

Semantics and Processing of Weak and Strong Definites in
Colloquial Persian: Evidence from an Offline
Questionnaire

by
Pegah Nikravan

A thesis submitted to the Faculty of Graduate and Postdoctoral
Affairs in partial fulfillment of the requirements for the degree of

Master of Cognitive Science

in

Cognitive Science

Carleton University
Ottawa, Ontario

© 2014, Pegah Nikravan

Abstract

The main goal of the present study was to investigate the semantics and processing of three (in)definite markers used in the colloquial Persian. It was proposed that colloquial Persian morphologically realizes two definite markers, the null marker ‘ \emptyset ’ and the suffix ‘-e’. It was further proposed that these correspond to so-called “weak” and “strong” definites, respectively; the presuppositions of strong definites need to be satisfied by an explicit antecedent but the presuppositions of weak definites do not (Schwarz, 2009, 2013). It was also proposed that ‘ye...-i’ is an indefinite marker. This proposal is supported by introspective judgments, as well as by quantitative data from an off-line questionnaire study (building on Burkhardt, 2006; Hirotani & Schumacher, 2011) that measured the naturalness of sentences that only varied in NPs marked with ‘ \emptyset ’, ‘-e,’ ‘ye...-i’ in contexts which (i) an antecedent was available (Given contexts), (ii) an antecedent was unavailable but could be accommodated (Bridged contexts), and (iii) an antecedent was unavailable and accommodation was unlikely (New). As expected, there was an interaction between CONTEXT and MARKER and in particular the most natural text was the one in which ‘-e’ was used in a Given context (this was the only condition that satisfied all relevant constraints).

Acknowledgements

I would like to first thank my co-supervisors Dr. Masako Hirotsu and Dr. Raj Singh for all of their support, guidance and encouragements from the initial to the final stages of working on my thesis project and writing the thesis. I'm indebted to Dr. Raj Singh who has always patiently put up with me and supported me in the number of ways. I'm indebted to him for the background I've built in cognitive linguistics, semantics, and specifically in (in)definites. Also, his generous financial support during my master's has greatly helped me focus on my research. He also generously provided the financial sponsorship for my experimental study. I also would like to express my deepest gratitude to Dr. Masako Hirotsu for giving me the opportunity of working in a high level experimental environment, for providing me with the Lab facilities, for all of her generous and invaluable trainings, advice, feedback, and tips in doing research. I also would like to thank my committee members Dr. Erik Anonby and Dr. Kumiko Murasugi for their support and help and their invaluable advice and feedback. I also would like to thank Dr. Ana Arregui for drawing my attention to the weak/strong realization of definites cross-linguistically, and for her generous and insightful guidance. Thanks also go to Maryam Hosseini for the inspiring discussions regarding Persian definites. I'm indebted to many people who helped and supported me: Dr. Jo-Anne LeFevre, Dr. Robert West, Dr. Ida Toivonen, Adam, Amir, Azade, Aysegul, Babak, Chris, Ehsan, Gita, Krista, Liz, Maryam, Mona, Olivia, Roxana, Sara, Tabish, and all members of LLI Lab and LBL Lab at Carleton University. Last but not least, I would like to thank my parents, Moahmmad-Javad Nikravan and Zahra Rasuli-zade, my beloved husband Mehrdad Kalantar, and other members of my family for their unconditional love and support.

Table of Contents

Abstract	2
Acknowledgements	3
List of Tables	6
List of Illustrations	7
1 Chapter: Introduction	8
2 Chapter: Semantics and Processing of (In)definites	11
2.1 Familiarity theory of definites.....	11
2.2 Background: the processing of (in)definites	14
2.3 (In)definites in Persian	17
2.3.1 Background	17
2.3.2 Definite markers: ‘-e’ and ‘Ø’	19
2.3.3 Indefinite markers: ‘ye’ and ‘-i’	22
2.3.4 Strong and weak realization in Persian definites.....	23
2.4 Strong and Weak: background assumptions/predictions.....	25
2.5 Experimental hypothesis and predictions.....	28
3 Chapter: Offline Naturalness Sentence-Pair Experiment	31
3.1 Method: materials and procedures	31
3.1.1 Experimental stimuli	31
3.1.2 Participants.....	35
3.1.3 Procedures	36
3.2 Results	37
3.3 Discussion	40
4 Chapter: General Discussion and Conclusion	43
4.1 General discussion	43

4.2 Conclusion.....	43
Appendix A List of Abbreviations Used for Glossing the Examples.....	45
Appendix B Relatedness Noun-Pair Questionnaire.....	46
B.1 Instruction and sample stimuli for the relatedness noun-pair questionnaire	48
Appendix C Instruction and Sample Stimuli for the Naturalness Sentence-Pair Questionnaire.....	50
Appendix D Language Background Questionnaire.....	52
References	54

List of Tables

Table 1 Predictions Based on Appropriateness Conditions.....	27
Table 2 Experimental Predictions.....	29
Table 3 Results of Offline Questionnaire.....	36

List of Illustrations

Figure 1 Mean Ratings in Nine Conditions.....	39
--	-----------

1 Chapter: Introduction

In every-day conversations discourse participants construct a mental model for keeping track of all of the incoming and old information. (e.g., Chafe, 1976; Haviland and Clark, 1974). If the hearer trusts the speaker, he/she accepts the ‘*new*’ information *asserted* by the speaker and this information will become *common ground* (Grice, 1975; Lewis, 1979). The common ground (or *context*) can be thought as the set of propositions that is *presupposed* by the speaker for the purpose of the conversation (Stalnaker, 1978). It is not cooperative of the speaker to utter a sentence that requires the satisfaction of a presupposition that is not shared by the hearer (Stalnaker, 1978). One important component of maintaining ‘*discourse coherence*’ involves linking the information that the hearer is not expected to already know (*new*) to the information he/she is expected to know (*given*) (Haviland and Clark, 1974).

One way of distinguishing *given* from *new* information in languages is definite marking. For example in English the articles ‘*a*’ and ‘*the*’ are two such expressions: a sentence like ‘*a teacher_i is very nice*’ introduces a new teacher into the context and asserts that this teacher is very nice, and a sentence like ‘*the teacher_i is very nice*’ signals that the existence of a teacher has already been established in the context, and it expresses that this teacher is very nice (e.g., Haviland and Clark, 1974, Heim, 1982).

English definite noun phrases (NPs) like the *teacher_i* have been argued to behave like anaphors: ‘*the teacher_i*’ is an *anaphoric element* that co-refers with an already established discourse referent. For example, in a text like ‘*a teacher walked in and I couldn’t believe what the teacher was wearing*’, the definite ‘*the teacher*’ must co-refer with the indefinite ‘*a teacher*’: in formal syntax, they must be co-indexed. However, the

similarities between definites and anaphors break down in many cases, and in particular when the definite NP does not relate to an *'explicit antecedent'* in discourse. In Heim's (1982) influential treatment, definites are characterized as both anaphoric and presuppositional. Non-anaphoric uses of definites are accounted for with a mechanism of *'accommodation'* (unavailable to pronouns; for discussion see e.g., Heim, 1982; van der Sandt, 1992; Beaver & Zeevat, 2007).

However, many other languages morphologically realize at least two kinds of definites: those that are anaphoric, and those that are not (e.g., German, among others; cf. Schwarz, 2009, 2013). Schwarz (2009, 2013) calls anaphoric definites *'strong'* and non-anaphoric definites *'weak'*.

We present evidence that Persian also realizes a *strong/weak* distinction: there is a null definite marker, call it \emptyset (a morpheme that receives no phonology), and there is an overt definite marker, suffix *'-e'*¹. An offline naturalness rating experiment was conducted in Persian to test this proposal. Specifically, we extrapolate predictions about the appropriate use of strong/weak definites and indefinites in different kinds of context based on prior theoretical assumptions about the nature of anaphora, presuppositions, and discourse reference; we expect that these predictions should be cross-linguistically applicable to any language that realizes a strong/weak distinction. If Persian is such a language, then we expect the naturalness ratings to have certain identifiable patterns; as we will see, the expected patterns are attested, although the data also point to new puzzles that we hope to investigate in future work. The study contributes to the semantics of Persian (in)definite markers by formalizing these markers in terms of the strong/weak

¹ In object position these definite noun phrases would also be marked with suffix *-ra*.

distinction. In so doing, the study contributes to our understanding of the cross-linguistic typology of (in)definite markers, and it extends previous psycholinguistic studies of the definite/indefinite contrast to a language (Persian) that realizes a strong/weak distinction in its inventory of definite markers.

2 Chapter: Semantics and Processing of (In)definites

2.1 Familiarity theory of definites

Common intuition concerning the interpretation of definites is that they are anaphoric elements:

- (1) John saw a waiter_{*i*}.
- a. # A waiter_{*i*} was very nice.
 - b. The waiter_{*i*} was very nice.

*the waiter_{*i*}* (1b) is anaphorically related to the indefinite *a waiter_{*i*}* introduced earlier. This is the basis of the classical idea that definites refer to ‘*discourse referents*’ that are already *familiar* (e.g., Christopherson, 1939; Karttunen, 1976; see Heim, 1982 for a modern reformulation).

For the purposes of this work, when I use the term anaphora, or ‘anaphoric relations’, I follow common practice in modern semantics and mean it in the strict sense of co-indexation between NPs: an explicit antecedent X_i , and a subsequent anaphoric element Y that gets indexed with i , Y_i , so that X and Y refer to the same entity (i.e., X and Y co-refer; we ignore bound-variable anaphora here). The later term Y , like *the waiter_{*i*}* (1b), is further assumed to be a ‘variable’, namely, its meaning is assignment-dependent, and in particular its semantic value is assigned by the context, here from the expression X , *a waiter_{*i*}*.

Note that indefinites *cannot* be used anaphorically (1a); this is the basis of the classical idea that indefinites introduce *new* discourse referents. (e.g., Christopherson, 1939; see Heim, 1982 for a modern reformulation). Definite NPs like *the waiter* refer to

already *given* discourse referents that are known to be waiters, and indefinite *NPs* like *a waiter* introduce *new* discourse referents.

Definites are thus very much like pronouns (see also van der Sandt; 1992; Geurts, 1999):

- (2) John saw a waiter_i.
- a. He_i was very nice.

