Chapter 5

The specificity marker *-e* with indefinite noun phrases in Modern Colloquial Persian

Klaus von Heusinger
Universität zu Köln

Roya Sadeghpoor
Universität zu Köln

Persian has two indefinite markers, the prenominal *ye(k)* and the suffixed *-i*. Both forms express particular kinds of indefiniteness, as does their combination: for Modern Colloquial Persian, indefinites ending in *-i* express a non-uniqueness or anti-definite implication and behave similarly to *any* in English. *Ye(k)*, on the other hand, expresses an at-issue existence implication and behaves similarly to the English *a(n)* (Jasbi 2016). The combination of *ye(k)* and *-i* expresses an ignorance implication. Modern Colloquial Persian has the specificity marker *-e*, which can be combined with *ye(k) NP*, as well as with the combined form of *ye(k) NP-i*, but not with (solitary) *NP-i* (Windfuhr 1979; Ghomeshi 2003). In this paper, we investigate the function of the indefinite form when combined with the specificity marker *-e*, namely *ye(k) NP-e* and *ye(k) NP-e-i*. We present two pilot studies that tested our hypothesis, which is that the contrast between these two specific forms depends on whether the specificity is speaker-anchored, as for *ye(k) NP-e*, or non-speaker anchored, as for *ye(k) NP-e-i*. The results of the two studies provide weak support for this hypothesis, and provide additional evidence for the fine-grained structure of specificity as referential anchoring (von Heusinger 2002).

1 Introduction

Persian is a language with no definite marker and two indefinite markers. In Modern Colloquial Persian, the prenominal indefinite article *ye(k)* 'a(n)' marks

an NP as indefinite and expresses an existential entailment 'there is at least one N', as in (1), similar to a noun phrase with the indefinite article in English. In Modern Colloquial Persian, the suffixed (or enclitic) marker *-i* is interpreted as a negative polarity item (NPI), as in (2), similar to the English *any* (Jasbi 2014; Lyons 1999; Windfuhr 1979). Both indefinite markers can be combined into a complex indefinite, consisting of *ye(k) NP-i*, which is interpreted as a free-choice item, as in (3), or with a certain 'flavor' of referential ignorance, as in (4), similar to *some or other* in English (Jasbi 2016).[1]

(1) Emruz yek māšin tu xiābun didam.   *ye(k) NP* (existential)
    today a   car   at street saw.1SG
    'Today I saw a car on the street.'

(2) Māšin-i ro emruz tu xiābun didi?   *NP-i* (negative polarity item)
    car-i   rā today at street saw.2SG
    'Did you see any cars on the street today?'

(3) Context: Ali wants to play the lottery. Reza is explaining to him how it is played.

    Ye šomāreh-i ro entexāb kon   va   injā alāmat bezan.   *ye(k) NP-i*
    ye number-i rā choose do.2SG and here mark   do.2SG   (free choice)
    'Choose a number and mark it here.'

(4) Yek bače-i tu xiābun gom šode   bud.   *ye(k) NP-i* (ref. ignorance)
    ye child-i at street lost became.3SG was.3SG
    'A/some child was lost in the street.'

Modern Colloquial Persian has the optional suffix *-e*, which we take to express specificity. The literature assumes different functions of this suffix, such as a demonstrative, a definite, or a referential function (Windfuhr 1979: 40; Hincha 1961: 173-177; Lazard 1957: 163; Ghomeshi 2003: 67) or familiarity of the referent (Hedberg et al. 2009) as in the anaphoric noun *pesar-e* in (5):

(5) Emruz ye pesar va ye doxtar ro did-am. Pesar-(e) tās   bud.
    today a boy and a girl   rā saw-1SG boy-e   bald was.3SG
    'Today I saw a boy and a girl. The boy was bald.'

---

[1]Persian has a differential object marker *-ra/-ro/-a/-o* (generally glossed as *-rā* or as OM, DOM or ACC), which is obligatory with definite and specific direct objects, and optional with non-specific indefinite direct objects (Ghomeshi 2003; Karimi 2003; 2018; Lazard 1957; 1992; Windfuhr 1979).

The suffix *-e* is typically used with demonstrative and definite noun phrases, but it can also be combined with the indefinite constructions discussed above, which we take as evidence that it expresses specificity (or referential indefiniteness): (i) its combination with the indefinite marker *ye(k)*, i. e., *ye(k) NP-e*, as in (6), yields a specific reading; (ii) it cannot be combined with suffixed indefinite *-i*: **NP-e-i*, as in (7), due to the incompatibility of the specific function of *-e* and the free-choice function of *-i*; (iii) the specific marker *-e*, however, can be combined with the complex indefinite *ye(k) NP-e-i*, yielding a specific reading in (8), which is very similar to (6).

(6) **Ye pesar-e** injā kār mikone.
    a boy-e here work do.3SG
    'A specific boy works here.'

(7) Diruz **māšin-i** jolo-e dare xune didi?
    yesterday **car-i** front.of-E door.of home saw.2SG

    – *Na, man hič māšin-e-i nadidam.
       no I **any car-e-i** not.saw.1SG
    'Did you see any cars in front of the house door yesterday?'
    Intended reading: 'No, I didn't see any specific car.'

(8) Emruz **ye māšin-e-i** az pošt behem zad.
    today **a car-e-i** from behind to.me collided.3SG
    'Today a specific car collided into me from behind.'

These data, then, raises the following questions. First, what are the differences in the meanings of the three forms expressing indefiniteness in (1) through (4) in Modern Colloquial Persian? Second, what is the contribution of the marker *-e*? Does it express specificity or a different semantic pragmatic notion, such as referentiality, demonstrativeness, topicality, or partitivity? Third, what is the function of the marker *-e* with indefinite constructions, and, more specifically, what is the difference between the two (specific) indefinite constructions *ye(k) NP-e* and *ye(k) NP-e-i*? We assume the following functions of the three indefinite constructions (cf. Jasbi 2014; Lyons 1999; Windfuhr 1979): (i) the indefinite marker *ye(k)* signals a regular indefinite, i. e., it expresses an existential entailment, but does not encode specificity (like the English *a(n)*); (ii) the suffixed marker *-i* is a negative polarity item (like the English *any*); (iii) the combination of the two markers, resulting in *ye(k) NP-i*, shows an ignorance or free-choice implicature.

Second, we assume that the marker *-e* in Modern Colloquial Persian signals specificity in terms of "referential anchoring", in accordance with von Heusinger

(2002). An indefinite is referentially anchored if the speaker, or another prominent discourse referent, can readily identify the referent. This more fine-grained notion of specificity allows us to formulate our Hypothesis 1, about the semantic difference between the two indefinite constructions with the specificity marker, namely *ye(k) NP-e* and *ye(k) NP-e-i*; the specific indefinite construction *ye(k) NP-e* only reflects the intention of the speaker (or speaker-oriented specificity), while the form *ye(k) NP-e-i* only expresses the intention of another salient discourse participant (i. e., non-speaker-oriented specificity).

In §2, we provide a brief overview of the variety of indefinites found in different languages, as well as the ranges of different functions that indefinites can take. In particular, we focus on the contrast between speaker-oriented specificity and non-speaker-oriented specificity. In §3, we discuss the different functions of the indefinite markers in Modern Colloquial Persian and modify the approach of Jasbi (2016). In §4, we present some relevant data for the use of the marker *-e* in Modern Colloquial Persian, and in §5, we present the two pilot studies that addressed our hypotheses about the speaker-oriented specificity of these forms. Finally, §6 provides a discussion and a conclusion.

2 Indefinites in the languages of the world

2.1 Indefinite articles

Languages differ as to whether or not they mark indefinite noun phrases with special morphological means, such as indefinite articles. In Dryer's (2005) WALS sample, 57% of the languages do not have indefinite articles.

Among the 43% of languages that do have an indefinite marker, we find some that have more than one indefinite marker or article, which often expresses the

Table 1: Types of article systems (Dryer 2005)

Type of article system	Instances	Percentages
Indefinite word distinct from numeral for 'one'	91	19%
Numeral for 'one' is used as indefinite article	90	19%
Indefinite affix on noun	23	5%
No indefinite article but definite article	81	17%
Neither indefinite nor definite marker	188	40%
Total	473	100%

contrast between a specific reading, as in (9a), and a non-specific reading, as in (9b), from Lakhota, North America (Latrouite & Van Valin 2014: 405).[2]

(9) a. Wówapi waŋ   o⟨Ø-wá⟩le.   [olé 'look for']
         book   a[+specific] look.for⟨INAN-1SG.A⟩
         'I'm looking for a [particular] book.'

     b. Wówapi waŋží   o⟨Ø-wá⟩le.
         book   a[−specific] look.for⟨INAN-1SG.A⟩
         'I'm looking for a book [any book will do].'

