Chapter 7

The particular–characterizing contrast in Marathi and its historical basis

Ashwini Deo

The Ohio State University; The University of Texas at Austin

The two-copula and two-auxiliary systems instantiated in a number of Indo-Aryan languages can be best analyzed as lexicalizing particular vs. characterizing meanings. The data is drawn from Marathi [mar, 71,700,000 speakers] but the pattern is observed across different sub-groups of Indo-Aryan, and is possibly inherited from Sanskrit. I provide historical data that indicates that while this contrast is not grammatically categorical in Middle and Old Indo-Aryan, it appears to be present through the interpretational possibilities for the *bhū* copula and its cognates. I use this observed categoricalization of a copular contrast to reflect on whether the overt marking of the particular–characterizing contrast represents a change for the better or for the worse.

1 Introduction

The goal of uncovering systematic principles governing cross-linguistic variation in (the realization of) meaning has been strongly pursued in semantics over the past two decades. One result of this research has been the identification of recurring similarities in the meaning contrasts that get reliably encoded across diverse grammatical systems. Yet another issue that emerges from the same pursuit is that of variation with respect to how a universal inventory of model theoretic components is mapped onto lexical/functional items. Specifically, we find that certain elements of functional meaning that are covert in some languages may find overt realization in others. For instance, languages differ with respect to whether "event-in-progress" and "habitual/generic" meanings are expressed

by distinct aspectual markers (progressive and imperfective) or by a single imperfective marker (see Bybee 1994, Comrie 1976, Deo 2009, 2015, among others). Languages may also choose to lexicalize or keep covert the semantic contrast between "alienable possession" and "inalienable possession" (Clark 1978, Aristar 1996, Stassen 2009). This variation in languages has to do with whether salient semantic contrasts are individually packaged and lexicalized, or whether they are subject to contextual disambiguation. Given typological variation in this respect, one might ask the question: Is the "individualized packaging" strategy more complex than a "contextual disambiguation" strategy?

In this paper, I will consider a previously undescribed phenomenon – a morphosyntactic contrast in copulas/auxiliaries that is pervasive in several New Indo-Aryan languages. As I will show, restricting myself to the Marathi facts (with a brief nod to Hindi), the morphosyntactic contrast reflects a semantic distinction between particular and characterizing claims. When one considers the origin of this contrast, one finds that categoricality in the expression of this contrast is only to be found in the Modern New Indo-Aryan languages. While earlier stages of Indo-Aryan (Middle and Old Indo-Aryan) appear to show sensitivity to the semantic distinction between particular and characterizing claims, there is no specialized device for conveying particular claims in these systems. I suggest that the New Indo-Aryan languages may have transitioned into a strategy in which this contrast is categorically expressed as a secondary consequence of a change in their broader tense marking systems.

2 The phenomenon

Although they contain the same tensed form of the copula *be*, the (a) and (b) sentences of English in (1–3) are understood very differently with regard to their temporal reference.

(1)   a.   *The baby* is *tired. Let's get the bath ready.*

    b.   *The baby* is *tired by the time we pick him up from daycare. So let us start picking him up earlier.*

(2)   a.   *People* are *unhappy because they just raised the taxes.*

    b.   *People* are *unhappy when they are on diets.*[1]

(3)   a.   *Sam* was *asleep. He had had a long day.*

    b.   *Whenever Mary telephoned on a Friday, Sam* was *asleep.* (Partee 1984: 246)

---

[1]http://www.brainyquote.com/quotes/quotes/m/mireillegu530744.html

The (a) sentences intuitively seem to be about a particular salient time – either the time of utterance or a salient past time. The (b) sentences, on the other hand, do not seem to make reference to any particular time, but rather describe a larger situation, extending over an indefinite interval, characterized by a predictable recurrence of relevant episodes under certain circumstances. For instance, while (1a) conveys that the baby is tired at the time that the sentence is uttered, (1b) conveys that (almost) every time (within some larger contextually understood interval) at which the baby is picked up from daycare is a time when the baby is tired. Similarly, (2a) conveys that people are unhappy at utterance time because of a rise in taxes, while (2b) conveys that there is a tendency for individuals to be unhappy during the times that they are on a diet.

Marathi morphosyntactically distinguishes between these two uses of the English tensed copula. In both the past and the present tenses, the language uses distinct copular paradigms to express the senses corresponding to (a) and (b). These are glossed as COP1 and COP2 respectively. To compare with English, note that (4a), which describes a single episode of the baby being tired, contains a form of COP1 *āhe*, contrasting with (4b), where the presence of the COP2 form *asta* signals that a recurring generalization over episodes is being described. Similar contrasts hold between the COP1 and COP2 sentences in (5) and (6), with (6) exemplifying the past referring copular forms.

(4) a. bāḷ   thaklel-a ***āhe***
         baby.NOM.N.SG tired-N.SG COP1.PRES.3SG
         'The baby is tired.'

     b. bāḷ   sandhyākāḷī thaklel-a ***as-ta***
         baby.NOM.N.SG evening.LOC tired-N.SG COP2-PRES.3N.SG
         'The baby is tired in the evenings.'

(5) a. karvāḍh-i   muḷe   lok-a   dukkhi ***āhe-t***
         tax.increase-OBL because people-NOM.N.PL unhappy COP1.PRES-3PL
         'The people are unhappy because of the tax increase.'

     b. ḍāyeṭ kar-ṇāri   lok-a   dukkhi ***as-tāt***
         diet do-PART.N.PL people-NOM.N.PL unhappy COP2-PRES.3PL
         'Dieters (lit. diet-doing people) are unhappy.'

(6) a. rām   sandhyākāḷī dukkhi ***hotā***
         Rām.NOM.M.SG evening.LOC unhappy COP1.PAST.3M.SG
         'Rām was unhappy in the evening.'

     b. rām   sandhyākāḷī dukkhi ***as-āycā***
         Rām.NOM.M.SG evening.LOC unhappy COP2-PAST.3M.SG
         'Rām used to be unhappy in the evenings.'

Hindi exhibits the same semantic contrast but the morphosyntactic devices used to convey the contrast are slightly different. Hindi has a single copular element and expresses characterizing claims by periphrastically combining the imperfective participial form of this copula with the tensed form. For example, (7a), which describes a single episode of the baby being happy, contains the simple tensed copular form *hɛ*, contrasting with (7b), where the periphrastic imperfective construction *hotā hɛ* signals that a recurring generalization over episodes is being described.

(7)   a.   bacchā   khush ***hɛ***
             baby.NOM.M.SG happy COP.PRES.3SG
             'The baby is happy.'

       b.   bacchā   shām=ko   khush ***hotā***   ***hɛ***
             baby.NOM.M.SG evening=DAT happy COP.IMPF.M.SG COP.PRES.3SG
             'The baby is happy in the evenings.'

Copulas are commonly taken to be the carriers of tense/aspect/modality distinctions without any additional lexical semantic contribution. In many languages (including English and Marathi), copulas in non-verbal clauses are identical to auxiliaries in verbal clauses, enabling further articulation of TAM distinctions in the linguistic system. For instance, in its auxiliary function *be* is used in the realization of the progressive aspect (*is/was/will be V-ing*) as well as the prospective aspect (*is/was/will be going to V*) in English. The two Marathi copular paradigms behave similarly in that they are used as auxiliaries in marking the progressive and perfect aspects in the linguistic system, yielding contrasts in interpretation that remain covertly expressed in English.[2]

Consider the contrast in (8a–8b) and (8c–8d). While the aspectual morphology remains the same in each pair, the interpretation of each member is clearly distinct. Example (8a) asserts that the reference time is contained in an event of John smoking, while (8b) conveys that all/most contextually relevant times (within some larger stretch of time) are contained in an event of John smoking. In (8c), the reference time is understood to be located after the time of an event of John's making dinner and setting the table. The example in (8d), in contrast conveys that in general, the time of my return is located after the time of an event of John's making dinner and setting the table.