However, an apparent problem for this analysis is the well-known observation that, unlike pronouns, definites do not always have an explicit antecedent in discourse:

- (3) John_k went to a bar_i last night.
- a. # He_j argued with him for a while.
- b. A bouncer_j argued with him for a while.
- c. The bouncer_j argued with him for a while.

(Singh et al., 2014)

In (3) there is no discourse referent for *he_j* or *the bouncer_j* to pick up (note that the indexing is such that *he_j* is intended to not co-refer with *John_k*). These kinds of examples seem problematic for the familiarity theory of definites: similar to the indefinites like *a bouncer*, definites like *the bouncer* also can be used for introducing a *new* referent in discourse. In order to deal with such cases, proponents of the familiarity theory follow Lewis, (1979), and suggest that hearers can accommodate the antecedent that is demanded by the definite NP. For example, in (3), upon hearing *the bouncer_j*, the listener simply amends the context by assuming that there was a (contextually distinguished)

bouncer at the bar, which provides the required antecedent for *the bouncer_i*. Thus, a familiarity theory of definites together with an *accommodation strategy* is one prominent line of analysis for English (Lewis, 1979; Heim, 1982).²

However, the picture is more complicated when we look at a typologically broader range of languages. For example, Schwarz (2009, 2013) suggests that in many languages there are at least *two* definite morphemes: one that typically serves the anaphoric function as in (1) and one that is typically used non-anaphorically.

According to Schwarz (2009), Robert (2003) first distinguished between two kinds of familiarity as *weak* and *strong* familiarity: *weak* familiarity is when a referent becomes *accessible* to the discourse participants via “contextual existence entailment” – where the existence can be satisfied by an accommodation inference -- and *strong* familiarity is when the definite is anaphoric to a preceding linguistic expression (Schwarz, 2009).

The variations related to the (in)definite NPs in various languages raise an important question that whether the underlying mental processing involved in the comprehension of (in)definite NPs are similar/different cross-linguistically. In the next section, I will review some of the experimental findings on the (in)definiteness processing in English and other languages, and then will get back to the weak/strong realization in definites, by proposing that this distinction is morphologically realized in the colloquial Persian as well.

² Pronouns do not readily admit accommodation (for discussion, see e.g., Beaver and Zeevat, 2007; Heim, 1982; van der Sandt, 1992; Geurts, 1999)

2.2 Background: the processing of (in)definites

Underlying processing of definite and indefinite NPs have been investigated using event-related potential methodology (ERP)-- a highly time sensitive neurological method-- mostly by discussing two ERP components that are shown to be related to the *contextual integration* in sentence processing, known as N400 and P600. N400, a negativity with peak latency near 400 ms after the onset of the target NP, was found in response to the semantic abnormalities and contextual incoherence appearing in a sentence (e.g., Kutas & Hillyard, 1980). It has been shown in numerous ERP studies that the amplitude of the N400 is reduced as a function of semantically related words and repetition priming (e.g., Kutas and Federmier, 2000; Holcomb 1988, 1993; Rugg, 1985, and Ruz et al., 2003). P600, positive deflection near 600 ms after the onset of the target NP, has been characterized as an index for the increased processing demands caused by syntactic abnormalities or possibly other integration complexities. (e.g., Burkhardt, 2007)

Schumacher 2009 investigated the processing of the definite and indefinite NPs, using a set of Context-Target sentence pairs in German, designed as follows:

(4) a. CONTEXT:

Given: *Peter has recently visited **a speaker** in Munich.*

Bridged: *Peter has recently visited **a lecture** in Munich.*

New: *Peter has recently met H*

b. TARGET: (following each of the contexts)

*'He said that {**the/a speaker**} had been very nice.'*

annah in Munich.

As shown in (4) above, context sentences manipulated the degree of salience of the a definite or indefinite NP in the target sentence: the (in)definite NP either had an explicit antecedent in the context (Given context), or was related to an implicit antecedent (Bridged), or had no antecedent in the context at all (New context-- another study was previously done by Burkhardt (2006), with a similar design only the targets did not include indefinite (NPs). Both studies (Burkhardt, 2006; Schumacher, 2009) showed that N400 amplitude was decreased as an effect of *contextual accessibility* of an antecedent-- for the definite NPs in Burkhardt, 2006, and for both definite and indefinite NPs in Schumacher, 2009: the NPs with New context registered the most effect, followed by the Bridged and then Given (N400: New > Bridged > Given); P600 was also observed for the definite NPs with the Bridged and New contexts, but not with the Given contexts, and for the indefinites in all three types of contexts. This finding supported the previous literature on *new-given* strategies (e.g., Chafe 1976; Clark and Haviland, 1974; Haviland and Clark, 1974) that integration of the *new* information is more costly than *given* information. Note that, following Clark and Haviland, 1974, *given* information is the information that the listener is expected to already know (like definite NPs with the Given context as in (4a)); whereas, *new* information is not expected to be already known by the listener. The information for which there is no direct antecedent in discourse is considered as *new* information (Haviland and Clark, 1974)--like the definite NPs with Bridged or New contexts in (4).

The findings illustrated above, supported previous theoretical views on (in)definites, suggesting that *anaphoric* definite NPs (Given) like *the waiter* in (1b), *presuppose* the *existence* of an *explicit* antecedent in discourse, whereas the indefinite

NPs cannot be used anaphorically, (Christopherson, 1939; Heim, 1982), rather they introduce and establish a *new* referent in discourse (Heim, 1982; Karttunen, 1976). Additionally, their findings (Burkhardt, 2006; Schumacher, 2009; Hirotani and Schumacher, 2011) on definites with an implicit antecedent in discourse (Bridged), like *the bouncer* in (3c), suggested that definite NPs with Bridged contexts, “share properties with both Given and New” (specifically by finding the reduced N400 in Given and Bridged, but not in the New, and finding P600 in Bridged and New, but not in Given).

Using the same paradigm of design, Hirotani and Schumacher, (2011), investigated the comprehension of *given-new* information by Japanese speakers. Japanese, which does not have an overt definite marker, is a ‘topic-prominent’ language; and informational *accessibility* and salience are closely related to *topicality*. In Japanese, *wa*-- a topic marker -- marks a NP as a salient entity or the center of attention which is considered as *given* information. Although Japanese lacks an overt definite marker, but, interestingly, the neurological pattern found was consistent with the previous findings in Gemran (Burkhardt, 2006; Schumacher, 2009), providing cross-linguistical evidence supporting the distinction of two types of processing: when there is an explicit antecedent in the discourse (*given* information) versus where there is not (*new information*), corresponding to the searching memory and *linking/dependence relation* versus searching memory and *bridging/independence*. The cost in the bridging relations can be accounted by “establishing new discourse referents via accommodation” (Heim, 1982). However, the distinctions between “establishing a new referent via accommodation”, (i.e., in definite NPs with the Bridged contexts), definite NPs with the Bridged contexts, compared to the “establishing new discourse referents” in indefinite NPs was not clear.

More specifically, it is not clear whether the processing cost in the Bridging contexts, is a *repair-strategy* cost or is because of introducing a new referent. (For discussion see Singh, et al., 2014).

In the present study, the same paradigm of design as Schumacher and Hirotsu (2011), (CONTEXT: Given/Bridged/New vs. MARKER: Definite/Indefinite), was used in an offline questionnaire to investigate the (in)definite marking in Persian, using an offline questionnaire. However, in addition to definites versus indefinites, we investigated the distinction between strong versus weak definites as well.

In the next section, I will first discuss the (in)definite markers in the colloquial Persian, then I will discuss the weak/strong realization in Persian definites, and finally I will review the experimental hypotheses and predictions, which are built on the theoretical and experimental background reviewed in this chapter.

2.3 (In)definites in Persian

2.3.1 Background

The language under the discussion in the present study is standard Modern Persian, a null subject language with SOV word order (e.g., Ghomeshi, 2003, 2006, 2013) which is a subcategory of the Indo-European languages (Windfuhr, 1979), and the official language of Iran. Persian has two distinct registers of written and spoken (also known as formal and informal) forms. (e.g., Ghomeshi, 2003, 2006; Hamadani, 2011). Written Persian is the official language used in the written texts (e.g., books, formal letters, newspapers, etc.); and the spoken Persian is the colloquial language used in the daily conversations by the Persian speakers. Since we are concerned with understanding

of the pragmatics and semantics of the (in)definites normally used by Persian speakers and the processing of these structures in discourse comprehension, the primary object of the study is the colloquial Persian. The colloquial Persian varies in different areas of Iran. The dialect under the discussion in the present thesis is the register spoken in and around Tehran--known as “standard Persian” (Ghomeshi, 2003, 2006). The Persian examples and the judgments provided related to the (in)definiteness marking in this chapter, [those that do not cite the previous work], are based the descriptive data gathered from my interactions with native Persian speakers.

It is commonly assumed that there is no overt definite marker in Persian (e.g., Ghomeshi, 1997, 2003; Karimi, 1999). However, although, Persian in general lacks a definite article comparable to the English *the* (Karimi, 1999), but there are multiple linguistic structures in Persian that arguably signal definiteness, (e.g., object marker *rā*³, possessive marker *-eš*, bare nouns, and the suffix *-e* (specific to the colloquial Persian)). Although (in)definite marking in the written Persian has been extensively studied by Persian scholars, there is a lack of knowledge on those structures that only exist in the colloquial Persian in both (Sahraei, 2011). In the next section, I will discuss some of the complexities regarding the (in)definite marking in Persian briefly, by focusing on the most common (in)definite NPs used in the colloquial Persian in the subject positions.⁴

³ Persian object marker *rā*, which has been extensively studied by Persian scholars, has been argued to encode definiteness. (e.g., see Browne, 1970; Geranpaye, 2000; Ghomeshi, 1997; Hosseini, 2013; Karimi, 1989, 1990; Sadeghi, 1970).

⁴ In the present study, the topic of particular interest with regard to definiteness marking is the suffix *-e*, and the *null* marker. So I will focus only on the subject *NPs* to avoid the complications of the ‘object marker’ *-ra*.

2.3.2 Definite markers: ‘-e’ and ‘Ø’

Subject NPs in Persian (both spoken and written Persian) may appear without any article:

- (5) dær bāz=šod⁵
door open=became.
‘The door opened.’
(Mahootian, 1997)

A bare noun⁶ as in (5) signals that the noun phrase is definite (e.g., Ghomeshi, 1997, 2003). In addition to the null-marked nouns, there is also a stressed suffix, *-e*, that appears on nouns and, when a noun is marked with *-e*, it is interpreted as definite.

