these continuations, participants typically wrote about why the object referent needed the transferred item. For instance, one participant wrote:

(64) Can Mehmet-e   arac-ı-nı   ödünç ver-di. **O** . . . iş-e   geç kalmış-tı,
       Can Mehmet-DAT car-ACC-POSS lend   give-PST **he** . . . work-DAT late be.late-PST
       otobüs-le zaman kaybetmek iste-mi-yor-du.
       bus-INS   time   lose   want-NEG-PROG-PST
       "Can lent his car to Mehmet. **He** . . . was late to work and did not want to lose time
       taking the bus.'

For Catalan, participants wrote continuations that are similar to (64). Mayol (2018) suggests that, in a finer-grained categorisation, such continuations could be coded under *conveying purpose* rather than Explanation. Findings of Experiment 2 differ from Mayol's in that, while she reports an object bias for null prompts (71%), not many object continuations for this prompt is found in Experiment 2 (16% of total Elaborations in null prompt). This suggests that Explanation, despite being influenced by verb-type, is not as strongly subject biased as Elaboration. Moreover, for Occasion and Results relations, there was a shift towards mentioning the object. This is consistent with the claims that these two relations are end-state related and are, thus, biased for goal antecedents (Kehler et al., 2008). Yet, Result does not seem to be as object biased as Occasion relations in Source-Goal verbs. A closer look revealed that Results in the null prompt lead to more subject related continuations (67%) while in the overt prompt they yield more object related ones (61%).

4.2.4   Experiment 2: Summary of Results

After reviewing the results, we can now now attempt to answer the research questions posed in Section 4.2.1. The first question addressed the factors that influence the interpretation of the forms. The results indicate that both semantically and grammatically driven factors affect interpretive preferences. In general, there were more continuations related to the subject antecedent. An overall main effect of verb type was found, this implies that verb meaning plays a role in interpretation. The interaction between the null and the free prompt demonstrated that the appearance of the null pronoun yields more subject references. Therefore, I conclude that null pronouns in Turkish appear to behave similarly to overts in English and nulls in Catalan and are subject to both grammatically and semantically/pragmatically driven factors. In addition, null pronouns in Turkish are in partial similarity to those of Japanese whereby their appearance leads to an increase in subject re-mentions yet they are sensitive to semantic-biases.

Overt pronouns in Source-Goal contexts were resolved towards the object in around 50% of the total continuations. Yet, when contrasted with the null prompt, the rate of subject continuations did not reach significance. Therefore, despite often resolved towards non-subject antecedents, overts do not show a clear preference for interpretation.

It appears that verb type does not determine the use of referential forms; participants used nulls to refer to the subject, whereas they preferred a proper name in reference back to the object. This remained constant for both verb types. Given that there was an increase of object biased interpretations of the null pronoun in Source-Goal verbs when compared to the Goal-Source, interpretation and production of nulls appear to be sensitive to different factors. Whether this is the case for interpretation and production of overts remains unclear since their use was very rare, occurring only seven times in the total data despite displaying an almost 50/50 preference for interpretation. Thus, it seems that there is no strong division of labour between overts and nulls.

Conforming to the observations made by Rohde and Kehler (2014) among others, the data revealed an overwhelming bias for goal entities in subsequent mentions. Although the source antecedent was still preferred in many continuations when it was also the subject (as in Source-Goal verbs), there was a crucial difference between two verb types. Participants showed an increased tendency to re-mention object/goal antecedents in Source-Goal verbs. These results replicated Stevenson et al.'s (1994) who reported more Goal continuations following context sentences with Goal-Source verbs in pronoun prompt (84.6% Goal) than there were with Source-Goal verbs (49.0% Goal). Although more object biased continuations were found with overt prompts rather than free prompts, it is not possible to conclude that overts carry an object bias since the interaction for the second contrast between free and overt prompts did not reach significance (see Table 4.5).

With regard to the research question (iii), concerning the production of forms, nulls are preferred when referring back to the subject. As this pattern was observed for both verb types, I conclude that nulls are insensitive to semantically driven factors. In reference back to the object, participants predominantly used proper names. Overt pronouns were used only seven times and the reflexive only once in the free prompt. The rare occurrence of overts implies that the context does not support this form. As suggested by several studies, this form functions as a marker for similarity or contrast (Azar et al., 2016). Accordingly, including a contrastive connective, such as *fakat* 'whereas', in the prompt instead of a full stop may have induced more overt pronouns in the data.

As for question (iv) on coherence relations, participants overwhelmingly linked clauses with Occasion and Explanation, then with Elaboration and Result. When prompts were analysed, a shift towards using more Explanations in overt prompts was found, and null prompts displayed a similar distribution to free prompts. Data from coherence relations replicated previous findings on Explanations and Elaboration to show a subject bias (Kehler et al., 2008: among others). Yet, there were still quite few cases of object reference for Explanations. Occasion yielded more object re-mentions in Source-Goal contexts as expected

for end-state bias (Kehler et al., 2008). On the other hand, participants referred either to the subject or the object in Result relations.

4.3   General Discussion

In this chapter, findings from two experiments in Turkish were reported. Although there are several studies that empirically test pronoun interpretation in IC or ToP contexts in Turkish (D. Özge et al., 2018; 2017), there does not exist a study that investigates production of forms in these context. First, Experiment 1 investigated production of pronouns in IC contexts and examined whether and how reference production is influenced by grammatical roles of the antecedents and/or verb type. Experiment 2 analysed both interpretation and production of pronouns in ToP contexts to examine whether both are sensitive to the same set of factors. For this experiment, prompt type was manipulated and prompts with an overt or null pronoun or without any pronouns were analysed, these pronouns are referred to as the overt prompt, the null prompt and the free prompt, respectively.