Moroccan Arabic provides a three-way system of indefinite marking: (i) bare nouns are not marked for specificity, as in (10a); (ii) a specific indefinite article *wahed-l*, composed of the numeral 'one' and the definite article, as in (10b); (iii) a non-specific indefinite article *shi*, derived from the word for 'thing', as in (10c) (from Fassi-Fehri 2006; see Brustad 2000: 26-31 for other Arabic dialects):

(10) a. Meryem bgha-t   te-t-zewwej b-**muhami**   wa-layenni waldii-haa
          Maryam wanted-F to-F-marry   with-**lawyer** but   parents-her

          ma bghaw-eh-sh   / wa-layenni ma lqa-t-u-sh.
          not wanted-him-neg / but   not met-her-him-NEG
          'Maryam wanted to marry **a lawyer** but her parents don't like
          him/but she has not met **one** yet.'

      b. Meryem bgha-t   te-t-zewwej b-**wahed r-rajel**   wa-layenni
          Maryam wanted-F to-F-marry   with-**one the-man** but

          ma lqa-t-u-sh.
          not met-her-him-NEG
          'Maryam wanted to marry a (specific) man but she hasn't found
          him/(*one).'

      c. Meryem bgha-t   te-t-zewwej b-**shi**   **rajel** wa-layenni
          Maryam wanted-F to-F-marry   with-**some man** but

          ma lqa-t-u-sh.
          not met-her-him-NEG
          'Maryam wanted to marry a (non-spec.) man but she hasn't found
          one/(*him).'

---

[2] Abbreviations: A 'actor', INAN 'inanimate'.

We will argue in this paper that Modern Colloquial Persian not only exhibits the specific vs. non-specific contrast, as in Lakhota and Moroccan, but also allows us to morphologically mark a more fine-grained structure of specificity, namely whether the specific indefinite is oriented to the speaker or to some other prominent discourse referent within the context.

2.2 Speaker- vs. non-speaker-oriented specificity

German, like English and other languages, has just one indefinite article (11a). However, it has other means of marking the specificity or referentiality of an associated noun phrase. While the regular indefinite in (11a) allows for both a wide- and a narrow-scope reading of the indefinite, the indefinite demonstrative in (11b) clearly signals a referential reading and forces a wide-scope reading:

(11) a. Jeder Student sagte ein Gedicht von Pindar auf.
          'Every student recited a poem by Pindar.'

      b. Jeder Student sagte dieses Gedicht von Pindar auf.
          'Every student recited this$_{indef}$ poem by Pindar.'

Many languages also have special adjectives that can induce different degrees of specificity. Ebert et al. (2013: 31) discuss the differences between the German adjectives *ein bestimmter* and *ein gewisser*, both of which the authors translate as 'a certain', even though the English translation does not reflect the subtle differences in meaning of the German adjectives. Their main observation is that both adjectives force the indefinite noun phrase to scope over the intentional verb *suchen* 'search' (12a-b), while the regular indefinite also allows for the narrow-scope reading, as in (12c):

(12) a. Peter sucht eine bestimmte CD / zwei bestimmte CDs / bestimmte CDs.
          Peter searches a BESTIMMT CD / two BESTIMMT CDs / BESTIMMT CDs.
          'Peter is looking for a certain CD / two certain CDs / certain CDs.'

          ∃ > SEARCH

      b. Peter sucht eine gewisse CD / zwei gewisse CDs / gewisse CDs.
          Peter searches a GEWISS CD / two GEWISS CDs / GEWISS CDs.
          'Peter is looking for a certain CD / two certain CDs / certain CDs.'

          ∃ > SEARCH

      c. Peter sucht eine CD / zwei CDs / CDs.
          Peter searches a CD / two CDs / CDs.
          'Peter is looking for a CD / two CDs / CDs.'   SEARCH > ∃, ∃ > SEARCH

The authors claim that the main difference between *ein bestimmter* and *ein gewisser* has to do with the bearer of the referential intention of that indefinite. For *ein gewisser*, only the speaker of the sentence can have that referential intention. For *ein bestimmter*, in contrast, the speaker or some other salient discourse agent, such as the subject of the sentence, can have this intention. This can be shown by the incompatibility of *ein gewisser* with speaker ignorance in (13b). The most natural reading of (13a) is that Peter knows which CD, but the speaker does not. So, the speaker only reports the assertion that there is some source (e. g., the subject) that has this referential intention.

(13)   a. Peter sucht eine bestimmte CD, aber ich weiß nicht, welche.
            'Peter is looking for a BESTIMMT CD, but I do not know which one.'

        b. Peter sucht eine gewisse CD, #aber ich weiß nicht, welche.
            'Peter is looking for a GEWISS CD, #but I do not know which one.'

We can rephrase Ebert et al.'s observation in terms of "referential anchoring" in von Heusinger (2002; 2011; see also Onea & Geist 2011). The idea is that specific indefinites are anchored to the discourse referent that holds the referential intention about the identity of the referent. In a default case, indefinites are anchored to the speaker of the utterance. However, they can also be anchored to some other salient discourse referent, such as the subject of the sentence or other (implicit) referents. (For more on the notion of salience or prominence in discourse, see von Heusinger & Schumacher 2019.) We use this notion of speaker-oriented specificity vs. non-speaker-oriented specificity to account for the differences between the two specific indefinite constructions in Modern Colloquial Persian. That is, we will draw parallels between the two specific indefinite constructions in Modern Colloquial Persian and the contrast found for the German specificity adjectives *ein gewisser* vs. *ein bestimmter*.

3 Types of indefinites in Persian

Persian is a language with two dominant registers, spoken and written Persian, both of which have informal and formal forms that are very distinct (Jasbi 2014; Lazard 1957; 1992; Modarresi 2018; Nikravan 2014; Windfuhr 1979). The language that we investigate in this paper is Standard Modern Colloquial Persian. The function of the indefinite marker varies with register; the specificity marker *-e* is only used in Modern Colloquial Persian. In this section, we provide a brief overview of the way definiteness is expressed, the different indefinite forms in Standard Written Persian, and the use and function of indefinite forms in Modern Colloquial Persian.

3.1 Definiteness in Persian

Persian does not have a definite article, but it has two markers for indefiniteness (see the next section). To express definiteness, then, Persian typically uses bare noun phrases. This holds for different kinds of definite noun phrases. The definite in (14a) is a familiar definite, (14b) is a typical bridging definite, (14c) shows a unique definite, and (14d) is an example of generic use.

(14)   a.   Anne yek xune xarid. **Xune** labe marze kešvare.
             Anne a   house bought **house** on   edge.of country.be.3SG
             'Anne bought a house. The house is at the border of the country.'

       b.   Anne rafte   bud   ye marāseme arusi.   **Arus** xeyli
             Anne went.3SG AUX.3SG a   ceremony marriage **bride** very

             xošgel   bud.
             beautiful was.3SG
             'Anne went to a wedding. The bride was very beautiful.'

       c.   **Māh**   xeyli rošan mideraxše.
             moon very bright PROG.shine.3SG
             'The moon shines very brightly.'

       d.   **Dianāsor** 60 milion sāle qabl monqarez šode.
             dinosaur 60 million year ago extinct   became.3SG
             'Dinosaurs became extinct 60 million years ago.'

There is controversy among scholars as to whether, in Persian, bare nouns are inherently definite (Krifka & Modarresi 2016), or underspecified with respect to definiteness and genericity (Ghomeshi 2003). Although it is not clear whether or not the non-specific indefinite nature of bare nouns can be detached from their generic (kind) reading, Dayal (2017) argues, using Hindi as an example, against the view that bare nouns are ambiguous and can have either a definite or an indefinite reading. She concludes that bare singulars in articleless languages like Hindi are definite and not indefinite (specific/non-specific), and that their apparent indefiniteness is construction-specific or restricted to bare plurals. Šimík & Burianová (2020) claim that in Czech, bare NPs, where they are indefinite, cannot be specific. Rather, bare NPs are either definite or indefinite non-specific, which is in line with Dayal's argument. Šimík & Burianová (2020), finally, annotate bare nouns for (in)definiteness, and their findings suggest that the definiteness of a bare noun is affected by its absolute position in the clause, and that indefinite bare NPs are unlikely to occur in clause-initial position (see also Borik

et al. 2020 [this volume]). Note that this is also applicable to Persian: Persian bare nouns can express a non-definite reading, as in (15a) with a kind-reading of 'book', or a definite-reading of 'book' as in (15b). Note that a bare noun in the pre-verbal direct object position is typically interpreted as pseudo-incorporated (in the sense of Massam 2001) as in (15c), while a definite reading must be signaled by the object marker *-rā* as in (15d) (see Modarresi 2014 for an analysis of bare direct objects in Persian):

(15)   a. Roo miz   ketābe.
         on   table book.be.3sɢ
         'There is a/some book on the table.'

       b. Ketāb roo mize.
          book   on   table.be.3sɢ
          'The book is on the table.'

       c. Ali ketāb xarid.
          Ali book bought.3sɢ
          'Ali bought book/books.'

       d. Ali ketāb-rā xarid.
          Ali book-rā bought.3sɢ
          'Ali bought the book.'