---

[2]It is interesting that the distinction between the two copulas is neutralized in the non-finite part of the system. There are no non-finite forms of COP1 and only COP2 forms are available in non-finite contexts regardless of whether the claim is particular or characterizing.

(8)　a.　*John* is smoking *in the common room (right now).*　　　*(episodic)*

　　　b.　*John* is *always/often* smoking *in the common room.*　　*(characterizing)*

　　　c.　*John* has made *dinner and* set *the table (right now).*　　*(episodic)*

　　　d.　*By the time I return, John* has made *dinner and* set *the table.*
　　　　(characterizing)

In English, the contrast between episodic/particular and characterizing interpretations of verbal periphrases is facilitated by the presence of quantificational adverbial material – lexical expressions in (8b) and clausal material in (8d). In Marathi, while such disambiguating material may be present, the episodic vs. characterizing readings are clearly and obligatorily disambiguated by the choice of auxiliary – glossed AUX1 and AUX2. The sentence in (9a) obligatorily conveys that a smoking event is ongoing at reference time and can never be used (even with overt quantificational adverbs) to express something like (8b).[3] In fact, AUX1 is unacceptable in sentences that contain overt quantificational adverbs. Example (9b) correspondingly has only a characterizing interpretation, even in the absence of quantificational adverbs.

(9)　a.　John　　　　　sigreṭ　　　　***pī-t***　　　　***āhe***
　　　　John.NOM.M.SG cigarette.NOM.F.SG drink-IMPF.PART AUX1.PRES.3SG
　　　　'John is smoking a cigarette.'　　　　　　　　　　*(episodic)*

　　　b.　John　　　　　*(*nehmi/kadhi-kadhi*)* sigreṭ
　　　　John.NOM.M.SG. always/sometimes　　cigarette.NOM.F.SG
　　　　pī-t　　　　***as-to***
　　　　drink-IMPF.PART AUX2-PRES.3M.SG
　　　　'John is always/sometimes smoking cigarettes.'　　*(characterizing)*

The same distinction is made with the perfect aspect, where the disambiguation between episodic and characterizing readings is effected by the choice of the tense auxiliary.

(10)　a.　John=ne svayampāk　***banav-lelā***　***āhe***
　　　　John=ERG meal.NOM.M.SG make-PERF.M.SG AUX1.PRES.3SG
　　　　'John has made dinner.'　　　　　　　　　　*(episodic)*

　　　b.　John=ne svayampāk　***banav-lelā***　***as-to***
　　　　John=ERG meal.NOM.M.SG make-PERF.M.SG AUX2-PRES.3M.SG
　　　　'(By the time I return), John has made dinner.'　　*(characterizing)*

---

[3]The imperfective participle+AUX1 periphrasis is the general exponent of the progressive aspect in Marathi. The imperfective participle optionally inflects for gender and number in this periphrasis.

This paper investigates this particular type of split copula/auxiliary system found in Marathi. In fact, this appears to be a genetic feature, since several Indo-Aryan languages, including Hindi, Gujarati, and Ahirani, also exhibit this abstract pattern differing only with respect to the exponents that realize it (as briefly shown for Hindi).[4] The observed pattern demonstrates that semantic distinctions in the interpretations of tensed sentences that are only covertly made in some linguistic systems (e.g. the English copula/auxiliary *be*) can be teased apart systematically due to how markers of temporal reference are lexicalized in the systems of other languages. The division of labor effected by this two-copula/auxiliary system in Indo-Aryan has not yet been described as a possible pattern in either the typological or the semantic literature, making its study particularly interesting from the typological perspective as well. The question for cross-linguistic semantic variation presented by the observed system can be framed as follows: why do some linguistic systems employ the same grammatical device to convey both particular and characterizing claims while other linguistic systems obligatorily signal this difference with distinct devices? Related to the question of semantic variation is a diachronic question: how does a contrast such as the one found in Marathi and other Indo-Aryan languages morphosyntactically emerge in languages? Focusing on this latter question here, I will suggest that the categorical nature of this contrast in Marathi arises from changes in the tense marking system in the transition from Middle Indo-Aryan to the Early New Indo-Aryan languages. In a nutshell, the Middle Indo-Aryan system is aspectually based and lacks the morphosyntactic means to mark the past-present distinction. In Late Middle Indo-Aryan, a new tense auxiliary emerges with specific properties: it presupposes contextually salient intervals and anchors the proposition to the utterance world. This innovation effects a contrast in the copula/auxiliary system in which the contrast between particular and characterizing claims gets obligatorily expressed.

The rest of this paper is organized as follows. §3 describes in detail the interpretations associated with the two paradigms in Modern Marathi. I will only consider the effect of the contrast in copular clauses without any quantificational adverbs since this is enough to show that there are clear differences in interpretation that are obtained with individual-denoting vs. kind-denoting subjects and stage-level vs. individual-level predicates in combination with the relevant forms.

---

[4]Bangla and Oriya (Mahapatra 2009) also exhibit multi-copula/auxiliary systems, but the presence of more than two such elements (as well as the possibility of zero-copula constructions) in these languages yields a different pattern than the one present in the more commonly attested two-copula systems. A thorough investigation of these patterns must be left for later investigation.

In §4, I present data from Epic Sanskrit and Middle Indo-Aryan to show that this contrast, while not identically manifested at these stages, is already partly realized by the presence of a copula that overwhelmingly occurs with characterizing readings. In §5, I discuss the changes with respect to their effect on the overall complexity of the system and conclude.

3 The Marathi facts

The relevant paradigms of the present and past tense copulas/auxiliaries are given in Tables 1 and 2. The forms exhibit agreement along the morphological categories of person and number, and also in many cases, gender. Gender-based contrast has been noted within each person-number cell in which it occurs, in the order masculine/feminine/neuter. The data presented henceforth contains examples only in the present tense since the facts are largely comparable for the past tense cases.

Table 1: Present tense copula/auxiliary paradigms of Marathi

	COP1		COP2	
	SG (M/F/N)	PL (M/F/N)	SG (M/F/N)	PL (M/F/N)
1	*āhe*	*āhot*	*asto/aste*	*asto*
2	*āhes*	*āhāt*	*astos/astes*	*astā*
3	*āhe*	*āhet*	*asto/aste/asta*	*astāt*

Table 2: Past tense copula/auxiliary paradigms of Marathi

	COP1		COP2	
	SG (M/F/N)	PL (M/F/N)	SG (M/F/N)	PL (M/F/N)
1	*hoto/hote*	*hoto*	*asāyco/asāyce*	*asāyco*
2	*hotās/hotis*	*hotā(t)*	*asāycās/asāycis*	*asāycā(t)*
3	*hotā/hoti/hota*	*hote/hotyā/hoti*	*asāycā/asāyci/* *asāyca*	*asāyce/asāycyā/* *asāyci*

3.1 Copular clauses

We now examine non-verbal predicational copular clauses, which consist of a subject, a non-verbal adjectival or nominal element or a postpositional phrase, and the relevant copula. The organization of the data is by the syntactic type of the subject – names and bare nominals, and within each category, by the episodicity of the predicate – i.e. whether it is most naturally construable as a stage-level or individual-level property.