The results from both experiments revealed that when participants had to select which referential form to use, they did so either by employing a null pronoun or a proper name. Only a few overt pronouns were observed for Experiment 1 and for the free prompt in Experiment 2. In previous studies of Turkish, the use of overt pronouns appeared to be immensely restricted, and they were reported to be an uncommon strategy in reference tracking (Azar et al., 2016). This suggests the function of these forms is not to signal the prominence of a referent in a discourse. Otherwise, we should have observed more instances of this form instead of the proper name (or presumably instead of the null). When the few occurrences of overt *o* are analysed closely, they are found to be used mainly to refer to non-subject referents. Yet, they are not preferred across the board. Though at a different rate, this was also reported to be the case for Catalan (Mayol, 2018). It is not exactly clear why overts are widely dispreferred in Turkish for these contexts. One possible reason could be that their function is to act as an indicator of similarity or contrast, and only when participants want to signal such a function are they used. It is also presumable that the IC and ToP contexts do not favour such discourse status of contrast or similarity by default. Therefore, other contexts, for instance, sentences with connectives such as *whereas* and *although* (for contrast) or *also* and *too* (for similarity), must be studied in order to investigate whether the use of overts can be facilitated by contextual factors.

Analysing the pronominal form in terms of the selected antecedent, the data indicated that while participants re-mentioned the subject referent in Object-Experiencer and Goal-Source verbs, they re-mentioned the object referent in Subject-Experiencer and Source-Goal verbs. The rate at which these referents were mentioned was higher for Experiment 1 as there were still many subject references in Source-Goal verbs. Turkish, nevertheless, appears to corroborate with a cross-linguistic tendency to re-mention Goal and Stimulus referents.

Moreover, pro-drop was a widely employed strategy in reference back to the subject antecedent while a proper name was preferred for the non-subject antecedent. This distribu-

tion was constant irrespective of the verb type in Experiment 2. In Experiment 1, however, more nulls were found in reference back to the object when it was the stimulus (i.e. Subject-Experiencer verbs). This was also reported for null pronouns in Japanese (Ueno & Kehler, 2016), Italian Fedele (2016), and Catalan (Mayol, 2018) Given that connectives are reported to influence biases, this could have resulted from the fact that all prompts in this experiment were because-clauses and/or that the pronominal dependencies were intra-sentential. In particular, the connective may have reinforced the stimulus bias, and made it a more prominent referent to be re-mentioned in a more reduced form, i.e. as a null pronoun.

As for factors affecting interpretation, recall that there was an effect of prompt type for the first contrast, null vs. free prompts (reported in Table 4.5), and null prompts yielded more subject mentions. Yet, there was still some variation in re-mentions, and null pronouns in Source-Goal contexts showed a significantly milder bias towards the subject. This suggests that interpretation might be influenced by both the pronoun bias and semantic-pragmatic driven factors as found for English overt pronouns (Rohde et al., 2006) and Korean null pronouns (Kim et al., 2013; Kwon & Sturt, 2013; Song & Kaiser, 2020). Moreover, Ueno and Kehler (2016) note that "[i]n every experiment that we are aware of ... [t]he occurrence of a pronoun always contributes a subject bias". Though the percentages in the data suggest a parallel pattern, the interaction between overt and free prompts did not reach significance. Had there been a larger number of participants, a meaningful interaction may have been found. Therefore, the data reported here are inconclusive.

In a pronoun resolution study in Turkish with because-clause prompts (similar to the ones used in Experiment 1), D. Özge et al. (2017) demonstrated that a null pronoun significantly increases the subject preference only for SE verbs, and nulls were widely resolved towards the non-subject. In contrast, null pronouns in OE verbs were readily resolved towards the subject. In the case of overts with SE verbs, they were not interpreted as referring back to the subject but to the object instead. Later on, in another study, D. Özge et al. (2018) investigated Source-Goal ToP contexts and reported a similar distribution of null and overt pronouns: while nulls' appearance increased subject references when used with the connectives *and* and *so*, overts tended to be resolved towards the non-subject referent regardless of the coherence marker. They concluded that while both grammatical and semantic factors affect the interpretation of null pronouns, overt pronouns are influenced solely by grammatical factors. When production was tested, it is found that null pronouns can in fact be used to refer to non-subject referents but only in Experiment 1 where the prompts included the connective *because*, and the pronominal dependency was intra-sentential. On the other hand, overts were rarely used for non-subjects. Thus, these results point towards the idea that there is an asymmetry between pronoun interpretation and production as proposed by Rohde and Kehler (2014), especially for overt pronouns. They are interpreted as referring to the object, yet they are not used for reference to non-subject antecedents in the absence of a connective. Further analyses are, nevertheless, required to test whether these observations are meaning-

ful, and additional linear models needs to be performed in order to examine whether the Bayesian model, or other models, can adequately predict the observed tendencies.

Last but not least, for coherence relations, which were only analysed for the data from Experiment 2, many Occasion and Explanation relations were observed which was in line with the findings of previous studies on ToP contexts (Kehler et al., 2008: *inter alia*). Even though participants were only instructed to provide continuations to the story, and the task did not involve an implicit question following the prompt, they overwhelmingly chose to write a continuation providing an answer to the question *What happened next?* (Occasion) and *Why?* (Explanation). A fair share of Elaboration continuations were also observed. This relation answers the question of *How?*. Consistent with Kehler et al. (2008), Elaboration and Explanation were mostly subject biased and Occasion was goal biased. On the other hand, the Result relation did not show a clear bias to a specific referent. This relation, instead, was influenced by the prompt type.

Chapter 5

Conclusions

In this thesis, pronoun production and interpretation were analysed with an aim to extend the current knowledge of the production and interpretation of pronouns in Turkish.

The psycholinguistic literature on anaphora has concentrated on the third person singular pronoun in English, and many studies have attempted to identify the factors that determine how this pronoun is interpreted. It has been argued that the most prominent entity in the discourse is selected as an antecedent for the pronoun. As discussed in Chapter 2, a number of factors, including grammatical role, topichood, thematic role, information structure etc., have been argued to influence a referent's degree of prominence.

Yet, when researchers examined production, an interesting picture emerged. The referent that participants tend to mention next in the discourse was not necessarily mentioned with a pronoun. From this, Rohde and Kehler (2014) proposed that there exists a disparity between pronoun production and interpretation. They subsequently claim that different factors influence interpretation and production: interpretation is sensitive to semantic/pragmatic and structural factors whereas production is only sensitive to structural ones. This difference is a mere polarity not a complete asymmetry and is formalised probabilistically by Bayes' rule (see Section 2.4.3 for a detailed presentation of the model).