3.2 Indefiniteness in Standard Written Persian

Standard Written Persian has the suffixed[3] indefinite marker *-i*, which has quite a large range of functions, and the independent lexeme *ye(k)*, which derives from the numeral *yek*, but behaves like a regular indefinite article.[4] Both forms can be combined, yielding three different indefinite configurations: *ye(k) NP*, *NP-i*, and *ye(k) NP-i*. For Standard Written Persian, the suffixed *-i* has indefinite readings,

---

[3]There is some controversy as to whether *-i* is suffixed or enclitic. Herein we follow the works of Ghomeshi (2003); Hincha (1961); Karimi (2003); Paul (2008). This does not affect our analysis in any way.

[4]Ghomeshi (2003: 64-65) shows that the indefinite article *ye(k)* is different from the numeral *yek*. The former can appear without a classifier (i), which is obligatory for numerals, as in (ii) (see also Bisang & Quang 2020 [this volume] for Vietnamese), and the indefinite article can also appear with plurals, as in (iii).

   (i) ye-(ta) ketab   (ii) se-*(ta)   ketab   (iii) ye ketab-ha-i
        a-(cʟ) book   three-*(cʟ) book   a book-pʟ-ɪɴᴅ
        'a book'   'three books'   'some books'

including readings that undergo negation and other operators. The use of *ye(k)* is thought to express the typical "cardinal" reading of indefinites. There is no clear delimitation of the function of the combined form *ye(k) NP-i*.

Windfuhr (1979) considers *NP-i* to have three functions: (i) as *-i* of 'unit', the construction has similar functions as *a(an)* in English; (ii) as *-i* of indefiniteness, the construction is very similar to what Jasbi (2016) describes as 'antidefinite', similar to 'any' or 'some'; (iii) as demonstrative *-i*, the construction appears with relative clauses.[5] Toosarvandani and Nasser (2017) report that some traditional (Lambton 1953) as well as contemporary linguists (Ghomeshi 2003) assume that the indefinite determiner *yek+NP* and the suffixed *NP-i* can be equivalent in positive, assertive contexts, see example (16) (mainly in non-contemporary or more literary usages); however, Toosarvandani and Nasser (2017) provide examples that show a difference in distribution and meaning between the two constructions, mainly in negative, non-assertive contexts, see examples (17) and (18). In the following, the two indefinites' similarities and differences are discussed.

Since *-i* is a suffix, it can occur with quantifiers. In fact, when universal quantifiers such as *har* ('every') and *hich* ('no') are present, the suffixed *-i* usually accompanies the NP. Lyons (1999: 90) states that the "suffix *-i* semantically marks the noun phrase as non-specific or arbitrary in reference and is approximately equivalent to *any* in nonassertive contexts and *some…or other* in positive declarative contexts". Ghomeshi (2003: 64-65) argues that the two forms partly overlap, but that the suffixed *-i* has a wider range of application. She does not discuss the combined form, however. Paul (2008: 325) argues that *-i* has the function of "picking out and individuating entities". He argues that this function should be kept separate from specificity and referentiality. Hincha (1961: 169-170) assumes that *ye(k)* expresses an individualized entity, while *-i* signals an arbitrarily chosen element of a class. Modarresi (2014: 16-19) focuses on the differences between bare nouns in an object position, and *ye(k) NP* and *NP-i* objects. The latter both introduce discourse referents and show scopal effects, while the bare noun does not. We cannot do justice to the whole discussion on indefinites in written Persian, but we try to summarize the main, and hopefully uncontroversial, observations in Table 2.

Semantically, *yek NP-i* can express existence and signals that the referent is arbitrarily chosen (Lyons 1999). Pragmatically, it can show a speaker's indifference or ignorance, or a free-choice implication (Jasbi 2016). In written form, the three indefinites behave similarly in positive declarative contexts, as shown in (16a-c).

---

[5]There is an ongoing discussion as to whether the use with relative clauses is a use of the suffixed article or a different morpheme (see discussion in Ghomeshi 2003: 65).

Table 2: Definite and indefinite constructions in Standard Written Persian

Function	Form	Positive context	Negative context
definite	*NP*	uniqueness and familiarity	wide scope
indefinite – expressing 'cardinality' or 'existential entailment'	*ye(k) NP*	cardinal reading, existential 'a/one N'	variation between wide and narrow scope
indefinite, existential	*NP-i*	existential/'one of a class' 'some or other'	NPI-narrow scope-'any'
indefinite – expressing 'cardinality' or 'existential entailment' and speaker's indifference or ignorance	*ye(k) NP-i*	existential/arbitrary in reference/speaker's indifference similar to German: '(irgend) jemand'	wide scope

(16) Context: There were three books. Ali bought one of them.
    a. Ali ketāb-i xarid.
        Ali book-i bought.3sG
        'He bought a/some book.'
    b. Ali yek ketāb xarid.
        Ali a   book bought.3sG
        'He bought a book.'
    c. Ali yek ketāb-i xarid.
        Ali a   book-i bought.3sG
        'Ali bought a/some book.'

Considering negation, *ye(k) NP-i* takes a wide scope over negatives and questions, while *NP-i* takes a narrow scope in the same context, and *ye(k) NP* can take variable scope (Toosarvandani & Nasser 2017: 8-9; Modarresi 2014: 26-30). The acceptability of a wide scope under negation with different indefinites is illustrated in (17) and (18). Context (17) forces a narrow-scope reading for the in-

definites, which is available for *NP-i* in (17a) and *ye(k) NP* in (17b), but not for *ye(k) NP-i* in (17c). The context in (18) strongly suggests a wide-scope reading, which is not available for *NP-i* in (18a), but possible for *ye(k) NP* in (18b), and for *ye(k) NP-i* in (18c). (Note that the wide-scope reading goes hand in hand with the object marker *-rā*.)

(17) Context: There were three possible books I could buy. I didn't buy any of them.

    a. Man **ketāb-i** naxaridam.
        I   **book-i** not.bought.1sG
        'I didn't buy any books.'   ¬ > ∃

    b. Man hattā **ye ketāb** ham naxaridam.
        I   even **a book** also not.bought.1sG
        'I didn't buy any book.'   ¬ > ∃

    c. #Man **ye ketāb-i** ro naxaridam.
        I   **a book-i** rā not.bought.1sG
        Intended: 'I didn't buy any book.'   *¬ > ∃

(18) Context: There were three possible books that I could buy. I bought two of them but not the third.

    a. #Man **ketāb-i** naxaridam.
        I   **book-i** not.bought.1sG
        'There is a book I didn't buy.'   *∃ > ¬

    b. Man **ye ketāb** ro naxaridam.
        I   **a book** rā not.bought.1sG
        'There is a book I didn't buy.'   ∃ > ¬

    c. Man **ye ketāb-i** ro naxaridam.
        I   a **book-i** rā not.bought.1sG
        Intended: 'There is a book I didn't buy.' [6]   ∃ > ¬

As shown in (18a), *NP-i* takes wide scope neither under negation nor under questions (similar to NPIs). However, in positive contexts (written form), it behaves similarly to simple indefinites and can have an existential or numerical implication.

---

[6](18c) is felicitous in the written variety with DOM 'rā' whereas it is not felicitous in Modern Colloquial Persian.

3.3 Indefiniteness marking in Modern Colloquial Persian

One of the main distinctions between the system of indefinite forms in the written vs. spoken register is the semantic role of suffixed *-i*. In the written register, *-i* is a common way of marking an indefinite NP, whereas in colloquial Persian, *yek NP* is common and *-i* is very restricted as it is used as an NPI. Jasbi (2016: 246) categorizes the indefinite markers in his native Tehrani colloquial Persian into three main categories: simple, complex, and antidefinite. He illustrates their difference in the following table:

Table 3: Definite and indefinite constructions in Modern Colloquial Persian (Jasbi 2016: 246)

	Type	Form	Example	Translation
Definites	Bare	*NP*	*māshin*	the car
Indefinites	Simple	*ye NP*	*ye māshin*	a car
	Antidefinite	*NP-i*	*māshin-i*	~ a/any car
	Complex	*ye NP-i*	*ye-māshin-i*	~ some car or other

Jasbi calls *ye(k) NP* a simple indefinite because it behaves similarly to *a(n)* in English and carries an existential inference ($|\llbracket NP \rrbracket| \geq 1$). On the other hand, *NP-i* entails an antidefinite interpretation, meaning that it rejects any set that can have a unique inference ($|\llbracket NP \rrbracket| \neq 1$) and can have a non-existential implication ($|\llbracket NP \rrbracket| = 0$). Therefore, the respective set either is empty or contains more than one element. Now, the complex indefinite *ye(k) NP-i* has an anti-singleton implication ($|\llbracket NP \rrbracket| > 1$), which is compositionally derived from the existential inference and the anti-uniqueness condition. The summary of the semantic differences proposed by Jasbi (2016: 251) is provided in Table 4.