(Ghomeshi, 2003; Ghomeshi, 2006; Samiiian, 1983):

- (6) a. doxtær-e umæd.
girl-DEF came.
‘The girl came.’
(Ghomeshi, 2003)

Null-marking, as in (5), is commonly assumed as a way of signalling definiteness in Persian (e.g., Ghomeshi, 1997, 2003; Geranpaye, 2000; Jabari and Rezaei, 2010; Mahootian, 1997); and in the recent literature *-e* is commonly assumed as an optional

⁵See Appendix A for the list of abbreviations used for glossing the Persian examples, following Roberts (2003).

⁶I will treat definite bare nouns as null-marked nouns; and will assume that null-marking corresponds to insertion of a null morpheme (Ø).

colloquial definite marker (e.g., Samiian, 1983; Ghomeshi, 2003, 2006); however, there are controversial views on the semantics and function of *-e*. (e.g., Afzali, 2012). The suffix *-e* has not been studied that much and this can be because this phenomenon is not part of the written Persian (Milanian, 1971). According to Windfuhr, (1979), Andreas (unpublished notes, 1875), who was a German Iranologist, was the first person to recognize *-e* as a definite marker in the colloquial Persian, and suggested that it functions as a “demonstrative marker”. Hedberg et al., (2009), also suggest that Persian suffix *-e* has a demonstrative function, by arguing that in contrast to the English *the*, *-e* does not refer to the entities that are not *familiar* in the discourse--namely it cannot be used *non-anaphorically* in the sense we use in this thesis. However, based on observing frequent cases of marking Bridged definites, with *-e*, I argue that, similar to *the*, *-e*-marked NPs can be also used for introducing a *new referent* in discourse--for example in the cases like *a bar_i /the bouncer_j* in (3)--where no explicit antecedent is given. I will get back to this later in the chapter.

Following, Samiian’s influential work in (1983), Ghomeshi (2003) suggests that *-e* has number of distinctive properties similar to the English *the*. For example, it cannot attach to anything that is already of category D (e.g., proper names, pronouns, and NPs containing possessors):⁷

(8) hæ sæn-æ ro did-æm.

⁷ Ghomeshi (2003) suggests that *-e* takes as an input something of category [N⁰] and turns it into something of category D⁰ --similar to a pronoun or proper name:

[N⁰ + e_{Def}] --> D⁰

She also suggests that the major difference between the *-e* and the English definite marker *the*, [from a syntactic point of view], is that *-e* can only be attached to unmodified common nouns. (see Samiian 1983 for more discussions about *-e*).

#Hasan-DEF OM saw-1SG

'I saw the Hasan'.

(Ghomeshi, 2003)

However, I have observed many cases in the conversational uses by the Persian speakers, in which this rule is violated. Xorasani, (1950), also suggests that *-e* can attach to the proper nouns. *-e* has been traditionally interpreted as “the noun mentioned earlier” (Afzali, 2012; Ghomeshi, 2003; Hedberg, et al., 2009). Ghomeshi, (2003), suggests that the meaning of *-e*-marked nouns are discourse-related: “anaphoric in the sense of pointing back to something already discussed” (Ghomeshi, 2003). However, there are many complexities related to the *-e* that need to be accounted for; such as being used in the contexts without an explicit antecedent (Bridged cases), as mentioned above.⁸

Further, assuming that there are at least two definite markers in the spoken Persian (*-e* and \emptyset), this question still remains unanswered that what are the semantics/pragmatics conditions under which Persian speakers decide to choose either of these markers to convey definiteness. I will use the weak/strong framework mentioned earlier in the chapter, in order to analyze the behavior of *-e* and \emptyset . I will propose that *-e* and \emptyset systematically map onto two strong (call it anaphoric) and weak (call it presuppositional) uses of definite NPs, and both are systematically distinguished from the indefinite markers that cannot be used neither anaphorically, nor presuppositionally. In the next section, I will briefly discuss indefiniteness marking in Persian, then will get back to the *weak/strong* realization in Persian definite NPs.

⁸ Also: its interaction with *Ezafe* construction, *-e*, demonstratives, *in* (this) and *un* (that), and the numeral *ye*.

2.3.3 Indefinite markers: ‘*ye*’ and ‘*-i*’

There are three different ways of encoding indefiniteness in Persian (Ghameshi, 2006):

(9)

a. māšin-i	b. <i>ye(k)</i> māšin	c. <i>ye(k)</i> māšin-i
car-IND	one car	one car-IND
‘a car’	‘a car’	‘a car’

The enclitic *-i*, as in (9a), is usually considered as the main indefinite marker in the formal (written) Persian.⁹ In the colloquial Persian, however, the numeral *ye(k)*¹⁰, (meaning one), as in (9b), is used to mark indefiniteness. *ye* can occur with *-i* or replace it entirely. The combination of *ye* and *-i* together, as in (9c), also (10) below, is assumed to be the most common way of marking indefiniteness in the colloquial Persian (Ghameshi, 2003; 2006).

(10) ye ketāb-i nevešt.
one book-IND wrote.
‘He/She wrote a book.’

(Ghameshi, 2006)

⁹ *-i* can be used in the colloquial Persian as well; however in colloquial, it indicates that there will be an upcoming relative clause in the sentence, without which the meaning of the sentence will be incomplete (Ghameshi, 2003, 2006).

¹⁰ In the colloquial Persian, the last consonant of *yek* is dropped and it is pronounced *ye* (Ghameshi, 2003, 2006).

However, similar to Persian definite NPs, there seem not to be a unified theory accounting for all different ways of indefiniteness marking in Persian.¹¹

2.3.4 Strong and weak realization in Persian definites

In some cases, like (11) below, only the *-e*-marked definite NP (*man-e* in (11b)) is felicitous:

(11) a. *bā ye z'aen_i o mærd-i_j tū mæšhæd āšnā šod-æm.*

With one woman_i and man-IND_j in Mashhad familiar became-1SG

b. {#ye mærd-i_i/ #mærd∅_i/ mærd-e} xeyli bæd=æxlāq būd.

{#ye man-i_i/ #man∅_i/ man-e} very bad=manner was

'I met a woman_i and [a] man_j in Mashhad.'¹² {#A /#The_{WEAK}/The_{STRONG}} man was very mean.'

Given that an explicit antecedent is available in the discourse (*a teacher* in (11) below), this can be considered as a realization of strong definites.¹³

¹¹ (e.g., the semantic of *-i* has been discussed under 'topicality', 'referentiality' and 'specificity' (see Ghomeshi, 2003; Karimi, 1999); also *ye* has been identified as a 'cardinal numeral' (Hamedani, 2009); a 'prenominal counterpart to *-i* (Ghomeshi (2003); and as encoding cardinality in both *ye* and *-i*, Lyons (1999).

¹² Persian is a pro-drop language (e.g., Ghomeshi, 2003, 2006), it is usually preferable to drop the subject noun when it is mentioned in the previous sentence, rather than repeating it. In (11b), two nouns are used in the context sentence (*a woman* and [*a*] *man*), so it sounds natural to repeat one of the nouns in the subject of the next sentence. Also, the judgment regarding the preference of using *-e* over \emptyset would be the same, if it was {*a woman*} and {*a man*}, instead of a {*woman* and [*a*] *man*} as it is now.

¹³ The null-marked and *ye...i* marked NPs are both worse than the *-e* marked NPs, but we make no claims here about the relative badness of these inappropriately used NPs. We hope to provide quantitative measures of appropriateness in the experiment; we return to this in the discussion.

However, *the sun* in (12), which can be considered as an example of *non-anaphoric* cases (not having an explicit antecedent) cannot be marked by *-e*:

- (12) a. {#ye xoršid-i/ #xoršid-e/ xoršidØ_i} dāre mi-deraxš-e.
 b. {#A sun-i/ #sun-e/ sunØ_i} has DUR-shine.PER.3SG
 ‘{#A / #The_{STRONG} / The_{WEAK}} sun_i is shining.’

However, the appropriate uses of *-e* and the *null* marker, *Ø*, cannot be reduced to the anaphoric versus non-anaphoric distinction. For example, there are cases of *non-anaphoric* uses of definites that can naturally be marked with both *-e* and *null*:

- (13) a. dirūz ræft-æm ye restūrān-i_i qæzā xord-æm.
 Yesterday went-1SG a restaurant-IND_i food ate-1SG
 b. {(?)#ye āšpæz-i_j/ āšpæzØ_j/ āšpæz-e_j} xeyli xoš=æxlāq būd.
 {(?)#a cook_j/ cookØ_j/ cook-e_j} very good=manner was
 ‘I went to a restaurant yesterday. {(?)#a cook/ the_{STRONG} / the_{WEAK} cook} was very

In (13) the context *I went to a restaurant yesterday* does not establish a discourse referent corresponding to a cook, nor does it entail the existence of a cook. However, both *-e* and *Ø* can be naturally used. As illustrated in the previous sections, contexts like (13) have been discussed in discussions of English *the* and are commonly referred to as *bridging* in the psycholinguistic literature (Burkhardt, 2006; Heim, 1982; Singh et al., 2014; Scumacher and Hirotni, 2011). Unlike Given contexts, which establish an explicit antecedent in the preceding sentence, Bridged contexts merely suggest *an implicit*

antecedent to which subsequent definites may refer (so-called ‘indirect referent with association’: Clark, 1975). Thus, the appropriateness of both the strong and weak definite in (13) would follow from the assumption that bridging inferences can be made to relate definites to implicit antecedents in the prior discourse. For current purposes, we follow the literature in assuming that bridging or accommodation is appropriate when the context is plausible enough to support an implicit antecedent (see Singh et al., 2014 for references and discussion). Following Heim (1982), if we consider Bridged as a special case of a context in which an explicit antecedent is not provided, then we need to account for such cases as (10) *the sun*, that cannot be naturally used as marked by the strong marker in Persian.

2.4 Strong and Weak: background assumptions/predictions

We assume that the choice between SDef and WDef is governed by appropriateness condition and violations of such conditions will lead to judgments of unnaturalness, and also greater processing difficulty. In the following we elaborate our assumption regarding the appropriateness condition for both strong and weak definite markers, and also how this conditions are/are not satisfied in either of the three discussed contexts (Given, Bridged, New) and then, will discuss how this analysis can predict the behavior of *-e* and *null* in Persian.