Recent studies showed that the behaviour of referential forms across different languages are far from uniform, and that two forms classified under the same grammatical label (e.g. personal pronouns, null pronouns) may show distinct characteristics. They may be differentially affected by different factors depending on the variety of forms available in the inventory of a given language. This divergence can be observed even between typologically similar languages (Ueno & Kehler, 2016). As shown by several authors, in pro-drop languages the appearance of overt forms is conditioned by structural as well as pragmatic and semantic properties.

By conducting two experiments employing the passage completion paradigm, the present thesis attempted to investigate how null and overt pronouns behave in Turkish. In line with previous studies in Turkish, nulls were resolved towards the subject antecedent (D. Özge et al., 2018; 2017). As for their use, null pronouns seem to be the default form in reference

back to the subject antecedent. This provides support to the idea that null pronouns in Turkish can be taken as the counterpart of unstressed overt pronouns in English. In Experiment 2, null forms were used for subject antecedents in ToP contexts regardless of verb type. However, in Experiment 1, verb-type proved to have an effect on the choice of form: when referring to the object, more null pronouns were used with SE verbs than with OE. In addition, implicit causality verbs showed a strong bias towards re-mentioning non-subject antecedents. These results suggest that null pronouns in Turkish can also be used to refer to non-subject referents, and that these forms are modulated by semantically/pragmatically driven factors.

I have to note that production data from the two experiments are not fully comparable given that while Experiment 1 investigates production of pronouns within sentence boundaries, Experiment 2 investigates it across sentence boundaries. Therefore, it would be useful to analyse production in ToP contexts with prompts including sentence connectives, to better understand the specific biases contributed by the coherence markers.

Moreover, overt pronouns were observed scarcely in both experiments. These forms have been previously observed to function as signals of topic status (Azar et al., 2016; Enç, 1986). It is possible that the design of the experiments did not allow for such functions. To test this supposition, topichood should be examined independent of subjecthood. The experiments in this thesis conflated these two notions. Analysing referents within longer contexts may make it possible to further investigate the use of overts. Rendering information status by scrambling word order or by manipulating the recency of mentions, could reveal the use of overt pronouns in Turkish.

Another notable outcome from Experiment 2 was that overts did not appear to significantly prefer an entity whereas nulls increased the number of continuations about subject antecedents. This argues against a strong form of *division of labour* between nulls and overts. As found by Konuk and von Heusinger (To appear), the distribution of overts and nulls in Turkish does not seem to occur in a complementary fashion. Similar patterns were observed for various other languages. For instance, while nulls showed a strong subject bias in Catalan Mayol (2018) and Korean (Song & Kaiser, 2020), overts were not necessarily object biased, instead they were also used in reference back to subject antecedents. Likewise Japanese nulls and overts did not conform to a complementary distribution, instead both were subject biased (Ueno & Kehler, 2016). Thus, recent cross-linguistic studies appear to provide support for form specific approaches, as argued by Kaiser and Trueswell (2008). Yet, given that the overt pronoun was used very infrequently, I refrain from making any further claims on the (in)existence of a division of labour in Turkish.

Taken together, the results from Experiment 2 seem to support the claims by the Bayesian Model of pronoun resolution (Kehler et al., 2008). There was an asymmetry between interpretation and production: while null pronouns were resolved towards subject or non-subject referents, they were not used to refer to non-subjects. A similar pattern was found for overt pronouns: while they tended to be resolved towards non-subjects (though

this interaction did not reach significance), they were not used for this purpose at all. More-over, pronoun resolution was influenced not only by syntactic but also by semantic and prag-matic factors, as argued by the Bayesian Model. Participants shifted away from interpreting the pronouns as referring to subject due to verb semantics, and Source-Goal verbs yielded significantly more continuations about the non-subject antecedents. Even though pronoun resolution in different languages show different degrees of sensitivity to various factors, these results replicate the data from Korean (Kim et al., 2013; Kwon & Sturt, 2013; Song & Kaiser, 2020), Japanese (Ueno & Kehler, 2016), Italian (Fedele, 2016), and Catalan (Mayol, 2018), that both syntactic and semantic/pragmatic factors affect interpretation of pronouns.

Although this thesis solely scrutinised subjecthood as a structural factor, studies in-vestigating the influence of various other factors, like linear order and information structure, are necessary to create sufficient models that predict re-mention effects. It is also crucial to investigate how manipulations may affect re-mention bias by providing context rather than testing sentences in isolation. Only when analysing pronouns in more natural environments, would it be sensible to talk about factors influencing their behaviour.

In addition, results from Experiment 1 and 2 indicated that re-mention is sensitive to the previously discussed structural and semantic/pragmatic factors. This raises a questions as to whether analyses of written and spoken corpora would reveal a similar pattern. While corpora were analysed for a number of languages (e.g. Bosch et al., 2003; Portele & Bader, 2016: German), I am not aware of any existing work that informs this question for Turkish.

Following previous works, this thesis explored the behaviour of 3rd person singular pronouns (overt and null). Yet, this is only a small part of a larger referential system. For instance, the plural *they* may also be ambiguous in contexts whereby three or more referents are introduced. It is yet to be answered whether the plural *they* behaves similarly to the 3rd person singular pronouns. Also, the question of what guides speakers to choose a noun instead of a more informative form, such as a noun modified with an adjective, remains an issue for further research (for a recent overview on informativeness and referential choice see Davies, Arnold, Davies, & Arnold, 2019).

Whilst beyond the scope of this paper, it is also worth considering how pronouns are processed on-line, and at what point of processing the interpretive biases are manifested. Only by carrying out online studies, can one obtain a fuller picture on the behaviour of referential forms, in terms of predictive and interpretive mechanisms.

To sum up, only by an amalgamation of various factors, both syntactic and semantic-pragmatic, can theories and models adequately account for the interpretation patterns re-ported in the literature. One thing that is clear is that further studies are needed in order to assess whether a synthesis of several factors is necessary to account for production of such forms. Or whether the weight of factors, presumably for both processes, vary cross-linguistically for pronouns in languages with divergent referential systems as suggested by Kaiser and Trueswell (2008).