Table 4: Cardinality implications for definites and indefinites in Modern Colloquial Persian (Jasbi 2016: 251)

	Type	Form	Cardinality
Definite	Bare	*NP*	$|\llbracket NP \rrbracket| = 1$
Indefinite	Simple	*ye NP*	$|\llbracket NP \rrbracket| \geq 1$
	Antidefinite	*NP-i*	$|\llbracket NP \rrbracket| \neq 1$
	Complex	*ye NP-i*	$|\llbracket NP \rrbracket| > 1$

To summarize, the function of the different indefinite markers in Standard Written Persian is controversial, and their function in Standard Colloquial Persian requires more investigation. Based on Jasbi's (2016) semantic characterization (see Table 4) and the examples discussed above as well as in the subsequent sections, we assume that the form *ye(k) NP* corresponds to the unmarked indefinite, the form *NP-i* only appears with negation, in conditionals, and in questions, and the combined form *ye(k) NP-i* expresses a speaker's ignorance or indifference.

4 The specificity marker *-e* in Modern Colloquial Persian

Modern Colloquial Persian has the suffix *-e*, which can optionally combine with bare, i. e., definite, noun phrases, demonstrative noun phrases, and indefinite noun phrases. With bare noun phrases, *-e* is assumed to express a demonstrative or definite function (Windfuhr 1979: 40; Lazard 1957: 163; Ghomeshi 2003: 67; Toosarvandani & Nasser 2017; Jasbi 2020a). Hincha (1961: 173-177) summarizes the distributional properties of *-e*: it is always optional — there are no conditions that makes its use obligatory. If used, it is always accented and attached directly to the stem. It stands in opposite distribution to the plural suffix *-hā*, i. e., either *-hā* or *-e* can be used, but not both, which leads Hincha (1961: 175) to assume that both suffixes share some features and express some contradictory features, such as number. Ghomeshi (2003: 68) adds that *-e* "cannot attach to anything already of category D", such as proper names, pronouns, and noun phrases containing possessors. It cannot combine with the suffixed marker *-i*, but as we will discuss below, it can combine with the complex *ye(k) NP-i*. With indefinite noun phrases, the suffix signals specificity. In the following, we first provide an overview of specific definite contexts that license the use of the marker, and then provide data on the possible combination of the marker with indefinite constructions.

4.1 Specificity marker *-e* with definites

Modern Colloquial Persian can express (certain kinds of) definiteness by means of the marker *-e*, which is absent in Standard Written Persian (Windfuhr 1979: 50; Ghomeshi 2003). The function of *-e* is described as demonstrative, definite, determinative, or referential. Hincha (1961: 176) assumes that *-e* signals that the NP refers to one particular or individualized entity ("Einzelgegenstand"). There is no comprehensive study of this marker.

There is an interesting distribution of *-e* with the unmarked bare noun. Nikra-van (2014) argues that there is a functional difference between unmarked noun phrases, on the one hand, and noun phrases marked with *-e* on the other. The former express weak definiteness and the latter strong definiteness, as is found in other languages with two definite articles (see Schwarz 2013). Strong forms are used in anaphoric and situational contexts, while weak forms appear in en-cyclopedic, unique, and generic contexts. This contrast is illustrated in (19).

(19) Emruz yek pesar va   yek doxtar ro didam.   Pesar??(-e) ro mišnāxtam.
      today a   boy   and a   girl   rā saw.1sg boy??(-e) rā knew.1sg
      'Today I saw a boy and a girl. I knew the boy.'

In (19), *pesar* 'boy' is anaphoric and much more acceptable with the marker *-e* than without it. Consequently, it is argued that in contexts where an explicit antecedent is present, the strong definite is used. Other scholars propose that *-e* marks familiarity of the associated referent (Hedberg et al. 2009). The results of a questionnaire presented in Nikravan (2014) seem to indicate that there is a marginal effect of *-e* towards a familiarity reading. However, it is unclear from her presentation whether the effect is statistically reliable or not. The results also show that the use of *-e* is optional, as in (19).

The use of *-e* with different types of definite noun phrases (see (14) above) provides further evidence that (i) the use of *-e* is optional and (ii) *-e* can only be used with referential definites, i. e., anaphorically used definites, as in (20a), and definites in bridging contexts, as in (20b). The use of *-e* is ungrammatical for unique definites, as in (20c), and generic uses, as in (20d).

(20) a. Anne yek xune xaride.   **Xune(-he)** labe marze kešvare.
          Anne a   house bought.3sg **house(-he)** on   edge.of country.be.3sg
          'Anne bought a house. The house is at the border of the country.'
      b. Anne rafte   bud   ye marāseme arusi.
          Anne went.3sg aux.3sg a   ceremony marriage

          Arus(-e) xeyli xošgel   bud.
          bride(-e) very beautiful was.3sg
          'Anne went to a wedding. The bride was very beautiful.'
      c. **Māh(*-e)**   xeyli rošan mideraxše.
          moon(*-e) very bright prog.shine.3sg
          'The moon shines very brightly.'

    d. **Dianāsor(*-e)** 60 milion sāle qabl monqarez šode.
        dinosaur(*-e) 60 million year ago extinct   became.3sɢ
        'Dinosaurs became extinct 60 million years ago.'

The referential function of *-e* can also be shown in the contrast between a referential and an attributive reading of a definite NP (Donnellan 1966; Keenan & Ebert 1973). Sentence (21) strongly suggests an attributive or non-referential reading of the noun *barande-ye* 'the winner, whoever the winner will be'. In this reading, the use of *-e* is ungrammatical, which confirms the assumption that *-e* signals referentiality, in the sense that the hearer, as well as the speaker, can uniquely identify the referent of the noun phrase.

(21) **Barandeye(*-e)** in mosābeqe yek safar be ālmān   migirad.
      winner(*-e).of this competition a   trip   to Germany get.3sɢ
      'The winner of this competition (whoever he/she is) will get a trip to Germany.'

Therefore, we can conclude that the function of *-e* is to mark referentially strong definites, i. e., definites that refer to a discourse referent that was explicitly or implicitly introduced into the linguistic context.

4.2 The suffix *-e* with indefinites

The specificity marker *-e* can combine with two of the three indefinite configurations, as in the examples (6)-(8) above, repeated here as (22)-(24).

(22) **Ye(k) pesar-e** injā kār   mikone.
      a   **boy-e**   here work do.3sɢ
      'A (specific) boy works here.'

(23) Diruz   **māšin-i** jolo-e   dare   xune didi?
      yesterday **car-i**   front.of-ɛ door.of home saw.2sɢ

      – *Na, man hič **māšin-e-i** nadidam.
         no I   any **car-e-i**   not.saw.1sɢ
      'Did you see any cars in front of the house door yesterday?'
      Intended: 'No, I didn't see any specific car.'

(24) Emruz **ye māšin-e-i** az   pošt   behem zad.
      Today **a** **car-e-i**   from behind to.me smashed.3sɢ
      'Today some/other (specific) car smashed me from behind.'

The form *NP-i* cannot combine with *-e*. We speculate that this is due to a conflict of the referential meaning of *-e* and the NPI-meaning of *NP-i* in Modern Colloquial Persian.[7]

However, both forms with the indefinite article *ye(k)* can combine with *-e*, yielding *ye(k) NP-e* and *ye(k) NP-e-i*, respectively. With both indefinite constructions, the marker *-e* signals referential and wide-scope readings of the indefinites. The regular indefinites *ye doxtar* in (25a) and *ye doxtar-i* in (25c) allow for (i) a wide-scope and (ii) a narrow-scope reading with respect to the universal quantifier. However, the forms *ye doxtar-e* in (25b) and *ye doxtar-e-i* in (25d) only allow for a wide-scope, referential, or specific reading. We find the same contrast for indefinites in sentences with verbs of propositional attitudes, as in (26). The *-e* marked indefinites can only take a wide scope with respect to the intensional operator *mixad* 'to want'.

(25) a. Hame pesar-hā bā   ye doxtar raqsidan.
          all   boy-PL with a girl   danced.3PL
          (i) 'There is a girl such that every boy danced with her.'
          (ii) 'For every boy, there is a different girl, such that, that boy dances with her.'

      b. Hame pesar-hā bā   ye doxtar-e raqsidan.
          all   boy-PL with a girl-e   danced.3PL
          (i) 'There is a girl such that every boy danced with her.'

      c. Hame pesar-hā bā   ye doxtar-i raqsidan.
          all   boy-PL with a girl-i   danced.3PL
          (i) 'There is a girl such that every boy danced with her.'
          (ii) 'For every boy, there is a different girl, such that, that boy dances with her.'

      d. Hame pesar-hā bā   ye doxtar-e-i raqsidan.
          all   boy-PL with a girl-e-i   danced.3PL
          (i) 'There is a girl such that every boy danced with her.'

(26) a. Ali mixād   bā   ye doxtar dust   še.
          Ali want.3SG with a girl   friend become.3SG
          (i) 'Ali wants to make friends with a specific girl.'
          (ii) 'Ali wants to make friends with a girl/whoever she may be.'