One way for interpreting the meaning of the definite NPs, following Frege, (1892), is that they are not anaphoric at all. Rather, *the waiter* (1) for example presupposes that in the context there is a distinguished waiter, *x*, and *the waiter* refers to *x* (see e.g., Heim and Kratzer, 1998).

If we assume that *weak* definite (WDef) is a ‘Fregean definite’, and hence, it has an existence and a uniqueness presupposition, namely this presupposition is already presupposed, not asserted as part of the content of definite NPs, then we seem to be able to account for the weak definites in Persian. Strong definite NP (SDef) can be assumed as having the existence and uniqueness presupposition of WDef, but in addition it also requires that this presupposition would be satisfied by an explicit discourse antecedent. Thus we can think of strong (S) and weak (W) as two distinct properties of definite NPs, when they can both be satisfied in SDef, namely: SDef= [W+ S].

For purposes of this work, I will sometimes call these presuppositional/anaphoric requirements of W and S background requirements, since they place constraints on what the background context has to be like for W and S to be appropriate.

In the GIVEN context: the presupposition will be satisfied by the witness of the indefinite a waiter, so the background requirements of both W and S are satisfied in GIVEN.

In the BRIDGED contexts, however, neither W nor S are their background requirements satisfied. However, we predict that accommodation is possible for both WDef and SDef because accommodation is appropriate here given that context supports the background requirements of each very well, even though it does not technically satisfy them.

In the NEW context neither of the background requirements are satisfied. The competition between SDef and WDef can be accounted for by Heim's (1991) “Maximize Presupposition!” Principle: the (in)definite marker whose background requirements are stronger should be used in CONTEXTS in which these requirements are satisfied. Thus,

this predicts that WDef/SDef should be used instead of Indef, when the presupposition of WDef/SDef is satisfied (GIVEN and BRIDGED). Specifically, the BRIDGED contexts very strongly support the implicit antecedents referred to by the definites. Following Beaver (2001), speakers and hearers are constantly having to guess the right context, and in such cases, when the “better” guess is a context that entails the relevant presupposition, this might trigger a soft pressure (assuming Maximize Presupposition!) to use definites (with easy bridging) instead of indefinites.

Similarly, SDef should be used instead of WDef in contexts that (S’s background requirement is stronger than W’s, and the context satisfies it) (GIVEN).

This pressure might also make the participants disprefer Indef in BRIDGED contexts because such contexts might sometimes get very close to satisfying the presupposition of WDef.

Since SDef is designed for anaphoric use, we might expect WDef to ‘compete’ with SDef as a way to express definiteness in Given contexts. In such contexts, we might expect SDef to ‘win’, because it is most naturally anaphoric. This would be a general case of sentences with ‘stronger’ presuppositions being preferred to ones with ‘weaker’ presuppositions when the stronger presupposition is satisfied in the context; this is Heim’s (1991) Maximize Presupposition! principle (cf. the contrast between the sun is shining and #a sun is shining).

Table 1. Predictions based on appropriateness conditions. To summarize, our analysis predicts the behavior of the strong and weak definites for each of the CONTEXTS (Given, Bridged, New) as following:

GIVEN Explicit antecedent: S + W: satisfied S: the most natural (Heim, 1982)

	(IND violated, Heim, 1982)	W not as good as S (MPP, Heim, 1991)
		Indef should be the less natural
BRIDGED	Implicit antecedent: S, W, violated but accommodation is expected	S and W the most natural * IND the less natural (MPP, Heim, 1991)
	(IND satisfied, but maybe violated because of the accommodation for S and W)	S vs. W: because of having different presuppositions, an effect is expected, but we are not sure how.
NEW	Neither explicit nor antecedent: S, W not satisfies and accommodation is not expected. IND not violated.	IND the most natural, S and W the less natural S vs. W: S is expected to be worst, since stronger presuppositions are violated.

2.5 Experimental hypothesis and predictions

The experimental hypotheses and predictions are built on the predictions of the behavior of weak/strong definites, and indefinites, illustrated above.

It is hypothesized that (i): *-e* marked NPs (SDef) function as strong definites (S), (ii): null marked NPs (WDef) function as weak definites (W), and (iii): *ye...-i* marked NPs (Indef) function as indefinites (IND). The identification of three NP types with the MARKER (SDef, WDef, or Indef) makes clear predictions in the context of our experiment, which we identify below:

It is expected that the naturalness of both SDef and WDef would be a function of how good the background presuppositions are satisfied, namely a function of how accessible an antecedent is, as illustrated above. Thus, it is predicted that SDef and WDef would be rated the best in the contexts in which an explicit antecedent is available (GIVEN); they would be not as good as in the GIVEN context in the context in which the antecedent is not available but can be easily inferred (BRIDGED); and they are the worst in the context in which an antecedent is neither directly or indirectly available (NEW). Also, we predict Indef to be less natural in the GIVEN context, compared to BRIDGED and NEW. We expect the Indef to be rated the best in NEW if the presupposition of W (i.e., uniqueness and existence of the reference of a target NP) is satisfied, in which case the rating for Indef-BRIDGED is expected to be decrease, (and this consequently makes the Indef to be the best, compared to the GIVEN and BRIDGED) (Heim, 1982, 1991).

As for WDef versus SDef, we expect a penalty for using WDef in the GIVEN contexts. That is, (i) it is not expected that the GIVEN-WDef would be rated much more natural than BRIDGED-WDef, (ii) we expect the SDef to be rated more natural than WDef in GIVEN. This hypothesis does not predict any difference between WDef and SDef in the BRIDGED and hence the prediction is left open for empirical research. However, by relying on the intuitive judgment, I predicted that SDef would be rated more natural than WDef in the BRIDGED. As for WDef/SDef versus Indef, we expect the Indef to be rated more natural than WDef/SDef in NEW, and less natural in GIVEN and BRIDGED.

The predictions for MARKER and CONTEXT interactions, as illustrated above, are summarized in Table 2 below, compared in all nine conditions. The conditions with

more checkmarks (‘✓’) are expected to be rated more natural and the conditions with more cross marks (‘✗’) are expected to be less natural.

Table 2. Experimental predictions.

CONTEXT	MARKER		
	SDef (-e)	WDef (\emptyset)	Indef (ye...-i)
Given	✓ ✓ ✓	✓ ✓ (?)	✗ ✗
Bridged	✓ ✓ (✓)	✓ ✓	✗
New	✗ ✗ ✗	✗ ✗	✓ (?)

3 Chapter: Offline Naturalness Sentence-Pair Experiment

An offline questionnaire was conducted to test the predictions put forward in Section 2.5 above.

3.1 Method: materials and procedures

3.1.1 Experimental stimuli

Experimental stimuli were created with the following design:

(14) CONTEXT

[TP[N] [VP [PP [P NP, *Indef-N*]] [V, *nonverbal element +LV-PSPT + be.PST.3SG*]]]

a. GIVEN

šimā	bā	ye=xānānde-i	hæm=sæfær	šod-e	būd.
Shima	with	one=singer-IND	co=travel	became-PSPT	was

‘Shima had traveled with a singer’.

b. BRIDGED

šimā	æz	ye=konsert-i	didæn	kærd-e	būd.
Shima	of	one=concert-IND	visiting	did-PSPT	was

‘Shima had visited a concert’

c. NEW

šimā	be	ye-rūznāme-i	negāh	kærd-e	būd.
Shima	to	one-newspaper-i	look	did-PSPT	was

‘Shima had looked at a newspaper’.

(15) TARGET

(For each of a, b, and c: syntactic structure, followed by Persian, gloss, and translation)

a. *First 4 words (relative clause)*

[CP [RC [CP [D + N + C, {*un=tor=i* + *REL*, *ke*}] VP [N + V, *mi-say.PRES.3SG*]]]...

un=tor=i ke šimā mi-ge,
That=way-IND that shima DUR-says
'According to Shima,...

b. *Target NP*

...[VP [NP [D + N, ({SDef/WDef/Indef}N)]]]

{*xānānde-he/ Øxānānde/ ye xānānde-i*}
{SDef-singer/ WDef-singer/ Indef-singer}
...{*the_{STRONG/WEAK} singer/ a singer*}...

c. *The rest of the sentence:*

VP [{5-7 words, mostly: *PP PP*} + V {(mi-) + PSPT + (bud.PSPT) + (SU Agr)}]]

{*æz pænj sālegi rūye səhn*} -- -- būd-e [=æst]
of five years on scene -- -- was-PSPT [=is]
...had been on the scene from the age of five'.

As illustrated above, three context sentences (14) (a. Given, b. Bridging, c. New) preceded the target sentence (15). The stimuli were constructed from 24 lexical sets (or items), resulting in a total of 216 context-target sentence pairs. For each stimulus item, to provide an explicit and implicit antecedent for the target NP in the Given and Bridged contexts, respectively, sentences varied in their object NPs and the verbs associated with them. Target sentences were not randomly paired with the context sentences. Rather, each target sentence was individually chosen in correspondence to the context sentence, to which it was associated, based on the intuitive judgment. Hence, the target sentence was the same for each stimulus item. This was to prevent the unwanted effect caused by

incongruent context-target sentences.¹⁴ The Context-Target words in the sentence pairs with the Bridged contexts, (like *a concert* in (14b) and *a/the singer* in (15), were assumed to be semantically associated to one another. To control for the plausibility of presupposition accommodation, Clark’s classification of the different types of “bridging relations” between a definite NP and its antecedent in discourse, suggested in his “bridging” paper (1975), was used as a measure for choosing the associated word pairs; all word pairs were chosen such that the Context word (i.e., the words providing the *implicit antecedent*) would be considered an “indirect referent by association or characterization” (Clark, 1975) in relation to the Target word associated with it. More specifically, the Context words represented a location or a situation, and the Target words, represented an individual or inanimate object, which had a *necessary or optional salient* role (Clark, 1975), or was a *necessary or optional salient* part of the context represented by the Context word. The final words pairs used in the stimuli, were chosen from the highly rated word pairs from a pilot relatedness noun-pair questionnaire conducted prior to the main study. (See Appendix B for details regarding the relatedness noun-pair questionnaire). For the purpose of having context sentences with consistent structures, New context sentences (i.e., the contexts that assumingly provided neither an explicit nor implicit antecedent for the (in)definite NPs in the target sentences) also consisted of an indefinite NP, similar to the Given and Bridged. In contrast to the Given and Bridged contexts, however, the context NPs used in the New were chosen such that they would be “neutral” in relation to the target NPs; namely, on the one hand, they were

¹⁴ This explains the variations of the predicates of the target sentences (15c) in each stimuli item.

chosen such that they would not be semantically associated with any of the Target NPs, so the existence of a reference for the Target NPs would not be predictable by the context word, and background requirement for the definite NPs would not be likely to be satisfied. On the other hand, they were not intentionally chosen such that the target NPs would be implausible in such context, so they could provide a plausible context for the indefinite targets. This would prevent having incongruent Indefinite target NPs and New context (i.e. Indef-NEW condition).