3.1.1 Names

With stage-level predicates, the two copulas contrast particular temporally delimited claims (COP1) with habitual generalizations (COP2). In (11a) the use of COP1 conveys that the property of being busy or angry holds of Anu at the utterance time. In (11b), the use of COP2 obligatorily conveys that over some indefinite interval of time, there are recurring, regularly instantiated episodes of Anu being busy or angry.[5]

(11)   a.   anu   vyasta/cidleli ***āhe***
            Anu.NOM.F.SG busy/angry   COP1.PRES.3SG
            'Anu is busy/angry (right now).'

      b.   anu   vyasta/cidleli ***as-te***
            Anu.NOM.F.SG busy/angry   COP2-PRES.3.F.SG
            'Anu is generally busy/angry.'

Sentences (12a) and (12b) provide examples in which the main predicate is a locative prepositional phrase, another instance of stage-level predication. The observation is identical: the two copulas contrast in whether the assertion pertains to the utterance time or conveys some generalization that holds at some larger interval including the utterance time.

(12)   a.   anu   gharā-t   ***āhe***
            Anu.NOM.F.SG house.OBL-in COP1.PRES.3SG
            'Anu is in the house (right now).'

      b.   anu   gharā-t   ***as-te***
            Anu.NOM.F.SG house.OBL-in COP2-PRES.3F.SG
            'Anu is generally in the house (e.g. when the postman comes by.)'

---

[5]The use of the latter adjective does not convey that Anu is an angry person but rather that Anu is often/regularly found in a state of anger.

When the main predicate is individual-level and denotes a relatively perma-nent, intrinsic property of an individual, only COP1 is acceptable as shown in the contrast between (13a) and (13b). The use of COP2 introduces the sort of oddity that is associated with the use of quantificational adverbs with individual-level predicates (Kratzer 1995, Chierchia 1995, Magri 2009 among others).[6] It conveys that Anu habitually or generally has the property of being cowardly, tall, or in-telligent, which is infelicitous because it tends to give rise to a scalar inference that this property only holds discontinuously in time, i.e. that it is possible that there are times when Anu is not cowardly, tall, or intelligent.

(13) a. anu   ghābraṭ/unca/huśār   **āh-e**
           Anu.NOM.F.SG cowardly/tall/intelligent COP1.PRES.3SG
           'Anu is cowardly/tall/intelligent.'

      b. # anu   ghābraṭ/unca/huśār   **as-te**
           Anu.NOM.F.SG cowardly/tall/intelligent COP2-PRES.3F.SG
           'Anu is (habitually) cowardly/tall/intelligent.'

3.1.2 Bare nominal subjects

Like many languages without determiners, Marathi allows both bare singular and bare plural arguments, and, in subject position, these may be understood either as making reference to unique, contextually salient entities or as mak-ing reference to kinds. In Dayal (1999, 2004), Veneeta Dayal makes a convincing case for Hindi, using arguments from scopal (non-)interaction that bare singulars are not ambiguous between indefinite and definite interpretations in Hindi (and other determiner-less languages). The Marathi facts closely parallel the Hindi facts and I will investigate the range of readings of bare nominals in Marathi only in the context of copular clauses here.

3.1.2.1 Stage-level predicates

Consider the examples in (14a) and (14b), which contain bare singular subjects and stage-level predicates. These sentences are most naturally interpreted as de-scribing the properties of the contextually most salient dog in the utterance con-text – i.e. the bare nominal has a directly referential use – like an NP with the definite article in English.

---

[6]This is the observation that a sentence like *John is generally/always/often/sometimes intelligent* is understood as deviant or unacceptable without context.

(14) a. kutrā   thaklelā/bhukelā **āhe**
          dog.NOM.M.SG tired/hungry   COP1.PRES.3SG
          'The dog is tired/hungry.'

    b. kutrā   thaklelā/bhukelā **as-to**
          dog.NOM.M.SG tired/hungry   COP2-PRES.3M.SG
          'The dog is generally tired/hungry.'

With locative predicates, the pattern remains the same: the contrast lies in whether the property of being in the house is said to hold of the most salient dog in the utterance context, at the utterance time (COP1) or more generally over an indefinite interval that contains the utterance time (COP2).

(15) a. kutrā   gharā-t   **āhe**
          dog.NOM.M.SG house.OBl-in COP1.PRES.3SG
          'The dog is in the house (right now).'

    b. kutrā   gharā-t   **as-to**
          dog.NOM.M.SG house.OBL-in COP2-PRES.3M.SG
          'The dog is generally in the house.'

In both (14) and (15), the bare nominal subject has most naturally a directly referential reading – its referent is understood to be an entity that is most salient in the utterance context (in the actual world at utterance time).

However there is another non-referential reading of bare nominals that arises with the use of COP2. For illustration, consider (16b), which contains COP2. Here, the bare singular *rāṅgoḷī* does not refer to any particular contextually salient entity at utterance time in the actual world, but rather to the rangoli that gets drawn every day by Anu in front of her door.[7]

(16) *Context: My friend is telling me about her sister Anu, who draws elaborate rangoli motifs in front of her house every day. She says:*

    a. anu   roj   rāṅgoḷī   kāḍh-te
          Anu.NOM.F.SG everyday rangoli.NOM.F.SG draw-IMPF.PRES.3F.SG
          'Anu draws a rangoli motif every day.'

    b. rāṅgoḷī   dārā-samor   **as-te**
          rangoli.NOM.F.SG door.OBL-in.front.of COP2-PRES.3F.SG
          'The rangoli (that she draws) is in front of the (main) door.'

---

[7]Rangoli (Marathi *rāṅgoḷī*) is a traditional art form in which decorative patterns are created on the floor using materials such as colored rice, dry flour, colored sand, or flower petals.

In the given context, (16b) conveys that for each day d within some indefinite interval overlapping with the utterance time, the unique rangoli r_d that Anu draws on d, is located in front of the main door. It is infelicitous to follow up (16a) with (17), which contains COP1, since the bare nominal *rāṅgolī*, in this case, can only be taken to refer to the contextually salient rangoli *at utterance time*.

(17) # rāṅgolī   dārā-samor   ***āhe***
         rangoli.NOM.F.SG door.OBL-in.front.of COP1.PRES.3SG
         'The rangoli is (right now) in front of the (main) door.'

In addition to the directly referential and non-referential readings described above, bare singular subjects may also be understood as kind-denoting. For instance, (18) has two salient readings: on the definite referential reading of the nominal, it may describe the general coordinates of a specific clock salient in the utterance context (for instance, the one my uncle gave me for my birthday). On the other reading, the sentence describes a generalization about where clocks in general tend to be located (i.e. as a claim about the kind clocks).

(18) ghaḍyāḷ   bhinti-var ***as-ta***
         clock.NOM.N.SG wall.OBL-on COP2-PRES.3N.SG
         definite referential: 'The clock [my uncle gave me] is (generally) on the wall.'
         kind: 'A clock (in general) is on a wall (rather than on the floor).'

In a slightly different context, the bare nominal in (18) can also be interpreted non-referentially. For example, in a context such as the one below, the bare nominal refers to the unique clock in each room in John's hotel that lacks a mantelpiece, and not to a unique entity in the utterance context.

(19) *Context: John is describing the organization of the rooms in his hotel to his*
         manager. In each room, the time-piece is placed on the mantelpiece above
         the fireplace, if there is one. If there is none, the time-piece is hung on the
         wall above the bed. John says to his manager: "When there is no
         mantelpiece..."
         ghaḍyāḷ   bhinti-var ***as-ta***
         clock.NOM.N.SG wall.OBL-on COP2-PRES.3N.SG
         definite non-referential: 'The clock (in a room without a mantelpiece) is (generally) on the wall.'