---

[7]The occurrence of *-e* with *NP-i* is not possible with restrictive relative clauses (Ghomeshi 2003).

b. Ali mixād   bā   ye doxtar-e dust   še.
   Ali want.3SG with a girl   friend become.3SG
   (i) 'Ali wants to make friends with a specific girl.'

c. Ali mixād   bā   ye doxtar-i dust   še.
   Ali want.3SG with a girl-i   friend become.3SG
   (i) 'Ali wants to make friends with a specific girl.'
   (ii) 'Ali wants to make friends with a girl/whoever she may be.'

d. Ali mixād   bā   ye doxtar-e-i dust   še.
   Ali want.3SG with a girl-e-i   friend become.3SG
   (i) 'Ali wants to make friends with a specific girl.'

We take the distribution of *-e* discussed here as a good evidence that the marker encodes a specific or referential reading of the indefinite.[8]

4.3 Specificity marker and referential anchoring

Epistemic specific indefinites express the "referential intention" of the speaker. That is, the speaker signals with these expressions that he or she has already decided on the referent of the indefinite. Non-specific indefinites, on the other hand, assert the existence of an individual that falls under the descriptive content of the indefinite, but not a particular individual. The concept of epistemic specificity as speaker-oriented (or speaker-anchored) seems too narrow, however, as we also find (epistemic) specific indefinites where the speaker cannot identify the referent, but can recognize some other salient discourse participant. Therefore, von Heusinger (2002; 2019) proposes the concept of "referential anchoring", modeling the dependency of the referent of the indefinite from some other salient discourse referent or participant (typically the speaker, the subject

---

[8]Here we leave open what the exact semantics of the marker *-e* is. Hincha (1961: 176) describes it as "punctualization"; Jasbi (2020b) assumes that the marker *-e* creates a singleton set, thereby simulating wide-scope behavior. However, this approach would not explain why it can be used with certain definites and why it can be combined with the complex form *ye(k) NP-e-i*, as it would include the combination of a singleton and an anti-uniqueness condition. An alternative approach is to assume that the marker is interpreted as an indexed choice function (Egli & von Heusinger 1995; Winter 1997) that selects one element out of a set. This would explain the use with certain definites, and also the complementary distribution with the plural suffix *-hā*. Such an account could provide an explanation for the definiteness effect on bare nouns. (The value for the index of the choice function is provided by the local situation or the local discourse, but not by encyclopedic knowledge.) In the form *ye(k) NP-e*, the index is locally bound by the speaker, and for the form *ye(k) NP-e-i*, the index can also be bound by other salient discourse referents.

of the sentence, etc.). The discussion of the contrast between the specificity adjectives *ein gewisser* and *ein bestimmter* in §2.2 was explained along these lines: *ein gewisser* is speaker-oriented, while *ein bestimmter* is not obligatorily speaker-oriented, i. e., it can also be anchored to another salient agent in the discourse.

The two indefinite forms *ye(k) NP-e* and *ye(k) NP-e-i* are interpreted as specific or referential indefinites. We suggest that the difference between the two forms lies in the specificity orientation in epistemic contexts. It seems that the form *ye(k) NP-e-i* is less acceptable in general; however, we still find examples such as the following on Twitter:[9]

(27) Tanhāi vasate   pārke   Mellat nešastam, ye xānum-e-i dāre
      alone   middle.of park.of Mellat sitting.1SG a   woman-e-i AUX.3SG
      kenāram Qurān mixune.
      next.me Quran reading.3SG
      'I am sitting alone in the middle of Mellat Park and some woman is reading the Quran next to me.'

(28) Ye doxtar-e-i tu bašgāham hast,   ajab heykali   dāre!
      a girl-e-i   at gym.my   be.3SG what body-shape have.3SG
      'There is some girl at the gym I go to, that has a perfect body-shape!'

(29) Zošk ke   berim, ye sag-e-i ham hast   unjā, tāze   āšnā   šode
      Zoshk that go.1PL a   dog-e-i also be.3SG there recently familiar got.3SG
      bāhām.
      with.me
      'If we go to Zoshk, there is some dog there that made friends with me last time.'

(30) Ye dars   dāštam be   esme "Riāzi Pišrafte".   Unjā ye pesar-e-i
      a   course had.1SG with name "Math advanced" there a   boy-e-i
      bud   be   esme   Vahid ya Hamid.
      was.3SG with name.of Vahid or Hamid
      'I had a course called "Advanced Mathematics". There was some boy named Vahid or Hamid.'

---

[9]The first anonymous reviewer pointed out that all the Twitter examples (27)-(30) are speaker-oriented and would therefore contradict our hypothesis that the form *ye(k) NP-e-i* is non-speaker-oriented. We think that it is difficult to judge this without more context. Moreover, we believe that, in most of the examples, the speaker signals that he or she is not able or willing to reveal the identity of the indefinite NP. The main point of the Twitter examples is to show that these forms are in current use, which contradicts some assumptions made in the literature.

We propose that the basic function of the suffixed indefinite article *-i* in Modern Colloquial Persian is to signal speaker ignorance or indifference. Combining speaker ignorance with the epistemic specificity or referentiality might result in a semantic-pragmatic condition which we have termed non-speaker-oriented specificity (see discussion in §2.2 above). Therefore, we hypothesize that the difference between these two forms is the orientation or anchoring of the specificity relation. For *ye(k) NP-e*, we assume that the indefinite is referentially anchored to the speaker, i. e., the indefinite is speaker-oriented specific. The form *ye(k) NP-e-i*, in contrast, is referentially anchored to a discourse referent other than the speaker, i. e., it is non-speaker-oriented. We summarize this hypothesis in Table 5.[10]

Table 5: Specificity marker *-e* with different indefinite markers in Modern Colloquial Persian

Indefinite form	Cardinality (Jasbi 2016)	Pragmatic difference	Combination with specificity marker *-e*	Assumed function
ye(k)+NP	$\|\llbracket \text{NP} \rrbracket\| > 1$	normal indefinite marker like *a(n)*	acceptable	speaker-oriented specificity
NP-i	$\|\llbracket \text{NP} \rrbracket\| \neq 1$	with negations, conditionals, and questions	ungrammatical	–
ye(k)+NP-i	$\|\llbracket \text{NP} \rrbracket\| > 1$	speaker's indifference or ignorance is more likely	less acceptable	non-speaker -oriented specificity

Our hypothesis makes clear predictions about the acceptability of sentences containing these forms in contexts that express a speaker orientation vs. a non-speaker orientation of the indefinite. We assume that the indefinite *ye ostād-e* expresses a speaker orientation, which predicts that the continuation (31i) is coherent, while the continuation (31ii) is incoherent. For the indefinite *yek ostād-e-i* in sentence (32), we assume a non-speaker-orientation, which predicts that continuation (32i) is not felicitous, while continuation (32ii) is.

---

[10]Our second reviewer asks whether we assume a compositional semantics, which would provide an independent function for each marker, or whether we assume just one function for the whole construction. For a compositional approach, see Jasbi (2016) for the indefinite forms, and footnote 8 in this chapter, on the choice function approach to the specificity marker *-e*. However, we have not yet developed full semantics for these configurations.

(31) Sārā **ye ostād-e**   **ro** xeyli dust dāre.
      Sara **a   professor-e rā** very like have
      'Sara likes a specific professor very much.'

   (i) Man midunam kudum ostād.
        I   know.1SG which professor
        'I know who he is.'

   (ii) #Vali nemidunam   kudum ostād.
         but not.know.1SG which   professor
        'But I don't know which professor.'

(32) Sārā **ye ostād-e-i**   **ro** xeyli dust dāre.
      Sara **a   professor-e-i rā** very like have
      'Sara likes some specific professor very much.'

   (i) #Man midunam kudum ostād.
         I   know.1SG which professor
         'I know who he is.'

   (ii) Vali nemidunam   kudum ostād.
        but not.know.1SG which   professor
        'But I don't know which professor.'

We can summarize this prediction in Table 6 with the expected acceptability of the continuation.[11]

---

[11]The second reviewer also suggested that we test the examples (31)-(32) without the specificity marker *-e*, as in (31′) and (32′). The reviewer reported that his or her informants would accept the continuations (i) and (ii) for both sentences, but that the informants expressed a preference for (31′ii) and (32′i), which would be the opposite of the expectation expressed for (31)-(32). We agree that both continuations are good for both sentences, but we do not share their preferences. We do not have any predictions with respect to (31′) and (32′). Note that both (31)/(32), and (31′)/(32′), have the direct object marker *-rā*, which is assumed to express specificity by itself. We cannot go into details about the difference between the function of *-e* and *-rā* here; however, our test items had examples with and without *-rā*.

(31′) Sārā **ye ostād**   **ro** xeyli dust dāre.   (i) Man midunam kudum ostād.
       Sara **a   professor rā** very like have   I   know.1SG which professor
       'Sara likes a specific professor very much.'   'I know who he is.'

(32′) Sārā **ye ostād-i**   **ro** xeyli dust dāre.   (ii) Vali nemidunam   kudum ostād.
       Sara **a   professor-i rā** very like have   but not.know.1SG which professor
       'Sara likes some specific professor very much.'   'But I don't know which professor.'