The syntactic structure of all of three context sentences was the same, as illustrated in (14). Each of the context sentences had seven words, and consisted of a proper name followed by a prepositional phrase (PP), and ended with a past perfective compound verb. Compound verbs, consisting of a nonverbal element and a verbal element, account for the majority of verbs in Persian (e.g., see Robert, 2003). The verbal elements used in the compound verbs of the context sentences consisted of past perfective light verbs (LV), which were among the most frequent LVs used in Persian, as identified by Dehdari (2006). Target sentences were also controlled for their syntactic structures: each target sentence consisted of nine to eleven words (and seventeen to twenty five syllables, started with a relative clause (RC)(*according to Shima in (15)*), followed by an (in)definite NP (i.e., the Target word), and then by the rest of sentence, as in (15). The reason for adding a relative clause in the beginning of the target sentences was that Persian is a subject-drop language (Ghomeshi, 2003, 2006), hence, it is preferable to drop the subject when it is mentioned in the preceding sentence. This could cause a repetition penalty for the conditions with the Given context, where the (in)definite subject phrase in the target sentence had an identical referent (providing the

explicit antecedent) in the context sentence. To prevent this error, the same relative clause was added at the beginning of all the target sentences in order to make them longer and create a gap between the occurrence of the target NP and its antecedent. There were some repetitions across 24 stimuli items: the verbs used in the Given and Bridged contexts were repeated once across items. The New contexts were repeated three times across 24 items.

Context-Target sentence pairs were distributed by constraint-randomization. The sentence pairs were distributed into nine sets, each of which consisted of 24 Context-Target sentence pairs from different stimulus items. (i.e., there were no two Context-Target pairs in the same set that came from the same lexical set.) No sentence pairs providing the same conditions followed one another more than three times in a sequence. The distribution of the repetitions of the New contexts, and the verbs of the Given and Bridged contexts, was controlled such that no sentence pairs with the repeated contexts followed one another.

3.1.2 Participants

34 monolingual native Iranian Persian speakers (15 female) between the ages of 23 and 38 (Mean: 29.8; Median: 30) participated in this study. Participants were recruited by posting recruitment announcements, and were paid for their participation. All participants had completed their primary and secondary education in Persian in Iran, and either had a graduate degree or were a graduate student (49%: Ph. D, 38%: master's, 5% bachelor, 8%: other advanced post graduate degrees). On average, participants stayed in Canada for 2 years, and they had not been outside of Iran for more than 8 years. All

participants considered themselves to be monolingual native Persian speakers and reported that they speak Persian on a daily basis. 15 out of 34 participants were originally from Tehran, and the rest were from 7 different cities in Iran.

3.1.3 Procedures

Prior to the experimental questionnaires, participants signed an informed consent form, filled out a form asking for their demographic information, and were also asked to fill out a language background questionnaire indicating their exposure to Persian and to any other languages. The language background questionnaire was created based on the measures of language exposures used by Wartenburger, et al., (2003). (See Appendix D for the language background questionnaire). Participants were then presented with two types of experimental questionnaires: naturalness sentence-pair questionnaire and noun-pair questionnaire. Since the naturalness sentence pair questionnaire is the focus of the present study, the participants were instructed to fill out the naturalness sentence-pair questionnaire (see the Experimental material section) before they were asked to fill out the noun pair questionnaire. Participants were first presented with the first set of the naturalness sentence-pair questionnaire, which consisted of the first 96 sentence pairs out of a total of 216 pairs. After a short break the second set consisting of the rest of the sentence pairs was given to the participants. In the naturalness sentence-pair questionnaire, participants were asked to read each context-target sentence pair carefully and rate how natural the target sentence was as the continuation of the context sentence, on a basis of 1 to 7 scale point where 1 being most natural and 7 being most unnatural. (See Appendix C for the instruction and sample stimuli for the sentence-pair

questionnaire). After finishing both sets of the naturalness questionnaire, participants were presented with the noun-pair questionnaire. In the noun-pair questionnaire they were asked to read each set of written word pairs and rate it for the relatedness between each of the nouns in a pair, on a basis of 1 to 5 scale point where 1 having no apparent connection and 5 having strong connection. (See Appendix B1 for the instruction and sample stimuli for the noun-pair questionnaire). The entire experimental session, including instructions and debriefing, took approximately 90 minutes.

3.2 Results

Table 2 presents the results of the offline questionnaire (the naturalness sentence-pair questionnaire) conducted. The results were analyzed in the following way. Although six participants filled out only the first part of the questionnaire (96 sentence pairs), the data for those participants were included in the final analyses presented below. The descriptive statistics for the data for those participants were conducted and the results showed that the pattern of the data was similar between them and the rest of the participants.

Table 3. Results of off-line questionnaire. Mean rating and its standard deviation (in parentheses) for nine conditions (1:most natural; 7: less natural)

MARKER			
CONTEXT	SDef (-e)	WDef (\emptyset)	Indef (ye...-i)
Given	1.67 (0.67)	2.28 (0.93)	4.74 (1.05)
Bridged	2.13 (0.67)	2.40 (0.72)	3.40 (0.76)
New	5.01 (0.93)	4.80 (0.89)	3.46 (0.98)

For the statistical analyses for the results, a repeated measures 3 x 3 ANOVA with the first factor CONTEXT (Given, Bridged, New) and the second factor MARKER (SDef, WDef, Indef) was carried out using error terms based on participant ($F1$) and item variability ($F2$). Following this, pairwise contrast tests were conducted using the Bonferroni correction. Only significant results are reported in what follows.

The results showed that there was a significant effect of CONTEXT ($F1(2, 66) = 141.18, p < .001, F2(2, 46) = 53.93, p < .001$), a significant effect of MARKER ($F1(2, 66) = 48.21, p < .001, F2(2, 46) = 34.66, p < .001$), and a significant interaction between CONTEXT and MARKER ($F1(4, 132) = 114.61, p < .001, F2(4, 92) = 70.73, p < .001$). As for the significant effect of CONTEXT, Given and Bridged contexts were, on average, rated better than New context (Given vs. New, $F1(1, 33) = 129.04, p < .001, F2(1, 23) = 69.22, p < .001$; Bridged vs. New, $F1(1, 33) = 240.25, p < .001, F2(1, 23) = 74.30, p < .001$). This result suggests that regardless of MARKER participants rated the target sentences with either an explicit (GIVEN) or implicit (BRIDGED) antecedent better than those without an antecedent in the prior context (NEW). For the significant effect of MARKER, the SDef was rated better than the WDef ($F1(1, 33) = 9.85, p = .01, F2(1, 23) = 6.55, p = .05$). The WDef was rated better than the Indef ($F1(1, 33) = 46.92, p < .001, F2(1, 23) = 40.32, p < .001$). The significant interaction between CONTEXT and MARKER is due to the difference in the rating patterns found for each CONTEXT for some of the MARKER conditions. As shown in Figure 1 below, in the GIVEN CONTEXT, the results differed significantly among the three MARKER conditions. That is, the SDef was rated better than the WDef ($F1(1, 33) = 14.36, p < .001, F2(1, 23) = 393.22, p < .001$) and the WDef was rated better than the Indef ($F1(1, 33) = 155.25, p < .001, F2(1, 23) = 326.16, p < .001$). This finding

supports the main experimental hypothesis concerning *-e*, \emptyset , *ye NP-i* with three distinct types of Marker: *-e* as a strong definite marker, \emptyset as weak, and *ye -i* as the indefinite marker. In the BRIDGED CONTEXT, a two way distinction for the MARKER was observed: both SDef and WDef were rated better than the Indef (SDef vs. Indef: $F1(1, 33)=78.49, p<.001, F2(1,23)=35.64, p<.001$; WDef vs. Indef: $F1(1,33)=42.64, p<.001, F2(1,23)=13.83, p<.001$).¹⁵ This result implies that when an implicit antecedent is provided, both SDef and WDef are more appropriate, compared to the Indef, indicating that *-e* and \emptyset signal definiteness and *ye-i* signals indefiniteness. Finally, the NEW CONTEXT showed the pattern of the results similar to the Bridged Context though the direction of the rating results was opposite: as predicted, the Indef was rated better than both SDef and WDef (Indef vs. SDef: $F1(1, 33)=64.00, p<.001, F2(1, 23)=35.64, p<.001$; Indef vs. WDef: $F1(1, 33)=68.89, p<.001, F2(1, 23)=13.82, p<.001$). This result shows that *ye -i* is more appropriate when no explicit or implicit antecedent is given, providing additional evidence for identifying *-e* and \emptyset as definite markers and *ye...-i* as an indefinite marker.

¹⁵ There was a significant difference in the rating results between the Bridged SDef and WDef by participants ($F1(1, 33)=19.44, p<.001$), however, not by items ($F2(1, 23)=3.57, p=.21$).

