Chapter 1

The particular–characterizing contrast in Marathi and its historical basis

Ashwini Deo
The Ohio State University

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 arguably inherited from Sanskrit. I provide historical data that indicates that while this contrast is not grammatically categorical in the older stages of Marathi (and Middle and Old Indo-Aryan), it is present through the interpretational possibilities for the bhu copula and its cognates. I use this observed categorization 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-gram1994; comrie76; deo2009c; deo2015-sp among
Languages may also choose to lexicalize or keep covert the semantic contrast between “alienable possession” and “inalienable possession” (Clark 1978; Arista 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 or a “contextual disambiguation” strategy?

In this paper, I will consider a previously un-described 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, 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 the older stages of Marathi and 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 referential 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.

(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 \text{cop1} and \text{cop2} respectively. To compare with English, note that 4a, which describes a single episode of the baby being tired contains a form of \text{cop1} āhe, contrasting with 4b where the presence of the \text{cop2} form asta signals that a recurring generalization over episodes is being described. Similar contrasts hold between the \text{cop1} and \text{cop2} sentences in 5 and 6, with 6 exemplifying the past referring copular forms.

\begin{enumerate}
\item \text{a.} bāḷ thaklel-a āhe
\begin{align*}
\text{copy} & \quad \text{tired} & \quad \text{cop1} \text{.} \\
\text{nominative} & \quad \text{nominative} & \quad \text{3rd person singular} \\
\text{baby} & \quad \text{tired} & \quad \text{be}
\end{align*}
\text{The baby is tired.}
\item \text{b.} bāḷ sandhyākāḷī thaklel-a as-ta
\begin{align*}
\text{copy} & \quad \text{tired} & \quad \text{cop2} \text{.} \\
\text{locative} & \quad \text{nominative} & \quad \text{3rd person singular} \\
\text{baby} & \quad \text{evening} & \quad \text{be}
\end{align*}
\text{The baby is tired in the evenings.}
\item \text{a.} karvāḍh-i mule lok-a dukkhi āhe-t
\begin{align*}
\text{tax.increase} & \quad \text{because} & \quad \text{people} & \quad \text{unhappy} & \quad \text{cop1} \text{.} \\
\text{oblique} & \quad \text{nom. pl.} & \quad \text{nom. pl.} & \quad \text{cop1} \text{.} & \quad \text{3rd person plural}
\end{align*}
\text{The people are unhappy because of the tax increase.}
\end{enumerate}

\footnote{The glosses used in this paper are as follows. \text{pres.part} = Imperfective Participle; \text{part} = participle; \text{impf} = imperfective aspect; \text{pst} = past tense; \text{pres} = present tense; 1 = First person; 2 = Second person; 3 = Third person; \text{nominative} = nominative; \text{acc} = accusative; \text{obl} = oblique; \text{ins} = instrumental; \text{dat} = dative; \text{gen} = genitive; \text{loc} = locative; \text{voc} = vocative; \text{ptcl} = particle; \text{m} = masculine; \text{f} = feminine; \text{n} = neuter; \text{sg} = singular; \text{pl} = plural; \text{inf} = infinitive; \text{neg} = negation marker; \text{act} = active voice marker; \text{pass} = passive voice; \text{ger} = gerund; \text{excl} = exclusive clitic; \text{poten} = potential mood; \text{inter} = interrogative particle; \text{rel} = relative pronoun; \text{correl} = correlative pronoun}
Dieters (lit. diet-doing people) are unhappy.

Rām was unhappy in the evening.

Rām used to be unhappy 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.

Consider the contrast between 7a-7b and 7c-7d. While the aspectual morphology remains the same in each pair, the interpretation of each member is clearly distinct. 7a asserts that the reference time is contained in an event of John smoking, while 7b conveys that every/most contextually relevant times (within some larger stretch of time) are contained in an event of John smoking. In 7c, the reference time is understood to be located after the time of an event of John’s making dinner and setting the table. 7d 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.