Table 6: Prediction of type of epistemic specificity of *-e* marked indefinites

Indefinite form	Epistemic type	Example	Context	Predited acceptability
ye(k) NP-e	speaker-specific	(31)	(i) speaker knowledge	very good
ye(k) NP-e	speaker-specific	(31)	(ii) speaker ignorance	#
ye(k) NP-e-i	non-speaker-specific	(32)	(i) speaker knowledge	#
ye(k) NP-e-i	non-speaker-specific	(32)	(ii) speaker ignorance	good

5 Empirical evidence for speaker orientation of specific noun phrases

In this section, we present two pilot acceptability studies that tested the predictions outlined above. In the first pilot, we used eight sentences, which we continued with either (i), a context that was only coherent with a speaker-oriented specific reading or (ii), a context that was only coherent with a non-speaker-oriented specific reading. The results show that simple indefinites with *ye(k) NP-e*, regardless of their specificity orientation, are more acceptable than complex indefinites, but there were no clear effects of specificity orientation. We assume that our results might reflect a mix-up between different degrees of animacy in the included indefinites. Therefore, we conducted a second pilot study with only human indefinites and a different design; as well as simple sentences and their speaker-oriented vs. non-speaker-oriented continuations, we also presented sentences that clearly signaled speaker ignorance in order to test whether informants can distinguish between different specificity orientations. The results of the second study not only confirm that speakers are capable of making this distinction, but also provide some support for our claim that the simple indefinite *ye(k) NP-e* is speaker-oriented, and the complex indefinite *ye(k) NP-e-i* is non-speaker-oriented.

5.1 Experiment 1

Our hypothesis H1 is that in Modern Colloquial Persian *ye(k) NP-e* always functions as speaker-specific ('gewiss NP'), while *ye(k) NP-e-i* can only function as non-speaker-oriented. In order to test this hypothesis, we conducted a pilot questionnaire with speakers of Modern Colloquial Persian. We used a simple sentence,

as seen in (33), with simple indefinites with the marker *-e* (*yek doktor-e*), as well as complex indefinites with the marker *-e* (*yek doktor-e-i*). The first sentence with the critical item (*yek doktor-e* or *yek doktor-e-i*) is continued with either (i) an assertion that the speaker had knowledge of the referent, or (ii) a statement signaling the ignorance of the speaker. That is, continuation (i) strongly forces a speaker-specific reading and continuation, while (ii) forces a non-speaker-specific reading. Note that we did not test indefinites without the marker *-e*, as we assume that there is ambiguity between a specific and non-specific interpretation.[12]

(33)   a. Mona bā   **yek doktor-e** ezdevaj   karde.
           Mona with **a**   **doctor-e** marriage did.3sg
           'Mona married a doctor.'

       b. Mona bā   **yek doktor-e-i** ezdevaj   karde.
           Mona with **a**   **doctor-e-i** marriage did.3sg
           'Mona married a doctor.'

           (i) Man midunam kudum doktor.
                I   know.1sg which doctor
                'I know which doctor he is.'

           (ii) Vali man nemidunam   kudum doktor.
                but I   not.know.1sg which doctor
                'But I do not know which doctor he is.'

5.1.1 Participants and experimental technique

Twenty male and female participants participated in the study. Their native language was Persian and they had lived all or most of their lives in Iran. Their ages varied between 25 and 67. In terms of educational level, six participants had high school diplomas, ten had bachelor's degrees, and four had master's degrees. Participants read Persian written texts for at least one hour a day, and they spoke/heard Persian all or most of the day.

The study followed a 2x2 design with two different indefinite forms: (a) *ye(k) NP-e* and (b) *ye(k) NP-e-i* and two continuations: (i) "I do know who/which" for

---

[12]In half of the examples the critical indefinite was the direct object, as in (31), and a different argument in the other half, as in (33). We found that this alternation had no significant effect, even though we added the differential case marker *-rā* in the direct object instances. It is unclear what additional function this marker performs (see the discussion in the last footnote). We also balanced for animacy, see the discussion below and Figure 2.

the speaker-oriented epistemic specificity and (ii) "I do not know who/which" for the non-speaker-oriented epistemic specificity. The assumption was that all forms are epistemically specific, as in Table 6 above. We had eight different sentences and created four lists using a Latin square design, so that each participant heard one sentence and two conditions each. Probable factors which might intervene with the evaluation, such as animacy, position of NP in the sentence (direct object/indirect object), and direct/indirect speech, were equally present in all items.

As we were testing Modern Colloquial Persian, i. e., spoken Persian, we read out the sentences to our participants at least once and asked them to evaluate the sentence on a scale from 1 for "completely acceptable" to 7 for "completely unacceptable" on the answer sheet, where they were also able to read the test sentence themselves.

5.1.2 Results

We observed that participants complained (even verbally) about the appearance of *-e* in *ye(k) NP-e-i* in both speaker-specific and non-speaker-specific readings. This is also reflected in the acceptability scores. We summarize the pilot questionnaire with 20 participants in Table 7, together with the expected acceptability.

Table 7: Effect of *-e* as specificity marker of indefinites on the kind of epistemicity (1 = very good; 7 = very bad)

Indefinite form	Epistemic type	Mean value	Acceptability	Predicted acceptability
ye(k) NP-e	speaker-specific	2.725	good	very good
ye(k) NP-e	non-speaker-specific	3.025	good	#
ye(k) NP-e-i	speaker-specific	4.675	weak	#
ye(k) NP-e-i	non-speaker-specific	4.425	weak	good

Overall, we see that the form *ye(k) NP-e* was more acceptable than the form *ye(k) NP-e-i*, which confirms the intuition reported above. However, we also see that *ye(k) NP-e* performed well in both conditions (speaker- and non-speaker-specific), which went against our hypothesis. The judgment for the non-speaker-specificity condition is marginally weaker. The form *ye(k) NP-e-i* was clearly weak-er; however, there is only a marginal difference between speaker-oriented
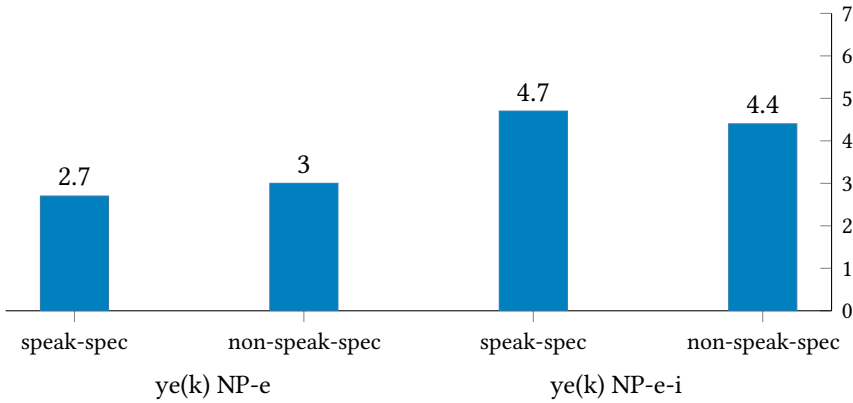
Figure 1: Acceptability (1 = very good; 7 = very bad) of simple and complex indefinites in non-speaker and speaker-oriented specificity contexts
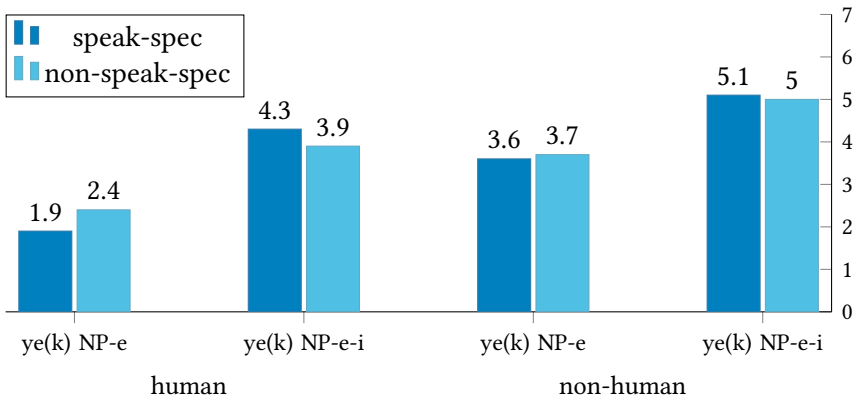

Figure 2: Acceptability (1 = very good; 7 = very bad) of simple and complex indefinites for human and non-human noun phrases in non-speaker and speaker-oriented specificity contexts

specificity (slightly weaker) and non-speaker-oriented specificity. Interestingly, when distinguishing between human and non-human indefinites, as in Figure 2, we see that the non-human indefinites were less acceptable than the human indefinites. Furthermore, when looking at the human indefinites we can see that the simple indefinites (*ye(k) NP-e*) were rated as slightly better in the speaker-specificity condition than in the non-speaker conditions (1.85 vs. 2.35). Complex indefinites (*ye(k) NP-e-i*), on the other hand, were slightly better in the non-speaker-specificity condition than in the speaker-oriented specificity condition (3.9 vs. 4.3).