Bare plural subjects differ from bare singulars in that the kind interpretation is much more easily available for clauses in which they occur regardless of copula or predicate type. The sentence in (20) contains a stage-level predicate, COP1, and a bare plural subject, *kāmgār* 'workers'. On one reading, it is a claim about the worker-kind at utterance time; we might be talking about workers all over the world (or in the US) working in exploitative conditions without job security, on the verge of a world-wide revolution.[8] But it can also be read as a claim about a contextually salient plural entity – for instance, the group of workers that works at an air-conditioning plant that is planning to close shop and declare bankruptcy. This is the directly referential reading of the bare nominal.

(20) kāmgār   asantuṣṭa   ***āhe-t***
     worker.NOM.M.PL discontented COP1.PRES-3PL
     kind: 'Workers (in general) are discontented (right now).'
     definite referential: 'The workers (working at the air-conditioning plant right now) are discontented (right now).'

With COP2 and a stage-level predicate, as in (21), the sentence is understood to report a generalization obtaining over an indefinite interval containing the utterance time. However, the content of the generalization depends on how the bare plural is interpreted. It may refer to the kind, it may refer to the contextually salient plural entity in the utterance context, e.g. the workers that work at the air-conditioning plant right now, or it may pick out (possibly different) groups of workers across different times – this is the definite non-referential reading.

(21) kāmgār   asantuṣṭa   ***as-tāt***
     worker.NOM.M.PL discontented COP2-PRES.3PL

     kind: 'Workers (in general) are (generally) discontented.'
     definite referential: 'The workers (who are working at the air-conditioning plant right now) are (generally) discontented.'
     definite non-referential: 'The workers (whoever happen to work at the air-conditioning plant at a given time) are (generally) discontented.'

---

[8]This kind reading is also available when the subject is a bare singular but it is a little more difficult to access. There is no number morphology on the subject, but singular/plural reference is inferred through agreement marking on the copula.

3.1.2.2 Individual-level predicates

With individual-level predicates and COP1, both bare singulars and bare plurals, are preferentially interpreted as referring to singular or plural entities that are salient in the utterance context rather than to the kind.

(22)   a.   kāmgār   ghābraṭ/unca/huśār   ***āhe***
             worker.NOM.M.SG cowardly/tall/intelligent COP1.PRES.3SG
             definite referential: 'The worker (at that air-conditioning plant) is cowardly/tall/intelligent.'
             kind: '?A worker (in general) is cowardly/tall/intelligent.'

       b.   kāmgār   ghābraṭ/unca/huśār   ***āhe-t***
             worker.NOM.M.PL cowardly/tall/intelligent COP1.PRES-3PL
             definite referential: 'The workers (at that air-conditioning plant) are cowardly/tall/intelligent.'
             kind: '?Workers (in general) are cowardly/tall/intelligent.'

However, the kind reading of bare nominals becomes available with individual-level predicates and COP1 given suitable context and supporting linguistic information. For instance, in a context in which one is contrasting workers in this age with workers of previous eras, one may use COP1 to describe "the worker of today" (in contrast to that of yesteryear) as being intelligent – (23a). Similarly, (23b), which contains a bare plural and COP1, is fully acceptable in a context where the evolutionary potential of donkeys is under consideration and one considers the possibility of intelligence determining genes mutating to make donkeys stupid.[9]

(23)   a.   āj-cā   kāmgār   jāsti huśār   ***āhe***
             today-GEN.M.SG worker.NOM.M.SG more intelligent COP1.PRES.3SG
             kind: 'The worker of this age (lit. today) is more intelligent.'

       b.   gāḍhav-e   huśār   ***āhet***
             donkey-NOM.N.PL intelligent COP1.PRES-3PL
             kind: 'Donkeys, as a kind, are intelligent (right now).'

---

[9]Of course, the definite referential reading is available for both examples in (22). In a context in which we are talking about the particular worker that has come in today to help with cleaning the machines, the singular expression, *ājcā kāmgār* can refer to this specific worker. Similarly, the bare plural *gāḍhav-e* can refer, in the right context, to my pet donkeys. It is the kind reading that is somewhat difficult to access with COP1, but can be made available given contexts such as those above.

With cop2 and individual-level predicates like *huśār* 'intelligent,' definite referential readings are unavailable with bare nominals (both singular and plural). This can be illustrated with the example in (24).

(24)   *Context: One/two of the workers at the air-conditioning plant fix(es) a problem with the cooling mechanism in an ingenious way. I praise his/their ingenuity, remarking to the manager:*
   \# kāmgār   huśār   ***as-to/as-tāt***
      worker.nom.m.sg/pl intelligent cop2-pres.3m.sg/pl
   definite referential: 'The worker(s) (who fixed the problem) is/are intelligent.'

In such contexts, where it is clear that the bare nominal must refer to a singular/plural entity that is salient in the utterance context, speakers always choose cop1 and reject cop2. In attempting to construe bare nominals in cop2 sentences as definite referential expressions, speakers encounter the same oddity observed with names in (13b) – that the property holds of a contextually salient singular or plural entity discontinuously in time.

Both definite non-referential and kind readings are possible with bare nominals when combined with cop2 and individual level predicates. (25) and (26) contain examples of the contexts in which the definite non-referential and kind readings of bare nominals arise respectively.

(25)   *Context: One/two of the smarter worker(s) is/are assigned to work over-time each month to keep the machinery in working order and repair malfunctions. Because the workers are smart and already very familiar with the machinery, this system proves more efficient than calling outside expertise to service the machines. I explain this system to the manager saying:*
   kāmgār   huśār   ***as-to/as-tāt***   mhaṇun
   worker.nom.m.sg/pl intelligent cop2-pres.3m.sg/pl therefore
   kām   lavkar āṭap-t-a
   work.nom.n.sg fast   finish-impf.pres.3n.sg
   definite non-referential: 'The worker(s) (that get assigned to the job) is/are intelligent (smart) (and) so the work gets done faster.'

(26) *Context: I am explaining to my students in a class on Labor Dynamics that they should always be transparent in their interactions with labor unions and try to understand their point of view.*[10]

kāmgār   prāmāṇik ***as-tāt***   mhaṇun tumhi suddhā
worker.NOM.M.PL honest   COP2-PRES-3M.PL therefore you   also
prāmāṇik as-āva
honest   be-POT.N.SG

kind: 'Workers (in general) are honest, therefore you should also be honest (in your interactions with them).'

3.2 The generalization

The distribution of the two (present tense) copulas detailed in the previous section is summarized below in Table 3. The terms D-ref., D-non-ref., and kind-ref. stand for the definite referential, definite non-referential, and kind-referring readings of bare nominals.

What is immediately apparent through the table is that the type of copula influences the range of readings available to bare nominal subjects. Specifically, COP1 forces a referential interpretation of bare nominals (i.e. the nominal must pick out an individual (singular, plural, or kind) in the actual world at utterance time). COP2, on the other hand, is unacceptable (oddity inducing) when neither the subject denotation nor the predication of the property of the subject denotation may be construed as variable across time. Intuitively, the meaning of COP1 sentences seems to depend on the valuation of the embedded predication at utterance time in the actual world, while the meaning of COP2 sentences seems to require consideration of the valuation of the embedded predication at times beyond the utterance time. In other words, the morphosyntactic devices COP1 and COP2 allow Marathi to distinguish between descriptions whose interpretation is anchored to the utterance time i_0 and the utterance world w_0 on the one hand and those that lack such anchoring on the other.

This contrast is typologically interesting since, as far as I know, it has not been described as being the basis of a multiple-copula system in any language (family). As reported in the introduction, the distinction is wide-spread in the New Indo-Aryan languages. While the choice of devices may differ, all languages systematically disambiguate interpretations that are anchored to the utterance world and

---

[10]The sentence has a kind reading with the adjective *intelligent* as well, but it was difficult to construct a context in which such a sentence could be uttered without there also being some bias in the context that workers are not intelligent. The change from *intelligent* to *honest* is in order to avoid invoking such a bias.