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\textsuperscript{7b} and clausal material in\textsuperscript{7d}. 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 aux\textsubscript{1} and aux\textsubscript{2}. \textsuperscript{8a} 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\textsuperscript{7b}. In fact, aux\textsubscript{1} is unacceptable in sentences that contain overt quantificational adverbs. \textsuperscript{8b}, correspondingly has only a characterizing interpretation, even in the absence of quantificational adverbs.

\begin{enumerate}
\item[a.] \textit{John} cigarette \textit{pi-t} \textit{āhe} \\
John.NOM cigarette.NOM drink-impf.part aux\textsubscript{1}.pres.3sg \\
John is smoking a cigarette. \hspace{2cm} (episodic)
\item[b.] \textit{John} (nehmi/kadhi-kadhi) cigarette \textit{pi-t} \\
John.NOM always/sometimes cigarette.NOM drink-impf.part \\
as-to aux\textsubscript{2}.pres.3m.sg \\
John is always/sometimes smoking cigarettes. \hspace{2cm} (characterizing)
\end{enumerate}

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.

\begin{enumerate}
\item[a.] \textit{John-ne} svayampāk \textit{banav-lelā} \textit{āhe} \\
John.erg meal.nom.m.sg make-perf.m.sg aux\textsubscript{1}.pres.3sg \\
John has made dinner. \hspace{2cm} (episodic)
\item[b.] \textit{John-ne} svayampāk \textit{banav-lelā} as-to \\
John.erg meal.nom.m.sg make-perf.m.sg aux\textsubscript{2}.pres.3m.sg \\
(By the time I return), John has made dinner. \hspace{2cm} (characterizing)
\end{enumerate}

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.\textsuperscript{3} The observed pattern demonstrates that semantic distinctions in the interpretations of

\textsuperscript{3} The imperfective participle+aux\textsubscript{1} periphrasis is the general exponent of the progressive aspect in Marathi. The imperfective participle optionally inflects for gender and number in this periphrasis.

\textsuperscript{4} Bangla and Oriya (mahapatra2009) 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 at-
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 that 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. Section 3 describes in detail the interpretations associated with the two paradigms. 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 obtain with individual-denoting vs. kind-denoting subjects and stage-level vs. individual-level predicates in combination with the relevant forms. 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 is only compatible with characterizing readings. In §5, I discuss the changes with respect to their effect on the overall complexity of the system and conclude.

---

tested two-copula systems. A thorough investigation of these patterns must be left for later investigation.
3 The Marathi facts

The relevant paradigms of the present and past tense copulas/auxiliaries are given in Table 1 and Table 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

<table>
<thead>
<tr>
<th>COP1</th>
<th>COP2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PL (M/F/N)</td>
</tr>
<tr>
<td>1</td>
<td>āhe</td>
</tr>
<tr>
<td>2</td>
<td>āhes</td>
</tr>
<tr>
<td>3</td>
<td>āhe</td>
</tr>
</tbody>
</table>

Table 2: Past tense copula/auxiliary paradigms of Marathi

<table>
<thead>
<tr>
<th>COP1</th>
<th>COP2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PL (M/F/N)</td>
</tr>
<tr>
<td>1</td>
<td>hoto/hote</td>
</tr>
<tr>
<td>2</td>
<td>hotās/hotis</td>
</tr>
<tr>
<td>3</td>
<td>hotā/hoti/hota</td>
</tr>
</tbody>
</table>

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 (10a), the use of cop1 conveys that the property of being busy or angry holds of Anu at the utterance time. In (10b), 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.

(10)  

(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.

(11a) and (11b) 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.

(11)  

(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.3.f.sg 
Anu is generally in the house (e.g. when the postman comes by.)

When the main predicate is individual-level and denotes a relatively permanent, intrinsic property of an individual, only cop1 is acceptable as shown in the contrast between (12a) and (12b). The use of cop2 introduces the sort of oddity that is associated with the use of quantificational adverbs with individual-level predicates (kratzer95; chierchia1995; magri2009 among others). It conveys that Anu habitually or generally has the property of being cowardly, tall, or intelligent,

---

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.

6 This is the observation that a sentence like John is generally/always/often/sometimes intelligent is understood as deviant or unacceptable without context.
which is infelicitous because it seems 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.

(12) a. \textit{anu} \textit{ghābraṭ/unca/huśār} \textit{āh-e}
\textit{Anu.NOM.F.SG cowardly/tall/intelligent COP1.PRES.3SG}  
Anu is cowardly/tall/intelligent.

b. \#\textit{anu} \textit{ghābraṭ/unca/huśār} \textit{as-te}
\textit{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 making reference to kinds. In dayal1999; dayal2004 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.

Stage-level predicates

Consider the examples in (13a) and (13b), which contain bare singular subjects and stage-level predicates. These sentences are most naturally interpreted as describing the properties of the contextually most salient dog in the utterance context – i.e. the bare nominal has a directly referential use – like an NP with the definite article in English.

(13) a. \textit{kutrā} \textit{thaklelā/bhukelā} \textit{āhe}
\textit{dog.NOM.M.SG tired/hungry COP1.PRES.3SG}  
The dog is tired/hungry.

b. \textit{kutrā} \textit{thaklelā/bhukelā} \textit{as-to}
\textit{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).