5.1.3 Discussion

Our first pilot study shows that complex indefinites with the marker *-e* are less acceptable than simple indefinites with the marker *-e*. Animacy is also an important factor: our study demonstrates that human indefinites were more acceptable than non-human indefinites. However, the predicted contrast between simple and complex indefinites in speaker- vs. non-speaker-oriented specificity contexts was not shown to be significant. We surmise that this contrast might be more pronounced with human indefinites, which led us to design a second pilot experiment.

5.2 Experiment 2

In order to test the hypothesis that the two specific indefinites in Modern Colloquial Persian differ with respect to the referential anchoring of the indefinite, i. e., in the specificity orientation, in the second study, we focused on human indefinites. Additionally, we included some examples that provided contexts that signaled speaker ignorance in the first sentence. These examples were used to test whether participants were sensitive to the speaker- vs. non-speaker-orientation.

5.2.1 Design

Experiment 2 was conducted to test for a feature which is only present in spoken colloquial Persian, namely the *-e* marker with indefinite NPs. It followed the same 2x2 design with four lists as the first pilot study. There were 24 items consisting of 12 test and 12 filler items in each list. The experimental stimuli consisted of two sentences for each item. Since the feature under investigation was simple vs. complex indefinites with the marker *-e*, the first sentence contained an indefinite noun phrase either with *yek NP-e* or *yek NP-e-i*. The second sentence forced either a speaker-specific reading of the indefinite in the first sentence, or a non-speaker-specific reading.

In the speaker-specific continuation, we asserted the knowledge of the speaker about the identity of the referent of the indefinite. In the non-speaker-specific continuation, we asserted the ignorance of the speaker about the identity of the referent, thus forcing a non-speaker-specific reading.

(34)   Critical items for Experiment 2

      a. Simple indefinite (**yek NP-e**) + speaker-specific continuation

           Sara emruz az **ye vakil-e** vaqte mošāvere gerefte.
           Sara today from **a lawyer-e** appointment consulting took.3sɢ

Man ham ba   vakil-e   čandinbar   kar   kardam.
I   also with lawyer-e several.time work did.1sɢ

Kareš   xeyli xube.
work.his very good.be.3sɢ

'Sara had a consultation appointment with a lawyer today. I have also consulted with the lawyer. His work is very good.'

b. Simple indefinite (*yek NP-e*) + non-speaker-specific continuation

Sara emruz az   **ye vakil-e**   vaqte   mošāvere gerefte.
Sara today from **a   lawyer-e** appointment consulting took.3sɢ

Migan vakil-e   maroofe   vali man čiz-i   azaš
say.3ᴘʟ lawyer-e known.be.3sɢ but I   thing-ɪɴᴅᴇꜰ from.him
nemidunam.
not.know.1sɢ

'Sara had a consultation appointment with a lawyer today. They say that this lawyer is well known, but I do not know anything about him.'

c. Complex indefinite (*yek NP-e-i*) + speaker-specific continuation

Sara emruz az   **ye vakil-e-i**   vaqte   mošāvere gerefte.
Sara today from **a   lawyer-e-i** appointment consulting took.3sɢ

Man ham ba   vakil-e   čandinbār   kār   kardam.
I   also with lawyer-e several.time work did.1sɢ

Kareš   xeyli xube.
work.his very good.be.3sɢ

'Sara had a consultation appointment with a lawyer today. I have also consulted with the lawyer, several times. His work is very good.'

d. Complex indefinite (*yek NP-e-i*) + non-speaker-specific continuation

Sara emruz az   **ye vakil-e-i**   vaqte   mošāvere gerefte.
Sara today from **a   lawyer-e-i** appointment consulting took.3sɢ

Migan vakil-e   maroofe   vali man čiz-i   azaš
Say.3ᴘʟ lawyer-e known.be.3sɢ but I   thing-ɪɴᴅᴇꜰ from.him
nemidunam.
not.know.1sɢ

'Sara had a consultation appointment with a lawyer today. They say that this lawyer is well known, but I do not know anything about him.'

The test items also differed in their constructions: eight items had a third person subject (proper name), as in (34), and four other items were constructions that showed a greater distance from the speaker, namely two items of the type "They say...", as in (35), and two items of the form "I heard...", as in (36). Note that we provide only the a-condition with *ye(k) NP-e* and the specific continuation, as in (34a).

(35)   "They say..." construction:

Migan **ye moalem-e** tu madrese Tizhushān hast   ke   bečeha
say.3PL **a   teacher-e** in school   Tizhushan be.3SG that student.PL

azaš   xeyli mitarsan.   Man ham bāhāsh 4ta dars   daštam va
from.her very frighten.3PL I   also with.him 4CL course had.3SG and

oftadam.
failed

'They say there is a teacher in Tizhushan school that every student is afraid of. I also have had four courses with him and failed them all.'

(36)   "I heard..." construction:

Šenidam **ye pesar-e** hast   tu in   mahale
heard.1SG **a   boy-e** be.3SG in this neighborhood

ke   vase doxtara mozāhemat ijad   mikone.
who for   girl.PL harassment make do.3SG

Man mišnasameš az   vaghti bače bud.
I   know.1SG   him.from when   child was.3SG

'I have heard that there is a boy in this neighborhood who harasses girls. I have known him since he was a child.'

5.2.2 Results of Experiment 2

There was strong agreement in relation to the filler/control items, with marginal differences between participants (< 0.8 points). The results of the test items can be summarized as follows. Firstly, in contrast to Experiment 1, Figure 3 does not show a clear preference for simple indefinites. Rather, both types were rated very similarly. Secondly, we clearly see that the contexts which signaled speaker ignorance ("They say... ", "I heard...") preferred non-speaker-oriented specificity continuations. It shows that participants were aware of this contrast.

A more detailed inspection of the neutral contexts in Figure 3 reveals a slight preference for the simple indefinite *ye(k) NP-e* in speaker-oriented specificity
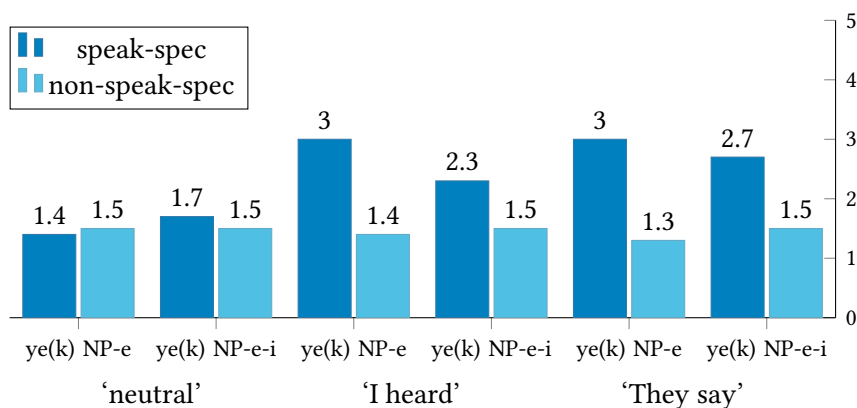
Figure 3: Acceptability (1 = very good; 5 = very bad) of simple and complex indefinites for human noun phrases in non-speaker and speaker-oriented specificity contexts, across types of constructions

contexts (1.39) vs. non-speaker-oriented specificity contexts (1.54), while the complex indefinite *ye(k) NP-e-i* was rated slightly better in non-speaker-oriented specificity contexts (1.5) vs. speaker-oriented specificity contexts (1.67).

In summary, the direct comparison in neutral contexts between the simple and the complex indefinite with the marker *-e* does not provide significant contrasts. It only suggests a preference of the simple indefinite for speaker-oriented specificity, while the complex indefinite prefers non-speaker-oriented specificity. However, constructions with "They say..." or "I heard...,", which clearly encode non-speaker-oriented specificity, show a preference for the complex indefinite. This supports our hypothesis for the difference between the two specific indefinites.

6 Summary and open issues

Persian has two indefinite markers, prenominal *ye(k)* and suffixed *-i*. Both forms express particular kinds of indefiniteness, as does their combination. For Modern Colloquial Persian, indefinites with *-i* express a non-uniqueness or anti-definite implication, and behave similarly to the English *any*. *Ye(k)*, on the other hand, expresses an at-issue existence implication and behaves similarly to the English *a(n)*. Finally, the combination of *ye(k)* and *NP-i* expresses an ignorance implication (Jasbi 2016). The specificity marker *-e* can be combined with *ye(k) NP* and with the combined form *ye(k) NP-i*, but not with (solitary) *NP-i* (Windfuhr 1979;

Ghomeshi 2003). Based on these semantic functions and on the comparison of the two specificity adjectives *ein gewisser* and *ein bestimmter* in German, we hypothesized that the difference between the interpretation of the two indefinites lies in the anchoring of the indefinite either to the speaker or to some other salient discourse referent; the simple indefinite *ye(k) NP-e* is interpreted as a speaker-oriented-specific referent. The complex indefinite *ye(k) NP-e-i* is interpreted as a non-speaker-specific referent.