Table 3: Readings associated with copular clauses in Marathi

Subject NP	COP1		COP2	
	SLP + COP1	ILP + COP1	SLP + COP2	ILP + COP2
names	episode in actual world at UT	property in actual world overlapping with UT	generalization over episodes across time (invariant NP referent)	*induces oddity*
bare nominals D-ref.	episode in actual world at UT	property in actual world overlapping with UT	generalization over episodes across time (invariant NP referent)	*induces oddity*
bare nominals D-non-ref.	*unavailable*	*unavailable*	generalization over episodes across time (variable NP referent)	generalization over properties of individuals across time (variable NP referent)
bare nominals kind-ref.	episode involving the kind in actual world at UT	property of the kind in the actual world overlapping with UT	generalization over episodes involving kind-instances across time	generalization over properties of the kind across time

time from those that are not. The question we turn to next is: How/when does this semantic contrast become morphosyntactically expressed in Indo-Aryan languages in a categorical way? To answer this question, I will consider facts from Epic Sanskrit (Old Indo-Aryan) and Prakrit and Apabhraṃśa (Middle Indo-Aryan).

4 Historical basis of the contrast

4.1 Old Indo-Aryan (Epic Sanskrit)

Old Indo-Aryan, like most Ancient Indo-European languages, inherits two PIE "be" verbs – *as* (PIE $*h_1es$) and *bhū* (PIE $*b^h\mu eh_2$). The precise distribution of the two forms in Epic Sanskrit is not well-established, but there are some observed environments in which each copula occurs.[11] For instance, *as* is the tensed element of choice in existential clauses but it can also be used in predicational clauses. In both constructions, it can be used to make both particular and characterizing claims. *bhū*, in contrast, as a copular expression, only appears in predicational clauses and in those structures, appears to be compatible only with non-referential, characterizing readings.[12] Notice that this distribution, in which one copular element tends to have non-referential characterizing readings in its stative uses, yields a partially articulated contrast between particular and characterizing claims. To the best of my knowledge, this interpretive contrast between the two copular expressions has not been explicitly described for Sanskrit, although, in §5, I point out that this appears to be a contrast instantiated in Old English as well, potentially pointing to a tendency inherited from the PIE system. Consider the examples below from the Mahabharata, one of the two Epic Sanskrit texts, that illustrate this distribution.[13]

---

[11]A reviewer observes that using Epic Sanskrit as the only source for investigating the distribution of the two copular forms in Old Indo-Aryan is an unfortunate choice. The reasoning is that "the Epic Sanskrit corpus is not coherent and in quintessence it is not even Old Indo-Aryan. It varies massively diachronically and diatopically." It is indeed conceivable that the writers of the Sanskrit Epics, are, in fact, native speakers of a language with a Middle Indo-Aryan type system, leading to peculiarities of Middle Indo-Aryan entering the Epic corpus. We know that the Middle Indo-Aryan Prakrits were the vernacular languages in the region at least since 300 BCE (based on Aśokan inscriptions). However, the limited goal of this paper is to establish that the semantic contrast observed in Marathi (and other new Indo-Aryan languages) is not present in the same categorical way in the diachronically prior systems although there is a tendency to associate the *bhū* copula with characterizing claims. I therefore take showing the existence of such tendential data in the Epic corpus to be sufficient for this goal.

[12]*bhū* also has an inchoative use where it corresponds to a verb like 'become' or 'happen'. This is an eventive use of the verb, which I will be ignoring for the purposes of this paper.

[13]The Mahabharata text is attributed to a single author Vyāsa but is usually understood to be a compiled text (dateable to ~100BCE) with interpolations from multiple authors.

(27) a. madadhīn-o   **'-si**   pārthiva
me.dependent-NOM.M.SG **as**-PRES.2.SG king.VOC.SG
'O King, you are dependent on me.' (Mbh. 1.78.35b)

 b. abhijāt-o   **'s-mi**   siddh-o
high.born-NOM.M.SG **as**-PRES.1.SG accomplished-NOM.M.SG
's-mi   na **as-mi**   kevalamānuṣ-aḥ
as-PRES.1.SG NEG **as**-PRES.1.SG ordinary.man-NOM.M.SG
'I *am* of high birth, I *am* accomplished, I *am* not an ordinary man.'
(Mbh. 12.28.7a)

 c. y-e   **s-anti**
Those.REL-NOM.PL **as**-PRES.3.SG
vidyātapasopapann-ās   te-ṣāṃ
knowledge.ascetism.possessed-NOM.M.PL those.CORREL.GEN.M.PL
vināśa-ḥ   prathamaṃ tu kār-ya-ḥ
destruction-NOM.M.SG first   PTCL do-POTEN-NOM.M.SG
'Those who *are* possessed of knowledge and ascetic virtue, their
destruction should be undertaken first.' (Mbh.3.99.19c)

In (27a) and (27b), both predicational copular clauses, the copula *as* is used to convey that the referent (the addressee and the speaker of the utterance context, respectively) has the relevant property at utterance time in the utterance world. In (27c), a more involved sentence, describes a resolution arrived at in terms of a course of action to be undertaken at utterance time. The subject referent, in this case, is the set of all individuals in the utterance world at utterance time, who are possessed of knowledge and ascetic virtue. Crucially, the sentence does not express a general claim about how such people are to be treated across all situations.

In (28a) and (28b), both existential clauses, we see that the *as* copula is used to make characterizing claims. Example (28a) makes the generalization that kingless kingdoms have neither rain nor Gods while (28b) asserts the existence of cowardly and brave men across different indices of evaluation – not only at utterance time.

(28) a. arājak-eṣu   rāṣṭr-eṣu   na **as-ti**   vṛṣṭi-r   na
king.less-LOC.PL kingdom-LOC.PL NEG **as**-PRES.3.SG rain-NOM.SG NEG
devatā-ḥ
God-NOM.PL
'In kingdoms without a king, there *is* no rain and no Gods.' (Mbh.
1.99.41a)

b. ***s-anti***   vai   puruṣ-āḥ   śūr-āḥ   ***s-anti***
as-PRES.3.PL PTCL man-NOM.M.PL brave-NOM.M.PL **as**-PRES.3.PL
kāpuruṣ-ās   tathā
coward-NOM.M.PL likewise
'There *are* brave men, and likewise those that *are* cowards.' (Mbh.
5.3.2a)

When we study the distribution of the *bhū* copula, it appears that its stative
uses involve only characterizing readings, as in the examples in (29).[14] The sentence in (29a) describes what is said in the code – a guideline to be followed not
just at utterance time but more generally. The example in (29b) describes the
defining properties of the Rākśasa women kind and the clause containing the
bhū copula predicates the property of being many-formed to the kind.