Chapter 6

References

Aksu-Koç, A., & Nicolopoulou, A. (2015). Character reference in young children's narratives: A crosslinguistic comparison of English, Greek, and Turkish. *Lingua*, *155*, 62–84. doi: 10.1016/j.lingua.2014.04.006

Alonso-Ovalle, L., Fernández-Solera, S., Frazier, L., & Clifton, C. (2002). Null vs. Overt Pronouns and the Topic-Focus Articulation in Spanish. *Rivista di Linguistica*, *14*(2).

Ariel, M. (1988). Referring and accessibility. *Journal of Linguistics*, *24*, 65–87. doi: 10.1017/S0022226700011567

Ariel, M. (1990). *Accessing noun phrase antecedents*. London/Newyork: Routledge.

Ariel, M. (2001). Accessibility theory: An overview. *Text representation: Linguistic and psycholinguistic aspects*, *8*(8), 29. Retrieved from `https://www.jbe-platform.com/content/books/9789027297679-hcp.8.04ari`

Arnold, J. E. (1998). *Reference form and discourse patterns* (Doctoral Dissertation). Stanford University.

Arnold, J. E. (2001). The effect of thematic roles on pronoun use and frequency of reference continuation. *Discourse Processes*, *31*(2), 137–162. doi: 10.1207/S15326950DP3102_02

Arnold, J. E. (2010). *How speakers refer: The role of accessibility*. doi: 10.1111/j.1749-818X.2010.00193.x

Arnold, J. E., & Griffin, Z. M. (2007). The effect of additional characters on choice of referring expression: Everyone counts. *Journal of Memory and Language*, *56*(4), 521–536. doi: 10.1016/j.jml.2006.09.007

Azar, Z., Backus, A., & Özyürek, A. (2016). Pragmatic relativity: Gender and context affect the use of personal pronouns in discourse differentially across languages. In A. Papafragou, D. Grodner, & D. Mirman (Eds.), *The 38th annual meeting of the cognitive science society (cogsci 2016)* (pp. 1295–1300). Austin : Cognitive Science Society. Retrieved from `http://hdl.handle.net/2066/159293`

Azar, Z., Backus, A., & Özyürek, A. (2019). General- and Language-Specific Factors Influence Reference Tracking in Speech and Gesture in Discourse. *Discourse Processes*, *56(7)*, 553–574. doi: 10.1080/0163853X.2018.1519368

Bader, M., & Portele, Y. (2019). The interpretation of German personal pronouns and d-pronouns. *Zeitschrift für Sprachwissenschaft*, *38*(2), 155–190. doi: 10.1515/zfs-2019-2002

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi: 10.18637/jss.v067.i01

Bittner, D., & Kuehnast, M. (2012). Comprehension of intersentential pronouns in child German and child Bulgarian. *First Language*, *32*(1-2), 176–204. doi: 10.1177/0142723711403074

Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review*, *89*(1), 1–47. doi: 10.1037/0033-295X.89.1.1

Bock, K., & Levelt, W. (1994). Language production: grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). Academic Press.

Bosch, P., Rozario, T., & Zhao, Y. (2003). Demonstrative pronouns and personal pronouns: German der vs er. In *Proceedings of the 2003 eacl workshop on the computational treatment of anaphora.*

Bott, O., & Solstad, T. (2014). From Verbs to Discourse: A Novel Account of Implicit Causality. In *Psycholinguistic approaches to meaning and understanding across language. studies in theoretical psycholinguistics* (Vol. 44, pp. 213–251). Springer. doi: 10.1007/978-3-319-05675-3_9

Brennan, S. E., Friedman, M. W., & Pollard, C. J. (1987). A centering approach to pronouns. In *Proceedings 25th annual meeting of the association for computational linguistics* (pp. 155–162). Stanford, CA: Association of Computational Linguistics.

Brown, R., & Fish, D. (1983). The psychological causality implicit in language. *Cognition*. doi: 10.1016/0010-0277(83)90006-9

Cameron, R. (1992). *Pronominal and null subject variation in Spanish: Constraints, dialects, and functional compensation* (Doctoral dissertation). Retrieved from https://repository.upenn.edu/dissertations/AAI9227631

Carminati, M. N. (2002). *The processing of Italian subject pronouns* (Doctoral Dissertations Available from Proquest.). University of Massachusetts Amherst.

Chafe, W. (1976). Givenness, Contrastiveness, Definiteness, Subject, Topics, and Point of View. In C. N. Li (Ed.), *Subject and topic* (pp. 25–56). New York: Academic Press Inc.

Clark, H. H., & Sengul, C. J. (1979). In search of referents for nouns and pronouns. *Memory & Cognition*, *7*(1), 35–41.

Colonna, S., Schimke, S., & Hemforth, B. (2012). Information structure effects on anaphora resolution in German and French: A crosslinguistic study of pronoun resolution. *Linguistics*, *50*(5), 991–1013. doi: 10.1515/ling-2012-0031

Colonna, S., Schimke, S., & Hemforth, B. (2014). Information Structure and Pronoun Resolution in German and French: Evidence from the Visual-World Paradigm. In (pp. 175–195). Springer, Cham. doi: 10.1007/978-3-319-05675-3_7

Comrie, B. (1997). Pragmatic Binding: Demonstratives as Anaphors in Dutch. *Proceedings of the Twenty-Third Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Pragmatics and Grammatical Structure*, 50–61. doi: 10.3765/bls.v23i1.1281

Cowles, H. W., Matthew, W., & Kluender, R. (2007). Linguistic and cognitive promi-

nence in anaphor resolution: topic, contrastive focus and pronouns. *Topoi*(26), 3–18. doi: 10.1007/s11245-006-9004-6

Cozijn, R., Commandeur, E., Vonk, W., & Noordman, L. G. (2011). The time course of the use of implicit causality information in the processing of pronouns: A visual world paradigm study. *Journal of Memory and Language*, *64*(4), 381–403.

Crawley, R. A., Stevenson, R. J., & Kleinman, D. (1990). The Use of Heuristic Strategies in the Interpretation of Pronouns. *Journal of Psycholinguistic Research*, *19*(4), 245–264.