In two pilot acceptability tasks, we tested these two indefinites in two contexts, one that suggested a speaker-specific interpretation of the indefinite, and a second that suggested a non-speaker-specific interpretation. The first study provided some support for our hypothesis, but we also found that type of indefinite and animacy had a significant effect on interpretation. We therefore designed a second pilot study with only human indefinites. Additionally, we inserted constructions with "I heard..." and "They say...", which clearly suggest a non-speaker-oriented specificity. The results of the second study do not show a preference for the simple indefinite. However, they provide some evidence that, in neutral contexts, the simple indefinite is more acceptable with speaker orientation, and the complex with non-speaker orientation. Still, the evidence is very weak. Finally, in contexts that encode speaker ignorance ("They say...", "I heard..."), the complex indefinite was slightly more acceptable than the simple indefinite, which supports our original hypothesis.

In summary, we have seen that the complex system of indefinite marking in Modern Standard Persian provides a fruitful research environment for learning more about the formal marking of subtle semantic and pragmatic functions of noun phrases, such as specificity and the referential anchoring of nominal expressions.

Acknowledgments

References

Bisang, Walter & Kim Ngoc Quang. 2020. (In)definiteness and Vietnamese classifiers. In Kata Balogh, Anja Latrouite & Robert D. Van Valin, Jr. (eds.), *Nominal anchoring: Specificity, definiteness and article systems across languages*, 15–49. Berlin: Language Science Press. DOI: 10.5281/zenodo.4049679.

Borik, Olga, Joan Borràs-Comes & Daria Seres. 2020. Preverbal (in)definites in Russian: An experimental study. In Kata Balogh, Anja Latrouite & Robert D. Van Valin, Jr. (eds.), *Nominal anchoring: Specificity, definiteness and article systems across languages*, 51–80. Berlin: Language Science Press. DOI: 10.5281/zenodo.4049681.

Brustad, Kristen E. 2000. *The syntax of spoken Arabic*. Washington, D.C.: Georgetown University Press.

Dayal, Veneeta. 2017. Determining (in)definiteness in the absence of articles. In Vera Hohaus & Wanda Rothe (eds.), *Proceedings of TripleA 3. Fieldwork perspectives on the semantics of African, Asian and Austronesian languages*, 85–89. Tübingen: Universitätsbibliothek Tübingen.

Donnellan, Keith S. 1966. Reference and definite descriptions. *The Philosophical Review* 75(3). 281–304.

Dryer, Matthew S. 2005. Indefinite articles. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures*. Oxford: Oxford University Press. https://wals.info/chapter/38.

Ebert, Christian, Cornelia Ebert & Stefan Hinterwimmer. 2013. The interpretation of the German specificity markers *bestimmt* and *gewiss*. In Cornelia Ebert & Stefan Hinterwimmer (eds.), *Different kinds of specificity across languages* (Studies in Linguistics and Philosophy 92), 31–74. Heidelberg/Berlin: Springer.

Egli, Urs & Klaus von Heusinger. 1995. The epsilon operator and e-type pronouns. In Urs Egli, Peter Pause, Christoph Schwarze, Arnim von Stechow & Götz Wienold (eds.), *Lexical knowledge in the organization of language*, 121–141. Amsterdam/Philadelphia: John Benjamins.

Fassi-Fehri, Abdelkader. 2006. A short note on Moroccan specifics. unpublished.

Ghomeshi, Jila. 2003. Plural marking, indefiniteness and the noun phrase. *Studia Linguistica* 57(2). 47–74.

Hedberg, Nancy, Emrah Görgül & Morgan Mameni. 2009. On definiteness and specificity in Turkish and Persian. In Frédéric Mailhot (ed.), *Proceedings of the 2009 annual conference of the Canadian Linguistics Association*, 1–15. Ottawa: Canadian Linguistic Association.

Hincha, Georg. 1961. Beiträge zu einer Morphemlehre des Neupersischen. *Der Islam* 37(1-3). 136–201.

Jasbi, Masoud. 2014. Semantics of differential object marking in Persian. Qualification Paper, University of Stanford.

Jasbi, Masoud. 2016. Three types of indefinites in Persian: Simple, complex, and antidefinite. In Mary Moroney, Carol-Rose Little, Jacob Collard & Dan Burgdorf (eds.), *Proceedings of SALT 26*, 244–263. Linguistic Society of America.

Jasbi, Masoud. 2020a. The meaning of the Persian object marker rā: What it is not, and what it (probably) is. In Richard K. Larson, Sedigheh Moradi & Vida Samiia (eds.), *Advances in Iranian linguistics*, 119–135. Amsterdam: John Benjamins.

Jasbi, Masoud. 2020b. The suffix that makes Farsi nouns unique. In Richard K. Larson, Sedigheh Moradi & Vida Samiia (eds.), *Advances in Iranian linguistics*, 107–118. Amsterdam: John Benjamins.

Karimi, Simin. 2003. On object positions, specificity, and scrambling in Persian. In Simin Karimi (ed.), *Word order and scrambling*, 91–124. Oxford: Blackwell.

Karimi, Simin. 2018. Generative approaches to syntax. In Anousha Sedighi & Pouneh Shabani-Jadidi (eds.), *The Oxford handbook of Persian linguistics*. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780198736745.013.7.

Keenan, Edward L. & Karen Ebert. 1973. A note on marking transparency and opacity. *Linguistic Inquiry* 4(3). 421–424.

Krifka, Manfred & Fereshteh Modarresi. 2016. Number neutrality and anaphoric update of pseudo-incorporated nominals in Persian (and weak definites in English). In Mary Moroney, Carol-Rose Little, Jacob Collard & Dan Burgdorf (eds.), *Proceedings of SALT 26*, 874–891. Linguistic Society of America.

Lambton, Ann K. S. 1953. *Persian grammar*. Cambridge: Cambridge University Press.

Latrouite, Anja & Robert D. Van Valin Jr. 2014. Referentiality and telicity in Lakhota and Tagalog. In Doris Gerland, Christian Horn, Anja Latrouite & Albert Ortmann (eds.), *Meaning and grammar of nouns and verbs*, 401–426. Düsseldorf: Düsseldorf University Press.

Lazard, Gilbert. 1957. *Grammaire du persan contemporain*. Paris: Klincksieck.

Lazard, Gilbert. 1992. *A grammar of contemporary Persian, English translation*. Costa Meca, CA: Mazda Publishers.

Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.

Massam, Diane. 2001. Pseudo noun incorporation in Niuean. *Natural Language and Linguistic Theory* 19(1). 153–197.

Modarresi, Fereshteh. 2014. *Bare nouns in Persian: Interpretation, grammar and prosody.* Berlin: Humboldt-Universität zu Berlin. (Doctoral dissertation).

Modarresi, Yahya. 2018. Sociolinguistics. In Anousha Sedighi & Pouneh Shabani-Jadidi (eds.), *The Oxford handbook of Persian linguistics.* Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780198736745.013.13.

Nikravan, Pegah. 2014. *Semantics and processing of weak and strong definites in colloquial Persian: Evidence from an offline questionnaire.* Ottawa: Carleton University. (MA thesis).

Onea, Edgar & Ljudmila Geist. 2011. Indefinite determiners and referential anchoring. *International Review of Pragmatics* 3(2). 194–227.

Paul, Daniel. 2008. The individuating function of the Persian "indefinite suffix". In Simin Karimi (ed.), *Aspects of Iranian linguistics*, 309–328. Cambridge: Cambridge Scholars Publishing.

Schwarz, Florian. 2013. Two types of definites cross-linguistically. *Language and Linguistics Compass* 7(10). 534–558.

Šimík, Radek & Markéta Burianová. 2020. Definiteness of bare NPs as a function of clausal position: A corpus study of Czech. In Tania Ionin & Jonathan Eric MacDonald (eds.), *Formal Approaches to Slavic Linguistics 26: The Urbana-Champaign meeting 2017*, 343–361. Ann Arbor, MI: Michigan Slavic Publications.

Toosarvandani, Maziar & Hayedeh Nasser. 2017. Quantification in Persian. In Denis Paperno & Edward L. Keenan (eds.), *Handbook of quantifiers in natural language: Volume II*, 665–696. Berlin: Springer.

von Heusinger, Klaus. 2002. Specificity and definiteness in sentence and discourse structure. *Journal of Semantics* 19(3). 245–274.

von Heusinger, Klaus. 2011. Specificity. In Klaus von Heusinger, Claudia Maienborn & Paul Portner (eds.), *Semantics. An international handbook of natural language meaning*, vol. 2 (HSK 33.2), 1025–1058. Berlin: Mouton De Gruyter.

von Heusinger, Klaus. 2019. Indefiniteness and specificity. In Jeanette Gundel & Barbara Abbott (eds.), *The Oxford handbook of reference*, 146–167. Oxford: Oxford University Press.

von Heusinger, Klaus & Petra B. Schumacher. 2019. Discourse prominence: Definition and application. *Journal of Pragmatics* 154. 117–127.

Windfuhr, Gernot L. 1979. *Persian grammar. History and state of its study*. The Hague: Mouton Publishers.

Winter, Yoad. 1997. Choice functions and the scopal semantics of indefinites. *Linguistics and Philosophy* 20(4). 399–467.