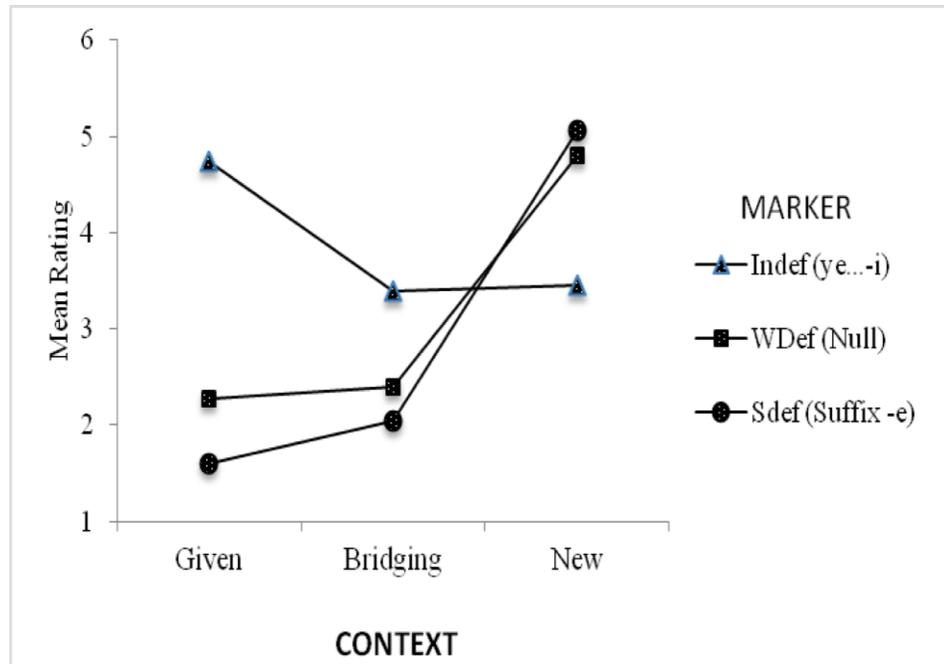


Figure 1. Mean ratings in nine conditions. (1:most natural; 7: less natural)

3.3 Discussion

As predicted, by identifying *-e* with being a strong definite marker, *-e*-marked NPs (SDef) were rated more natural than WDef when an explicit antecedent was available (GIVEN), and *-e*-marked NPs were rated more natural in GIVEN contexts than in either BRIDGED or NEW contexts. In addition, even though the background requirements for WDef were satisfied in the GIVEN context but not in the BRIDGED context, the ratings for WDef were not better in GIVEN than in BRIDGED (unlike SDef; see Figure 1). We assumed that the two distinct types of definite markers in Persian can be distinguished by the types of background requirements they presuppose: anaphoric elements (e.g., strong definite NPs (SDef)) would require both anaphoric requirements (S) and presuppositional requirements (W) (i.e., uniqueness and existence (Frege, 1982) to be satisfied; and weak elements (W) only would require the presuppositional

requirements (W) to be satisfied. The Maximize Presupposition! Principle (Heim, 1991) predicts that definites that impose stronger presuppositions on context are used in contexts in which these requirements are satisfied. The GIVEN contexts in which an explicit antecedent is provided would satisfy both W and S background requirements, but Maximize Presupposition! penalizes use of W (S should be used). Thus the only condition that satisfies all constraints is use of S in GIVEN; use of S and W in BRIDGED violates their background requirements (but these can be rescued to an extent by bridging inferences), and use of W in GIVEN violates Maximize Presupposition! We hope this study contributes to providing a quantitative measure of the severity of these violations; currently, most semantics work on the topic provide a simple binary “appropriate or not” judgment.

The pattern found for both WDef and SDef was consistent with the regular use of definite NPs, being rated more natural in GIVEN compared to NEW, indicating that they impose background presuppositions on the context, and second, accommodate the presupposition, if plausible. (Burkhardt, 2006; Heim, 1982; Hirotani and Schumacher, 2009; Singh et al., 2014).

Additionally, the opposite pattern for Indef is consistent with the use of indefinite markers, assuming that they cannot be used anaphorically (Karttunen, 1976; Heim, 1982).

Our analysis did not predict that SDef would be rated significantly better than WDef in the Bridged context. At the moment it is not entirely clear how to account for this. It is clear that neither was particularly inappropriate, and in particular the appropriateness of S suggests that a new referent is quite easily established via

accommodation. This suggests that we are correct in rejecting the view that *-e* functions as a demonstrative marker (e.g., Hedberg et al., 2009), but it does raise the question of why accommodation is easy with anaphoric definites even though accommodation for other kinds of anaphoric elements is hard (e.g., Heim, 1982; van der Sandt, 1992). We leave this as a question for future work.

4 Chapter: General Discussion and Conclusion

4.1 General discussion

There are many more complexities and remaining questions to answer regarding the (in)definite marking in Persian in general, and specifically regarding the function and processing of the suffix *-e*.

A more sensitive methodology like ERP might be able to help clarify some of the effects we found in the naturalness questionnaires. Specifically, the strong versus weak distinction showed an interesting interaction with the Bridged context. It was found that both strong and weak definites are somewhat acceptable in the Bridged context, but the underlying processes that facilitate bridging/accommodation are likely to be different, assuming that strong definites involve anaphora in addition to presuppositions. Finding different ERP components for these two kinds of inference might in turn shed light on whether English *the* shares more in common with strong or weak definites; as it stands, English *the* seems to possess qualities of anaphoric and non-anaphoric definites, suggesting that it might be ambiguous between a strong/weak.

4.2 Conclusion

Our findings supported the predicted three-way distinction between the naturalness of NP-e, ØNP, and ye NP-i, as strong definite, weak definite, and indefinite. These results can hopefully shed light on understanding the semantics and processing of the weak and strong definite NPs in the other languages as well. There is still a lot to discover related to the semantics and processing of the (in)definite NPs in Persian. The present study only scratches the surface and hopefully opens a new door for more

research on (in)definites in colloquial Persian, from a semantic/pragmatics/discourse processing point of view. In particular, we hope that the quantitative measures found here provide a useful supplement to intuitive binary acceptability judgments in formulating theories of the factors governing the appropriate use of definite and indefinite markers.

Appendix A List of Abbreviations Used for Glossing the Examples

C	complementizer	PREP	preposition
CP	complementizer phrase	PRES	present tense
DEF	the ‘definite marker’, (both ‘-e’ and ‘Ø’)	PSPT	past participle
DP	determiner phrase	PST	past
DUR	durative prefix (<i>mi-</i>)	REL	relative pronoun (<i>ke</i>)
IND	the ‘indefinite marker’ (<i>-i</i>)	SG	singular
IPFV	imperfective	SU agr	subject agreement
LV	the ‘light verb’	V	verb
N	noun	VP	verb phrase
NP	noun phrase	1SG	first person singular
OM	the object marker (<i>rā</i>)	2SG	second person singular
P	preposition	3SG	third person singular
PP	prepositional phrase	1PL	first person plural
PFV	perfective	2PL	second person plural
POSS	the possessive marker (<i>-eš</i>)	3PL	third person plural

The specific abbreviations used for distinguishing the three (in)definite markers investigated in the study:

Indef	the colloquial indefinite noun phrase/marker (‘ye NP-i)
SDef	the strong definite noun phrase/marker (‘NP-e’)
WDef	the weak definite noun phrase/marker (ØNP)

Appendix B Relatedness Noun-Pair Questionnaire

In addition to the main questionnaire (naturalness sentence pair questionnaire, participants were given a questionnaire consisting of 200 Persian noun pairs (140 experimental items and 60 filler items). They were asked to rate the noun pairs based on how closely each of a pair of words were semantically related, on a 5 point scale with 1 having no apparent connection and 5 having strong connection. The aim was to provide a database of Persian noun pairs with different types of semantic relations. A pilot version of the noun pair questionnaire was distributed among Persian speakers prior to the main experiment. The distribution of the noun pair questionnaire was continued along with conducting the main experiment as well.

100 of the experimental word pairs consisted of an inanimate word (a location or event) and an individual who had a *necessary or optional salient role* (Clark, 1975) in the context represented by the other word. The remaining experimental items consisted of semantically related inanimate word pairs one of which was assumed to be *salient* in the context of the other word. 20 of the inanimate word pairs which were assumed to have *part-whole* relation (e.g., ‘*otāq-sæqf*’ ‘room-ceiling’) (Clark, 1975) were categorized separately from the other 20 inanimate word pairs (e.g., *tavllod-keyk* ‘birthday-cake’). Filler items (60 noun pairs) also were divided into the following two groups: 30 *inanimate* noun pairs with no intended association, and 30 animate-inanimate noun pairs. Half of the animate words were chosen from the animate words used in the experimental item paired with the random and unrelated inanimate words. The other half were animate words which were paired with some randomly chosen inanimate words from the experimental items.

Different noun-pair categories were pseudo-randomized such that there were not more than four experimental and filler items in on order. And the noun pairs in each of the five categories did not follow one another for more than three times. The critical Context-Target words used in Bridging contexts of the sentence pair questionnaire (24 word pairs) were chosen from the highly related word-pairs from the *experimental category* used in the pilot relatedness noun-pair questionnaire.

B.1 Instruction and sample stimuli for the relatedness noun-pair questionnaire (The English translation follows in the next page)

شماره کاربری: تاریخ:

دستور عمل: لطفا جفت کلمات نشان داده شده در جدول زیر را بر اساس این که چقدر به هم ارتباط دارند به شرح زیر از ۱ تا ۵ امتیاز بندی کنید:

- امتیاز ۱ برای دو کلمه‌ای که شما ارتباط آشکاری نمی‌بینید، یا این که فکر می‌کنید کاملاً به هم بی‌ارتباط هستند.
- امتیاز ۲ برای دو کلمه‌ای که ارتباط ضعیفی با هم دارند.
- امتیاز ۳ برای دو کلمه‌ای که ارتباط متوسط با هم دارند.
- امتیاز ۴ برای دو کلمه‌ای که ارتباط آشکار با هم دارند.
- امتیاز ۵ برای دو کلمه‌ای که ارتباط نزدیک و قوی با هم دارند.

برای مثال دو کلمه ی "پدر" و "مادر" و همچنین دو کلمه ی "کنسرت" و "آهنگساز" می‌توانند به عنوان جفت کلماتی که ارتباط خوبی با هم دارند، در نظر گرفته شوند. در حالی کلمات "آشپزخانه" و "کنسرت" یا "خودکار" و "آش" به نظر می‌رسد که با هم بی‌ارتباط باشند. برای هر جفت کلمه‌ای که در جدول نشان داده شده است، لطفاً دور عددی که مورد نظر دارید و فکر می‌کنید از همه مناسب‌تر است یک دایره بکشید. به عنوان مثال، من به دو کلمه پدر و مادر امتیاز پنج را می‌دهم (دارای ارتباط قوی)، و برای تولد و هدیه امتیاز ۴ را در نظر می‌گیرم (دارای ارتباط آشکار).