Davies, C., Arnold, J. E., Davies, C., & Arnold, J. E. (2019). Reference and Informativeness. *The Oxford Handbook of Experimental Semantics and Pragmatics*, 473–493. doi: 10.1093/oxfordhb/9780198791768.013.6

de la Fuente, I. (2015). *Putting pronoun resolution in context: The role of syntax, semantics, and pragmatics in pronoun interpretation* (Doctoral dissertation, Université paris Diderot). Retrieved from `https://tel.archives-ouvertes.fr/tel-01535977/document`

Dinçtopal-Deniz, N. (2009). Anaphora in Turkish. *Linguistics in the Big Apple: CUNY/NYU Working Papers in Linguistics.*(1999).

Drummond, A. (2013). *Ibex Farm.* Retrieved 2020-09, from `http://spellout.net/ibexfarm/`

Enç, M. (1986). Topic Switching and Pronominal Subjects in Turkish. In D. Slobin & K. Zimmer (Eds.), *Studies in turkish linguistics* (pp. 195–208). Amsterdam: John Benjamins.

Erguvanlı-Taylan, E. (1986). Pronominal versus Zero Representation of Anaphora in Turkish.. doi: 10.1075/tsl.8.12tay

Featherstone, C. R., & Sturt, P. (2010). Because there was a cause for concern: An investigation into a word-specific prediction account of the implicit-causality effect. *Quarterly Journal of Experimental Psychology*, *63*(1), 3–15. doi: 10.1080/17470210903134344

Fedele, E. (2016). *Discourse level processing and pronoun interpretation* (Doctoral Dissertation). University of Southern California.

Féry, C., & Krifka, M. (2008). Information structure: Notional distinctions, ways of expression. In *Unity and diversity of languages* (pp. 123–135). Amsterdam: John Benjamins Publishing Company. doi: 10.1075/z.141.13kri

Filiaci, F. (2010). Null and overt subject biases in Spanish and Italian: A cross-linguistic comparison. In C. Borgonovo, M. Español-Echevarría, & P. Prévost (Eds.), *Selected Proceedings of the 12th Hispanic Linguistics Symposium* (pp. 171–182).

Foraker, S., & McElree, B. (2007). The role of prominence in pronoun resolution: Active versus passive representations. *Journal of Memory and Language*, *56*(3), 357–383. doi: 10.1016/j.jml.2006.07.004

Frederiksen, J. R. (1981). Understanding anaphora: Rules used by readers in assigning pronominal referents. *Discourse Processes*, *4*, 323–347.

Fukumura, K., & Van Gompel, R. P. G. (2010). Choosing anaphoric expressions: Do people take into account likelihood of reference? *Journal of Memory and Language*, *62*(1), 52–66.

Garvey, C., & Caramazza, A. (1974). Implicit Causality in Verbs. *Source: Linguistic*

Inquiry, *5(3)*, 459–464.

Gelormini-Lezama, C. (2018). Exploring the Repeated Name Penalty and the Overt Pronoun Penalty in Spanish. *Journal of Psycholinguistic Research*, *47*(2), 377–389. doi: 10.1007/s10936-017-9529-4

Gernsbacher, M. A., & Hargreaves, D. J. (1988). Accessing sentence participants: The advantage of first mention. *Journal of Memory and Language*, *27*(6), 699–717. doi: 10.1016/0749-596X(88)90016-2

Gernsbacher, M. A., Hargreaves, D. J., & Beeman, M. (1989). Building and accessing clausal representations: The advantage of first mention versus the advantage of clause recency. *Journal of Memory and Language*, *28*(6). doi: 10.1016/0749-596X(89)90006-5

Givón, T. (1983). Topic continuity in discourse. In T. Givón (Ed.), *Topic continuity in discourse: A quantitative, cross-language study.* Amsterdam: John Benjamins. doi: 10.1075/tsl.3.01giv

Givón, T. (1992). The grammar of referential coherence as mental processing instructions. *Linguistics*, *30*(1), 5–56. doi: 10.1515/ling.1992.30.1.5

Godoy, M. C., Weissheimer, J., & Mafra, M. A. (2018). When grammar meets pragmatics: Subject preference and coherence relations in Brazilian Portuguese pronoun interpretation. *Journal of Portuguese Linguistics*, *17*, 1–22. doi: 10.5334/jpl.197

Goikoetxea, E., Pascual, G., & Acha, J. (2008). Normative study of the implicit causality of 100 interpersonal verbs in Spanish. In *Behavior research methods.* doi: 10.3758/BRM.40.3.760

Göksel, A., & Kerslake, C. (2005). *Turkish: A Comprehensive Grammar.* London: Routledge.

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, *17*(3), 311–347. doi: 10.1207/s15516709cog1703_1

Gračanin-Yuksek, M., Lago, S., Şafak, D. F., Demir, O., & Kırkıcı, B. (2017). The Interaction of Contextual and Syntactic Information in the Processing of Turkish Anaphors. *Journal of Psycholinguistic Research*, *46*(6), 1397–1425. doi: 10.1007/s10936-017-9502-2

Grice, H. P. (1975). Logic and conversation. In *Syntax and semantics 3: Speech arts.*

Grober, E. H., Beardsley, W., & Caramazza, A. (1978). Parallel function strategy in pronoun assignment. *Cognition*, *6*(2), 117-133. doi: 10.1016/0010-0277(78)90018-5

Grosz, B., Weinstein, S., & Joshi, A. (1995). Centering: A Framework for Modeling the Local Coherence of Discourse. *Computational linguistics*, *21*, 203–225.

Gundel, J. (2010). Reference and accessibility from a givenness hierarchy perspective. *Materials Science and Technology*, *26*(11), 148–168.

Gundel, J., Hedberg, N., & Zacharski, R. (1993). Cognitive Status and the Form of Referring Expressions in Discourse. *Language*, *69*(2), 274–307. Retrieved from `http://www.jstor.org/stable/416535`

Gundel, J., & Mulkern, A. (1998). Quantity implicatures in reference understanding. *Pragmatics and Cognition*, *6*, 21–45. doi: 10.1075/pc.6.1-2.05gun

Hartshorne, J. K. (2013). What is implicit causality? *Language, Cognition and*

Neuroscience, *29*, 804–824. doi: 10.1080/01690965.2013.796396

Hartshorne, J. K., Sudo, Y., & Uruwashi, M. (2012). Are Implicit Causality Pronoun Resolution Biases Consistent Across Languages and Cultures? *Experimental psychology*, *60*, 1–18. doi: 10.1027/1618-3169/a000187

Hobbs, J. (1979). Coherence and coreference. *Cognitive Science*, *3*(1), 67–90.