(29) a. sakhībhartā   hi   dharm-eṇa   bhartā
friend.husband.NOM.M.SG PTCL code-INS.M.SG husband.NOM.M.SG
bhava-ti   *śobhane*
bhu-PRES.3.SG beautiful.VOC SG
'O beautiful one, the husband of a friend, according to the code, *is*
also one's husband.' (Mbh. 1.78.20c)

b. sadyo   hi   garbha-ṃ   rākṣas-yo labh-ante
immediately PTCL embryo-ACC.SG R-NOM.F.PL receive-PRES.3.PL
prasav-anti   ca kāmarūpadhar-āś   ca eva
give.birth-PRES.3.PL and desire.form.holding-NOM.F.PL and PTCL
bhav-anti   bahurūpiṇ-aḥ
bhu-PRES.3.PL many.formed-NOM.F.PL
'The Rākśasa women give birth the very day they conceive, and being
able to assume any form at will, they *are* many-formed.'
(Mbh.1.143.32a-c)

What is crucial is that the verbs *as* and *bhū* do not seem to stand in free variation in the Epic Sanskrit system. While *as* is used to make both particular and

---

[14]I searched through John Smith's electronic text version of of the Bhandarkar Oriental Research
Institute's critical edition of the Mahabharata (1933–1966) for instances of the 3rd singular and
plural present indicative forms of *bhū* – *bhavati* (362 occurences) and *bhavanti* (114 occurences).
Examining the first fifty occurrences among the results from each set did not yield a clear instance where *bhu* appeared as the main predicate in a copular clause, was interpreted statively
with present reference, and gave rise to a particular claim. The electronic edition used is available at https://bombay.indology.info/mahabharata/statement.html.

characterizing claims, there is a strong tendency for the *bhū* copula to NOT be used in clauses that describe a state determined by the valuation of the embedded predication at utterance time in the actual world – in other words, the uses that I examined occur in clauses that convey generalizations. I take these facts to suggest that there is some evidence in Old Indo-Aryan for a dedicated device for expressing non-referential characterizing claims but there is no clear-cut division of labor between the two copulas of the kind one sees in Marathi and other New Indo-Aryan languages.

4.2 Middle Indo-Aryan

4.2.1 Maharashtri Prakrit

The changes from the inflectional system of verbal contrasts in Old Indo-Aryan to the relatively morphologically impoverished inflectional system of Middle Indo-Aryan have been described in terms of "erosion" or "simplification", primarily because many of the rich conjugational paradigms and the semantic categories expressed were lost in Middle Indo-Aryan (Bloch 1965, Beames 1966 [1872–1879], Bubenik 1998, 1996, Pischel 1981 [1900], Vale 1948, Masica 1991 and others). Of the several changes in the expression of tense-aspect-modality distinctions,[15] critical is the loss of a morphosyntactic distinction between the past and present tenses. This subsection describes the resulting contrasts in the re-organized aspectually based system in which temporal reference is established through contextual cues (Deo 2012).[16] The examples in (30) and (31) illustrate the basic Middle Indo-Aryan pattern, which forms the backdrop for an innovation in Late Middle Indo-Aryan. In (30a), the Old Indo-Aryan Present paradigm, glossed IMPF, has present reference, while in (30b), the same paradigm has past reference.

---

[15]The Middle Indo-Aryan tense/aspect system inherits only the Present, the Perfective Participle, and the Sigmatic Future paradigms from Old Indo-Aryan. The rich system of past tense markers is lost. Pischel (1981 [1900]), on the basis of careful textual study, reports that the Imperfect, the Aorist, and the Perfect occur in Middle Indo- Aryan texts only as a few scattered forms for a few verbs. From among the past-referring forms of Epic Sanskrit, only the perfective participial paradigm remains and it is used regularly to refer to past time events in Middle Indo-Aryan.

[16]In brief, the reorganization is as follows: The Old Indo-Aryan Present tense realizes a tenseless imperfective and is compatible with both present and past imperfective reference. The Old Indo-Aryan Present Participle is also starting to be used in this function. The Old Indo-Aryan Past Participial form in *-ta* realizes the perfective aspect and is used to refer to completed events. Therefore, by default, the use of this form leads to past temporal reference. However, this form may also be used systematically to describe future eventualities.

(30) a. nipphala-ṃ   duma-ṃ   pakkhiṇ-o   vi
       fruitless-ACC.N.SG tree-ACC.N.SG bird-NOM.M.PL also
       paricchaya-nti
       abandon-IMPF.3.PL
       'Even birds *abandon* a fruitless tree.' (VH.DH 31.24-25)   *imperfective*
       present reference

   b. *tato aham   aṇṇayā kayāi   āyariya-giha-rukkha-vāḍiyā-e*
       then i.NOM.SG other   some time teacher-house-tree-garden-LOC.F.SG
       joga-m   ***kare-mi***
       yoga-ACC.M.SG do-IMPF.1.SG
       'Then, sometimes, I *would perform* Yoga in the orchard at my
       teacher's house.' (VH:DH 37.1)   *imperfective past reference*

In (31), a set of consecutive sentences reports part of a past episode about a monkey who entered a mountain cave and mistook some sticky liquid tar to be water. The example in (31a) describes a past event using perfective marking while (31b) and (31c) also with past reference, describe past activities using the temporally unmarked imperfective.

(31) a. te-ṇa   palāyamāṇ-eṇa   purāṇakuv-o
       that-ERG.M.SG running-ERG.M.SG old.well-NOM.M.SG
       taṇadabbhaparichinn-o   ***diṭ-ṭho***
       weed.grass.covered-NOM.M.SG notice-PERF.M.SG
       'That running one *noticed* an old well
       covered with weed and grass.' (VH.KH. 8.6)   *perfective past reference*

   b. tattha ayagar-o   mahākā-o   vidāriyamuh-o
       there python-NOM.M.SG gigantic-NOM.M.SG open.mouthed-NOM.M.SG
       gāsiukām-o   tam   purisam
       swallow.desiring-NOM.M.SG that-ACC.M.SG man-ACC.M.SG
       avaloe-i
       observe-IMPF.3.SG
       'There a giant python, baring its mouth, eager to eat, *observed* the
       man.' (VH.KH. 8.9)   *imperfective past reference*

   c. sapp-ā   bhīsaṇ-ā   aṣiukām-ā
       snake-NOM.M.PL fearsome-NOM.M.PL eat.desiring-NOM.M.PL
       ciṭṭha-nti
       stand-IMPF.3.PL
       'Fearsome snakes, eager to bite, *stood* (in the well).' (VH.KH. 8.9)
       imperfective past reference

To summarize, the core distinctions made in the Middle Indo-Aryan aspecto-temporal system are as follows:

(32)
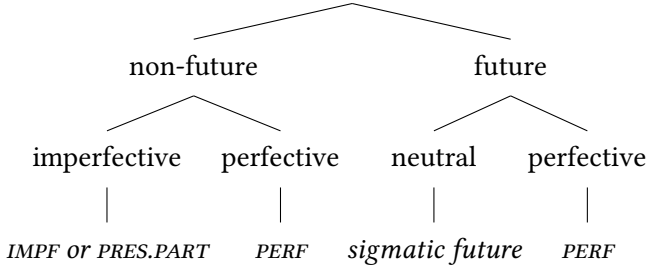
The Early New Indo-Aryan aspecto-temporal system is systematically built up from this aspectual core through periphrastic constructions based on tense auxiliaries, already visible in Middle Indo-Aryan (Bubenik 1996, 1998, Kellogg 1990 [1893], Beames 1966 [1872–1879], Chatterji 1975 [1926], and others).

4.2.2 Apabhraṃśa

A key change in the Late Middle Indo-Aryan verbal system, specifically Apabhraṃśa, involves the loss of the Old Indo-Aryan *as* copula.[17] Additionally, we see the introduction of a new tense auxiliary, based on the verb *acch* 'sit' that establishes temporal reference with respect to speech time (Turner 1936). This auxiliary presupposes contextually salient reference intervals and anchors the embedded property to the utterance world. The present tense auxiliary paradigms of several Indo-Aryan languages are cognate forms of this original auxiliary, as seen below.[18]

- *acchai > āhe*   Marathi
- *acchai > ahai > hai*   Hindi
- *acchai > chai > che*   Gujarati
- *acchai > āchi*   Bangla

---

[17]It has been proposed that the loss of *as* is purely morphological; it is an athematic verb from the second conjugation, which is lost as part of the simplification of the verbal system (von Hinüber 2001: 293).