(14) 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 13 and 14, 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 15b, which contains cop2. Here, the bare singular rāṅgoli does not refer to any contextually salient entity at utterance time in the actual world, but rather to the rangoli that gets drawn everyday by Anu in front of her door.

(15)  

   Context: My friend is telling me about her sister Anu, who draws elaborate rangoli motifs in front of her house everyday. She says:

   a. anu roj rāṅgoli kāḍh-te
   Anu.NOM.F.SG everyday rangoli.NOM.F.SG draw-IMPF.PRES.3F.SG
   Anu draws a rangoli motif everyday.

   b. rāṅgoli 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.

In the given context, 15b 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 15a with 16, which contains cop1, since the bare nominal rāṅgoli, in this case, can only be taken to refer to the contextually salient rangoli at utterance time.

(16) #rāṅgoli 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.

---

7 Rangoli (Marathi rāṅgoli) 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.
The particular–characterizing contrast in Marathi and its historical basis

In addition to the directly referential and non-referential readings described above, bare singular subjects may also be understood as kind-denoting. For instance, (17) 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.).

(17)  ghadyāḷ  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 (17) 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, not to a particular entity in the utterance context.

(18)  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…”

ghadyāḷ  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. (19) 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. But it can also be read as a claim about a contextually

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.
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.

(19) \(kāmgār\ asantuṣṭa\ āhe-t\)

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 \(\text{cop2}\) and a stage-level predicate, as in (20), 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.

(20) \(kāmgār\ asantuṣṭa\ as-tāt\)

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.

**Individual-level predicates**

With individual-level predicates and \(\text{cop1}\), both bare singualrs 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.

(21) a. \(kāmgār\ ghābrat/unca/huśār\ ahe\)

worker.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. \(kāmgār\ ghābrat/unca/huśār\ ahe-t\)

worker.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 yesteryears) as being intelligent – 22a. Similarly, 22b, 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.

(22)  

a. āj-cā kāmgār jāsti huśār ahe  
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).

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 23.

(23)  

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

---

9 Of course, the definite referential reading is available for both examples in 21. 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.
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 12b – 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. 24 and 25 contain examples of the contexts in which the definite non-referential and kind readings of bare nominals arise respectively.

(24) **Context:** One/two of the smarter worker(s) is/are assigned to work overtime 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:

\[
\text{kāmgār } \text{huśār as-to/as-tāt } \text{mhaṇun} \\
\text{worker.NOM.M.SG/PL intelligent COP2-PRES.3M.SG/PL therefore} \\
\text{kām } \text{lavkar āṭap-t-a} \\
\text{work.NOM.N.SG fast finish-IMPF.PRES.3N.SG}
\]

*definite non-referential:* The worker(s) (that get assigned to the job) is/are intelligent (smart) (and) so the work gets done faster.

(25) **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

\[
\text{kāmgār } \text{prāmāṇik as-tāt } \text{mhaṇun tumhi suddhā} \\
\text{worker.NOM.M.PL honest COP2-PRES-3M.PL therefore you also} \\
\text{prāmāṇik as-āva} \\
\text{honest be-POT.N.SG}
\]

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

---

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.
The particular-characterizing contrast in Marathi and its historical basis

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.

<table>
<thead>
<tr>
<th>Subject NP</th>
<th>cop1</th>
<th>cop2</th>
</tr>
</thead>
<tbody>
<tr>
<td>names</td>
<td>SLP+cop1</td>
<td>SLP+cop2</td>
</tr>
<tr>
<td>episode in actual world</td>
<td>property in actual world overlapping with UT</td>
<td>generalization over episodes across time (invariant NP referent)</td>
</tr>
<tr>
<td>at UT</td>
<td></td>
<td>induces oddity</td>
</tr>
<tr>
<td>bare nominals</td>
<td>ILP+cop1</td>
<td>ILP+cop2</td>
</tr>
<tr>
<td>D-ref.</td>
<td>episode in actual world overlapping with UT</td>
<td>generalization over episodes across time (invariant NP referent)</td>
</tr>
<tr>
<td>unavailable</td>
<td>unavailable</td>
<td>generalization over properties of individuals across time (variable NP referent)</td>
</tr>
<tr>
<td>bare nominals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D-non-ref.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>unavailable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bare nominals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kind-ref.</td>
<td>episode involving the kind in actual world at UT</td>
<td>property of the kind in the actual world overlapping with UT</td>
</tr>
</tbody>
</table>
What is immediately apparent through the table is that the type of copula influences the range of readings available to bare nominal subjects. Specifically, \( \text{cop1} \) 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). \( \text{cop2} \), 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 \( \text{cop1} \) sentences seems to depend on the valuation of the embedded predication at utterance time in the actual world, while the meaning of \( \text{cop2} \) sentences seems to require consideration of the valuation of the embedded predication at times beyond the utterance time. In other words, the morphosyntactic devices \( \text{cop1} \) and \( \text{cop2} \) 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 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 (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 – \( \text{as} \) (PIE *\( h_1 \)es) and \( \text{bhu} \) (PIE *\( b^h \)uh\(_2\)). 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. For instance, \( \text{as} \) 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. \( \text{bhu} \), in contrast, as a copular expression, only appears in predicational clauses and in those structures, is only compatible with non-referential, charac-
The particular-characterizing contrast in Marathi and its historical basis