Hobbs, J. (1990). *Literature and cognition.* Cambridge University Press.

Järvikivi, J., Van Gompel, R. P., Hyönä, J., & Bertram, R. (2005). Ambiguous pronoun resolution contrasting the first-mention and subject-preference accounts. *Psychological Science*, *16*(4), 260–264. doi: 10.1111/j.0956-7976.2005.01525.x

Kaiser, E. (2010). Investigating the Consequences of Focus on the Production and Comprehension of Referring Expressions. *International Review of Pragmatics*, *2*(2), 266–297.

Kaiser, E. (2011). On the Relation between Coherence Relations and Anaphoric Demonstratives in German. *Proceedings of the 2010 Annual Conference of the Gesellschaft für Semantik. Sinn und Bedeutung 15*, 337–351.

Kaiser, E., & Trueswell, J. (2008). Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes*, *23*, 709–748. doi: 10.1080/01690960701771220

Kameyama, M. (1995). *Indefeasible Semantics and Defeasible Pragmatics* (Vol. 356). Retrieved from `http://arxiv.org/abs/cmp-lg/9506016`

Keating, G. D., Jegerski, J., & VanPatten, B. (2016). Online processing of subject pronouns in monolingual and heritage bilingual speakers of Mexican Spanish. *Bilingualism: Language and Cognition*, *19*(1), 36–49.

Keating, G. D., Vanpatten, B., & Jegerski, J. (2011). Who was walking on the beach?: Anaphora resolution in Spanish heritage speakers and adult second language learners. *Studies in Second Language Acquisition*, *33*(2), 193–221. doi: 10.1017/S0272263110000732

Kehler, A. (2002). Coherence, reference, and the theory of grammar. In *CSLI lecture notes series.*

Kehler, A., Kertz, L., Rohde, H., & Elman, J. L. (2008). Coherence and coreference revisited. *Journal of semantics*, *25*(1), 1–44.

Kehler, A., & Rohde, H. (2013). A probabilistic reconciliation of coherence-driven and centering-driven theories of pronoun interpretation. *Theoretical Linguistics*, *39*(1-2), 1–37. doi: 10.1515/tl-2013-0001

Kehler, A., & Rohde, H. (2019). Prominence and coherence in a Bayesian theory of pronoun interpretation. *Journal of Pragmatics*, *154*, 63–78. doi: j.pragma.2018.04.006

Kerslake, C. (1987). Noun phrase deletion and pronominalisation in Turkish. In H. E. Boeschoten & L. T. Verhoeven (Eds.), *Studies on modern turkish* (pp. 91–104). Tilburg: Tilburg University Press.

Kieras, D. E. (1979). *The relation of topics and themes in naturally occurring technical paragraphs* (Tech. Rep.). Tucson: Arizona University Department of Psychology.

Kim, K., Theres, G., & Schafer, A. J. (2013). *Effects of event-structure and topic/focus-marking on pronoun reference in Korean* [Poster presented at the 26th Annual CUNY Conference on Human Sentence Processing]. Columbia, SC. Retrieved from `http://www2.hawaii.edu/~aschafer/CUNY2013`

`-KitaekKim-final-poster.pdf`

Konuk, G., & von Heusinger, K. (To appear). Discourse prominence in Turkish: The interaction of grammatical function and semantic role. In I. Khomchenkova, Y. Sinitsyna, & S. Tatevosov (Eds.), *Proceedings of the 15th Workshop on Altaic Formal Linguistics (WAFL 15)*. Cambridge, MA: MIT Working Papers in Linguistics.

Koornneef, A. W., & Van Berkum, J. J. (2006). On the use of verb-based implicit causality in sentence comprehension: Evidence from self-paced reading and eye tracking. *Journal of Memory and Language*, *54*(4), 445–465. doi: 10.1016/j.jml.2005.12.003

Kornfilt, J. (2008). Asymmetries between pre-verbal and post-verbal scrambling in turkish. In J. Sabel & M. Saito (Eds.), *The free word order phenomenon* (pp. 163–180). De Gruyter Mouton.

Kornfilt, J. (2018). NP versus DP: Which one fits Turkish nominal phrases better? *Turkic Languages*, *22*, 155–166.

Kwon, N., & Sturt, P. (2013). Null pronominal (pro) resolution in Korean, a discourse-oriented language. *Language and Cognitive Processes*, *28*(3), 377–387. doi: 10.1080/01690965.2011.645314

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

Mayol, L. (2018). Asymmetries between interpretation and production in Catalan pronouns. *Dialogue and Discourse*, *9*(2), 1–34. doi: 10.5087/dad.2018.201

Mayol, L., & Clark, R. (2010). Pronouns in Catalan: Games of partial information and the use of linguistic resources. *Journal of Pragmatics*, *42*(3), 781–799.

Özge, D., & Evcen, E. (2020). Referential form, word order and emotional valence in turkish pronoun resolution in physical contact events. *Discourse meaning: The view from Turkish*, 165–188.

Özge, D., Evcen, E., Kesici, A., & Köse, E. (2018). *Pronoun resolution in Turkish Transfer-of-Possession Verbs* [Paper presented at the Second International Conference "Prominence in Language"]. Universität zu Köln, 12 July 2018.

Özge, D., Hartshorne, J., & Snedeker, J. (2017). *Salience of experiencer versus stimulus in Turkish psych verbs: A pronoun resolution study* [Poster presented at 30th Annual CUNY Conference on Human Sentence Processing, 30 March–1 April, 2017]. Cambridge, MA. Retrieved from `https://d2dg4e62b1gc8m.cloudfront.net/pub_pdfs/CUNY2017_OzgeHartshorneSnedeker-V2.pdf`

Özge, U., Özge, D., & von Heusinger, K. (2016). Strong indefinites in Turkish, referential persistence, and salience structure. *Empirical Perspectives on Anaphora Resolution*, 1–18. doi: 10.1515/9783110464108-009

Öztürk, B. (2005). Case, Referentiality and Phrase Structure. In *Linguistik aktuell* (Vol. 77). Amsterdam: John Benjamins Publishing Company.