پدر مادر ارتباط قوی ۱...۲...۳...۴...۵ بدون ارتباط آشکار

تولد هدیه ارتباط قوی ۱...۲...۳...۴...۵ بدون ارتباط آشکار

لطفا کلمات را به همان ترتیب داده شده بخوانید. لطفاً هیچ کلمه ای را بدون پاسخ نگذارید و یا به کلمات پیشین بازنگردید. لطفاً توجه داشته باشید که این امتیازبندی بر اساس شهود شخصی شما به عنوان یک فارسی زبان می‌باشد. بنابراین نیازی به فکر کردن عمیق درباره کلمات نمی‌باشد و می‌توانید امتیازبندی را بر اساس اولین حدس و شهود خود اعمال کنید. در هر زمان در هنگام پاسخ دادن به این پرسش نامه، اگر مایل به ادامه نبودید، می‌توانید انصراف دهید.

اگر سوال یا ابهامی درباره این پرسش نامه دارید لطفاً با الان با پژوهشگر در میان بگذارید.

با تشکر فراوان از همکاری شما

۰	کنسرت	خواننده	ارتباط قوی ۱...۲...۳...۴...۵ بدون ارتباط آشکار
۱	جشنواره	مجری	ارتباط قوی ۱...۲...۳...۴...۵ بدون ارتباط آشکار
۲	میوه فروشی	سخنران	ارتباط قوی ۱...۲...۳...۴...۵ بدون ارتباط آشکار
۳	آشپزخانه	اجاق گاز	ارتباط قوی ۱...۲...۳...۴...۵ بدون ارتباط آشکار
۴	دانشگاه	شهریه	ارتباط قوی ۱...۲...۳...۴...۵ بدون ارتباط آشکار
۵	تیراندازی	قربانی	ارتباط قوی ۱...۲...۳...۴...۵ بدون ارتباط آشکار
۶	آرایشگاه	سبب‌زمینی	ارتباط قوی ۱...۲...۳...۴...۵ بدون ارتباط آشکار
۷	تزیینات	کفاش	ارتباط قوی ۱...۲...۳...۴...۵ بدون ارتباط آشکار
۸	اسباب‌کشی	تلویزیون	ارتباط قوی ۱...۲...۳...۴...۵ بدون ارتباط آشکار
۹	رستوران	آشپز	ارتباط قوی ۱...۲...۳...۴...۵ بدون ارتباط آشکار
۱۰	تولد	کیک	ارتباط قوی ۱...۲...۳...۴...۵ بدون ارتباط آشکار

ID#: _____

Date: _____

Instructions: According to your intuitions, please rate the noun-pairs shown in the table below for how closely they are related. The rating is on the basis of a 5-point scale, with "1" having a very strong connection, "3" having an average connection, "4" having a weak connection, and "5" having no apparent connection.

For example, the word-pairs "father-mother" and "concert-conductor" can be evaluated as word pairs that are closely related. However, word pairs "kitchen-concert" and "soup-pen" do not seem to be related to one another.

For each word-pair given in the table, please circle the number that you feel is the most appropriate. For example, I would rate the word-pair "father-mother" as "1" (a very strong connection) and "birthday- present" as "2" (an apparent connection), as follows:

Father Mother Very strong connection 1 --- 2 --- 3 --- 4 --- 5 No apparent connection
Birthday Present Very strong connection 1 --- 2 --- 3 --- 4 --- 5 No apparent connection

Please follow the given order of the noun pairs. We ask you try not to skip or go back to previous words. Note that you are asked to rate the words according to your first instinct, so please do not worry about your judgment and do not think too deeply about it. If you do not feel comfortable with continuing the study, please feel free to withdraw at any point during the study.

If you have any questions, please feel free to ask now.

Thank you.

1	Concert	Conductor	Very strong connection 1--- 2 --- 3 --- 4 --- 5 No apparent connection
2	Kitchen	Pilot	Very strong connection 1--- 2 --- 3 --- 4 --- 5 No apparent connection
3	Guest	Party	Very strong connection 1--- 2 --- 3 --- 4 --- 5 No apparent connection
5	Hospital	Violin	Very strong connection 1--- 2 --- 3 --- 4 --- 5 No apparent connection
6	Flight	Airport	Very strong connection 1--- 2 --- 3 --- 4 --- 5 No apparent connection

Appendix C Instruction and Sample Stimuli for the Naturalness Sentence-Pair

Questionnaire

Following is a sample of the instruction given to the participants for the naturalness sentence pair questionnaire, followed by a few sample stimuli of what participants saw in the questionnaires. Please see the experimental stimuli section (3.1.1) for the English description of the stimuli.

L1 Date: _____

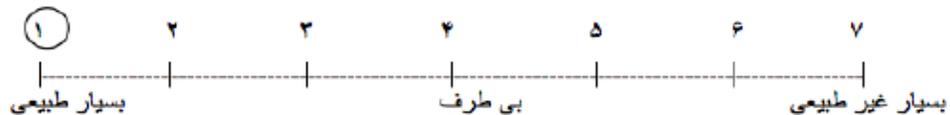
No: _____

دستور عمل پرسش نامه ی شامل جملات: لطفا دستور عمل را با حوصله و دقت بخوانید، و اگر ابهامی بود از پژوهشگر بپرسید.

در این پرسش نامه از شما خواسته می شود که یک سری جفت جملات را در ذهن خود خوانده و درباره ی آنها قضاوت کنید. لطفا بنا به شهود خود نسبت به زبان فارسی، به هر دو جمله ای که به شما نشان داده شده است، امتیازی بین "۱" و "۷" نسبت بدهید. این امتیاز بندی بر این اساس است که اگر در مکالمه ای که میان دو نفر در جریان است، دو جمله ی مربوطه را بشنوید، آیا جمله ی دوم (جمله ب) می تواند ادامه ای مناسب و طبیعی برای جمله ی اول (جمله ی الف) باشد یا خیر. امتیاز "۱" نشان دهنده ی «ادامه ی بسیار طبیعی» و امتیاز "۷" نشان دهنده ی «ادامه ی بسیار غیرطبیعی» می باشد. لطفا دور عددی که فکر می کنید از همه مناسب تر است خط بکشید. به عنوان مثال، من به مورد زیر امتیاز "۱" را می دهم، به این دلیل که فکرمی کتم جمله ی دوم «اونظوری که زهرا می گه، به قایقرانی تو مسابقات غرق شده بوده»، می تواند ادامه ی قایل قبولی برای جمله ی اول «سارا به یه روزنامه ای نگاه کرده بود»، باشد:

الف. زهرا به یه روزنامه ای نگاه کرده بود.

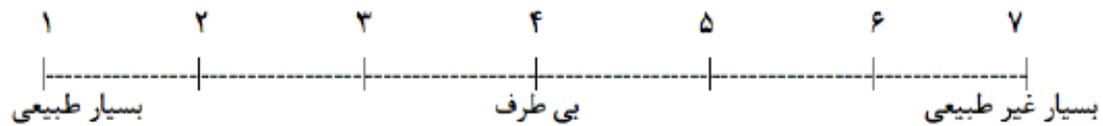
ب. اونظوری که زهرا می گه، به قایقرانی تو مسابقات غرق شده بوده.



1

مهرداد با به دوستی بیرون رفته بود.

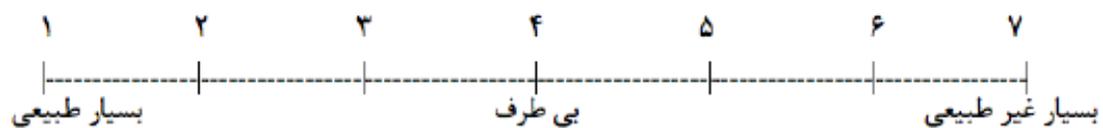
اونطوری که مهرداد می‌گه، به فیلمبرداری همینطوری از مردم فیلمبرداری می‌کرده.



2

شیرا تو به کتابخونه ای مطالعه کرده بود.

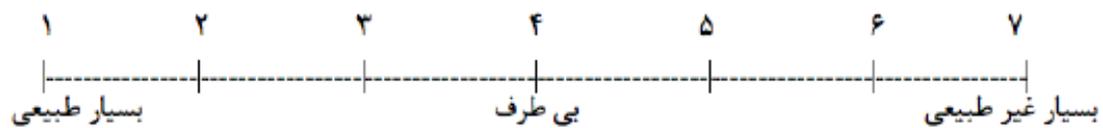
اونطوری که شیرا می‌گه، به منجمی اسمی همه ی ستاره‌ها رو می‌دونسته.



3

سپیده با به گزارشگری همسفر شده بود.

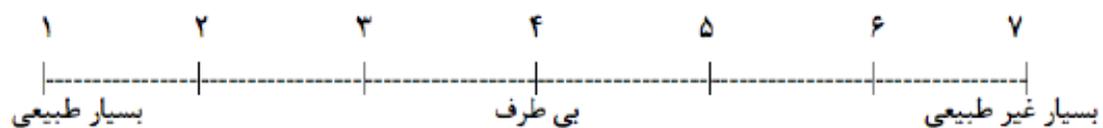
اونطوری که سپیده می‌گه، به گزارشگری تو خیابون همینطور با مردم مصاحبه می‌کرده.



4

سپیده به به اسباب کتبی کمک کرده بود.

اونطوری که سپیده می‌گه، یخچال یرت شده بوده وسط خیابون



Appendix D Language Background Questionnaire

(In Persian, followed by the English translation in the next page)

فرم سابقه ی زبانی

لطفا جدول زیر را پر کنید. در این جدول زیر پرسش‌هایی از شما پرسیده شده است که در مورد میزان در ارتباط بودن شما با زبان فارسی یا هر زبانی دیگری در طول زندگی به هر طریقی در معرض آن بوده اید می‌باشد. در ستون سمت چپ جدول مقطع‌های سنی که در تطابق با مقطع های مختلف تحصیلی می‌باشند نوشته شده است.