[18]I note here that in its early uses, the *acch* auxiliary is not always associated with present tense reference. It is compatible with both past and present temporal reference and uniformly conveys that a given state holds at reference time. In Early New Indo-Aryan languages, however, this form is exclusively used to convey present temporal reference.

In Late Middle Indo-Aryan, this auxiliary appears in copular clauses as illustrated below in (33). All examples here come from the *Paumacariu* of Svayambhudeva, a key Apabhraṃśa verse text from ~800CE.

(33) a. deva deva   ki-u   j-eṇa   mahāra
       Lord.VOC.SG do-PERF.M.SG that.REL-INS.M.SG great.sound.NOM.M.SG
       acch-ai   mattahatthi   airāva
       acch-PRES.3.SG musth.elephant.NOM.SG Airāvata.NOM.SG
       'Lord, the one who made a great sound, he is the elephant Airavata in musth season.' (PC 1.11.3.4)

    b. ***acch-ahi***   suha.dukkha.karamviya
       acch-PRES.2.SG pleasure.pain.engrossed.PERF.M.SG
       '(You) are engrossed with pleasure and pain.' (PC 2.33.5.2)

    c. ***acch-ai***   kailāsa-ho   uvari sāhu
       acch-PRES.3.SG Kailash-GEN.SG on   sage-NOM.M.SG
       'There is a sage on the Kailasa mountain.' (PC 1.13.2.6)

The cognate of the *bhū* copula retains its properties observed in Old Indo-Aryan – in its stative uses, it gives rise to non-referential characterizing readings, as the examples in (34) illustrate.

(34) a. sappurisa   vi   cañcalacitta   ***ho-nti***
       good.man.NOM.M.PL even unsteady.mind.NOM.M.PL bhu-PRES.3.PL
       'Even good men are fickle-minded.' (PC 2.22.10.7)

    b. sāsu-a   ***ho-nti***   viruāriya
       mother-in-law-NOM.F.PL bhu-PRES.3.PL cruel.NOM.F.PL
       'Mothers-in-law are cruel' (PC 1.19.4.8)

    c. has-iu   purandar-eṇa   are māṇava   devasamāṇa
       laugh-PERF.M.SG Purandara-INS.M.SG O   human.VOC.SG God.equal
       ho-nti   kiṃ   dāṇava
       bhu-PRES.3.PL INTER demon.NOM.M.PL
       'Purandara (Indra) laughed: "O human, are the Gods equatable with the demons?"' (PC. 1.8.8.8)

I speculate here that the Late Middle-Aryan innovation of a tense auxiliary built on *acch* 'sit' that anchors the embedded predication to the utterance world and contextually salient reference intervals, in fact leads to a reorganization of the copular/auxiliary system. The idea is as follows: the original Old Indo-Aryan

system had a dedicated device for making non-referential characterizing claims (the *bhū* copula) but the *as* copula was underspecified and could be used in both particular and characterizing senses. This underspecified copular element was independently lost for morphological reasons in Middle Indo-Aryan. The innovation of a new tense auxiliary, anchored to the utterance world and time, in the Late Middle Indo-Aryan system facilitated the "hardening" of the soft contrast between particular and characterizing claims found in the older systems. Specifically, I hypothesize that the innovated copular element was incompatible with characterizing claims, which led to a complementary distribution between the two copular forms in the domain of copular clauses.[19] There is much that is unknown about the precise development of tense marking in Indo-Aryan and a fuller understanding of the pathway associated with the copular contrast in New Indo-Aryan rests on that. I leave this investigation to future research, noting only that it is not the innovation of morphosyntactic tense marking per se that leads to the categorical realization of the contrast, but rather the particular temporal reference features associated with the auxiliary forms.

5 Concluding thoughts

Towards the beginning of this paper, I distinguished between morphosyntactic strategies in which salient contrasts within a semantic domain are overtly expressed or individually packaged vs. those in which such contrasts remain morphosyntactically unexpressed with the distinct meanings disambiguated in context. From the perspective of the questions that underpin this volume, one might ask if an individualized packaging strategy is more or less complex than a contextual disambiguation strategy. That is, are semantic contrasts retrievable by means of form-dependent strategies "better" than contrasts retrievable only in context (e.g. Modern English)? Depending on the answer, the emergence of the categorical contrast between particular and characterizing claims in a linguistic system may be seen as a change for the worse. There are at least two considerations. On the one hand, successful use of the two-copula strategy observed in Marathi (and mirrored in other Indo-Aryan languages) involves the acquisition of two semantically distinct paradigms for temporal reference. While the two-form system guarantees communicative success, the acquisition process is rendered more complex since the distribution of two forms relative to a semantic

---

[19]This contrast is still realized in some languages (e.g. Hindi and Bangla) by the cognates of *acch* and *bhu*, while in other languages, other lexical items and paradigms get recruited to realize the same semantic contrast.

domain must be learned. On the other hand, in a language that does not lexicalize the particular-characterizing contrast and has only a single form to convey distinct meanings (such as English), hearers must be contextually attuned so that the intended temporal interpretation is indeed retrieved reliably in a given context.

In the Indo-Aryan case, we observe a transition from a partially context-dependent strategy of meaning recovery to a form-dependent strategy of meaning recovery. While evaluating whether this transition is a worsening or complexification of the system, it is necessary to keep in mind that the categorical or "hardened" contrast between particular and characterizing meanings does not just emerge *spontaneously* in Indo-Aryan from the original soft contrast. Rather, it is situated within changes in the larger landscape of Middle Indo-Aryan temporal reference. The Middle Indo-Aryan system lacked dedicated devices corresponding to the present and the past tenses – a distinct morphosyntactic impoverishment in comparison to the Old Indo-Aryan system with three past-referring categories. In such a system, imperfective and perfective clauses are temporally underspecified and contextual cues are critical to the retrieval of information regarding their temporal reference. The development of the copular contrast described here was concomitant with the emergence of overt marking of tense distinctions via innovated tense markers. These tense markers (as illustrated here by the present tense forms) had the right anchoring properties – they anchor the embedded predication to the utterance time and the utterance world. The categorical marking of particular vs. characterizing claims became possible only after such devices were available in the linguistic system.[20] Thus, the transition to what might appear, on the formal metric, a more complex strategy of meaning packaging and recovery in the copular domain, turns out to be a consequence of a far more general change in the language – the development of overt morphosyntactic realization for the present and past tenses. If one argues that the development of basic temporal distinctions in the tense-aspect system does not constitute complexification, then the status of concomitant effects of this development becomes somewhat less clear.