Notice that this distribution in which one copular element is restricted to non-referential characterizing readings yields a partially articulated contrast between particular and characterizing claims. Consider the examples below from the Mahabharata, one of the two Epic Sanskrit texts, that illustrate this distribution.

(26)  
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  kevalamānuś-ah  
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-ṣām  
knowledge.ascetism.possessed-nom.m.pl  those.correl.gen.m.pl  
vināśa-ḥ  prathamam 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 26a and 26b, 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. 26c, 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.

---

11 bhu 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.
12 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.
In [27a] and [27b], both existential clauses, we see that the \textit{as} copula is used to make characterizing claims. \textbf{27a} makes the generalization that kingless-kingdoms have neither rain nor Gods while \textbf{27b} asserts the existence of cowardly and brave men across different indices of evaluation – not only at utterance time.

\begin{enumerate}[label=(\alph*)]
\item \textit{arājak-eṣu rāṣṭr-eṣu na as-ti vrṣṭi-r na}
\begin{flushright}
king\textless\text{LOC}\text{.PL kingdom\textless\text{LOC}\text{.PL NEG as\textless\text{PRES}\text{.3}SG rain\textless\text{NOM}\text{.SG NEG devatā-ḥ}}
\end{flushright}
\begin{flushright}
God\textless\text{NOM}\text{.PL}
\end{flushright}

In kingdoms without a king, there \textit{is} no rain and no Gods. (Mbh. 1.99.41a)
\item \textit{s-anti vai puruṣ-āḥ śūr-āḥ s-anti as-PRES.3.PL PTCL man-NOM.PL brave-NOM.PL as-PRES.3.PL kāpuruṣ-ās tathā coward-NOM.PL likewise \begin{flushright}
There are brave men, and likewise those that are cowards. (Mbh. 5.3.2a)
\end{flushright}
\end{enumerate}

When we study the distribution of the \textit{bhu} copula, we find that its stative uses only give rise to characterizing readings, as in the examples in [28]. \textbf{28a} describes what is said in the code – a guideline to be followed not just at utterance time but more generally. \textbf{28b} describes the defining properties of the Rākṣasa women kind and the clause containing the \textit{bhu} copula predicates the property of being many-formed to the kind.

\begin{enumerate}[label=(\alph*)]
\item \textit{sakhībhartā hi dharm-eṇa bhartā}
\begin{flushright}
friend\textless\text{husband}\textless\text{NOM}\text{.M.SG PTCL code\textless\text{INS}\text{.M.SG husband\textless\text{NOM}\text{.M.SG bhava-ti śobhane bhu-PRES.3.SG beautiful.VOC SG \begin{flushright}
O beautiful one, the husband of a friend, according to the code, \textit{is} also one’s husband. (Mbh. 1.78.20c)
\end{flushright}
\item \textit{sadyo hi garbha-ṃ rākṣas-yo labh-ante}
\begin{flushright}
immediately PTCL embryo\textless\text{ACC}\text{.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 bahurūpin-āḥ bhu-PRES.3.PL many.formed-NOM.F.PL \begin{flushright}
The Rākśasa women give birth the very day they conceive, and being
\end{flushright}
\end{flushright}
\end{enumerate}
able to assume any form at will, they are many-formed. (Mbh.1.143.32a-c)

What is crucial is that there are no uses of the bhu copula in clauses that describe a state determined by the valuation of the embedded predication at utterance time in the actual world – all uses occur in clauses that convey generalizations. That is, 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 (bloch65; beames66; bubenik98; bubenik96; pischel00; vale48; masica91 and others). Of the several changes in the expression of tense-aspect-modality distinctions, critical is the loss of a morphosyntactic distinction between the past and present tenses. In deo2012 I provided a detailed description of the re-organized aspectually based system in which temporal reference is established through contextual cues. A few examples in 29 and 30 illustrate the basic pattern, which forms the backdrop for an innovation in Late Middle Indo-Aryan. In 29a, the Old Indo-Aryan Present paradigm, glossed IMPF, has present reference, while in 29b, the same paradigm has past reference.