در مقاطع زمانی نشان داده شده در زیر، آیا با هیچ زبان (های) دیگری به غیر از فارسی به هر طریقی در ارتباط بوده اید؟				سن
اگر پاسخ شما بله می‌باشد:			خیر	
منبع این ارتباط چه بوده است (برای مثال: مدرسه، آموزشگاه، شغل، اوقات فراغت، مطالعه، فیلم، و غیره)	حدوداً چند ساعت در روز با این زبان (ها) در ارتباط بوده اید؟	چه زبانی(هایی)؟		
				۶-۰ سال
				۱۱-۷ سال
				۱۴-۱۲ سال
				۱۷-۱۵ سال
				۲۱-۱۸ سال
				۲۸-۲۲ سال
				۲۸ سال تا الان

Instruction: Please fill out the table below. The table is intended to collect information concerning your exposure to Persian and any other languages that you have been involved with during your life. Different age ranges corresponding to the crucial educational levels in Iran (e.g., 7-11 years old: elementary school, 12-14 years old: middle school, etc.) are listed in the most left column of the table. (Please note that you can skip any questions you do not feel comfortable with answering.)

Age	During each of the time periods indicated below, were you exposed to any language(s) other than Persian?			
	No	If Yes,		
		Which language(s)?	Approximately how many hours a day were you exposed to the language(s)?	What was the source of your exposure? (e.g., school work, conversation with friends, part time job, hobbies, etc.)
0-6 years				
7-11 years				
12-14 years				
15-17 years				
18-21 years				
22-28 years				
28 years until now				

References

- Afzali, P. (2012). *Marking definiteness in Farsi and English by Farsi speaking EFL learners* (Master's thesis). Linköping University, Sweden.
- Beaver, D. (2001). *Presupposition and assertion in dynamic semantics*. Stanford, CA: CSLI Publications.
- Beaver, D., & Zeevat, H. (2007). Accommodation. *The Oxford Handbook of Linguistic Interfaces*, 503-538. Oxford: Oxford University Press.
- Browne, W. (1970). More on definiteness markers: Interrogatives in Persian. *Linguistic Inquiry*, 1(3), 359-363.
- Burkhardt, P. (2006). Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related brain potentials. *Brain and Language*, 98(2), 159-168.
- Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Charles N. (Eds.). *Subject and Topic*, 25-55. New York, NY: Academic Press.
- Christophersen, P. (1939). *The articles: a study of their theory and use in English*. Copenhagen: Einar Munksgaard.
- Clark, H. H. (1975). Bridging. In Schank, R. C., & Nash-Webber, B. L. (Eds.). *Theoretical Issues in Natural Language Processing*.
- Clark H.H. & Haviland, S. E. (1974). Psychological processes in linguistic explanation. In D. Cohen (Ed.). *Explaining linguistic phenomena*. Washington, DC: Hemisphere Publication Corporation.
- Frege, Gottlob. (1892). Über Sinn und Bedeutung (On sense and reference), *Zeitschrift*

für Philosophie und philosophische Kritik, 25-50.

Geranpayeh, A. (2000). The acquisition of the English articles by Persian speakers.

Edinburgh Working Papers in Applied Linguistics, 10, 37-51.

Ghomeshi, J. (2006). NP-anaphora and focused possessors in parallel architecture.

Linguistics, 44(4), 721-748.

Ghomeshi, J. (1999). Review of: Persian by Mahootian, Sh., & Gebhart, L.

Anthropological Linguistics, 41(4), 574-577.

Ghomeshi, J. (1997). Topic in Persian VPs. *Lingua*, 102, 133-167. Copenhagen:

Munksgaard.

Ghomeshi, J. (2001). Control and thematic agreement. *Canadian Journal of Linguistics*,

46(1/2), 8-40.

Ghomeshi, J. (2003). Plural marking, indefiniteness, and the noun phrase. *Studia*

Linguistica, 57(2), 47-74.

Ghomeshi, J. (2008). Markedness and bare nouns in Persian. In Karimi, S. Samiiian, V.,

& Stilo, D. (Eds.), *Aspects of Iranian Linguistics*, 85-112. Cambridge Scholars

Publishing.

Ghomeshi, J. (2013). The syntax of pragmaticalization. *Proceedings of the 2013 Annual*

Conference of the Canadian Linguistic Association.

Grice, H.P. (1975). Logic and conversation. In Cole, P., & Morgan J. L. (Eds.). *Syntax*

and Semantics, 3, 41-58. New York, NY: Academic Press.

Gundel, J. K., Hedberg, N. and Zacharski, R. (1993). Cognitive status and the form of

referring expressions in discourse. *Language*, 69, 274-307.

Hamedani, L. (2011), *The function of number in Persian* (Doctoral dissertation),

University of Ottawa, ON, Canada.

- Haviland, S. E., & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, 13(5), 512–521.
- Hawkins, J. A. (1978). *Definiteness and indefiniteness: A study in reference and grammatically prediction*. London: Croom Helm.
- Hedberg, N. Gorgulu, E., & Mameni, M. (2009). On definiteness and specificity in Turkish and Persian. *Proceedings of the 2009 Annual Conference of the Canadian Linguistic Association*.
- Heim, I. (1982) *The semantics of definite and indefinite noun phrases* (Doctoral dissertation). University of Massachusetts, Amherst, MA.
- Heim, I. (1991). Artikel und definitheit (articles and definiteness). In von Stechow, A., & Underlinch, D. W. (Eds.). *Semantics: An International Handbook of Contemporary Research*, 487-535. Berlin: Walter de Gruyter.
- Heim, I. Kratzer, A. (1998). *Semantics in generative grammar*. Blackwell.
- Hirovani, M. Schumacher, P. (2011). The role of context and topic marking during discourse processing in Japanese: Evidence from ERPs. *Journal of Neurolinguistics*, (24), 276-292.
- Holcomb, P. (1988). Automatic and attentional processing: An event-related brain potential analysis of semantic priming. *Brain and Language*, 35(1), 66-85. Academic Press.
- Holcomb, P. (1993). Semantic priming and stimulus degradation: Implication for the role of N400 in language processing. *Psychophysiological research*, 30, 47-61.
- Hosseini Fatemi. M. (2013). *The semantics of the Persian object marker –ra*. (Master's

- thesis). Carleton University, Ottawa, ON.
- Burkhardt, P. (2007). The P600 reflects costs of new information in discourse memory. *Neuroreport*, 18(17), 1851-1854.
- Karimi, S. (1989). *Aspects of Persian syntax, specificity, and the theory of grammar* (Doctoral dissertation). University of Washington, DC.
- Karimi, S. (1990). Obliqueness, specificity, and discourse functions. *Linguistic Analysis* 20, 139-191.
- Karimi, S. (1999). A note on parasitic gaps and specificity. *Linguistic Inquiry*, 30(4), 704-713.
- Karimi, S. (1999). Specificity effect: Evidence from Persian. *The Linguistic Review*, 16(2), 125-141
- Karttunen, L. (1976). Discourse referents. In McCully, J. (Eds.). *Syntax and Semantics, Notes from Linguistic Underground*, 363-385. New York, NY: Academic Press.
- Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, 4(12), 463- 470.
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207(4427), 203-205.
- Lewis, D. K. (1979). Scorekeeping in a language game. *Journal of Philosophical Logic*, 8(1), 339-359.
- Lyon, C. (1999). *Definiteness*. Cambridge: Cambridge University Press.
- Mahootian, Sh., Gebhardt, L. (1997). *Persian*. New York, NY: Routledge Press.
- Milanian, H. (1971). Zabansenasi va ta'rif-e zaban (Linguistics and the definition of Language). *FvZ*, 2, 9-21.

- Nieuwland, M. S., & Van Berkum, J. J.A. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience*, 18(7), 1098-1111.
- Rezaei, M. J., & (2010). Jabbari, A. A. The acquisition of definiteness feature by Persian L2 learners of English. *The Journal of Teaching Language Skills*, 2(2), 123-154.
- Robert, C. (2003). Uniqueness in definite noun phrases. *Linguistics and Philosophy*, 26(3), 287-350.
- Rugg, M. D. (1985). The effects of semantic priming and word repetition on event-related potentials. *Psychophysiology*, 22(6), 642-647.
- Rugg, M. D., & Coles M. GH. (1995). *Electrophysiology of mind: Event-related brain potentials and cognition*, New York, NY: Oxford University Press.
- Ruz, M. Madrid, E. Lupianez, J., & Tudela, P. (2003). High density ERP indices of conscious and unconscious semantic priming. *Cognitive Brain Research*, 17, 719-731.
- Sahraei, R. (2011). Definiteness in Persian. *Persian grammar and literature*, 45, 129-157.
- Sadeghi, A. A. (1970). ra[^] dar zaban-e Farsi-ye emrouz (Ra[^] in modern Persian). *Journal of the Faculty of Literature and Humanities*, 9-22, Tabriz University, Iran.
- Samiian, V. (1983). *Origins of phrasal categories in Persian: an X-bar analysis* (Doctoral dissertation). University of California, Los Angeles, CA.
- Schumacher, P. B. (2009). Definiteness marking shows late effects during discourse processing: evidence from ERPs. *Lecture Notes in Artificial Intelligence*, 5847, 91-106.

- Schwarz, F. (2009). *Two types of definites in natural language* (Doctoral dissertation), University of Massachusetts. Amherst, MA.
- Schwarz, F. (2013). Two kinds of definites cross-linguistically. *Language and Linguistics Compass*, 7(10), 534-559.
- Shokouhi, H., & Kipka, P. (2003). A discourse study of Persian ra. *Lingua*, 113, 953-966.
- Singh, R. Fedorenko, E. Mahowald, K., & Gibson, E. (2014). Accommodating presuppositions is inappropriate in implausible contexts, *Cognitive Science* (To appear).
- Stalnaker, R.C. (1974). Pragmatic presupposition. In Munitz, M. K., & Unger. P. K. (Eds.). *Semantics and Philosophy*, 197-213. New York, NY: NYU Press.
- Van der Sandt, R. (1992). Presupposition projection as anaphora resolution. *Journal of Semantics*, 9(4), 333-377.
- Wartenburger, I. Heekeren, H. R. Abutalebi, J. Cappa, S. F. Villringer, A., & Perani, D. (2003). Early setting of grammatical processing in the bilingual brain. *Neuron*, 37(1), 159-170.
- Windfuhr, G. (1979). *Persian Grammar: History and state of its study*. Berlin: Mouton Publishers.
- Xorasani, A. (1950). Nesane-ye ma'refe va nakere (Definite and indefinite markers), *Danes*, 2, 37-53.