Pickering, M., & Majid, A. (2007). What are implicit causality and consequentiality? *Language and Cognitive Processes*. doi: 10.1080/01690960601119876

Portele, Y., & Bader, M. (2016). Accessibility and referential choice: Personal pronouns and d-pronouns in written German. *Discours. Revue de linguistique, psycholinguistique et informatique. A journal of linguistics, psycholinguistics and computational linguistics*(18).

Portele, Y., & Bader, M. (2020). *Semantic and structural bias in the interpretation of German personal and demonstrative pronouns.* Open Science Framework. doi: 10.17605/OSF.IO/FE5JV

Prince, E. (1981). Toward a taxonomy of given-new information. *Radical Pragmatics*, 223–255.

Pyykkönen-Klauck, P., & Järvikivi, J. (2010). Activation and Persistence of Implicit Causality Information in Spoken Language Comprehension. *Experimental psychology*, *57(1)*, 5–16. doi: 10.1027/1618-3169/a000002

R Core Team. (2017). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R -project.org/`

Reinhart, T. (1981). Pragmatics and linguistics: Analysis of sentence topics. *Philosophica*, *27*, 53–94.

Rohde, H. (2008). *Coherence-driven effects in sentence and discourse processing* (Unpublished doctoral dissertation). UC San Diego.

Rohde, H. (2019). Pronouns. In C. Cummins & N. Katsos (Eds.), *The Oxford Handbook of Experimental Semantics and Pragmatics* (pp. 452–474). doi: 10 .1093/oxfordhb/9780198791768.013.21

Rohde, H., & Kehler, A. (2014). Grammatical and information-structural influences on pronoun production. *Language, Cognition and Neuroscience*, *29*(8), 912–927. Retrieved from `http://dx.doi.org/10.1080/01690965.2013.854918`

Rohde, H., Kehler, A., & Elman, J. L. (2006). Event Structure and Discourse Coherence Biases in Pronoun Interpretation. In R. Sun (Ed.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society.* Mahwah, NJ: Lawrence Erlbaum Associates.

Rohde, H., Kehler, A., & Elman, J. L. (2007). Pronoun Interpretation as a Side Effect of Discourse Coherence. *Proceedings of the Annual Meeting of the Cognitive Science behaviour*, *29*(29), 276–281.

Rosa, E. C., & Arnold, J. E. (2017). Predictability affects production: Thematic roles can affect reference form selection. *Journal of Memory and Language*, *94*, 43–60. doi: 10.1016/j.jml.2016.07.007

Sauermann, A., & Gagarina, N. (2017). Grammatical Role Parallelism Influences Ambiguous Pronoun Resolution in German. *Frontiers in Psychology*, *8*, 1205. Retrieved from `https://www.frontiersin.org/article/10.3389/ fpsyg.2017.01205`

Schumacher, P. B., Dangl, M., & Uzun, E. (2016). Thematic role as prominence cue during pronoun resolution in German. In A. Holler & K. Suckow (Eds.), *Empirical perspectives on anaphora resolution* (pp. 121–147). Berlin: de Gruyter Mouton.

Sheldon, A. (1974). The role of parallel function in the acquisition of relative clauses in english. *Journal of verbal learning and verbal behavior*, *13*(3), 272–281.

Shoji, S., Dubinsky, S., & Almor, A. (2017). The Repeated Name Penalty, the Overt Pronoun Penalty, and Topic in Japanese. *Journal of Psycholinguistic Research*, *46*(1), 89–106. doi: 10.1007/s10936-016-9424-4

Silva-Corvalán, C. (1977). *A discourse study of word order in the Spanish spoken by Mexican-Americans in West Los Angeles* (Master's Thesis). University of

California, Los Angeles.

Smyth, R. (1994). Grammatical Determinants of Ambiguous Pronoun Resolution. *Journal of Psycholinguistic Research*, *23*(3).

Solstad, T., & Bott, O. (2017). Causality and Causal Reasoning in Natural Language. *Oxford Handbook of Causal Reasoning*.

Song, J., & Kaiser, E. (2020). Effects of discourse factors on the interpretation of Korean null pronouns in subject and object position. *Proceedings of the Linguistic Society of America*, *5*(1), 655. doi: 10.3765/plsa.v5i1.4751

Stevenson, R. J., Crawley, R. A., & Kleinman, D. (1994). *Thematic Roles, Focus and the Representation of Events* (Vol. 9) (No. 4). doi: 10.1080/01690969408402130

Stevenson, R. J., Knott, A., Oberlander, J., & McDonald, S. (2000). Interpreting pronouns and connectives: Interactions among focusing, thematic roles and coherence relations. *Language and Cognitive Processes*, *15*(3), 225–262. doi: 10.1080/016909600386048

Turan, Ü. D. (1996). *Null vs. overt subjects in Turkish discourse: A centering analysis* (Doctoral dissertation). IRCS Technical Reports Series.

Ueno, M., & Kehler, A. (2016). Grammatical and pragmatic factors in the interpretation of Japanese null and overt pronouns. *Linguistics*, *54*(6), 1165–1221. doi: 10.1515/ling-2016-0027

Van Berkum, J. J., Koornneef, A. W., Otten, M., & Nieuwland, M. S. (2007). Establishing reference in language comprehension: An electrophysiological perspective. *Brain Research*, *1146*, 158–171.

Verhoeven, E. (2014). Thematic prominence and animacy asymmetries. Evidence from a cross-linguistic production study. *Lingua*, *143*, 129–161. doi: 10.1016/j.lingua.2014.02.002

von Heusinger, K., & Schumacher, P. B. (2019). Discourse prominence: Definition and application. *Journal of Pragmatics*, *154*, 117–127. doi: 10.1016/j.pragma.2019.07.025

Walker, M., Iida, M., & Cote, S. (1996). Japanese discourse and the process of centering. *Computational Linguistics*, *20*, 193–232.

Winograd, T. (1972). Understanding natural language. *Cognitive Psychology*, *3*(1), 1–191. doi: 10.1016/0010-0285(72)90002-3

Woolf, V. (2000). *The Waves* (K. Carabine, Ed.). Wordsworth Editions Limited.