To close the paper, I will use a counterpoint to underscore that the categorical copular pattern seen in synchronic Marathi (and other Indo-Aryan languages) crucially relies on particular historical facts about the Late Middle Indo-Aryan tense-aspect system. The counterpoint is Old English, which also inherited the

---

[20]To be clear, our earliest records of Marathi already show these tense auxiliaries with established function. In contrast, they are rather infrequent the Late Middle Indo-Aryan record. So it is likely that there might not be textual material that allows us to track the gradual development of these devices in the Indo-Aryan languages.

two PIE "be" verbs – *is* (PIE *h_1es) and *bið* (PIE *$b^h\underset{\cdot}{u}eh_2$). Petré (2013) (also citing prior research) shows convincingly that in Old English, *is* was mainly used for predicating present states of specific subjects, and in identifying clauses, while *bið* was used to encode future situations and generic statements, which are connected to future situations through their implication of future validity. The pattern, while not identical to that of Epic Sanskrit, is similar in that the language morphosyntactically distinguishes between particular claims and characterizing claims (Petré's) "generic statements". The two examples in (35) illustrate the functional distribution of *bið* in Old English.[21]

(35)   a.   *Hit **byð** dysig þæt man speca   ær þone he þænce*
         it   is   foolish that man speak.SUBJ ere then he think.SUBJ
         'It *be* foolish that a man speaks before he thinks.' (c1100. Prov 1 [Cox]: 2.2, via Petré 2013)

   b.   *Wið   stede   &   for gebinde,   heortes   hær   **beoð***
         against strangury and for constipation hart.GEN hair.PL are.IND.3.PL
         swiðe gode   mid   to smeocanne wifmannum
         very   good with to smoke   women.DAT
         'Against strangury and constipation, hairs of the hart *be* very good for women to fumigate with.' (c1025. Med 1.1 [de Vriend]: 3.16, via Petré 2013)

If Petré is right, what is striking about the evolution from Old English to Middle English is that this functional contrast, which is soft, but remarkably stable in Old English, is eroded by the the grammaticalization of a future construction *sceal beon* "shall be." One result of this innovation is a drastic redistribution of the two "be" verbs, and finally their merger (the situation in Modern English). The hardening of soft semantic contrasts of this sort thus appears to be entirely dependent on patterns of innovation and loss in the larger system of temporal/aspectual contrasts.

---

[21]Petré refers to a quantitative analysis by Kilpiö (1993), which suggests that the pattern is tendential rather than categorical – there is a soft rather than categorical contrast. According to Kilpiö, *is* is found with its typical semantic characteristics in no less than 86.4% of all instances in HC. For *bið*, its presence in the typical semantic domain associated with it, amounts to only about 56.7%.

Abbreviations

1	First person	INTER	interrogative particle
2	Second person	LOC	locative
3	Third person	M	masculine
ACC	accusative	NEG	negation marker
ACT	active voice marker	N	neuter
AUX1	auxiliary (episodic)	NOM	nominative
AUX2	auxiliary (characterizing)	OBL	oblique
COP1	copula (particular)	PART	participle
COP2	copula (characterizing)	PASS	passive voice
CORREL	correlative pronoun	PAST	past tense
DAT	dative	PL	plural
EXCL	exclusive clitic	POTEN	potential mood
F	feminine	PRES.PART	imperfective participle
GEN	genitive	PRES	present tense
GER	gerund	PTCL	particle
IMPF	imperfective aspect	REL	relative pronoun
INF	infinitive	SG	singular
INS	instrumental	VOC	vocative

References

Aristar, Anthony. 1996. The relationship between dative and locative: Kuryłowicz's argument from a typological perspective. *Diachronica* 13(2). 207–224. DOI: 10.1075/dia.13.2.02ari.

Beames, John. 1966 [1872–1879]. *A comparative grammar of modern Indo-Aryan languages of India.* Delhi: Munshiram Manoharlal.

Bloch, Jules. 1965. *Indo-Aryan from the Vedas to modern times.* Translated by Alfred Master. Paris: Adrien-Maisonneuve.

Bubenik, Vit. 1998. *Historical syntax of Late Middle Indo-Aryan (Apabhraṃśa).* Amsterdam: John Benjamins.

Bubenik, Vit. 1996. *The structure and the development of Middle Indo-Aryan dialects.* Delhi: Motilal Banarsidass.

Bybee, Joan. 1994. The grammaticization of zero: Asymmetries in tense and aspect systems. In William Pagliuca (ed.), *Perspectives on grammaticalization*, 235–254. Amsterdam: John Benjamins.

Chatterji, Suniti K. 1975 [1926]. *The origin and development of the Bengali language.* Calcutta: Calcutta University Press.

Chierchia, Gennaro. 1995. Individual-level predicates as inherent generics. In Manfred Krifka, Francis J. Pelletier, Greg Carlson, Alice ter Meulen, Gennaro Chierchia & Godehard Link (eds.), *The generic book*, 176–223. Chicago: University of Chicago Press.

Clark, Eve. 1978. Locationals: A study of "existential," "locative," and "possessive" sentences. In Joseph Greenberg (ed.), *Universals of human language, vol. 4: Syntax*, 85–126. Stanford, CA: Stanford University Press.

Comrie, Bernard. 1976. *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge University Press.

Dayal, Veneeta. 1999. Bare NPs, reference to kinds, and incorporation. In Tanya Matthews & Devon Strolovitch (eds.), *Proceedings of the 9th semantics and linguistic theory conference (SALT 9)*, 34–51. Ithaca, NY: Cornell University. DOI: 10.3765/salt.v9i0.2816.

Dayal, Veneeta. 2004. Number marking and (in)definiteness in kind terms. *Linguistics and Philosophy* 27(4). 393–450.

Deo, Ashwini. 2009. Unifying the imperfective and the progressive: Partitions as quantificational domains. *Linguistics & Philosophy* 32(5). 475–521. DOI: 10.1007/s10988-010-9068-z.

Deo, Ashwini. 2012. The imperfective–perfective contrast in Middle Indic. *Journal of South Asian Linguistics* 5. 3–33.

Deo, Ashwini. 2015. The semantic and pragmatic underpinnings of grammaticalization paths: The progressive to imperfective shift. *Semantics & Pragmatics* 8(14). 1–52. http://semprag.org/article/view/sp.8.14.

Kellogg, Samuel H. 1990 [1893]. *Grammar of the Hindi language*. 2nd edn. Delhi: Motilal Banarsidass.

Kilpiö, Matti. 1993. Syntactic and semantic properties of the present indicative forms of the verb to be in Old English. In Matti Rissanen, Merja Kytö & Minna Palander-Collin (eds.), *Early English in the computer age: Explorations through the Helsinki Corpus*, 21–32. Berlin: Mouton de Gruyter.

Kratzer, Angelika. 1995. Stage-level and individual-level predicates. In Gregory N. Carlson & Francis J. Pelletier (eds.), *The generic book*, 125–175. Chicago, IL: University of Chicago Press.

Magri, Giorgio. 2009. A theory of individual-level predicates based on blind mandatory scalar implicatures. *Natural Language Semantics* 17. 245–297.

Mahapatra, Bibhuti B. 2009. *Stage-level vs. individual-level predicates and the four copulas in Odia*. Central Institute of Indian & Foreign Languages, Hyderabad, India. (Doctoral dissertation).

Masica, Colin P. 1991. *The Indo-Aryan languages* (Cambridge Language Surveys). Cambridge: Cambridge University Press.

Partee, Barbara. 1984. Nominal and temporal anaphora. *Linguistics and Philosophy* 7. 243–286.

Petré, Peter. 2013. On the distribution and merger of *is* and *bið* in Old and Middle English. *Transactions of the Philological Society* 111. 301–325. DOI: 10.1111/j.1467-968X.2012.01320.x.

Pischel, Richard. 1981 [1900]. *A grammar of the Prakrit languages.* Translated from German by Subhadra Jha. Delhi, India: Motilal Banarsidass.

Stassen, Leon. 2009. *Predicative possession.* Oxford: Oxford University Press.

Turner, Ralph L. 1936. Sanskrit *â-kṣeti* and Pali *acchati* in Modern Indo-Aryan. *Bulletin of the School of Oriental Studies* 8(2/3). 795–812.

Vale, Ramchandra N. 1948. *Verbal composition in Indo-Aryan.* Poona, India: Deccan College Postgraduate & Research Institute.

von Hinüber, Oskar. 2001. *Das Ältere Mittelindisch im Überblick.* Vienna: Österreichische Akademie der Wissenschaften.