13 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. pischel00 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.

14 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 Participle 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.
(29) a. nipphala-ṃ duma-ṃ pakkhiṅ-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-vādiyā-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 30, 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. 30a describes a past event using perfective marking while 30b and 30c also with past reference, describe past activities using the temporally unmarked imperfective.

(30) a. te-ṇa palāyamāṇ-en-ena purāṇakuv-o
that-ERG.M.SG running-ERG.M.SG old.well-NOM.M.SG
tanadabbhaparichinn-o diṭ-ṭho
glass-covered-NOM.M.SG notice-PERF.M.SG
That running one noticed an old well covered with 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

15 Summarizing the core distinctions made in the Middle Indo-Aryan aspec-to-temporal system:

<table>
<thead>
<tr>
<th>non-future</th>
<th>future</th>
</tr>
</thead>
<tbody>
<tr>
<td>imperfective</td>
<td>perfective</td>
</tr>
<tr>
<td>IMPF OR PRES.PART</td>
<td>PERF</td>
</tr>
<tr>
<td>neutral</td>
<td>sigmatic future</td>
</tr>
<tr>
<td>PERF</td>
<td>PERF</td>
</tr>
</tbody>
</table>

The Early New Indo-Aryan aspec-to-temporal system is systematically built up from this aspec-tual core through periphrastic constructions based on tense auxiliaries, already visible in Middle Indo-Aryan (bubenik96; bubenik98; kellogg93; beames66; chatterji70 a.o.).
The particular-characterizing contrast in Marathi and its historical basis

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

b. sappā bhisaṇā āsiukāmā
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

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 and the introduction of a new tense auxiliary, based on the verb *acch* ‘sit’ that establishes temporal reference with respect to speech time (turner1936). 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.¹⁶

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

In Late Middle Indo-Aryan, this auxiliary appears in copular clauses as illustrated below in (31). All examples here come from the *Paumacariu* of Swayambhudeva, a key Apabhraṃśa verse text from ~ 800CE.

(31) a. deva deva ki-u j-ena mahāra
Lord.voc.sg do-perf.m.sg that.correl-INS.M.SG great.sound.NOM.M.SG
acch-ai mattahatthi airāva
acch-PRES.3.SG musth.elephant.NOM.SG Airavata.NOM.SG

Lord, the one who made a great sound, he is the elephant Airavata in

¹⁶ 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.
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 bhu copula retains its distributional properties from Old Indo-Aryan — it only gives rise to non-referential characterizing readings in its stative uses as the examples in 32 illustrate.

(32) 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
Mother-in-laws 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 kim 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)

The upshot here is 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 leads to reorganization of the copular/auxiliary system. The original Old Indo-Aryan system did have a dedicated device for making non-referential characterizing claims (the bhu copula) but the as copula was underspecified and could be used in both particular and characterizing senses. It is the innovation in the Late Middle Indo-Aryan system that facilitates the “hardening” of a soft contrast between particular and characterizing claims that is ubiquitous in the New Indo-Aryan languages. 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, paradigms get recruited to realize the same semantic contrast. A full treatment of the rise of tense auxiliaries
The particular-characterizing contrast in Marathi and its historical basis

in Indo-Aryan that anchor the embedded predication to the utterance world and contextually salient reference intervals must be left to future research.

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 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 under-specified and contextual cues are critical to the retrieval of information regarding their temporal reference. The development of the copular contrast de-
scribed 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. 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 two PIE “be” verbs – *is* (PIE *h₁es*) and *bið* (PIE *bʰuh₂*). petré2013 (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 33 illustrate the functional distribution of *bið* in Old English.

(33) a. Hit byð dysig þaet man speca aer þone he þaence
   It is foolish that man speak.SUBJ ere then he think.SUBJ
   It be foolish that a man speaks before he thinks. (cl100. Prov 1 [Cox]: 2.2, via Petré 2013)

17 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.

18 Petré refers to a quantitative analysis by kilpio1993 which suggests that the pattern is tendential rather than categorical – there is a soft rather than categorical contrast. According to Kilpio, *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%.
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 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.