Appendices

Experiment 1

1. Öğretmen Ahmet'i üzdü çünkü... 'The teacher upset Ahmet because'

2. Ahmet öğretmene üzüldü çünkü... 'Ahmet was sorry for the teacher because'

3. Mehmet oyuncuyu şaşırttı çünkü... 'Mehmet surprised the player because ...'

4. Oyuncu Mehmet'e şaşırdı çünkü... 'The player was surprised by Mehmet because '

5. Şoför çocuğu kızdırdı çünkü... 'The driver annoyed the boy because ...'

6. Çocuk şoföre kızdı çünkü... 'The boy got mad at the driver because ...'

7. Ayşe kızı sevindirdi çünkü... 'Helen made the girl happy because'

8. Kız Ayşe'ye sevindi çünkü... 'The girl was happy for Ayşe because'

9. Adam Melih'i kandırdı çünkü... 'The man fooled Melih because'

10. Melih adama kandı çünkü... 'Melih got fooled by the man because'

11. Ressam Ece'yi korkuttu çünkü... 'The painter scared Ece because'

12. Ece ressamdan korktu çünkü... 'Ece was scared of the painter because'

13. Bekir balıkçıyı tiksindirdi çünkü... 'Bekir disgusted the fisherman because'

14. Balıkçı Bekir'den tiksindi çünkü... 'The fisherman disgusted by Bekir because'

15. Seyirci Berk'i utandırdı çünkü... 'The spectator embarrassed Berk because'

16. Berk seyirciden utandı çünkü... 'Berk was embarrassed by the spectator because'

17. Öykü aktörü etkiledi çünkü... 'Öykü impressed the actor/actress because'

18. Aktör Öykü'den etkilendi çünkü... 'The actor/actress was impressed by Öykü because'

19. Elif kadını büyüledi çünkü... 'Elif fascinated the woman because'

20. Kadın Elif'ten büyülendi çünkü... 'The woman was fascinated by Elif because'

Experiment 2

1. Ali Ahmet'e bir kitap verdi. 'Ali gave a book to Ahmet.'

2. Bekir Hasan'dan uhu aldı. 'Bekir got a glue from Hasan.'

3. Pınar Sevgi'ye bir şapka verdi. 'Pınar gave a hat to Sevgi.'

4. Lale Zeynep'ten bir bardak kahve aldı. 'Lale got a cup of coffee from Zeynep.'

5. Can Mehmet'e aracını ödünç verdi. 'Can lent his car to Mehmet.'

6. Onur Emir'den silgi ödünç aldı. 'Onur borrowed an eraser from Emir'

7. Simay Ipek'e şarj aletini ödünç verdi. 'Simay lent her charger to Ipek.'

8. İlke Beste'den bir kitap ödünç aldı. 'İlke borrowed a book from Beste.'

9. Esra Handan'a raporu teslim etti. 'Esra handed the report to Handan.'

10. Begüm Aleyna'dan mektubu teslim aldı. 'Begüm received the letter from Aleyna.'

11. Murat Hakan'a ödevi teslim etti. 'Murat delivered the assignment to Hakan.'

12. Kemal Engin'den hediye çekini teslim aldı. 'Kemal received the gift certificate from Engin.'

13. Neslihan İlayda'ya buzdolabını sattı. 'Neslihan sold her refrigerator to Ilayda.'

14. Mustafa Zeki'den bir kartpostal satın aldı. 'Mustafa bought a postcard from Zeki.'

15. Ümit Nejat'a telefonunu sattı. 'Ümit sold his phone to Nejat.'

16. Alara Berfin'den kulaklık satın aldı. 'Alara bought headphones from Berfin.'

17. Emir Kıvanç'a görevi devretti. 'Emir handed over the task to Kıvanç.'

18. Ömer Levent'ten anahtarları devraldı. 'Ömer took over the keys from Levent.'

19. Ecem Gülnur'a dükkanı devretti. 'Ecem handed over the shop to Gülnur.'

20. Ayşe Melike'den şirketi devraldı. 'Ayşe took over the company from Melike.'

21. Caner Kerem'e bir parça sakız uzattı. 'Caner handed Kerem a piece of gum.'

22. Gözde Derya'dan şekeri kaptı. 'Gözde snatched the candy from Derya.'

23. Nehir Aylin'e tuzu uzattı. 'Nehir passed the salt to Aylin.'

24. Kaan Emre'den oyuncağı kaptı. 'Kaan snatched the toy from Emre.'

25. Koray Canberk'e bir atkı verdi. 'Koray gave Canberk a scarf.'

26. Ayşe Melike'den iki yumurta aldı. 'Ayşe got two eggs from Melike.'

27. Yağmur Rüya'ya çakmak uzattı. 'Yağmur passed a lighter to Rüya.'

28. Doğa İlkay'dan topu kaptı. 'Doğa got the ball from Ilkay.'

29. Elif Bengisu'ya malzemeleri teslim etti. 'Elif delivered the materials to Bengisu.'

30. Kaya Can'dan aracı teslim aldı. 'Kaya has received the car from Can.'

Ehrenwörtliche Erklärung

Ich versichere hiermit, dass ich die vorliegende Arbeit selbständig und ohne Benutzung anderer als der angegebenen Quellen und Hilfsmittel verfasst habe. Wörtlich übernommene Sätze oder Satzteile sind als Zitat belegt, andere Anlehnungen, hinsichtlich Aussage und Umfang, unter Quellenangabe kenntlich gemacht. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen und ist nicht veröffentlicht. Sie wurde nicht, auch nicht auszugsweise, für eine andere Prüfungs- oder Studienleistung verwendet. Zudem versichere ich, dass die von mir abgegebenen schriftlichen (gebundenen) Versionen der vorliegenden Arbeit mit der abgegebenen elektronischen Version auf einem Datenträger inhaltlich übereinstimmen.

Statutory Declaration

I herewith declare that I have composed the present thesis myself and without use of any other than the cited sources and aids. Sentences or parts of sentences quoted literally are marked as such; other references with regard to the statement and scope are indicated by full details of the publications concerned. The thesis in the same or similar form has not been submitted to any examination body and has not been published. This thesis was not yet, even in part, used in another examination or as a course performance. Furthermore, I declare that the submitted written (bound) copies of the present thesis and the version submitted on a data carrier are consistent with each other in contents.

Place, Date: Signature: