Journal of South Asian Linguistics

Contents

Editorial Statement

We are very happy to present to you the fifth volume of the Journal of South Asian Linguistics. The volume consists of two papers: the paper by Ashwini Deo on aspectual contrasts in Middle Indic and the paper by Vasishth et al. on information structure and word order in Hindi-Urdu. These papers exemplify a fruitful trend in contemporary linguistics which combines traditional tools of linguistic analysis with data from a wider range of empirical domains. In the case of the first paper, the data comes from historical texts and in the second, the data comes from eyetracking studies. The combination of formal linguistic tools with new sources of data leads both papers to conclusions that challenge the existing consensus.

A careful examination of the semantics of tense/aspectual forms in Middle Indic leads Deo to the surprising conclusion that the forms hitherto analyzed as marking tense should actually be analyzed as marking aspect. This means that the aspect based system of Modern Indic goes back all the way to Middle Indic. Vasishth et al. study the role of clefting, word order and information structure in Hindi-Urdu using data from eyetracking. They find that the processing complexity introduced by syntactically complex structures such as clefts and non-canonical word orders can be mitigated by a processing preference for given-new orders over new-given orders. This leads the authors to the result that non-canonical word orders are not inherently harder to process; in fact in the appropriate context, non-canonical word orders can be processed more easily than canonical word orders.

Unfortunately, the 2012 issue does not include a descriptive paper. We would therefore like make a point of highly encouraging potential submissions that focus on a hitherto less described language and discuss interesting structural properties of this language.

Finally, we would like to thank Sebastian Sulger for continuing to maintain the journal website and Julika Winter for helping with the production of the volume. As ever we also thank Dikran Karagueuzian of CSLI Publications for his role as a continually outstanding and supportive publisher.

Rajesh Bhatt, University of Massachusetts, Amherst
Miriam Butt, University of Konstanz
Tracy Holloway King, eBay Inc.

The imperfective-perfective contrast in Middle Indo-Aryan

ASHWINI DEO, *Yale University*

ABSTRACT

This paper examines the distribution of two morphological paradigms inherited from Old Indo-Aryan in Middle and New Indo-Aryan languages – the Old Indo-Aryan Present (labeled PRES) and the Past Participle (labeled PERF). It is argued that these forms, contra standard assumptions, do not realize the present and past tenses, but rather the imperfective and perfective aspects with no tense specification. This hypothesis provides an explanation for the puzzling occurrences of the Present and the Past Participial forms with past and future reference in Middle Indo-Aryan. It also makes sense of some distributional patterns of these paradigms in New Indo-Aryan. This, in turn, supports the idea that the Middle Indo-Aryan proto-system that gave rise to the New Indo-Aryan languages was an aspect-based system with no present-past distinction.

1 Introduction

Sanskrit morphologically marks the contrast between the present and the past, and so do the New Indo-Aryan languages. The former system, however, is inflectional, while the latter is characterized by tense auxiliaries in combination with (mostly) non-finite perfective and imperfective forms. The basis of the innovated New Indo-Aryan verb system must lie in the properties of some intermediate system between New Indo-Aryan and Old Indo-Aryan — most plausibly, that of Middle Indo-Aryan. The goal of this paper is to investigate how the organization of the Late Middle Indo-Aryan tense-aspect system might bear on the analytic marking of tense and aspect in the New Indo-Aryan languages. The particular hypothesis that I explore is that Late Middle Indo-Aryan, unlike Old Indo-Aryan, did not morphologically contrast the present and the past tenses. Rather, the Late Middle Indo-Aryan system morphologized only the imperfective-perfective contrast in the non-future domain, the same aspectual contrast which underlies the innovated tensed periphrastic systems of most New Indo-Aryan languages.

The particular paradigms involved in this reorganization are the bold-faced forms in (1-a) and (1-b), which belong to the Old Indo-Aryan Present paradigm and the Past Participial paradigm respectively. The examples in (1) contain their cognates from the Middle Indo-Aryan period and come from the 6th century text, the *Vasudevahiṃḍī*. Throughout the paper, despite changes in their distribution, these two paradigms will be labeled and glossed PRES and PERF respectively in order to be able to track them over time.[1]

---

[1]Note that whenever a category represented by a gloss is capitalized, it refers to the standard name for an Indo-Aryan morphological paradigm with Indo-European cognates, and whenever it is not capitalized, the gloss stands for universal grammatical categories. PRES = Old Indo-Aryan Present; PERF = Old Indo-Aryan Perfective Passive Participle; PFCT = Old Indo-Aryan Perfect; IPFCT = Old Indo-Aryan Imperfect; AOR = Old Indo-Aryan Aorist; FUT = Old Indo-Aryan Future; PART = Imperfective Participle; IMPF = imperfective aspect; PST = past tense; PRS = present tense; IMP = imperative mood; CAUS = causative; 1 = first person; 2 = second person; 3 = third person; NOM = nominative; ACC = accusative; INS = instrumental; DAT = dative; GEN = genitive; LOC = locative; VOC = vocative; PTCL = particle; M = masculine; F = feminine; N = neuter; SG = singular; DU = dual; PL = plural; INF = infinitive; NEG = negation marker; ACT = active voice marker; PASS = passive voice; GER = gerund; EXCL = exclusive clitic.

(1) a. nipphala-ṃ   duma-ṃ   pakkhiṇ-o   vi **paricchaya-nti**
         fruitless-ACC.SG tree-ACC.SG bird-NOM.PL also abandon-PRES.3.PL
         Even birds *abandon* a fruitless tree. (VH.DH 31.24-25)
     b. **pat-to**   ya Seṇiyo rāyā   ta-m   paesa-m.
         reach-PERF.M.SG and S.NOM.SG king.NOM.SG that-ACC.SG place-ACC.SG
         And King Seniya *reached* that place. (VH.KH. 17.1)

Because they constitute the basic compositional building blocks of the early New Indo-Aryan system, it is crucial to establish the status of the PRES and PERF paradigms within the Middle Indo-Aryan tense/aspect system, i.e. the system that gives rise to the tensed New Indo-Aryan grammar. The claim advanced here is that the distribution and interpretation of the Middle Indo-Aryan forms in (1-a) and (1-b) and their cognates in New Indo-Aryan languages, best supports their analysis as aspectual operators rather than tense operators. That is, the verb form in (1-a) realizes imperfective aspect, rather than the present tense, while the verb form in (1-b) realizes the perfective aspect, rather than the past tense.

Within Indo-Aryan historical linguistics, this claim (especially regarding the paradigm represented by (1-a)) would be considered surprising. The implicit and explicit assumption in the Indo-Aryan literature has been that tense is a consistently expressed morphological category across Indo-Aryan diachrony. Although the loss of individual tense/aspect markers and paradigms of Sanskrit in Middle and New Indo-Aryan has been carefully documented in Indo-Aryan historical grammars (Pischel 1900; Beames 1872–79; Bloch 1914, 1965; Chatterjee 1926; Kellogg 1893; Singh 1980; among others), this loss has mostly been understood in terms of the loss of specific paradigms, rather than in terms of the reorganization of the larger tense/aspect system along aspectual lines. Masica (1991:262) observes that the category of aspect is at the heart of the New Indo-Aryan verbal system, citing Lienhard (1961:27) who suggests that the rebuilding of the New Indo-Aryan system proceeds by establishing aspectual distinctions, to which the refinements of tense (and mood) were only later added. While this view is correct, it must be pointed out that it has not been determined exactly when and how such an aspectual system emerges in Indo-Aryan diachrony. Moreover it has not been debated whether the verbal system at every stage of Indo-Aryan morphologically distinguishes between the present and the past tenses. This paper offers a reinterpretation of the Late Middle Indo-Aryan and Early New Indo-Aryan facts: these point to a broad trajectory from the overt realization of tense operators in Old Indo-Aryan, to the loss of overt tense marking in Late Middle Indo-Aryan and its later re-emergence in the form of innovated present and past tense auxiliaries in New Indo-Aryan. Crucially, the aspectual distinctions that are at the heart of the New Indo-Aryan system, as Masica (1991) says, are structurally inherited from the Late Middle Indo-Aryan system.

A brief note about future tense marking is in order here. Middle Indo-Aryan and some New Indo-Aryan languages (e.g., Gujarati, Marwari) inherit the sigmatic future paradigm from Sanskrit. Others innovate (e.g., Marathi, Bengali) future marking from periphrastic constructions and Sanskrit non-finite modal forms. However, I leave inherited and innovated future morphology out of the discussion in this paper, given the metaphysical and epistemological asymmetry between non-future and future meanings more generally. While the past and (to some degree) the present are factual and decided, any assertion about the future is accompanied with some degree of indeterminacy.[2] Correspondingly, future morphology is distinct from present/past morphology in that the former typically expresses both tense and modality. The present paper, therefore, restricts its scope to the past/present temporal opposition, leaving the integration of future marking to later research.

For readers unfamiliar with the basic diachrony of the Indo-Aryan languages, the table in (2) provides the timeline. The scope of this paper is restricted to the later Middle Indo-Aryan period (i.e. not the earlier dialects of Pāli or Ashokan Prakrits) since my primary concern here is to determine the tense/aspect properties of the proto-system that underlies the New Indo-Aryan grammars. Although a much more nuanced examination of the distinct diachronic layers of the Middle Indo-Aryan languages is essential to establishing the loss of the

---

[2]Futurity has to do with plans, intentions, obligations, and predictions, notions which all have to do with mood and modality and are inherently non-factual. Empirically, this raises the question of whether *any* future marking serves purely the purpose of expressing temporal oppositions within a language. Crosslinguistic surveys have revealed that forms that realize future time reference are often used atemporally and have functions associated with mood and modality, such as possibility or probability (e.g., Kiwai, Zapotec), intention (e.g., Garo, Zapotec, Pawnee), desire or volition (e.g., Goajiro, Quileute) (Ultan, 1978; Dahl, 1985; Bybee et al, 1994).

morphological category of tense in Middle Indo-Aryan, this is beyond the scope of the current paper. My goal here is simply to demonstrate that if we allow for the assumption that the Old Indo-Aryan tensed system, contrasting the present and the past tenses, was reorganized in Middle Indo-Aryan as an aspectual system, contrasting the imperfective and the perfective aspects, we have a better explanation for the distribution and interpretation of certain inherited tense/aspect forms of Late Middle Indo-Aryan and their cognates in New Indo-Aryan. Further, this assumption accounts for the innovation of the tense auxiliary based periphrastic paradigms of the New Indo-Aryan languages; these innovation patterns are mysterious if the verbal forms of the parent system are assumed to already encode tense information.

(2) **The Chronology**

TIMELINE	STAGE	LANGUAGE	SOURCE USED
1900 BCE–1100 BCE	I	**Early Vedic**	Ṛgveda (RV)
1000 BCE–350 BCE	I	**Late Vedic**	
100 BCE–400 CE	II	**Epic and Classical Sanskrit**	Mahābhārata, Rāmāyaṇa
300 BCE–500 CE	III	**Middle Indo-Aryan**	Vasudevahiṃdi (VH 500CE)
500 CE–1100 CE	III	**Apabhraṃśa**	Paumacariu (PC 800CE)
1100 CE–present	IV	**New Indo-Aryan**	native intuitions, fieldwork, grammars

(Approximate dates from Witzel (1999), Jamison and Witzel (2002), Alsdorf (1936)) [3]

The organization of the paper is as follows. In §2, I introduce the notions of tense and aspect in the context of tenseless languages. §3 briefly describes the relevant Old Indo-Aryan temporal paradigms, which demonstrate the well-established fact that Old Indo-Aryan contrasts the present and the past tenses morphologically. In §4, I provide evidence that the forms associated with the present and past tense categories in Middle Indo-Aryan are better interpreted as exponents of the imperfective and perfective aspects respectively. §5 provides further evidence from New Indo-Aryan languages in support of this claim.

2 Tense and Aspect

The terms tense and aspect are used in at least two distinct ways in typological and semantic literature. As *morphological categories*, they refer to grammaticalized, obligatorily encoded distinctions that express temporal properties of situations. As abstract *semantic categories*, they refer to temporal properties of propositions that may or may not have a morphological reflex in a given language. The semantic categories are universal in that they are implicated in both the grammar and the discourse of many languages and also find robust morphosyntactic expression in several unrelated languages. In the context of this paper, the claim is that Middle Indo-Aryan lacks tense as a *morphological category* and the information corresponding to the *semantic category* of tense is obtained via contextual cues and aspectual morphology.

Tense, on the traditional view, is a deictic temporal category that involves a precedence relation between the time of the situation described by a sentence and some deictic center, most often the speech time. Language-specific tense expressions are grammaticalized markers that constrain the location of situations in time with respect to the deictic center. Given an utterance, overt tense markers restrict its temporal reference, i.e. the time at which the eventuality description it contains can be understood to hold. Languages like English and Hindi contain obligatorily present overt morphological material that restricts the temporal reference of an utterance. Neither the English sentence in (3-a) nor the Hindi sentence in (3-b) can be interpreted as making reference to a time overlapping with or following the time at which they are uttered.

(3) a. John lov-**ed** Mary.
    b. Rām   Rādhā-ko cāh-tā   **thā**
       Rām.NOM Rādhā-ACC love-IMPF.M.SG PST.M.SG
       Rām loved Rādhā.

---

[3]These are approximate periods and the first of these, especially, only gives the broad window within which Northern and North-west India were settled (Jamison and Witzel 2002, p.6).

In tenseless languages, which lack such grammaticalized restrictors of temporal reference, discourse context, optional temporal adverbs, as well as the aspectual properties of predicates may serve to determine the temporal reference of utterances. A substantial literature has accumulated over the past decade investigating closely how grammatical, lexical, and contextual factors contribute to establishing present, past, and future reference in such languages (Bohnemeyer 2002; Bittner 2005, 2008; Lee & Tonhauser 2010; Tonhauser 2011). For instance, in a language like Yucatec Maya, (4-b) would be an appropriate answer to any one of the questions in (4-a). The reference time of (4-b) is thus constrained, not by overt tense morphology, but rather by contextual information or by the presence of optional overt temporal adverbs as exemplified in (4-c).

(4)   a.   What did you do yesterday?/What are you doing right now?/What will you do tomorrow?
       b.   Táan inw=óok'ot
            PROG A1SG=dance.INC
            I was/am/will be dancing.
       c.   Ho'leake' /Be'òoraa' /Sàamle' táan   inw=óok'ot
            yesterday now   tomorrow PROG A1SG=dance.INC
            Yesterday/Now/Tomorrow I was/am/will be dancing.[4]   (Tonhauser 2011)

Aspect is a term that has been employed to refer to properties relating to eventuality structure — lexical aspect, inner aspect, actionality, aktionsart, on the one hand, and to temporal relations on the other — grammatical or viewpoint aspect. The notion of lexical aspect makes reference to the telic–atelic opposition which is sensitive to whether a predicate has divisive/cumulative reference (atelic) or not (telic). Grammatical aspect has been construed as describing relations between the time at which a situation holds and some salient reference time without making any direct reference to the utterance time.[5] The fundamental opposition between imperfective and perfective aspect is between bounded and unbounded predicates (i.e. whether the reference interval contains the eventuality interval or whether it is contained in it).

The aspectual and temporal reference of clauses is determined by the interaction of their lexical aspect properties and overt morphological marking. In languages with a grammaticalized imperfective-perfective contrast, aspectual morphology obligatorily constrains the aspectual reference of clauses regardless of the telicity of the predicate in its scope, while in a language with grammaticalized tenses, tense morphology obligatorily constrains the temporal reference of clauses regardless of the (im)perfectivity of the predicates in its scope. More interesting is the interaction in languages that lack tense or aspect marking or both. First, in languages without grammatical aspect marking, lexical aspectual properties provide inferential defaults for grammatical aspectual reference (Smith 1991, 2008). Thus, in aspectually unmarked sentences, telic predicates are interpreted by default as perfective while atelic predicates are interpreted imperfectively (e.g., Bohnemeyer & Swift 2004 for German, Inuktitut, and Russian). Further, in the absence of tense marking, perfective predicates are by default interpreted as making reference to completed eventualities in the past while imperfective predicates tend to be interpreted as overlapping with speech time (Smith 2008; Bohnemeyer 2002 for Yucatek Maya; Shaer 2003 for West Greenlandic; Bittner 2005, 2008 for Kalaallisut). Contextual information, however, can always override these interpretive defaults.

Yet another way in which aspectual distinctions are implicated in temporal reference has to do with temporal anaphora. It has been recognized since Kamp (1979), Partee (1984) and Hinrichs (1986) that perfective (eventive) and imperfective (stative) sentences interact differently with reference time in narrative discourse: events occur within the reference time established in discourse, while states hold at the reference time. Aspect also affects the update of reference time. In narratives, sentences containing eventive predicates show a strong tendency to "push" reference time forward, so that the following sentence tends to be understood to hold at a later reference time, while sentences containing stative predicates do not do so.

This paper employs these basic ideas about how temporal reference can be conveyed in the absence of tense in order to establish that the past and present tenses are not morphosyntactically expressed categories in Middle Indo-Aryan. The larger goal, that of establishing the various ways of conveying temporal reference

---

[4]Glosses used in (4): A1SG = set A first person singular, PROG = progressive aspect, INC = incompletive status.

[5]Reference time (Reichenbach 1947) or Topic time (Klein 1994) refers to the temporal interval that is under discussion at any given point in a discourse.

in Middle Indo-Aryan, is far beyond the scope of this paper and must await further research. Before delving into the Middle Indo-Aryan data, §3 contains a brief description of the Old Indo-Aryan tense/aspect system, one which does morphosyntactically contrast past and present tenses.

3 The present-past distinction in Old Indo-Aryan

The Old Indo-Aryan verbal system consists of several paradigms marking distinct intersections of temporal, aspectual, and modal categories (Delbrück 1888; Whitney 1889, 1892; a.o.). The discussion here is restricted to the present and past tense forms of the indicative mood. The distribution of the relevant forms is summarized in Table 1. Each italicized form is the third person singular form of the tense-aspect configuration that it represents for the verb *gam* 'go'. The term below the form lists the label for the paradigm in the Indo-European tradition. The cell that a form occurs in indicates how the distribution of that paradigm may be best (although not perfectly) classified in terms of language-neutral semantic categories. Neutral aspect indicates that the paradigm is not aspectually specified and is compatible with both perfective and imperfective readings (despite the misleading name for the neutral past tense form — the Imperfect).

TENSE	ASPECT			
	neutral	**imperfective**	**perfective**	**perfect**
present		*gaccha-ti* Present		
past	*a-gaccha-t* Imperfect		*a-gā-t* Aorist	*ja-gā-ma* Perfect

TABLE 1 Present and Past tense forms in Old Indo-Aryan

3.1 Vedic

3.1.1 The Old Indo-Aryan Present

The Old Indo-Aryan Present paradigm (glossed PRES here and throughout the paper despite changes in its interpretation) morphologically realizes the present tense in Vedic. It is aspectually imperfective and gives rise to event-in-progress and continuous stative readings. In (5-a), the PRES form refers to an ongoing episode of axe-sharpening, temporally located by the adverbial *nūnám* 'now'. In (5-b), PRES marking occurs with a lexical stative predicate.

(5) a. **śíṣī-te**   nūnám paraśú-ṃ   suāyasá-ṃ
         sharpen-PRES.3.SG now   axe-ACC.SG iron-ACC.SG
         Now, he *is sharpening* his axe, made of iron. (RV 10.53.9c)

     b. tvám   hy àgn-e   divyá-sya   **rā́ja-si**
         you.NOM.SG PTCL A-VOC.SG heaven-GEN.SG reign-PRES.2.SG
         Agni, you (are the one who) *reigns* over the heaven. (RV 1.144.6a)

3.1.2 The Old Indo-Aryan Imperfect

Cognate to the Greek and Latin Imperfect, the Old Indo-Aryan Imperfect inflection consists of the prefix-like augment *a* marking past temporal location and the secondary person-number suffixes. Like in ancient Greek and Latin, the Imperfect realizes past tense but unlike them, it is aspectually neutral and allows for both imperfective (6-a) and perfective (6-b) interpretations.

(6) a. vŕ̥ṣno   vádhri-ḥ   pratimā́nam bū́bhūṣan   purutrā́   vr̥tró
         virile.GEN.SG emasculated-NOM.SG counterpart wanting to be everywhere V.NOM.SG
         a-śay-at   víasta-ḥ
         lie-IPFCT.3.SG dismembered-NOM.SG
         Emasculated yet wanting to be virile, thus Vrtra *lay* with scattered limbs dismembered.
         (RV 1.32.7.c-d)

b. **á-han**   áhi-m   ánu apás   tatard-a   prá vakṣáṇā
kill-IPFCT.3.SG dragon-ACC.SG up   water-ACC.PL open-PFCT.3.SG forth cavity.ACC.PL
   a-bhina-t   párvatā-nām
   cut-IPFCT.3.SG mountain-GEN.PL
   He *slew* the Dragon, then *opened up* the waters, and *cut* cavities through the mountains.
   (RV 1.32.1c-d)

3.1.3   The Old Indo-Aryan Aorist

For Proto-Indo-European, the Aorist has been reconstructed as the marker of perfective aspect in opposition to the imperfective Imperfect. Within Vedic, it has been notoriously difficult to establish this contrast based on the uses of the Imperfect and Aorist (Gonda, 1962:258-261; Delbrück, 1888; Hoffman, 1967).[6] As we have seen, the Imperfect has both imperfective and perfective readings which is why it is analysed as an aspectually neutral past tense in Vedic. The distribution of the Aorist is complex but it most frequently denotes culminated, completed events located in the past time as in (7-a.). The Aorist is also used in referring to an immediate past time, paralleling the 'recent past' use of the English Perfect, where the event denoted by the base predicate is interpreted as having occurred just before speech time (7-b).

(7)   a.   **nā́-tarī-d**   asya   sámṛti-ṃ   vadhā́nām
         NEG-bear-AOR.3.SG he-GEN impact-ACC.SG weapon-GEN.PL
         He *did not withstand* (failed to withstand) the impact of his weapons.
         (RV 1.32.6; translation from Kiparsky 1998:ex.3a)

      b.   idā́ hí   vo   dhiṣáṇā devy   áhn-ām   **á-dhāt**   pītí-ṃ
         now PTCL you.DAT.SG D.NOM.SG goddess.NOM.SG day-GEN.PL set-AOR.3.SG drink-ACC.SG
         sám   mádā   **a-gma-tā**   vaḥ
         together gladdening go-AOR.3.PL you.ACC.SG
         This day, now, the Goddess Dhiṣaṇā *has set* forth the drink for you. The gladdening draughts
         have reached (united with) you. (RV 4.34.1.c)

3.1.4   The Old Indo-Aryan Perfect

The Perfect, like the Aorist, is reconstructible for Proto-Indo-European as an aspectual category with result stative value (Renou 1925).[7] With a class of predicates (e.g., *vid* 'know', *cit* 'think', *sthā* 'stand'), the Perfect has a result stative interpretation and may be coordinated with the Present, which has present time reference. The perfect form of the verb *bhi* 'fear' is *bibhāya* and it is used in this context to refer to the state of having become scared, which holds at reference time (the present).

(8)   ká   īśa-te   tujyá-te   kó **bibhāya**
      who flee-PRES.3.SG rush-PRES.3.SG who fear-PFCT.3.SG
      Who *is fleeing* and *rushing*, who is afraid? (RV 1.84.17; translation from Kiparsky 1998:ex.6)

However, the Perfect also has a past eventive reading as illustrated by the examples in (9).

(9)   a.   ā́ **dad-e**   vas   trī́-n   yukt-ā́n
         to give-PFCT.1.SG you-DAT.PL three-ACC.PL yoked-ACC.PL
         I *received* three (chariots) in harness for you.
         (RV 1.126.5.a-b; translation from Kiparsky 1998:ex.14a)

---

[6]Hoffman (1967) has been able to demonstrate that within the sub-system of prohibitive injunctives, the augmentless Imperfect and Aorist forms correspond to an imperfective and perfective interpretation respectively, suggesting that the original PIE contrast might be visible only in this sub-system at the Vedic stage.

[7]For a full description of the uses of the Perfect, I refer the reader to Renou (1925) which is devoted to the Vedic Perfect and a more concise summary in Kiparsky (1998).

b.   urú   kṣáyā-ya   **cakrir-e**
    wide.ACC.SG dwelling-DAT.SG make-PFCT.3.PL
    [They conquered heaven, earth, and the waters] They *made* themselves a wide homeland.
    (RV 1.36.8.a-b; translation from Kiparsky 1998:ex.14c)

3.2 Epic Sanskrit

There are two main points of distinction between Vedic and the later Old Indo-Aryan Epic Sanskrit stages in the categories for present and past time reference according to existing grammatical descriptions.

a. The Imperfect, the Aorist, and the Perfect may be used interchangeably for past time reference and often occur together (Oberlies, 2003:152-154; Brockington, 1998:352; Speijer, 1886). The distinction between the Present and the three past tenses is nevertheless maintained.[8]

b. A new participial form — the PERF form — becomes available for referring to past, culminated events.

3.2.1 The PERF form in Epic Sanskrit

PERF is the label for the Indo-Aryan cognate of Proto-Indo-European deverbal, resultative, adjectival (participial) morphology with the *-to/*-no affix that attaches to verb roots.[9] This morphology is not part of the finite verbal paradigm of Vedic (which inflects for person-number-mood features) but overlaps with the verbal system due to its aspectual properties. In Vedic, the PERF morphology is stative and realizes the resultative aspect as seen in the example in (10).[10]

(10) **stīr-ṇáṃ**   te   barhí-ḥ   **su-tá**   indra   sóma-ḥ
      strew-PERF.N.SG you.DAT.SG B-NOM.N.SG extract-PERF.M.SG I.VOC.SG S-NOM.M.SG
      kr̥-tā́   dhān-ā́   át-tave te   hári-bhyām
      prepare-PERF.M.PL barley-NOM.M.PL eat-INF you-GEN.SG bay-horse-DAT.DU
      Strewn is the Barhis (grass) for thee; O Indra, *extracted* is the Soma. *Prepared* are the barley grains for thy two bay-horses to eat. (RV 3.35.7 (cited in Jamison, 1990:5))

The stative PERF form has a wider distribution in Epic Sanskrit (Oberlies 2003; Speijer 1886). The form exhibits an eventive reading and may refer to past culminated events.[11] Evidence for the availability of an eventive interpretation for PERF comes from its use with past-referring temporal adverbials, and coordination of PERF clauses with other past tense clauses. The examples in (11-a-b) show that the bare PERF morphology is compatible with past-time adverbials which locate the event (as opposed to a state) denoted by the PERF

---

[8]For the purpose of showing that the present and the past tenses are morphologically contrasted in Epic Sanskrit, it is not crucial to know the real distribution of the three forms. Specifically, the claim that I have to make pertains to Middle Indo-Aryan, which does not inherit any of them from Old Indo-Aryan. However, I want to point out that it is problematic to assume that the Imperfect, the Aorist, and the Perfect are interchangeable with no real distinction between them at the Epic Sanskrit stage. Moreover, as far as I know, it has not been substantiated through a close linguistic and statistical study of the distribution of the three forms. It is not clear whether the Imperfect, Aorist, and Perfect forms are available for every verb or whether there are semantic restrictions (or tendencies) for preferred paradigms for particular verbs. Further, while it is known that all three forms are restricted to past eventive interpretations, it is unknown whether all of them are also compatible with past stative interpretations (a highly unlikely possibility). This question can only be resolved through textual studies directed by semantically sophisticated research questions.

[9]This is cognate to the English past participial morphology *-ed/-en*.

[10]It has been claimed that the PERF morphology has an eventive (past time) interpretation in Vedic, but Jamison (1990) shows that PERF is uniformly stative at the earliest Vedic stage.

[11]PERF originates as a stative adjective and its complete inflectional paradigm is based on the nominal categories — number, gender, and case. As a sentential predicate, PERF agrees with the nominative marked theme argument in number, gender, and case. The construction is passive, so the agentive argument appears in the instrumental case. The nominative case forms of the PERF paradigm in all genders and numbers are the constitutive forms for the PERF paradigm when it gets incorporated into the verbal system of Old Indo-Aryan.

The PERF paradigm

PERSON	SG	DUAL	PL
MAS	ga-**taḥ**	ga-**tau**	ga-**tāḥ**
FEM	ga-**tā**	ga-**tau**	ga-**tāḥ**
NEU	ga-**tam**	ga-**te**	ga-**tāni**

predicate at a specific time in the past.[12]

(11)   a.   **purā**   **devayug-e**   ca eva   **dṛṣ-ṭaṃ**   sarv-aṃ   mayā
             formerly god.age-LOC.SG and PTCL see-PERF.N.SG everything-NOM.N.SG I-INS.SG
             vibho
             lord-VOC.SG
             Lord, formerly, in the age of the Deva (Gods), I *saw* everything. (Mbh. 3.92.6a)
       b.   **hṛ-tā**   gau-ḥ   sā   **tadā** t-ena   prapāta-s   tu   na
             steal-PERF.F.SG cow-NOM.F.SG that-NOM.F.SG then he-INS.3.SG fall-NOM.M.SG PTCL NEG
             tark-itaḥ
             consider-PERF.M.SG
             Then he *stole* that cow, but *did* not *consider* the fall (consequences). (Mbh. 1.93.27e)

Further, sentences with PERF-inflected predicates can be coordinated with the Imperfect (12-a), the Aorist
(12-b), and the Perfect (12-c), the three past tense paradigms attested in Epic Sanskrit. In each of the cases,
PERF is interpreted as referring to a past event and not a result-state.

(12)   a.   **yadā** tu   rudhir-eṇa   aṅg-e   **parispṛṣ-ṭo**   bhṛgūdvaha-ḥ   **tadā**
             when PTCL blood-INS.SG body-LOC.SG touch-PERF.M.SG great.energy-NOM.M.SG then
             a-budhya-ta   tejasvī. . .   ca idam **a-bravī-t**
             rouse-IPFCT.3.SG radiant.NOM.SG and this   say-IPFCT.3.SG
             And when the (preceptor Rama) of great energy, *was touched* in the body by the blood, then,
             the radiant one woke up, and. . . said this. (MB 12:3:10 a-d)

       b.   yadā pūrvaṃ **gata-ḥ**   kṛṣṇa-ḥ   śamārtha-ṃ   kaurav-ān prati na   ca taṃ
             when before   go-PERF.M.SG K-NOM.M.SG peace-ACC.SG K-ACC.PL to   NEG and that
             lab-dha-vān   kāma-ṃ   tato   yuddha-m   **a-bhū-d**   idam
             obtain-PERF-ACT.M.SG desire-ACC.SG therefore battle-NOM.SG be-AOR.3.SG this
             When, in the past, Kṛṣṇa *went* to the Kauravas for peace, he did not obtain that desired goal,
             and therefore, this battle *happened*. (Mbh. 9.62.2)

       c.   tayor aṇḍ-āni   **nidadh-uḥ**   prahṛṣ-ṭāḥ   paricārikā-ḥ. . .   tataḥ
             their egg-ACC.PL deposit-PFCT.3.PL joyous-PERF.F.PL maid-servant-NOM.F.PL then
             pañcaśat-e kāl-e   kadrūputr-ā   **viniḥsṛ-tāḥ**
             500-LOC.SG time-LOC.SG K.son-NOM.M.PL burst.out-PERF.M.PL
             The happy maidservants *deposited* their eggs. . . then after five hundred years, the sons of Kadru
             burst out. (Mbh. 1.14.13-14)

   (11) and (12) thus show that the participial form PERF expands in its distribution from Vedic to Epic
Sanskrit, overlapping in its functions with the Imperfect, Perfect, and the Aorist in referring to past time events
(also see Condoravdi & Deo 2008). This fact is particularly relevant for Middle Indo-Aryan, which inherits
only two of temporal/aspectual paradigms discussed thus far — the PRES and the PERF paradigms. The next
section is concerned with establishing the correct semantic categorization for these morphological paradigms.
Specifically, in the Indo-Aryan linguistic tradition, PRES and PERF are considered to be the markers of present
and past tense respectively. I will argue that, in fact, PRES and PERF realize the imperfective and perfective
aspects in Middle Indo-Aryan.

4   Middle Indo-Aryan

The changes from the inflectional system of verbal contrasts in Old Indo-Aryan to the relatively morpho-
logically impoverished inflectional system of Middle Indo-Aryan have been described in terms of 'erosion'

---

[12]In all the glosses involving PERF forms, gender information is given only for those NPs with which PERF agrees, because PERF
contrasts with other paradigms in agreeing with the nominative NP in number and gender.

or 'simplification', primarily because many of the rich conjugational paradigms and the semantic categories they expressed were lost in Middle Indo-Aryan (Bloch 1965; Bubenik 1996; Pischel 1900; Vale 1948; and others). The Middle Indo-Aryan tense/aspect system inherits only the PRES, the PERF, and the Sigmatic Future paradigms from Old Indo-Aryan.[13] The rich system of past tense markers is lost. Pischel (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.[14] From among the past-referring forms of Epic Sanskrit, only the PERF paradigm remains and it is used regularly to refer to past time events in Middle Indo-Aryan. Further, the distribution of the PRES paradigm appears to undergo an unexpected change from Old Indo-Aryan to Middle Indo-Aryan. PRES marks the imperfective present tense in Old Indo-Aryan; in Middle Indo-Aryan it extends to past time reference as well. How are these changes to be interpreted? What is the correct characterization of the Middle Indo-Aryan tense/aspect system?

4.1 The imperfective-perfective contrast

My interpretation of the above facts is as follows: The present-past opposition realized in Old Indo-Aryan by distinct present and past tense morphology is lost in Middle Indo-Aryan. Instead, the PRES and PERF paradigms realize the aspectual contrast between the imperfective and the perfective aspects. The extension of the PRES paradigm to past-time reference is not random or determined by narrative function, but rather is grammatically determined. PRES must always be interpreted imperfectively regardless of whether the intervals it refers to overlap with speech time or precede it. Similarly, I will show that there is no evidence that the PERF paradigm realizes an aspectually neutral past tense; rather it realizes the perfective aspect. This interpretation contrasts with the standard understanding about the semantic values for these two paradigms in Middle Indo-Aryan, which is the present *tense* and past *tense* respectively (Bloch 1914, 1965; Chatterjee 1926; Pischel 1900; Vale 1948; Singh 1980; and others). The table in (13) gives the two competing proposals for tense-aspect contrasts in Middle Indo-Aryan.

(13)

Paradigm	Received view	My proposal
PRES	*present tense*	*imperfective aspect*
PERF	*past tense*	*perfective aspect*

Despite classifying PERF and PRES as the past and present tenses respectively, none of the scholars listed above fails to document the ubiquitous use of the PRES paradigm for past time reference. As a clear example, consider Beames' (1872–79:101–102) discussion of the PRES form in the modern languages:

> "... in form, preserves clear traces of its origin, though, as in its abraded condition it now no longer indicates with sufficient clearness present time; it has wandered away into all sorts of meanings, and is given by grammarians under all sorts of titles. Considering the very vague meanings which it now expresses, especially in regard to the note of time, it has seemed to me that the Greek term "aorist" more accurately describes this tense in its modern usage than any other. The fact that it is a present, *no matter what additional indefinite meanings may be attached to it, is, however, necessary to be borne in mind*." [italics mine]

Pischel also observes that the past 'tense' is productively expressed either by the PERF or the PRES forms. Bloch (1914:247), in his study of Marathi, refers to the 'temporal indeterminacy' of the PRES morphology (by which he means its use in past situations) that has been inherited by modern Marathi from Middle Indo-Aryan. Chatterjee (1926:949-54) describes in detail the use of the PRES paradigm for the past in Old and Middle Bengali. In the next section, §4.2, I will offer arguments for my position based on Middle Indo-Aryan textual data from the archaic Mahārāṣṭrī text *Vasudevahiṃḍī* (cir. 500CE) and the later Apabhraṃśa

---

[13]Middle Indo-Aryan also inherits other non-finite participial forms (the potential participle and the imperfective participle) which are incorporated into the finite tense/aspect systems in New Indo-Aryan languages. However, the constructions that these forms participate in are innovated in Middle Indo-Aryan or in New Indo-Aryan and cannot be said to be directly inherited from Old Indo-Aryan.

[14]The single instance of the Imperfect retained in Middle Indo-Aryan is the Imperfect form of the verb *as* 'be' (Pischel 1900:421-22). The Aorist occurs relatively more frequently (Pischel 1900:422-24), while the Perfect is preserved only as an archaism for a few verbs. Bloch (1965:228-233) reaches the same conclusion.

text *Paumacariu* (Late Middle Indo-Aryan; cir. 800CE).[15]

A note about the sourcing of examples and their identification: For the *Vasudevahimḍī* (abbreviated as VH) I have used the edition published by Atmananda Jainagranthamala (vol 80–81) in 1929–30. The textual references against each example give the subsection and the page number followed by the line number where the example occurs. Thus, VH:DH 31.24 refers to the 24th line on page 31 of the *Vasudevahimḍī* in the section called the *Dhammillahimḍī*. For the *Paumacariu* (abbreviated as PC), I have used the H.C. Bhayani edition published by the Bharatiya Vidya Bhavan between 1953 and 1960. The text is available in searchable electronic format, input by Eva De Clercq at Ghent University. The sequence of numbers indicates the location of the example in the following form: chapter:subchapter:stanza:line. Thus PC 1.1.14.4 refers to the fourth line in the 14th stanza of the first subchapter of the first chapter in the *Paumacariu*.

4.2   The PRES-PERF opposition in Middle Indo-Aryan

In order to demonstrate that the PRES and PERF paradigms realize the imperfective-perfective aspectual contrast in Middle Indo-Aryan, and not the present-past tense contrast, one must show that the distribution of these paradigms is characterized by certain systematic properties. Specifically, one must show that:

(14)   a.   Unlike the present tense, the PRES paradigm is NOT constrained to present time reference but may also be used to make reference to past and future eventualities.

   b.   In its past uses, the PRES paradigm systematically has imperfective reference. Sentences containing lexically atelic predicates tend to appear with PRES inflection and telic predicates are interpreted as progressive or habitual/generic with PRES.

   c.   Unlike the past tense, the PERF paradigm obligatorily has perfective reference. Sentences containing lexically telic predicates tend to appear with PERF inflection.

   d.   The PERF paradigm is not constrained only to past time reference but may refer to culminated, completed eventualities (or their result states) obtaining in the past, present, or future.

If all these facts hold for the Middle Indo-Aryan stage, then the correct characterization of the Middle Indo-Aryan system must be in terms of an aspectual, rather than tense, contrast. In other words, only an imperfective aspect marker and not a present tense marker would be expected to be systematically constrained to imperfective reference in its past usage. Similarly, a perfective aspect marker, and not a past tense marker, would be constrained to perfective reference and be used to describe both events and result states. The correct characterization of the Middle Indo-Aryan system is thus dependent on whether the data really corresponds to the scenario in (14-a-d). The possibility that these paradigms have both aspectual and temporal value is ruled out here, at least as far as the present-past opposition is concerned. As will be shown, the fact that the PRES paradigm is compatible with past, present, and future reference indicates that it cannot be specified for present tense. Similarly, the PERF paradigm is compatible with both past, present, and future time reference in Middle Indo-Aryan, indicating a lack of tense specification.

4.3   Middle Indo-Aryan: PRES as imperfective

The Middle Indo-Aryan cognate of the Old Indo-Aryan Present paradigm can certainly convey present time imperfective reference. (15-a) is a generic imperfective sentence, while (15-b) contains a lexical stative predicate *jāṇ* 'know' and a habitual (passivized) predicate — both imperfective.

(15)   a.   nipphala-ṃ   duma-ṃ   pakkhiṇ-o   vi   **paricchaya-nti**
            fruitless-ACC.SG tree-ACC.SG bird-NOM.PL also abandon-PRES.3.PL
            Even birds *abandon* a fruitless tree. (VH.DH 31.24-25)

---

[15]The corpus of Middle Indo-Aryan literature is vast and spans a period of over a millenium. The two texts selected are part of the Jaina narrative literature and represent some of the most important published material in Middle Indo-Aryan (Jain 1981). The *Vasudevahimḍī* of Sanghadāsagaṇi Vācaka has been (rightly) claimed to be a very good specimen of archaic Mahāraṣṭrī Prākrit by Alsdorf (1936). The *Paumacariu* is one of the two important Apabhraṃśa texts written by Svayambhudeva (cir. 800 CE), the most celebrated of the Apabhraṃśa poets. It can be safely assumed that these texts together give a reasonable picture of the fundamental tense/aspect facts of the Middle Indo-Aryan system.

b.   e-eṇa   tumam na **jāṇa-si**   kiṃ pi kajja-ṃ   **kīr-ai**
       this-INS.SG you   NEG know-PRES.2.SG what PTCL use-NOM.SG do.PASS-PRES.3.SG
       Do you not *know* what use *is made* of this? (VH.DH 32.13)

On the other hand, PRES is often used to convey past time reference as well. Consider the short narrative in (16), from *Vasudevahiṃḍī*, which reports a past episode about a monkey who entered a mountain cave and mistook some sticky liquid tar to be water. He tried to drink it and got his face and hands caught in it (and ultimately perished in the cave). Some verbs are inflected with PRES while others are inflected with the PERF paradigm. Notice that the lexical predicates *reach* and *stick* are telic achievements and carry PERF inflection while the lexical predicates *flow*, *touch*, and *spread* are atelic and carry PRES inflection.[16]

(16)   a.   sa   …ekka-m   pavvayaguha-m **pat-to**
          he.NOM.SG one-ACC.SG cave-ACC.SG   reach-PERF.M.SG
          He *reached* a cave. (VH.KH 6.10)

      b.   tattha ya silājau-m   **parissava-ti**
          there   and bitumen-NOM.SG flow-PRES.3.SG
          There, some bitumen (tar) *was flowing* (from the walls of the cave). (VH.KH 6.10)

      c.   so…   jalaṃ ti   mannamāṇo… muha-ṃ   **chubbha-ti**
          he.NOM.SG water thus thinking   mouth-ACC.SG touch-PRES.3.SG
          Thinking it to be water, he *touched* (his) mouth to it. (VH.KH 6.11)

      d.   taṃ **baddha-ṃ**…   hatth-e   **pasār-ei**   te   vi **baddh-ā**
          it   stick-PERF.3.SG hand-ACC.DU spread-PRES.3.SG they also stick-PERF.M.PL
          It *got stuck*. (He) *spread* his hands. They also *got stuck*. (VH.KH 6.12)

(16-a-d) is representative of how the PRES and PERF inflected forms are interspersed throughout the *Vasudevahiṃḍī*. The set of sentences in (17) contains a short narrative fragment from the *Paumacariu*, the later Apabhraṃśa text, whose grammar is virtually identical to that of the *Vasudevahiṃḍī* with respect to the distributional facts of interest. The fragment clearly illustrates how the aspectual properties of the sentences within a narrative affect interpretation about temporal location. Specifically, PERF marked sentences tend to describe events and advance the reference time forward (Partee 1984; Dowty 1986; Hinrichs 1986). In contrast, the PRES marked sentences tend to describe unbounded states and are interpreted as holding at the reference time, rather than advancing it.

The preceding context in the text in (17) describes the beauty of Maruevi, the queen of a king called Nabhirāja to whom a glorious child, Ṛśabhadeva, is to be born. The narrative in (17) tells us that goddesses sent by Indra arrived (PERF) on Earth to serve her (17-a) and reached (PERF) her abode (17-b). The next five lines (entirely composed with PRES verbs) describe the events in progress after the goddess' arrival. They elaborate on how the goddesses were serving and entertaining Queen Maruevi. The final line of this stanza introduces the dream of Maruevi, carrying the narrative forward: as she was sleeping on her bed, she dreamt (PERF) a series of dreams.

(17)   a.   to   etthantare māṇavaves-eṃ   **ai-u**   dev-iu   indāeseṃ
          then later   human.form-LOC arrive-PERF.F.PL goddess-NOM.PL indra.command-INS.SG
          Then, later, at the command of Indra, the goddesses arrived there in human form. (PC 1.1.14.1)

---

[16]The elided material (…) here and elsewhere contains relative clauses and other modifiers that are not relevant to the aspectual structure of the sentences.

b. sapparivāra **ḍhukka**   tettahe sā maruevi   bhaḍārī   jettahe
   with.family reach.PERF there   she Maruevi.NOM.SG adorable.NOM.SG where
   With their families, they reached there where the adorable Maruevi was (located). (PC 1.1.14.3)

c. kā   vi   viṇou   kiṃ pi   **uppāy-ai**   **paḍh-ai**   **paṇacc-ai**
   some PTCL recreation some PRT make-PRES.3.SG study-PRES.3.SG dance-PRES.3.SG
   gāy-ai   **vāy-ai**
   sing-PRES.3.SG play.instrument-PRES.3.SG
   Someone made some sport, someone studied, someone danced, someone sang, someone played
   an instrument. (PC 1.1.14.4)

d. Someone offered betelnuts (PRES), someone offered ornaments (PRES), someone fanned with
   the fly-whisks (PRES), someone washed her feet... (PC 1.1.14.5-8)[17]

e. varapallaiṅk-e   pasuttiya-e   suviṇāvali   ** diṭṭh-ī**
   excellent.bed-LOC.SG sleeping.INS.SG dream.series-NOM.F.SG see-PERF.F.SG
   (Maruevi), sleeping on an excellent bed, saw a series of dreams. (PC 1.1.14.9)

This distribution of PRES and PERF forms in Middle Indo-Aryan is inexplicable on the assumption that
PRES encodes the present tense and PERF the past tense. If these forms provide information about temporal
location with respect to speech/coding time, it is strange that the sentences in (16) and (17) do not occur
with the same tense marking, since they all report eventualities located within a specific time in the past.
The use of PRES to make reference to past time eventualities has been observed for Indo-Aryan starting with
Sanskrit and has been described as the historical present function of PRES. The historical present refers to
a crosslinguistically well-attested use of the present tense in which eventualities occurring in the past are
presented as if they were occurring in the present in order to make the narrative more vivid.

Scholars of both Old Indo-Aryan and New Indo-Aryan have suggested that the use of PRES for past time
reference can be understood as arising from an idiosyncratic narrative device to vivify narrative description,
rather than grammatical properties of the Indo-Aryan tense/aspect system.[18] Consider, for instance Speijer's
(1886:244) observation that PRES is often used in relating past actions in Sanskrit. He labels this use the
Historical Present use and goes on to note (§327) that the most common employment of the historical present
is that of expressing facts when "going on". He suggests that this use of the present may derive from the
absence of an imperfective marker in the past tense, observing that PRES forms may be used even when the
surrounding context contains past marking. MacDonnell (1927:204) also notes that the historical present use
is much more common in Sanskrit than in English, especially "to express the durative sense that the Sanskrit
Imperfect lacks." For New Indo-Aryan, Beames (1872–79:107) notes the past-oriented use of PRES in Marathi
and notes especially for Bengali that it conveys a "historic present".

Thus, the few scholars who have observed that PRES tends to be used with imperfective reference still
maintain the historical present hypothesis for the distribution of PRES. §4.3.1 shows that this hypothesis
is untenable for Middle Indo-Aryan and that the PRES-as-imperfective analysis accounts for the facts much
better. That is, Speijer's and MacDonnell's observations about the use of PRES in a durative sense hold almost
categorically for the Middle Indo-Aryan stage.

4.3.1 The 'historical present' hypothesis

Cooper (1986:31) describes the historical use of the present tense as a rhetorical device to 'relocate discourse
to some past location.' In other words, the deictic center for temporal location, which is the speech/coding

---

[17]The unglossed text is as follows:
kā vi **de-i** *tamvolu sa-hattheṃ savvāharaṇu kā vi sahūvattheṃ (PC 1.1.14.5)*
pād-ai *kā vi camaru kama* **dhov-ai** *kā vi samujjalu dappaṇu* **dhov-ai** *(PC 1.1.14.6)*
ukkhaya-khagga kā vi **parirakkh-ai** *kā vi kiṃ pi akkhāṇa* **akkh-ai** *(PC 1.1.14.7)*
kā vi jakkhakaddameṇa **pasāh-ai** *kā vi sarīru tāhe* **saṃvāh-ai** *(PC 1.1.14.8)*

[18]I have been able to find no reference that explicitly describes the historical present function of PRES only for the Middle Indo-Aryan
grammatical system. Most such references occur in the context of the evolution of the New Indo-Aryan languages.

time by default, is shifted to the past in order to achieve a particular narrative effect. (18) provides an example of the historical present use of the English Present as a rhetorical device, in describing past-time eventualities.[19] The situation under discussion belongs to a historical moment in the past (July 1812), yet is narrated as if occurring in the present.

(18)   (07-28-1812) ... As the sun rises, Napoleon *sees* that the Russian army has withdrawn. Napoleon *gives up* on catching the Russian army. Napoleon and French army *enter* Moscow, peopled by only a few thousand Russians. Fires *break out* across Moscow, burn for four days, and *leave* the city in ruins.

There are two ways to determine whether the use of PRES for past-time reference in Middle Indo-Aryan is determined by non-semantic aspects of narrative structure or by the meaning of PRES:

a. by examining the class of predicates with which PRES typically occurs and its interpretation in context;
b. by examining whether the perspectival shift effected by the supposed historical present use of PRES is consistent within a narrative.

First, if the use of PRES for past time reference is a narrative device, then we expect that PRES should not be restricted to predicates of a particular aspectual class. Notice, for instance, that in (18), the English Present tense marking appears on atelic predicates (e.g., *see*) as well as eventive predicates (*give up*, *enter*) and gives rise to an eventive interpretation in both cases. That is, we interpret this discourse as narrating successive events that occurred in the past. Further, consecutive sentences, if they contain eventive predicates, advance reference time. So the entry into Moscow is understood to take place after Napoleon gives up on catching the Russian Army and the fires in Moscow are understood to take place following the entry into Moscow.

Second, we also expect on the historical present hypothesis that for a piece of narrative in which the deictic center has been relocated to a past location, the tense marking should remain consistent, assuming that all eventualities within that narrative are understood to overlap with the shifted perspectival center or the shifted 'present' of the narrative.

Neither of these expectations is met in the Middle Indo-Aryan corpus. When it refers to past time eventualities, PRES can only be interpreted imperfectively, appearing with lexical stative, progressive, and habitual/generic predicates. In particular, it does not exhibit an eventive reading (unlike the English Present in (18)). Further, narratives are not uniformly shifted to a past time location where all clauses — both eventive and stative — are inflected in the PRES paradigm. Within any given narrative, PRES-inflected forms are interspersed with PERF-inflected forms and seem to refer to ongoing situations rather than completed events.

Consider the narrative fragment in (19) from *Vasudevahiṃḍī*.[20]

(19)   a.   **pat-to**   ya Seṇiyo   rāyā   ta-m   paesa-m.
             reach-PERF.M.SG and S.NOM.SG king.NOM.SG that-ACC.SG place-ACC.SG
             And King Seniya *reached* that place. (VH.KH. 17.1)

       b.   **vand-io**   ṇe-ṇa   viṇay-eṇam
             greet-PERF.M.SG he-INS.SG humility-INS.SG
             He *greeted* (the monk) with humility. (VH.KH. 17.1)

---

[19]http://www.txdirect.net/users/rrichard/napoleo1.htm

[20]A short note about the glossing of ergative subject arguments for the Middle Indo-Aryan and later data is in order here. As has been well-noted, Old Indo-Aryan did not have an active, ergative construction (Andersen 1986, Butt 2001, Butt & Deo 2003). The original construction that gave rise to the ergative clause in the New Indo-Aryan languages was, in Old Indo-Aryan, a passive construction based on the PERF form with oblique agents in the instrumental case. It has not been established beyond debate that Middle Indo-Aryan transitive PERF clauses, such as those in (19-b) are ergative. I therefore uniformly gloss the subject argument in Middle Indo-Aryan as having instrumental rather than ergative case marking. For New Indo-Aryan, however, subjects of PERF clauses are glossed as ergative. Note that the verb *patta* (Skt. *prāpta*) is exceptional in that it requires nominative rather than instrumental/ergative marking on subjects.

    c.   **piccha-i**   ṇa-m   jhānaniccala-m
       gaze-PRES.3.SG that-ACC.SG meditation.unmoving-ACC.SG
       (He) *gazed* at the meditation-engrossed one. (VH.KH. 17.1)

    d.   **pat-to**   titthayarasamīva-m
       reach-PERF.M.SG monk.close-ACC.SG
       He *reached* (came) close to the monk. (VH.KH. 17.3)

The predicates in (19) *reach that place* and *greet the monk* are telic and have PERF inflection.[21] The predicate in (19-c) has PRES inflection. Suppose this PRES inflection does mark a perspectival shift and relocates the deictic center to the past reference time of the discourse in order to achieve a stylistic effect. Then it is unexpected that the very next sentence (19-d) should appear with PERF inflection, rather than continuing with the PRES marking that characterizes the previous sentence. The hypothesis that the present tense marker, PRES, performs a narrative historical present function when it refers to past time eventualities is untenable given this kind of distribution for PRES in discourse. Further, if we examine the aspectual class of the predicate in (19-c), we see that *gaze at the meditation-engrossed one* is an atelic (stative), non-eventive predicate. This coincidence, that PRES in its so-called historical present function, appears with the only stative predicate in this fragment begs for a more general account of PRES distribution.

    (20) offers yet another example where PRES appears with stative predicates without introducing any perspectival shift that is then maintained in the later discourse. The predicate *notice a well* in (20-a) is eventive and the verb is inflected in the PERF paradigm. The predicates in (20-b-c) *observe the man* and *stand* are lexically atelic and the verbs are inflected in the PRES paradigm.

(20)   a.   t-eṇa   palāyamāṇ-eṇa purāṇakuv-o   taṇadabbhaparichinn-o
       that-INST.SG running-INST.SG old.well-NOM.M.SG grass.covered-NOM.M.SG
       diṭ-ṭho
       notice-PERF.M.SG
       That running one *noticed* an old well covered with grass. (VH.KH. 8.6)

    b.   tattha ayagar-o   mahākā-o   vidāriyamuh-o   gāsiukām-o
       there python-NOM.SG gigantic-NOM.SG open.mouthed-NOM.SG hungry-NOM.SG
       ta-m   purisa-m   **avaloe-i**
       that-ACC.SG man-ACC.SG observe-PRES.3.SG
       There a giant python, baring its mouth, eager to eat, *observed* the man. (VH.KH. 8.8-9)

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

The *Paumacariu* exhibits a similar pattern that cannot be attributed to the shift of the deictic center. This is illustrated here with the fragment in (21).[22] The story here describes a prince Bhāmaṇḍala who was afflicted by passion for Sītā. The preceding context describes his condition and elaborates on how he displayed all the symptoms of a pining lover. In (21-a), the author concludes that Bhāmaṇḍala was indeed suffering (PRES) from the pain of separation which would not subside (PRES). Both sentences contain stative predicates. The following sentences (21-b-d) contain eventive predicates and the verbs, which relate successive events, are uniformly inflected with PERF: He stood up (PERF) like a lion, advanced (PERF) with his equipment and army,

---

[21]I am not assuming a one-to-one correspondence between the telicity of a predicate and its grammatical aspect inflection. Rather the idea is that in narratives where the primary purpose of the discourse is to report a sequence of events, telic predicates will tend to describe successive events and therefore exhibit a strong tendency to appear with PERF inflection. The appearance of atelic predicates with PRES inflection in a narrative context together with the surrounding context suggests that they must be imperfective rather than perfective.

[22]Some non-crucial parts of the contiguous text have not been included in this fragment in the interest of a briefer exposition.

and reached (PERF) the city of Viyaḍḍhapura. It is difficult to reconcile this pattern of the distribution of PRES and PERF forms with the 'historical present' hypothesis.

(21) a. **vāh-ijj-ai**   virahe-ṃ   dūsah-eṇa   ṇa
          afflict-PASS-PRES.3.SG pain of separation-INS.SG intolerable-INS.SG NEG
          phiṭṭ-ai   k-eṇa   vi   osah-eṇa
          subside-PRES.3.SG any-INS.SG PTCL medicine-INS.SG
          He *was afflicted* by the intolerable pain of separation; it *would not subside* with any medicine.
          (PC 2.22.5.6)

      b. ṇīsāsu   mue-ppiṇu dīhu-dīhu   puṇaravi **th-iu**   thakk-evi jema
          sigh.NOM.PL have-GER long.NOM.PL then   become-PERF.M.SG stand-GER like
          sīhu
          lion.NOM.SG
          He sat there giving deep sighs, then *stood up* like a lion. (PC 2.22.5.7)

      c. **ṇīsar-iu**   sa-sāhaṇu
          advance-PERF.M.SG with.equipment.NOM.SG
          He *advanced* forward with his equipment. (PC 2.22.5.8)

      d. **pat-tu**   viyaḍḍhapuru
          reach-NOM.SG viyaḍḍhapur.PERF.M.SG
          He *reached* Viyaḍḍhapura. (PC 2.22.5.9)

Thus, we see that unlike a present tense marker, PRES is not restricted to present time reference but may systematically be used to make reference to past eventualities. PRES forms refer to eventualities located in the past, not because of a perspectival shift driven by narrative/rhetorical goals, but rather, because the PRES paradigm realizes the (temporally unspecified) imperfective aspect in Middle Indo-Aryan. In the absence of a present-past opposition in the language, the Middle Indo-Aryan PRES has both past and present imperfective reference.

    The next section demonstrates that besides being used with lexically atelic predicates PRES also derives progressive and habitual/generic predicates. All the sentences in the following examples make reference to a past temporal interval and this is disambiguated by the surrounding textual material (which is not always transcribed here for brevity).

4.3.2 The imperfective readings of PRES

(22-a-b), taken from the *Vasudevahiṃḍī*, contain the lexically stative verbs *parivas* 'live' and *suṇ* 'hear' with PRES inflection.

(22) a. egam-mi   kira nayar-e   kā   vi   gaṇiyā   rūvavati
          one-LOC.SG some town-LOC.SG some PTCL courtesan.NOM.SG beautiful.NOM.SG
          guṇavati   **parivas-ai**
          skilled.NOM.SG live-PRES.3.SG
          In some town, there *lived* a beautiful and skilled courtesan. (VH.K. 4.12)

      b. **suṇ-anti**   ya bhayavay-o vayaṇa-ṃ. . .   dhammakahāsaṃsia-ṃ
          listen-PRES.3.PL and monk-GEN.SG speech-ACC.SG religious.story.filled-ACC.SG
          And they *listened* to the speech of the monk, filled with religious stories. (VH.K. 5.5-6)

    (23-a-b), taken from the *Paumacariu*, contain the lexically stative verbs *vas* 'live' and *icch* 'desire' with PRES inflection.

(23)   a.   **g-a**   vandaṇahatti-e   ta-ṃ   paesu   jahñ   **vas-ai**
              go-PERF.M.SG adoration.DAT.SG that-sc acc.sg place.ACC.SG where live-PRES.3.SG
              mahārisi   saccabhūi
              great.sage.NOM.SG Saccabhūi.NOM.SG
              He *went* to that place for adoration where the great sage Sacchabhūi *lived*.[23] (PC 2.22.7.1-2)

        b.   uvarambha   ṇa   **icch-ai**   daha-vayaṇu
              Uvarambha.NOM.SG NEG desire-PRES.3.SG ten-headed.NOM.SG
              The ten-headed one *did not desire* Uvarambhā. (PC. 1.15.15.8)

In (24-a-b) from *Vasudevahiṃḍī*, PRES appears on eventive verbs and gives rise to a habitual/generic reading. In (24-a), the predicates *give food-drink* and *offer a goat* are eventive, and the sentence receives a habitual reading. The predicate in (24-b) *perform Yoga* is also eventive with a habitual reading for the sentence.

(24)   a.   so   ya bambhaṇo   varisevaris-e   tam-mi   devayā-e
              he.NOM.SG and brahmin.NOM.SG year.year-LOC.SG that-DAT.SG deity-DAT.SG
              …anna-pāṇa-m   **d-ei**   chagala-m   ca **nived-eti**
              food.drink-ACC.SG give-PRES.3.SG goat-ACC.SG and offer-PRES.3.SG
              And that Brahmin, year after year, *used to give* food and drink and *used to offer* a goat to the deity. (VH:KH 29.20)

        b.   tato aham   aṇṇayā kayāi   āyariyagiharukkhavāḍiyā-e   joga-m
              then I.NOM.SG other some time teacher.house.tree.garden-LOC.SG yoga-ACC.SG
              kar-emi
              do-PRES.1.SG
              Then, sometimes, I *would perform* Yoga in the orchard at my teacher's house. (VH:DH 37.1)

The diachronically later counterpart of this use of PRES, from the *Paumacariu*, is illustrated in (25). (25-a) describes the life of people who had joined the reign of King Indra. The context preceding (25-b) describes the birth of Rāvaṇa and his childhood exploits. He used to pluck the tusks of elephants and sometimes touch the hoods of serpents with his bare hands. (25-b) recapitulates this part of his childhood and is interpreted habitually.

(25)   a.   kappu   ṇa **di-nti**   **ja-nti**   sirigāra-hñ   ā[24]   ṇa
              tax.NOM.SG NEG give-PRES.3.PL possess-PRES.3.PL wealth-INS.PL command-NOM.SG NEG
              kara-nti   viṇāhaikāra-hñ
              do-PRES.3.PL arrogance-INS.PL
              They *used to give* no tax, (although) they were wealthy; they *did not obey* the royal command, because they were arrogant. (PC 1.8.2.3)

        b.   āyae līlae rāmaṇu   **ram-ai**   ṇaṃ kālu   vālu
              these sports Rāmaṇa.NOM.SG play-PRES.3.SG like God.of.death.NOM.SG boy.NOM.SG
              hoe-vi   **bham-ai**
              become-GER roam-PRES.3.SG
              Rāvaṇa *would play* such sports; he *roamed* (around) like the God of death in the form of a boy. (PC 1.9.3.9)

---

[23]Here and elsewhere there is often no overt accusative marking on objects, a characteristic of the Late Middle Indo-Aryan period where the nominative and the accusative cases were syncretized. For instance, *paesu* carries no accusative marking but the demonstrative *tam* is so marked. The glossing method I have adopted here considers an expression to have accusative marking if some other expression in the NP contains overt accusative marking. If there is no other accusative marker in the NP, the NP is glossed as nominative in the absence of overt marking.

[24]from Skt. *ājñā*

PRES also gives rise to an event-in-progress reading. In (26-a), the sentence with the PRES inflected verb provides a 'temporal frame' (very much like the progressive) for the event of the spat-out betel leaf falling. The eventuality of going is seen as ongoing at the time of this event. In (26-b), the event of playing is construed as being ongoing at the time of the event of noticing. Both examples in (26) are from *Vasudevahiṃḍī*.

(26)   a.   so   ya ḍiṇḍī. . .   bhavaṇa-ssa   āsaṇṇeṇa **gaccha-ti**
           he.NOM.SG and worshipper.NOM.SG house-GEN.SG near   go-PRES.3.SG
           dhaṇasiriy-e   tambola-ṃ   nicchuḍha-ṃ   **paḍi-yaṃ**
           Dhaṇasiri-GEN.SG leaf-NOM.N.SG spat.out-NOM.N.SG fall-PERF.N.SG
           ḍiṇḍi-ssu-vvariṃ
           worshipper-LOC.SG-on
           And the worshipper *was going* from near that house. Dhaṇasiri's spat-out (betel)-leaf *fell* on the worshipper. (VH.D. 51.12-14)

       b.   so   vi laliyāgoṭṭhi-e   samaṃ gaṅgā-e   **khella-i**   t-eṇa   ya
           he-NOM.SG also friend.group-GEN.SG with   river-LOC.SG play-PRES.3.SG he-INS.SG and
           khellant-eṇa   patta.cchejja-ṃ   **di-ṭṭhaṃ**
           playing-INS.SG leaf.bed-ACC.N.SG notice-PERF.N.SG
           And he *was playing* by the river with his group of friends. And the playing one *noticed* the seat made from leaves. (VH.D. 58.18)

Paumacariu contains comparable examples given in (27). The bathing event in (27-a) is the backdrop against which the event of spotting the dead wasp occurs. As expected, the verb *bathe* is inflected with PRES while the verb *fall* is inflected with PERF. (27-b) occurs within the description of a battlefield that forms the backdrop for the narrative: warriors were rushing towards the sound of the elephants and the horses; they had placed the arrow on the bowstring and men were uttering the *huṃ* sound (a battlecry). The final line of the stanza locates an event against this temporal frame: streams of blood started to flow (PERF) like rivers from the temples of the great elephants.

(27)   a.   kīlaṇavāvih-e parimiuṇāri-hñ   **ṇhā-i** (. . .)   **ṇivaḍi-ya**   tā-su
           sport.pond-LOC surrounded.damsels-INS.PL bathe-PRES.3.SG fall-PERF.F.SG that.LOC.PL
           diṭṭhi   tahñ avasare jahñ   muu mahuyaru kamal-abbhantar-e
           gaze.FEM.SG that time   where dead wasp   lotus-inside-LOC.SG
           At the time when he *was bathing* in the sporting pond, surrounded by damsels, his gaze *fell* on a dead wasp inside a lotus. (PC 1.5.14.7-8)

       b.   jettahe dhaṇuhara   guṇagahiyasara   jettahe huṅkāra
           where warrior.NOM.PL string.placed.arrow.NOM.PL where huṃ.sound.NOM.SG
           mua-nti   ṇara
           utter.PRES.3.PL men.NOM.PL
           Where the warriors had placed the arrows over the bowstring, where men *were uttering* the *huṃ* battlecry. (PC. 4.66.2.9).

4.3.3 Futurate readings of PRES

Finally, PRES may also be used to make reference to future eventualities. Although the Middle Indo-Aryan stages corresponding to both the *Vasudevahiṃḍī* and the *Paumacariu* retain the Old Indo-Aryan Sigmatic Future, PRES sometimes has future reference in matrix clauses. Although this is more of an inheritance from the Old Indo-Aryan distribution pattern, rather than an innovation, it supports the general claim that PRES cannot be treated as a temporally restricted present tense encoding category. I only give examples from the later *Paumacariu* here but these are available quite robustly in the *Vasudevahiṃḍī* as well. In (28-a) the King asks Bhīma to take care of the kingdom while he himself plans to leave to take up monastic ordination (*pravrajya*). The boldfaced verbs carry PRES inflection but must be interpreted as having future reference. In

(28-b), the PRES inflected verb in both the subordinate and the matrix clauses has future temporal reference, clearly conveyed by the use of the temporal adverbial.

(28) a. pabhaṇ-iu   bhīmu   ho-hi   diḍhu   rajjah-o   haũ puṇu
          say-PERF.M.SG Bhīma.NOM.SG become-IMP.2.SG established kingdom-GEN.SG I   PTCL
          j-āmi   **th-āmi**   ṇiya-kajjaho
          go-PRES.1.SG take.up-PRES.1.SG self-business.GEN.SG
          He said to Bhīma: "Become firmly established in your kingdom. Now I *will proceed* forth and
          take up my own business. (PC 1.5.14.1)

      b. jai kallae   tāya   laṅkā-ṇayari   ṇa paisara-mi   to ṇiyaya-jaṇeri
          if tomorrow father.VOC Laṅkā-city.ACC.SG NEG enter-PRES.1.SG then self-mother.ACC.SG
          indāṇī   kara-yal-e   **dhar-ami**
          Indāṇī.ACC.SG hand-palm-LOC hold-PRES.1.SG
          O father, if I do not enter the city of Laṅkā tomorrow, then I *will place* my mother Indāṇī on the
          palm of my hand. (PC 1.7.12.9)

      c. paḍhamu kar-eppiṇu valivihāṇu   puṇu pacchae dhaṇaya-ho   **mal-ami**
          initially, do-GER   sacrifice.ACC.SG and then   Dhaṇaya.GEN.SG shatter-PRES.1.SG
          māṇu
          pride.NOM.SG
          After having first sacrificed you, then I *shall shatter* the pride of Dhaṇaya. (PC 1.10.8.9)

4.3.4   Summary

The assumption that PRES encodes the present tense fails to hold under closer scrutiny of the distribution of PRES in Middle Indo-Aryan texts. The textual facts suggest that PRES is better analyzed as the exponent of imperfective aspect with no tense specification. PRES-as-imperfective accounts for the past temporal reference expressed by PRES much more accurately than the 'historical present' hypothesis, which is inaccurate for two reasons. First, it does not explain the tendency for atelic predicates to appear with PRES marking in past referring contexts. Second, it does not explain why the perspectival shift, supposedly initiated by PRES, does not continue via PRES marking in contiguous sentences within a narrative. As §4.3.2 shows, PRES appears with both lexical stative and lexical eventive verbs; in the latter case, it gives rise to event-in-progress or habitual/generic readings. The more plausible generalization for Middle Indo-Aryan is that PRES-inflected sentences are interpreted imperfectively. The futurate readings available to PRES further strengthen the hypothesis that PRES does not directly encode tense information.

4.4   Middle Indo-Aryan: PERF as perfective

That PERF in Middle Indo-Aryan (already in later Old Indo-Aryan), allows reference to completed eventualities in the past is not a matter of debate. The question is whether PERF realizes an aspectually neutral past tense or whether it must be considered to be an exponent of the perfective aspect. If the latter, it must be further determined whether PERF is semantically *past perfective*, i.e. encoding perfective aspect restricted to past temporal reference or whether it carries no tense specification and allows for past interpretation only by inference in context. The predictions about the distribution of PERF are clear on either of these three hypotheses:

A. If PERF realizes the aspectually neutral past tense, it should be compatible with both eventive and stative past reference.
B. If PERF realizes the past perfective, it should be restricted to describing completed eventualities located in the past with respect to speech time.
C. If PERF realizes perfective aspect and carries no tense specification, it should be compatible with past, present, and future reference given the right contexts of interpretation.

The traditional view among Indo-Aryan grammarians on this matter was that PERF realizes the simple past tense (hypothesis A) although scholars have rightly noted that both PRES forms and the Present Participial forms are also used robustly in past referring clauses (Chatterjee 1926; Bloch 1965; Vale 1948; Sen 1953; 1960). Contemporary Indo-Aryan linguists tend to go with the more restrictive hypothesis B that limits the distribution of PERF to descriptions of past events and their result states obtaining in the present (Masica 1991; Bubenik 1998, 2007). On the basis of the distributional facts of Middle Indo-Aryan, I propose that the most accurate analysis of PERF is as a purely aspectual operator (hypothesis C). PERF carries no tense specification; i.e. it does not realize the combination of past tense and perfective aspect, but rather, perfective aspect simpliciter.

Evidence that PERF does not realize the simple past tense partly comes from the systematic distribution of PRES and PERF in narrative discourse presented in §4.3. There we saw that atelic and derived progressive and habitual predicates within a contiguous past narrative appear with PRES marking rather than PERF. This is entirely unexpected on the PERF-as-past hypothesis because if PERF encoded the aspectually neutral past tense, it would be compatible with both perfective and imperfective reference. The following sections provide more evidence that PERF morphology, while clearly showing the distribution of perfective marking, cannot be taken to directly encode past tense meaning. In §4.4.1, we see that sequences of PERF sentences are understood to describe sequences of events and result systematically in the advancement of the reference time — the discourse moves forward in time with each successive sentence. This type of narrative effect is associated with perfective aspect rather than the past tense. §4.4.2 shows that PERF sentences may also be interpreted as describing result states that hold in the past or the present. Finally, in §4.4.3, we see clear examples of PERF being used in matrix clauses to make reference to future events and result states, with future adverbial modifiers. Taken together, these facts considerably weaken the case for any account that analyzes PERF as encoding past tense meaning.

4.4.1 PERF-based sentences uniformly advance reference time

In a contiguous sequence of PERF sentences, each sentence describes a successive event and updates the reference time for the discourse to the time after the occurrence of the event. That is, an eventuality described by the later sentence is understood to occur after the eventuality described by the prior sentence. A representative example from the *Vasudevahiṃḍī* is given in the narrative fragment in (29). The main predicate in each of the sentences in (29) is a PERF-inflected form. The story describes the events before the sacrifice of a goat, beginning with the departure of the family (with their friends and relatives) to the sacrificial stake. Every following sentence is understood to describe an eventuality that took place later in time, each of them ordered with respect to each other.

(29) a. tato te   mittabāndhavasahiā. . .   **ga-yā**
          then they.NOM.PL friends.relatives.with.NOM.PL go-PERF.M.PL
          Then they *went* there with their friends and relatives.'

  b. chagalo   vi ya maṇḍe-uṃ tatth-eva **ni-o**
      goat.NOM.M.SG also and decorate-INF there-PTCL take-PERF.M.SG
      And the goat also *was taken* there to be decorated.

  c. gandhapupphamallapuyāvises-eṇa   ya **acchi-yā**   devayā
      sandal.flower.garlands.worship.ingredients-INS.SG and worship-PERF.M.PL god.NOM.M.PL
      The Gods *were worshipped* with sandalwood paste, flower garlands, the ingredients of worship.

  d. gharamahattar-ehi ya **bhaṇi-yam** chagala-o   āṇ-ijj-au
      house.elders-INS.PL and say-PERF.N.SG goat-NOM.SG bring-PASS-IMP.3.SG
      And the house elders *said*: "Let the goat be brought."

e. tato tassa putto. . .   chagalaya-m āṇe-uṃ **ga-to**
    then his   son.NOM.M.SG goat-ACC.SG bring-INF go-PERF.M.SG
    At that, his son . . . *went* to bring the goat. (VH:D 29.25-28)

Within this fragment, PERF sentences only have eventive reference and successive sentences describe succes-
sive events in the past.[25] The later text, *Paumacariu*, maintains this pattern, as is illustrated by the narrative in
(30).[26] The preceding context describes how the brother of the ten-headed one (the demon Rāvaṇa) entered the
country of Vaiśravaṇa and started fighting. This led to the sequence of events in (30). First, the populace went
(PERF) to the king to complain. The king got angry (PERF), sent (PERF) a messenger, who entered (PERF) the
ten-headed one's court. The messenger was welcomed somewhat (PERF) and he then commenced his speech.
Each PERF sentence in the narrative describes an event and moves the narrative forward temporally.

(30) a. **ga-ya**   paya   kūvar-eṃ   kou   **hū-u**
          go-PERF.F.SG populace.NOM.F.SG prince-LOC.SG anger.NOM.SG become-PERF.SG
          The populace *went* to the king; (he) *became* angry. (PC 1.10.7.6)

      b. **pes-iu**   Vayaṇalaṅkāra   dūu
          send-PERF.M.SG Vayaṇalaṅkāra.NOM.SG messenger.NOM.SG
          (He) sent a messenger named Vayaṇalaṅkāra. (PC 1.10.7.6)

      c. dahavayaṇaṭṭhāṇu   **paiṭṭh-u**   gampi
          ten.headed.place.NOM.SG enter-PERF.M.SG go-GER
          Having gone, he (Vayaṇalaṅkāra) entered the place (court) of the ten headed one. (PC 1.10.7.7)

      d. te-hi   mi **k-iu**   abbhutthāṇu   kiṃ pi
          he.INS.SG also do-PERF.M.SG welcome.NOM.SG some also
          He (the ten headed one) also made some welcome (to him). (PC 1.10.7.7)

      e. **pabhaṇ-iu**   sumāli.pahu   de-hi   kaṇṇu
          speak-PERF.M.SG Sumāli.lord.VOC give-IMP.2.SG ear.NOM.SG
          He (Vayaṇalaṅkāra) said: "Lord Sumāli, give your ear (listen carefully)." (PC 1.10.7.8)

4.4.2   The stative readings of PERF

One of the uses of PERF inherited from the Old Indo-Aryan period is to mark result states (Jamison 1990).
This use continues to obtain in Middle Indo-Aryan (Condoravdi & Deo 2008). PERF sentences, in addition
to allowing reference to completed eventualities, may also allow reference to the states resulting from the
culmination of prior eventualities. The temporal location of these states remains lexically unspecified and
gets disambiguated by the temporal context. If PERF is treated as specified for the past tense, i.e. as past
perfective morphology, the present and future oriented stative readings of PERF remain unexplained.

   The examples in (31) contain a short narrative from the older *Vasudevahiṃḍī*. The context is as follows:
the queen and her maidservant are standing at the window of the palace looking down at the street below. The
maidservant notices that her mistress has stood still with her eyes fixed on something. (30-a) is the maidser-
vant's thought described by the narrator. The PERF inflected form describes this state which is understood as
overlapping with speech time (from the maidservant's perspective). In (30-b), the PERF inflected forms are
from the perspective of the narrator and describe the actions of the maidservant. These sentences are eventive,
rather than stative, and have past temporal reference. The final instance of a PERF-form in (30-c) *nivesiyā* 'has
rested' is part of a sentence with present time reference. It describes a thought of the maidservant and asserts
that the mistress has rested her gaze on somebody at the coding/speech time.[27]

---

[25]The observation is of course more general. See VH:KH 3.10-17, VH:KH 7.7-11, VH:KH 23.8-12, VH:D 29.19-23, and VH:D.31.1-
8, VH:D. 34.18-25 as examples in support of the claims that PERF sentences have eventive reference and advance reference time.
[26]Also see PC 1.6.16, 1.7.13, and PC 1.16.14 as more illustrations of this pattern.
[27]Both anonymous reviewers point out that the states described by PERF in (31) appear to be more ongoing or progressive than result

(31)   a.   kim mann-e   devī   passamānī. . .   nicchal-acchī
            why think-PRES.1.SG lady.NOM.SG looking.NOM.SG unmoving-eyes.NOM.SG
            ṭhi-yā
            stand-PERF.F.SG
            I wonder why the watching lady *has stood* with an unmoving gaze?

       b.   tiy-e   vi   **avaloi-o**   **di-ṭṭho**   ya   ṇā-e   so
            she-INS.SG also look-PERF.M.SG notice-PERF.M.SG and she-INS.SG that.NOM.SG
            puriso   cakkhuramaṇo
            man.NOM.SG eye-beautiful.NOM.SG
            She (the maidservant) also *looked*, and she *noticed* that man, attractive to the eye.

       c.   **cinti-yam**   ca ṇā-e   asaṃsayam eyam-mi   puris-e   **nivesi-yā**
            think-PERF.N.SG and she-INS.SG undoubtedly this-LOC.SG man-LOC.SG rest-PERF.F.SG
            ṇā-e   diṭṭhi
            her-FEM.SG gaze.NOM.F.SG
            And she *thought*: "undoubtedly, she (the lady) *has rested* her gaze on this very man."
            (VH:K:9.7-9)

(32) contains a passage from the *Paumacariu* in which PERF denotes a result state that must be located in the *past* rather than the present. The context describes the setting out of the army towards the city Lanka at daybreak with the beating of drums. After the beating of drums, the army marched (32-a). (32-b) contains a PERF sentence but it describes the state of the army personnel. Some were mounted on elephants while some on horses. The mounting event had taken place earlier, before the reference time provided by (32-a). What holds at the reference time in the past is the result state of that event, which can be described by a PERF sentence. (32-c) again describes an event that shifts the reference time forward: the army surrounded the city they wished to conquer.

(32)   a.   **saṃcall-iu**   sāhaṇu   ṇiravasesu
            march-PERF.SG army.NOM.SG entire.NOM.SG
            The entire army *marched out*. (PC 1.7.13.2)

       b.   **ārūḍh-a**   ke vi ṇara   gayavar-esu   tura-esu   ke vi
            mount-PERF.PL some PTCL man.NOM.PL elephant-LOC.PL horse-LOC.PL some PTCL
            Some of the men *were mounted* on elephants, some on horses. (PC 1.7.13.2)

       c.   **pariveḍh-iya**   laṅkā-ṇayari   tehñ
            surround-PERF.F.SG Lanka.city.NOM.F.SG they.INS.PL
            They *surrounded* the city of Lanka. (PC 1.7.13.4)

These examples illustrate the result stative uses of PERF. Crucially, the temporal location of the result-state is not lexically specified by PERF but rather is determined by the surrounding context. These facts are incompatible with the PERF morphology carrying past tense specification. On the other hand, they cohere well with the categorization of PERF as perfective morphology with no tense specification.

---

states. However, this is simply an effect of the particular verbs appearing in the examples. Verbs like *stand* and *rest* can refer to both the event of coming to be in a posture and to the state of being in a particular posture. *ṭhiyā* 'has stood' and *nivesiyā* 'has rested' refer to the result of a standing or a resting event which is to be in a state of standing or resting. In English, this state tends to be described using the Progressive (*is standing, is resting her gaze*) or the Perfect Progressive (*has been standing, has been resting*). In Middle Indo-Aryan and in many of the New Indo-Aryan languages, these states are preferentially described using PERF morphology rather than PRES or other imperfective morphology. Regardless of the choice of marking, the fact remains that the states described in (31-a) and (31-c) must be construed as states that arise from a prior eventuality.

4.4.3　Eventive uses of PERF in future temporal contexts

The clincher for the lack of tense specification for PERF comes from the use of PERF forms in describing events that must be interpreted as being located after the speech/coding time. Such examples are not numerous but they are systematic in nature. Significantly, we do not find examples of such usage in the older *Vasudevahiṃḍī* but they are locatable in the *Paumacariu*.[28]

Consider the examples in (33). In (33-a), the population approaches the king and expresses their suffering from starvation. The people are not yet dead, but will be so if no measures are taken. The form used to express this future eventuality is PERF. In (33-b) is described the consequence of a present eventuality in the form of prophecy, which must necessarily take place in the future. In (33-c), the first PERF form denotes a current result state (resulting from the past eventuality of gathering), while the second PERF form appears with the temporal modifier *kallae* 'tomorrow' and denotes a future eventuality (a future battle).

(33)　a.　ekka divas-e **ga-ya**　　　paya　　　　kūvār-eṃ　　　deva-deva **mu-a**
　　　　　　one　day-LOC go-PERF.SG subject.NOM.SG prince-LOC.SG lord　　　die-PERF.PL
　　　　　　bhukkhāmār-eṃ
　　　　　　starvation-INS.SG
　　　　　　One day, the subjects went to the prince (and said): "O Lord, we *will die* from starvation."
　　　　　　(PC 1.2.8.2)

　　　b.　tuha sāsaṇu　　　　duhaṇāsaṇu　　　　　　　evahñ uṇṇai **caḍi-ya**　　　　j-eṃ
　　　　　　your reign.NOM.SG misery.destroying.NOM.SG now　high ascend-PERF.SG which-INS.SG
　　　　　　hont-eṇa　　pahavanteṇa　　jagu　　　　　saṃsār-eṇa　**paḍi-ya**
　　　　　　being-INS.SG influence-INS.SG world.NOM.SG cycle-INS.SG liberate-PERF.SG
　　　　　　Your rule is misery-destroying; now it has ascended to its heights. By this influence, the world
　　　　　　will be liberated from the cycle (of birth and death). (PC 1.3.8.10)

　　　c.　avarehi mi　sama **samāvaḍ-iu**　　　pekkhe-sa-hi kallae　　**abbhiḍ-iu**
　　　　　　others　also like　gather-PERF.M.SG see-FUT-2.SG tomorrow fight-PERF.M.SG
　　　　　　Others have also gathered. You will see, he *will fight* (you) tomorrow. (PC 2.30.1.8)

(34) makes the point even more strongly. It describes the tail end of the war between Ravana and Rama and lists the future-oriented predictions being made by Ravana as he vows to destroy Rama on the final day of the battle. The entire discourse is future-oriented and uses PERF and PRES forms. Either Ravana's or Rama's pride will have been shattered (34-a). Either Ravana's wife Mandoari will weep (PRES) or Rama's wife Sita will grieve (PRES) due to the death of one of the men (34-b). Either Ravana or Rama will enter the cremation fire (34-c). Either Ravana or Rama will take the path taken by Khara et al. (34-d) and either one of them will embrace victory (34-e). PERF used in (34-a), (34-c), (34-d) and (34-e) is clearly part of a future oriented discourse and must be interpreted as describing future events or result states.

(34)　a.　kallae　　taho　vi mahu vi ekkantaru　　　　　　jimva taho　jimva mahu
　　　　　　tomorrow he.GEN and I.GEN and one.difference.NOM.SG either he.GEN or　　I.GEN
　　　　　　bhag-gu　　　maḍappharu
　　　　　　shatter-PERF.SG pride.NOM.SG
　　　　　　Tomorrow there will be only one difference between him and me; either for him or for me, pride
　　　　　　will have been shattered. (PC 4.67.10.4)

---

[28]Note that the *Paumacariu* represents the Apabhraṃśa stage, closer to the New Indo-Aryan system and that the inherited Sigmatic Future is lost in most New Indo-Aryan languages. The use of PERF to express futurity in this text may possibly indicate an ongoing reorganization of the system of future marking as the sigmatic future falls out of use and new future paradigms get innovated. I leave a closer scrutiny and interpretation of the distribution of the Sigmatic Future and PERF in Apabhraṃśa to future research.

b. kallae   jimva mandoari   **rov-ai**   jimva jāṇai appāṇa **sov-ai**
   tomorrow either Mandoari.NOM.SG weep-PRES.3.SG or   Jāṇai herself grieve-PRES.3.SG
   Tomorrow either Mandoari (Mandodari) will weep or Jāṇai (Jānaki) will grieve. (PC 4.67.10.6)

c. kallae   huavaha-dhagadhagamāṇa-ho jimva so jimva haũ **ḍhukk-u**   masāṇa-ho
   tomorrow fire-burning-GEN.SG   either he or   I   enter-PERF.SG cemetary-GEN.SG
   Tomorrow, either he or I *will enter* the burning fire of the cemetery. (PC 4.67.10.8)

d. jima mañ   jimva t-eṇa   **ṇihāl-iu**   khara-dūsaṇa-samvukka-pahu
   either I.INS.SG or   he-INS.SG follow-PERF.SG khara-dūsaṇa-samvukka-path.NOM.SG
   Either he or I *will take* the path taken by Khara, Dūsaṇa and Samvukka. (PC 4.67.10.9)

e. jima mañ   jimva teṇa   **āliṅg-iya**   kallae   raṇ-e
   either I.INS.SG or   he.INS.SG embrace-PERF.SG tomorrow battlefield-LOC.SG
   jayalacchivahu
   victory.goddess.NOM.SG
   Either he or I *will embrace* the goddess of victory. (PC 4.67.10.10)

Finally, I note that Singh (1980:136) briefly discusses the use of PERF in future temporal contexts citing examples like those in (35) from texts other than the ones examined here. He offers no explanation for this distribution of PERF and only points to the fact that such use anticipates the future-oriented use of PERF in the New Indo-Aryan languages.

(35) a. iha **āga-ya**   jaṇaṇi-e   kahahi   vatta
         soon come-PERF.M.SG mother-DAT tell-FUT.1.SG news.ACC.SG
         (I) will soon come tell the news to mother. (*Paumasiricariu* 2.115)

      b. tam ajju   pavitti nisi-bhoyaṇu   **parihar-iu**   maī
         then today from   night-food.ACC.SG give.up-PERF.M.SG I.INS.SG
         Then from today, I will give up my night meal. (*Bhavisattakahā* 320.12)

4.4.4 Summary

The assumption that PERF carries temporal specification (encoding past tense or past perfective) fails to hold under closer scrutiny of its distribution in Middle Indo-Aryan texts. The textual facts suggest that PERF is better analyzed as the exponent of perfective aspect with no tense specification. That PERF is aspectually perfective is clearly seen from the fact that the past tense domain is systematically divided between PERF and PRES. Sequences of PERF sentences uniformly advance the reference time, carrying a narrative forward as illustrated by the examples in §4.4.1. The perfect-like result stative uses of PERF also support its aspectual semantics. Further the result states denoted by PERF can be temporally located both in the present and the past depending on context (§4.4.2), suggesting that PERF cannot carry past tense specification. Finally, §4.4.3 demonstrates that PERF is frequently used to make reference to future completed eventualities, a use that is entirely incompatible with past tense specification.

4.5 Middle Indo-Aryan and the "tense bias"

The empirical claim that Middle Indo-Aryan does not morphologize a contrast between the present and past tenses has never been made explicitly in the existing literature on Middle Indo-Aryan. However, as I have pointed out, grammars do not fail to mention the free use of the PRES paradigm in referring to past times, the durative sense associated with its past usage in Sanskrit (Speijer 1886; MacDonnell 1927), and to a much lesser extent the future-oriented use of PERF (Singh 1980). Why should it be that previous research on Middle Indo-Aryan has not brought to light the clear fact that the present-past opposition from Old Indo-Aryan gives way to an aspectual imperfective-perfective opposition in Middle Indo-Aryan? I believe that two factors could

have contributed to this. First, Middle Indo-Aryan has been first and primarily analyzed by speakers of tensed languages such as Germanic and New Indo-Aryan in the late nineteenth and early twentieth century. The notion that languages may not directly encode a crucial semantic category such as temporal location was foreign to these scholars (and to linguistics more generally) at the time that they were investigating Middle Indo-Aryan grammar. It is only recently that the "tense-bias" in the analysis of systems of temporal and modal expressions has been overcome, in large part due to the crosslinguistic semantic work of Bohnemeyer (2002, 2009), Bittner (2005, 2008), Tonhauser (2011) and others. The Middle Indo-Aryan facts add to this developing body of knowledge about the nature of tenseless systems crosslinguistically.

The second, Indo-Aryan-specific factor for why the tenselessness of the Middle Indo-Aryan system has never been noted in the literature is that Middle Indo-Aryan was always analyzed as a linguistic system intermediate between two tensed language systems — Old Indo-Aryan and New Indo-Aryan. Research on Middle Indo-Aryan has always been either backward-looking (evolution from Old Indo-Aryan) or forward-looking (examining the proto-New Indo-Aryan aspects visible in Middle Indo-Aryan). The tensed-lenses from which Middle Indo-Aryan was analyzed served to obfuscate the actual organization of the temporal-aspectual system of the language and perpetuated the assumption that Middle Indo-Aryan also realizes a morphological contrast between the present and the past tenses.

5 Evidence from New Indo-Aryan

Modern New Indo-Aryan languages retain some pieces of evidence that further support the claim that the PRES and PERF paradigms carry aspectual but no tense specification. New Indo-Aryan languages differ from Middle Indo-Aryan in that tense information (in most languages and almost always) is directly realized by present and past tense auxiliaries. The verb is inflected for imperfective or perfective aspect. As mentioned in the introduction, scholars like Masica (1991) and Lienhard (1961) have observed that the category of aspect is at the heart of the New Indo-Aryan verbal system. While the paradigm for perfective aspect is uniformly PERF (and its cognate variants) across New Indo-Aryan, the imperfective aspect is realized using one of two morphological exponents — PRES or the (cognate of) the Old Indo-Aryan Present Participle in *-ant* (glossed as PART here). Not all New Indo-Aryan languages inherit the PRES paradigm. In some languages, both PRES and the Present Participle are employed as markers of imperfectivity in different constructions (e.g., Gujarati). Before I turn to the New Indo-Aryan evidence, some discussion of this third key player in the Middle Indo-Aryan aspectual system is essential.

5.1 The Present Participle in *-ant* in Middle Indo-Aryan

In Late Middle Indo-Aryan, the present participle in *-ant* often occurs as the finite verb in main clauses. Sen (1953:112) reports that this form is compatible with reference in the past, the present, and the future. Some examples from the *Paumacariu* are given in (36). In (36-a), a character in the story lists the bad omens that appear as he is starting to prepare his army for war.[29] (36-b) narrates an episode in which the Jina Ṛśabha, the first of the Jaina Tirthankaras, after achieving supreme knowledge, is seated for a discourse before a divine and human crowd. The sentence is part of a larger description of the worship going on — people bowing, introducing themselves and their lineages, shouting slogans of victory and reciting hymns. The temporal reference is clearly to the past of the narrator's perspective or utterance time. (36-c) is part of a larger list of actions that Hanumāna vows to undertake on the battlefield against his opponent the following day. All of these actions, oriented towards the future, use the bare Present Participle.

(36) a. pekkh-u **phur-antu**   vāma loyaṇu
          look-IMP twitch-PART.M.SG left   eye.NOM.SG
          Look, the left eye *is twitching*. (PC 1.8.3.2)

---

[29]The twitching left eye is a standard bad omen; the list includes howling jackals, cawing crows, and a crying woman (PC. 1.8.3), all with the template: Look, X is happening.

b. sayala vi   jayajayakāru   **kar-antā**   sayala vi   thottasay-āñ
all   PTCL victory.slogan.ACC.SG do-PART.M.PL all   PTCL hymn.hundred.ACC.PL
paḍh-antā
recite-PART.M.PL
All of them *were chanting* slogans of victory; all of them *were reciting* hundreds of hymns.
(PC 1.3.7.8)

c. diḍha guḍa **toḍ-antu**   turaiṅgam-ā-hā paravalu   vali   **de-ntu**
strong rein break-PART.M.SG horse-PL-GEN   enemy.army.NOM.SG sacrifice give-PART.M.SG
vihaiṅgam-ā-hā
bird.PL-GEN
I will break the strong reins of the horses; I will give the enemy army in sacrifice to the birds.
(PC 4.58.7.7)

The distribution of the Present Participle in Middle Indo-Aryan includes its default use to express the habitual past tense and its use in both the antecedent and the consequent of counterfactual conditional clauses (Singh 1980; Sen 1953, 1960; Bubenik 1998). Discussing all these uses in detail is beyond the scope of this paper, but ultimately essential for establishing how this participial form becomes the building block for realizing imperfectivity in several New Indo-Aryan languages. The goal here is to simply introduce this form as an alternative imperfective marker in finite clauses, which, similar to the PRES paradigm, remains unspecified for tense, receiving temporal reference only in context.

5.2 Pawri: The Middle Indo-Aryan configuration

Pawri (ISO [bfb]; population 175,000) is a non-literate Indo-Aryan language spoken by the tribal Pawra community in parts of Northern Maharashtra and Madhya Pradesh.[30] Pawri is crucial from the perspective of the claim being made here for Middle Indo-Aryan, because unlike most standard New Indo-Aryan languages, it lacks obligatory present/past tense marking. In other words, Pawri seems to retain the archaic Middle Indo-Aryan pattern, as realized by the Present Participial paradigm discussed in §5.1. Pawri has been noted and described in Grierson's survey in his volume on the Bhili and Khandeshi dialects (vol. IX–III). The facts reported here are based on Deo (2006). The imperfective aspect in Pawri is realized by an extended variant of the Middle Indo-Aryan Present Participle — let us call this the Pawri Imperfective paradigm.[31] The Pawri Imperfective lacks temporal specification and is compatible systematically with both past and present temporal reference. The Pawri Imperfective paradigm for the verb *khā* 'eat' is in (37).

(37)   **Pawri Imperfective paradigm**

	SG	PL
M	khā-ta-lu	khā-ta-lā
F	khā-ta-li	khā-ta-lyā
N	khā-ta-la	khā-ta-le

The temporal reference of clauses containing the Pawri Imperfective is most often only recoverable from the context and, in some cases, from the presence of optional adverbs. The examples in (38) illustrate the use of Pawri Imperfective (glossed PART) in habitual contexts and with lexically stative verbs, while (39) gives some examples where the form gives rise to event-in-progress readings. (38-a) exemplifies the present habitual reading of the Pawri Imperfective. In (38-b), the second sentence refers to a past habit, but this is determined by the tense of the optional auxiliary *otu* 'was' in the preceding clause, not by any morphology

---

[30]All data for Pawri is based on my own findings during the fieldwork that I conducted between 2003-05 in Nandurbar district of Maharashtra. I would like to thank Gulabsingh Pawra, Bhaisingh Pawra, and Barfi Pawra (Pawri) of Maal village in Dhadgav tehsil for sharing their language, culture, and life with me and their careful and patient efforts to teach me aspects of their languages.

[31]Grierson (1907) speculates that these endings are older adjectival endings encoding number/gender information similar to the *-l* endings attested for the PERF form in Late Middle Indo-Aryan, Marathi, and the Eastern Indo-Aryan languages.

in the clause with the Imperfective itself. (38-c) contains an example of the Imperfective with a stative verb *roy* 'live' and past reference, supported by the presence of the adverb *pel* 'earlier'.

(38) a. chyu   kāyam ārhā-m   svotā-hā=j   **bāl-ta-lu**
          he.NOM always mirror-LOC self-ACC-EXCL look-PART-M.SG
          [**Context:** Describing a friend who is self-obsessed] He always *looks* at himself in the mirror.

      b. āgyāḍvāji   bānge-n   talapi   ot-u.   chyu   kāyam bāng
          āgyāḍvāji.NOM hemp-GEN addict.NOM be-PST.M.SG he.NOM always hemp-NOM
          pi-ta-lu
          drink-PART-M.SG
          Agyāḍvāji was a hemp addict. He *would* always *drink* hemp.

      c. chyi   pel   nandurbār **roy-ta-li**
          she.NOM earlier Nandurbar live-PART-F.SG
          Earlier, she *lived* in Nandurbar.

(39-a) describes an event in progress with present temporal reference. In (39-b), the temporal adverbial clause contains an imperfective-marked verb *rovtali*, while the main clause contains the perfective verb *lāgyu*, which receives a past time interpretation by default. The imperfective-marked verb is also interpreted in the past in this context, despite no overt expression of the past tense such as a tense auxiliary. It is the default reference of the PERF form in the main clause that provides the temporal reference for the imperfective form.[32] Likewise for (39-c.)

(39) a. Dhanirāyā, āpu   kāy kādav **khā-ta-lā**
          D-VOC   you-NOM.HON what mud-NOM eat-PART.M.PL
          Dhanirāyā, *are* you *eating* mud?

      b. mi   **rov-ta-li**   tevĩ mehe   seṇḍu   lāg-yu.
          I.NOM play-PART-F.SG then I-ACC.SG ball.NOM hit-PERF.M.SG
          When I *was playing*, a ball hit me.

      c. Vārirāya$_j$ jangalbāri-daryā-m **phir-ta-lu**.   tevĩ chyu$_i$ tināhā$_j$ **hād-yu**.
          V.NOM   forest-valleys-LOC wander-PART-M.SG then he.NOM he-ACC call-PERF.M.SG
          Vārirāya$_j$ *was wandering* in the forests-valleys. At that time, he$_i$ called out to him$_j$.

Pawri thus synchronically realizes a temporally unspecified imperfective marker, while all other surrounding languages (Marathi, Gujarati, and Hindi) are characterized by periphrastic constructions built on the Present Participle or the PRES paradigm with overt tense marking. This is not to say that Pawri does not have *any* morphological means of marking the past-present distinction. Pawri does have tense auxiliaries that are cognate to the auxiliaries of Gujarati; however, unlike in the other languages, these auxiliaries are not obligatory and are rarely expressed in discourse (except in non-verbal predicative clauses as in (38-b)).[33]

The distribution of the Pawri Imperfective is special from the viewpoint of Indo-Aryan languages. None of the standard, literate Indo-Aryan languages exhibit a system where temporal reference fails to be directly encoded in imperfective clauses. Yet if the findings from Late Middle Indo-Aryan, described in §4 and in §5.1 are accurate, then such a pattern can easily be interpreted as the retention of the archaic aspectual situation, the

---

[32]An anonymous reviewer wonders whether the default inference associated with PERF means that PERF is tense marked. This is clearly not what is being said here. Temporal location defaults for aspectually marked clauses only associate past temporal reference with perfective morphology in the absence of overriding contextual information (Smith 1991, 2008). If the context supports a non-past interpretation PERF may be interpreted as non-past. Default patterns of temporal reference do not provide evidence for considering PERF to be overtly tense-marked.

[33]Most of the examples with overt tense auxiliaries that I have for Pawri are elicited translations of Marathi or Gujarati sentences, rather than sentences from spontaneous discourse.

proto-system that gives rise to the surrounding tense-encoding systems. On this view, the Pawri Imperfective, because it lacks tense specification, is crucial to reconstructing the diachrony of the Indo-Aryan tense/aspect system. Further research into its precise organization can contribute to shedding light on the proto-system that evolved into the surrounding New Indo-Aryan languages.

5.3 PRES in Modern Gujarati

Although in the case of Gujarati, most finite clauses carry overt tense encoding, there persist some sub-domains of the tense-aspect system in which tense information is not obligatorily encoded. I describe here the distribution of the PRES paradigm, which may be used with both present and past temporal reference (an inheritance from its Middle Indo-Aryan use).[34] Imperfective aspect in Gujarati may be realized by either the PRES paradigm or the cognate of the Present Participle. The Present Participle is the general imperfective form which forms periphrastic constructions with tense and modal auxiliaries. The PRES paradigm, in periphrasis with the present tense auxiliary, realizes the present imperfective. In the modern language, tense auxiliaries are typically considered obligatory in indicative clauses, but they may be omitted in a small set of contexts.

Most commonly, in a sequence of clauses headed by a tensed clause (either present or past), the bare PRES form may be used to convey present or past temporal reference. In (40-a), for instance, the speaker describes her daily schedule with a series of habitual sentences. The temporal context (present) is introduced in the first clause by the tense auxiliary *che*. The verbs in the following clauses carry PRES inflection and the clauses lack any tense specification. The interpretation is uniformly present. (40-b) also contains a PRES form and has past reference that is determined by the presence of the temporal adverb *nānpaṇ-mā* 'in childhood.' In (40-c), the temporal reference of both the antecedent and consequent clauses comes purely from contextual knowledge that the speaker is referring to her childhood.

(40) a. māro divas   em **jā-e**   **ch-e**.   hũ savār-e   **uṭh-u**,
          my day.NOM thus go-PRES.3.SG PRS-3.SG I.NOM morning-LOC wake-PRES.1.SG
          nhā-u,   pachi pujā   **kar-u**,   pachi bajār-mã   **jā-u**
          bathe-PRES.1.SG then prayer.NOM do-PRES.1.SG then market-LOC.SG go-PRES.1.SG
          My day goes thus: I wake up, bathe, then pray, then go to the market . . .

      b. hũ   nānpaṇ-mā   sanskrit   **bhaṇ-u**
          I-NOM childhood-LOC sanskrit.NOM study-PRES.1.SG
          I used to study Sanskrit in my childhood.

      c. hũ   jyāre-jyāre Sanskrit   **bhaṇ-u**   tyāre-tyāre bā   ma-ne
          I.NOM whenever   sanskrit.NOM study-PRES.1.SG then   mother.NOM I-ACC.SG
          lāḍu   **āp-e**
          sweet.NOM give-PRES.3.SG
          Whenever I would study Sanskrit, my mother would give me a sweet.

The fact that bare PRES forms occur in finite clauses and receive temporal reference in context provides further evidence that PRES is not specified for tense. Given the facts of Middle Indo-Aryan, it is much more likely that this lack of tense specification is not an innovation in Gujarati but rather reflects an inheritance from Middle Indo-Aryan.

5.4 PRES in other New Indo-Aryan languages

The distribution of PRES in other standard New Indo-Aryan languages also supports the idea that its non-present uses are inherited from Middle Indo-Aryan rather than innovated within new Indo-Aryan. For in-

---

[34]The status of the PRES paradigm in the New Indo-Aryan languages appears to parallel the status of the Injunctive in Vedic (Kiparsky 2005). This form is unspecified for tense and, in the absence of tense auxiliaries, is compatible with both a present, a past, and (in some cases) a subjunctive/irrealis interpretation. The hypothesis that the proto-system for New Indo-Aryan articulated only an imperfective-perfective contrast and lacked the present-past tense distinctions is also parallel to the reconstruction of the Proto Indo-European system as primarily being aspectual in nature.

stance, Bloch (1914:245f) observes that PRES is used chiefly in Modern Marathi to denote an action repeated in the past (a habitual past tense). There is absolutely no reason for why the present tense paradigm from Old Indo-Aryan would evolve into a habitual past marker in Marathi unless the path involves a stage at which the paradigm ceases to encode tense and starts encoding purely aspectual meaning — the pattern claimed for Middle Indo-Aryan here. Bloch indirectly verifies the existence of such a path when he notes with examples that in Old Marathi (texts from *cir.* 1200CE) the same paradigm has "the sense of present, future, or past depending on context." Bloch's conclusion is that the temporal sense of PRES is extremely secondary in Marathi.

Beames (1872–1879:107) also notes that the PRES paradigm "has become vague and in modern times is often used in both a future and a past sense." In addition to the Marathi facts, he observes that in Punjabi and Hindi PRES is used as an indefinite future. Trumpp (1872) labels the PRES paradigm as the Sindhi Potential based on its future-oriented use. In Bengali, PRES is used to express habitual/generic present meaning (the language has distinct periphrastic progressive aspect marking) and is also used with no tense marking in past referring negative declaratives. The full details of the evolution of PRES in each New Indo-Aryan language and their theoretical implications for the development of tense in New Indo-Aryan deserve a much closer scrutiny. I only point out here that the diversity of uses exhibited by PRES in the modern languages, and particularly its restriction to certain non-overlapping contexts of use (e.g., only habitual past in Marathi vs. only indefinite future in Hindi and Punjabi) remains bewildering unless we assume that this distribution derives from an originally unified aspectually based parent system — the system proposed for Middle Indo-Aryan.

5.5   The future readings of PERF in New Indo-Aryan

I close by considering some uses of PERF in New Indo-Aryan languages which are incompatible with the hypothesis that PERF is specified for past tense in the modern languages. §4.4.3 has already shown that in Middle Indo-Aryan PERF may allow for future temporal reference, referring to events that are believed to occur with certainty. This use of PERF survives in the New Indo-Aryan languages — at least in Gujarati, Hindi, Marathi, and Bengali. The examples here are from Hindi (41) and Marathi (42) and based on native speaker intuitions.

(41)   āp   yahĩ ruki-ye.   maĩ   pāc miniṭ-me   **ā-yā**
       you.HON.NOM here wait-IMP.HON.2.SG   I.NOM.SG five minute-LOC come-PERF.M.SG
       You sit here. I *will come* back in five minutes.

(42)   a.   **Context:** Watching an election candidate's mediocre speech
       b.   hyā   varśi   hā   nakki   khāli **paḍ-lā**
            this.OBL year.LOC he.NOM definitely down fall-PERF.M.SG
            This year, he *will* definitely *fall down* (lose in the elections).

Additionally, PERF is the default form of the verb used in the antecedent of conditionals. The examples in (43) are from Hindi but the facts hold for Marathi as well.

(43)   a.   maĩ   bambai **ga-yā**   to   tumhārā kām   **kar-uṅgā**
            I.NOM Bombay go-PERF.M.SG then your   work.NOM.SG do-FUT.1.SG
            If I *go* (lit. went) to Bombay, I will do your work.

       b.   us-ne   agar paḍhāi   **k-ī**   to   vah   agli kakśā-mein **jā-egā**
            he-ERG if   studies.NOM.F.SG do-PERF.F.SG then he.NOM next class-LOC go-FUT.3.SG
            If he studies then he will go to the next grade.

These patterns of distribution provide further support to an aspectual rather than tense-specified meaning for PERF, in turn lending support to the hypothesis that PERF does not realize the past tense in Middle Indo-Aryan.

6 Conclusion

A closer look at Middle Indo-Aryan facts indicates that the traditional classification of two morphological paradigms inherited from Old Indo-Aryan must be revised. PRES, the Old Indo-Aryan present tense, and PERF, the Old Indo-Aryan past participial form, do not denote the present and the past tense in Middle Indo-Aryan, but rather realize the imperfective and the perfective aspectual categories respectively. The distribution of the two paradigms in narrative discourse in the archaic Mahārastrī text *Vasudevahiṃḍī* (cir. 500CE) and the later Apabhraṃśa text *Paumacariu* (cir. 800CE) demonstrate that modulo contextual factors, the two paradigms are compatible with past, present, and future temporal reference. Further evidence that this categorization of PRES and PERF is on the right track comes from the New Indo-Aryan languages. Pawri provides evidence in the form of imperfective morphology (the Present Participial paradigm) which lacks temporal specification. The similarity between the two diachronically related systems is more simply accounted for as inheritance rather than innovation within Pawri. Gujarati illustrates the temporally unspecified nature of PRES in some of its sub-domains where temporal reference may be disambiguated by context. Other New Indo-Aryan languages retain the use of PRES in non-overlapping subdomains that indicates an earlier wider distribution of the form. Finally, the use of PERF in immediate future and conditional contexts in the New Indo-Aryan languages echoes such attestations in Middle Indo-Aryan (§4.4.3) supporting the idea that these phenomena are inherited rather than innovated.

There is thus strong support for the hypothesis that Middle Indo-Aryan does not morphologize the contrast between past and present tenses, but relies on an aspectual contrast and contextual disambiguation to distinguish the temporal location of eventualities with respect to speech time. This pattern is not typologically uncommon and is instantiated in languages as diverse as Arabic, Navajo, and Chinese. If Middle Indo-Aryan is classified as a similar aspectual system lacking a present-past contrast, then the rise of periphrastic tensed constructions in New Indo-Aryan acquires a functional significance. The proto-system on which New Indo-Aryan is based lacks grammaticalized morphologically expressed tense information in finite clauses. The innovation of tense auxiliaries and periphrastic constructions involving these is a functionally motivated change in New Indo-Aryan that introduces tense information as an obligatory morphosyntactic element in the clause. If it is maintained that Middle Indo-Aryan always encoded the contrast between the past and present tenses morphosyntactically, then the introduction of periphrastic paradigms of tense-aspect marking that are central to the New Indo-Aryan languages remains unmotivated. The aspectual hypothesis thus makes better sense than the tense hypothesis of three types of facts: (a) the Middle Indo-Aryan distribution of PRES and PERF; (b) some puzzles in the distribution of PRES and PERF in New Indo-Aryan; and (c) the rise of periphrastic constructions with tense auxiliaries in New Indo-Aryan. In this way, it points out a promising direction for further systematic research in Middle Indo-Aryan and New Indo-Aryan tense/aspect diachrony.

References

Alsdorf, Ludwig. 1936. Vasudevahiṃḍī, a Specimen of Archaic Jaina Maharastri. *Bulletin of the School of Oriental Studies* 8(2/3):319–333.

Andersen, Paul Kent. 1986. Die *-ta* Partizipialkonstruktion bei Ashoka: Passiv oder Ergativ? *Zeitschrift für Vergleichende Sprachforschung* 99:75–95.

Beames, John. 1872-79. *A Comparative Grammar of Modern Indo-Aryan Languages of India.* Delhi: Munshiram Manoharlal. Republished 1966.

Bhayani, Harivallabh. 1998. *Gujarati Bhāṣānu Aitihāsik Vyākaraṇ.* Ahmedabad: Pārśva Publication.

Bhayani, Harivallabh Chunnilal, ed. 1953–1960. *Paumacariu of Kaviraja Svayambhudeva (A Pre-tenth Century Puranic Epic in Apabhramsa) vol. 1–3,* Singhi Jaina Series 34–36. Bombay, India: Bharatiya Vidya Bhavan.

Bittner, Maria. 2005. Future discourse in a tenseless language. *Journal of Semantics* 22:339–387.

Bittner, Maria. 2008. Aspectual universals of temporal anaphora. In S. Rothstein, ed., *Theoretical and Crosslinguistic Perspectives on the Semantics of Aspect*. Amsterdam: John Benjamins.

Bloch, Jules. 1914. *The Formation of the Marathi Language*. Delhi: Motilal Banarsidass. Translated by Dev Raj Chanana. Published 1970.

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

Bohnemeyer, Jürgen. 2002. *The Grammar of Time Reference in Yukatek Maya*. Munich: Lincom Europa.

Bohnemeyer, Jürgen. 2009. Temporal anaphora in a tenseless language. In W. Klein and P. Li, eds., *Expression of Time in Language*, pages 83–128. Berlin: Mouton de Gruyter.

Bohnemeyer, Jürgen and Mary Swift. 2004. Event realization and default aspect. *Linguistics and Philosophy* 27(3):263–296.

Brockington, John. 1998. *The Sanskrit Epics*. Leiden: Brill Academic Publishers.

Bubenik, Vit. 1996. *The Structure and the Development of Middle Indo Aryan Dialects*. Delhi: Motilal Banarsidass.

Bubenik, Vit. 1998. *Historical Syntax of Late Middle Indo Aryan (Apabhramsa)*. Amsterdam: John Benjamins.

Bubenik, Vit. 2007. Prākrits and Apabhraṁśa. In G. Cardona and D. Jain, eds., *The Indo-Aryan Languages*, pages 204–249. New York: Routledge.

Butt, Miriam. 2001. A Reexamination of the Accusative to Ergative Shift in Indo-Aryan. In *Time over Matter: Diachronic Perspectives on Morphosyntax*. Stanford: CSLI Publications.

Butt, Miriam and Ashwini Deo. 2001. Ergativity in Indo-Aryan. *KURDICA* 5(3). http://www.cogsci.ed.ac.uk/%7Esiamakr/kurdica5.html.

Bybee, Joan, Revere Perkins, and William Pagliuca. 1994. *The Evolution of Grammar. Tense, Aspect, and Modality in the Languages of the World*. Chicago: The University of Chicago.

Caturavijaya, Muni and Muni Punyavijaya, eds. 1929–30. *Vasudevahiṁḍī of Saṅghadāsagaṇī,* Atmananda Jaingranthamala vol. 80–81. Bombay, India: Jaina Atmanandasabha.

Chatterji, Suniti Kumar. 1926. *The Origin and Development of the Bengali Language*. George Allen and Unwin Ltd. Reprinted 1975 by D.Mehra, Rupa& Co., Calcutta.

Condoravdi, Cleo and Ashwini Deo. 2008. Aspect Shifts in Indo-Aryan. In *Proceedings of International Congress of Linguists (CIL 18) in Seoul*.

Cooper, Robin. 1986. Tense and discourse location in situation semantics. *Linguistics and Philosophy* 9:17–36.

Dahl, Östen. 1985. *Tense and Aspect Systems*. Oxford: Blackwell.

Dahl, Östen, ed. 2000. *Tense and Aspect in the Languages of Europe*. Berlin: Mouton de Gruyter.

Dave, Trimbaklal N. 1935. *A Study of the Gujarati Language in the 16th Century*. London: The Royal Asiatic Society.

Delbrück, Berthold. 1888. *Altindische Syntax*. Halle, Tübingen: Niemeyer, 2nd edn.

Deo, Ashwini. 2006. *Tense and Aspect in Indo-Aryan Languages: Variation and Diachrony*. Ph.D. thesis, Stanford University.

Dowty, David. 1986. The effects of aspectual class on the temporal structure of discourse: Semantics or pragmatics? *Linguistics and Philosophy* 9(1):37–61.

Gonda, Jan. 1962. *The Aspectual Function of the Rgvedic Present and Aorist*. S-Gravenhage, Mouton.

Grierson, George Abraham. 1907. *Linguistic Survey of India: Indo-Aryan Family, Central Group*, vol. IX–III. Delhi: Motilal Banarsidass. Reprinted 1967.

Hinrichs, Erhard. 1986. Temporal anaphora in discourses of English. *Linguistics and Philosophy* 9(1):63–82.

Hoffmann, Karl. 1967. *Der Injunktiv im Veda; eine Synchronische Funktionsuntersuchung*. Heidelberg: Carl Winter.

Jain, Jagdishchandra. 1981. *Prakrit Narrative Literature*. Delhi: Munshiram Manoharlal.

Jamison, Stephanie. 1990. The tense of the predicated past participle in Vedic and beyond. *Indo-Iranian Journal* 33(1):1–19.

Jamison, Stephanie and Michael Witzel. 2002. Vedic Hinduism. Available at http://www.people.fas.harvard.edu/~witzel/vedica.pdf.

Kamp, Hans. 1979. Events, instants, and temporal reference. In R. Bäurle, U. Eglis, and A. von Stechow, eds., *Semantics from Different Points of View*, pages 376–417. Berlin: Springer.

Kellogg, S. H. 1893. *Grammar of the Hindi Language*. Delhi: Motilal Banarsidass. Second edition, reprinted 1990.

Kiparsky, Paul. 1998. Aspect and event structure in Vedic. In R. Singh, ed., *Yearbook of South Asian Languages and Linguistics*, vol. 1. Delhi: Sage Publications.

Kiparsky, Paul. 2005. The Vedic injunctive: Historical and synchronic implications. *The Yearbook of South Asian Languages and Linguistics* pages 219–235.

Klein, Wolfgang. 1994. *Time in Language*. London: Routledge.

Lee, Jungmee and Judith Tonhauser. 2010. Temporal interpretation without tense: Korean and Japanese coordination constructions. *Journal of Semantics* 27(3):307–341.

Lienhard, Siegfried. 1961. *Tempusgebrauch und Aktionsartenbildung in der modernen Hindī*. Stockholm: Almqvist & Wiksell.

Macdonell, Arthur A. 1927. *A Sanskrit Grammar for Students*. Oxford: Oxford University Press. Third edition.

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

Oberlies, Thomas. 2003. *A Grammar of Epic Sanskrit*. Berlin and New York: Walter de Gruyter.

Partee, Barbara. 1984. Nominal and temporal anaphora. *Linguistics and philosophy* 7(3):243–286.

Pischel, Richard. 1900. *Grammatik der Prākrit-Sprachen*. Delhi, India: Motilal Banarsidass. Translated from German by Subhadra Jha, 1981.

Reichenbach, Hans. 1947. *The Elements of Symbolic Logic*. New York: Dover.

Renou, Louis. 1925. *La Valeur du Parfait dans les Hymnes Védiques*. Paris: E. Champion.

Sen, Sukumar. 1953. *Historical Syntax of Middle Indo-Aryan*. Calcutta: Linguistic Society of India.

Sen, Sukumar. 1960. *A Comparative Grammar of Middle Indo-Aryan*. Calcutta: Linguistic Society of India.

Shaer, B. 2003. Toward the tenseless analysis of a tenseless language. *Proceedings of Semantics of Under-represented Languages of the Americas (SULA)* 2:139–156.

Singh, Ram Adhar. 1980. *Syntax of Apabhramsa*. Calcutta: Simant Publications.

Smith, Carlota. 1991. *The Parameter of Aspect*. Dordrecht: Kluwer Academic Publishers.

Smith, Carlota. 2008. Time with and without tense. In J. Guéron and J. Lecarme, eds., *Time and Modality*, pages 227–249. Berlin: Springer.

Speijer, J.S. 1886. *Sanskrit Syntax*. Delhi: Motilal Banarsidass. Republished 1973.

Tonhauser, Judith. 2011. Temporal reference in Paraguayan Guaraní, a tenseless language. *Linguistics and Philosophy* 34.3:1–47.

Trumpp, Ernest. 1872. *Grammar of the Sindhi Language: Compared with the Sanskrit-Prakrit and the Cognate Indian Vernaculars*. London: Trübner and company.

Ultan, Russell. 1978. The nature of future tenses. In J. Greenberg, ed., *Universals of Human Language*, pages 83–123. Stanford: Stanford University Press.

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

Whitney, William Dwight. 1889. *Sanskrit Grammar*. Leipzig: Breitkopf und Hartel.

Whitney, William Dwight. 1892. On the narrative use of the imperfect and the perfect in the Brahmanas. *Transactions of the American Philological Association* 11:5–34.

Witzel, Michael. 1999. Substrate languages in Old Indo-Aryan (R̥gvedic, Middle and Late Vedic. *Electronic Journal of Vedic Studies (EJVS)* 5(1):1–67.

Woolner, Alfred C. 1975. *Introduction to Prakrit*. Delhi, India: Motilal Banarsidass. Reprinted 1975.

The role of clefting, word order and given-new ordering in sentence comprehension: Evidence from Hindi

SHRAVAN VASISHTH, RUKSHIN SHAHER, AND NARAYANAN SRINIVASAN,
University of Potsdam; University of Potsdam; Centre of Behavioural and Cognitive Sciences, India

ABSTRACT

Two Hindi eyetracking studies show that clefting a noun results in greater processing difficulty initially, due to the extra processing steps involved in encoding a clefted noun (e.g., for computing the exhaustiveness interpretation). However, this extra difficulty in encoding a clefted noun results in a processing advantage when the clefted noun needs to be retrieved later on in the sentence – the clefted noun is retrieved faster in subsequent processing compared to its non-clefted counterpart. We also show that given-new ordering yields a processing advantage over new-given order, but this is only seen after the whole sentence is processed, i.e., it may be a late effect that occurs after syntactic processing is completed. Finally, following up on work on German by Hörnig et al. (2005), we present evidence that non-canonical order can be processed more easily than canonical order given appropriate context.

1   Introduction

What is the functional motivation for elaborate devices such as clefting and given-new ordering? Although it is clear that they serve as information-structuring devices for effectively communicating a message to the hearer/reader, it is less clear how exactly this kind of restructuring impacts processing in real-time sentence comprehension. In this paper, we investigate the role that clefting, given-new ordering and word order play in facilitating online sentence comprehension. We investigate these issues in Hindi, a language that has not been studied extensively in psycholinguistics.

We begin by surveying some of the relevant literature relating to clefting, given-new ordering and word order variation, and then present two Hindi eyetracking studies that address several open issues relating to these information-structuring devices.

1.1   Given-new ordering and word order variation

Psycholinguistic research on language comprehension has shown that given-new ordering is a major information-structure (IS) device. A widely accepted finding is that ordering a given (e.g., previously mentioned) referent before a new referent facilitates processing (Clark and Haviland 1977, Yekovich et al. 1979). According to Clark and Haviland, listeners and speakers adhere to a *given-new contract*; when presented with a sentence, listeners compute what part of the sentence is the given information in order to find an antecedent for it in memory before integrating the new information into memory. The computation of given information is aided by cues such as focal stress, type of construction (passive, cleft, pseudocleft), etc.

Yekovich et al. (1979) demonstrated the given-new processing advantage using a comprehension time task. They showed that in sentences like *The shark attacked the diver near the reef*, the reading time for the entire

sentence was faster when the subject (*shark*), as opposed to the object (*diver*) was given in context (i.e., was previously mentioned).[1]

In a series of studies, Hörnig and colleagues uncovered several subtleties that extend the given-new generalization (see, e.g., Hörnig et al. 2005, Hörnig et al. 2006). They carried out a psycholinguistic investigation of spatial assertions; in a picture-verification task, participants first read two propositions P1 and P2 in sequence, and in doing so had to build a mental model of the relative spatial locations of three objects — call them A, B, and C (the propositions described the relative locations of these objects).[2] After participants signaled with a button-press that the two propositions had been processed, a picture containing A was shown, and after 1000 milliseconds another picture containing C was added to the display (in half the trials C was shown before A). Participants had to verify whether the objects' locations matched the descriptions in the two propositions.

A surprising finding was that if a proposition P1 like *A is to the left of B* is followed by another proposition P2 that describes the location of a third, new element *C* relative to the spatial location of the given element *B*, then P2 is processed faster if the locative expression containing *B* is topicalized, i.e., occurs first in the sentence. That is, processing is *easier* if P2 has a topicalized locative, *To the left of B is C*, and more difficult if the order is canonical, *C is to the left of B*.

Following other work, Hörnig et al. (2005) refer to the locative expression as the *relatum*, since the locative phrase relates the target referent (which in the above example is C) in the mental model (Johnson-Laird 1983) of the comprehender. They hypothesized that two constraints operate independently and additively to determine integration speed in comprehension: the relatum should be given (*relatum=given*), and the order of elements should be: given, then new (*given-new*). These hypotheses were confirmed in their picture verification experiments.

The work by Hörnig and colleagues motivates one of our research questions: does the given-new preference hold in a language like Hindi, and can we find evidence for the *relatum=given* constraint that Hörnig and colleagues found in German, which, like Hindi, is also a relatively free word-order language? These constraints are expected to be cross-linguistically applicable, but there appears to be little cross-linguistic evidence available beyond relatively well-studied languages like English and German.

1.2   Clefting

Apart from the given-new ordering and *relatum=given* constraints, focus also plays an important role in information structuring. It is well-known that listeners detect focused information more quickly than non-focused information (Cutler and Fodor 1979) and can also remember it better than non-focused information (Singer 1976). A dramatic example of the effect of focus is the disappearance of the Moses illusion under focus. The Moses illusion refers to the following phenomenon. When participants are asked, *How many animals of each kind did Moses take on the Ark?*, most respond *two*, even though they know that the person concerned here is Noah and not Moses. This illusion disappears when the incorrect information is focused by clefting the critical noun; e.g., when they are asked to judge the truth of a sentence like *It was Moses who took two animals of each kind on the Ark*. This suggests that it is easier to detect errors in focused information as opposed to non-focused information (Bredart and Modolo 1988, Bredart and Docquier 1989).

Like other types of focus, *it*-clefts also facilitate processing in online sentence comprehension tasks. In a probe recognition and naming task, Birch and Garnsey (1995) observed that clefted nouns are named faster and more accurately than non-clefted nouns, suggesting that it is easier to access the clefted material. Similarly, in an eyetracking study on pronoun resolution, Foraker and McElree (2007) found that clefted noun phrases were more available as antecedents for anaphoric pronouns than their non-clefted counterparts; they argued that this clefting advantage comes from activation boosts rather than special buffers for clefted elements.

---

[1]Note, however, that the given-new preference is not a global constraint on processing. For example, Frazier and Clifton (2004) showed that the given-new ordering preference is limited to certain constructions. Furthermore, several studies have shown that the given-new preference does not hold for premise integration in a spatial reasoning task (Ehrlich and Johnson-Laird 1982, Johnson-Laird 1983, Baguley and Payne 1998).

[2]In their experiments, the letters A, B, and C stand for names of animals.

Interestingly, *it*-clefts interact with the given-new ordering constraint: so-called focus-background clefts (Delin and Oberlander 1995) like *It was John who bought a book* generally partition the sentence into a new-information segment (the clefted noun) and a given-information segment (the predicate), resulting in a new-given ordering (cf. Delin and Oberlander 1995; for so-called topic-comment *it*-clefts, this generalization does not hold). As a consequence, in focus-background *it*-clefts the default ordering is not given-new but rather new-given. In such clefts, it is possible that the preferred order is new-given, not given-new.

One goal of the present study was to investigate whether the given-new ordering preference is reversed in cleft constructions. Another goal was to further investigate the effect that clefting has on processing, independent of the given-new ordering constraint. We discuss the background for this second goal next.

As mentioned above, clefted material has been argued to be more accessible in memory. One reason for this greater accessibility may be that the focused item achieves a higher activation in memory because the clefted item ends up being encoded more richly. In a clefted structure like *It was John who bought a book*, an exhaustiveness interpretation is implied: no one but John bought a book.[3] Processing the clefted noun would at least involve generating the exhaustiveness reading and associating it with the noun; these are probably the underlying processing steps that result in a richer encoding of the clefted noun in memory. Indeed, recent work by Hofmeister (2009) shows that this extra processing increases the activation of the clefted noun in memory, which in turn results in easier retrieval of the noun at a subsequent stage.

Thus, the evidence shows that clefting a noun makes it more accessible in memory at a subsequent stage. In addition, evidence also exists for increased encoding cost due to clefting; this comes from an eyetracking study conducted by Birch and Rayner (1997). They compared clefted sentences such as (1a) with a non-clefted sentence (1b).

(1)   a. It was the *suburb* that received the most damage from the ice storm.
      b. Workers in the *suburb* hurried to restore power after the ice storm.

They found twice as many regressive eye movements[4] to the first three words of the sentence in the clefted condition compared to the non-clefted one. In addition, the probability of regressing out of the clefted word was higher but did not reach significance (note however that in a similar experiment (their 3B) presented in Birch and Rayner (2010), regression probability did reach significance). Finally, longer re-reading times were seen on the clefted versus non-clefted word. Birch and Rayner suggest that this is evidence for the encoding cost arising from the processing events triggered by a cleft structure. As an aside, note that the sentences in (1) are not really minimal pairs: the clefted noun in (1a) is a subject, whereas the same noun in (1b) occurs inside a prepositional phrase adjunct of a subject. For this reason, it is reasonable to question whether reading times can be compared in this pair of sentences. In fact, in a subsequent study, Birch and Rayner (2010) found shorter re-reading times in the clefted condition (their Experiment 3B).

The above review of the literature shows that clefting induces an encoding cost at the clefted word, but also results in facilitation later on in the sentence. However, to our knowledge the presence of both effects — increased encoding cost followed by facilitation during subsequent processing — has only been found in a self-paced reading study conducted by Hofmeister (2009) and in an eye-tracking study by Birch and Rayner (2010) (although Birch and Rayner argue in this paper that clefting only facilitates processing, their Experiment 3B shows increased regression probability at the clefted noun versus the non-clefted one). Indeed, recent work by Morris and Folk (1998) and Ward and Sturt (2007) argues that clefting results only in an integration advantage but (contrary to the claims by Birch and Rayner 1997, and contrary to Hofmeister's 2009 and Birch and Rayner's 2010 findings) not a greater encoding cost.

These disagreements in the literature lead to the second goal of the research presented here. We set out to establish whether (i) only encoding cost is seen with clefts, consistent with Birch and Rayner (1997); (ii) only facilitation effects are seen, consistent with Ward and Sturt (2007) and Morris and Folk (1998); or (iii) both encoding cost and subsequent facilitation occur in cleft constructions, as suggested by Hofmeister (2009) and

---

[3]For online evidence from event-related potentials that the exhaustiveness reading associated with clefts is not truth-conditional, see Drenhaus et al. (2011).

[4]Regressions, or leftward eye-movements while reading, have been associated with increased processing effort; see, e.g., Frazier and Rayner (1982).

Birch and Rayner (2010). Our results turn out to be consistent only with Hofmeister's 2009 and Birch and Rayner 2010's findings.

1.3 Some facts about Hindi word order and clefts

As mentioned above, our major research questions for Hindi are the following: (i) is given-new order the preferred order; (ii) in focus-background clefts, is new-given order the preferred order; (iii) does the *relatum=given* constraint apply in Hindi; and (iv) what is the effect of clefting on sentence comprehension? Because it is difficult to conduct multiple eyetracking studies in India, we took the unusual step of including all these questions into one experiment. This admittedly results in a more complex design that is generally deprecated in psycholinguistic research; however, for logistical reasons it was necessary to include as many of these questions as possible in one experiment.

Before we turn to the experiments, it may be useful to summarize some relevant aspects of Hindi syntax. We consider cleft structures first and then Hindi word order as it relates to our experimental design.

As Srivastav (1991) points out, there are essentially three different ways to realize relative clauses in Hindi: left-adjoined relative clauses, embedded relatives, and right-adjoined relative clauses.[5]

(2) a. Left-adjoined relative clause
         jo   laRkii   khaRii   hai vo lambii hai
         who girl.F.Nom standing is   she tall   is
         'The girl who is standing is tall.'
     b. Embedded relative clause
         vo laRkii   jo   khaRii   hai lambii hai
         that girl.F.Nom who standing is   tall   is
         'The girl who is standing is tall.'
     c. Right-adjoined relative clause
         vo laRkii   lambii hai jo   khaRii   hai
         that girl.F.Nom tall   is   who standing is
         'The girl who is standing is tall.'

Hindi clefts superficially look like right-adjoined relative clauses. An example of a cleft is shown in (3a). However, it is clear that there is a difference between the two structures. For example, the right-adjoined version (2c) derives from the center-embedded relative clause (2b), but the cleft has no center-embedded correlate, see (3b).

(3) a. vo laRkii   hai jo   khaRii   hai
         that girl.F.Nom is   who standing is
         'It is that girl who is standing.'
     b. *vo laRkii   jo   khaRii   hai hai
         that girl.F.Nom who standing is   is
         '(Intended) It is that girl who is standing.'

Moreover, as mentioned above, the cleft construction induces an exhaustiveness reading (accompanied by narrow focus on the clefted noun) that is not present in the right-adjoined relative clause shown in (2c). This suggests that the cleft should be analyzed as a construction with the fixed form *N hai jo predicate*, rather than a right-adjoined relative clause derived from a center embedded relative clause. For our purposes, it is only important to note that the cleft construction in Hindi induces the same exhaustiveness interpretation that *it*-clefts do in English and other languages, and that unlike English the segment signaling clefting (*hai jo*, 'is who') occurs after the head noun.

Regarding word order, Hindi is well-known as a head-final language (see Gambhir 1981, Kidwai 2000, a.o.). As in many other languages, there is a subject-first preference; this fact is relevant for our studies, where we look at locative expressions like *X is to the left of Y*. The canonical word order in a locative expression

---

[5]The glosses used for Hindi-Urdu are follows: M=Masculine; F=Feminine; Nom=Nominative; Obl=Oblique; Gen=Genitive; PROG=Progressive; EMPH=Emphatic.

is grammatical subject, locative expression, and copula, see (4a). In the locative construction, several non-canonical orders are possible, such as locative inversion with canonical (4b) and non-canonical (4c) word order inside the locative phrase. In this paper we focus on the locative inversion shown in (4b); this is because one of the goals of the present paper is to compare our findings for the processing of non-canonical order with those of Hörnig et al. (2005). In this paper, we will refer to the first noun phrase in the canonical sentence (4a) as the subject noun, and the noun in the locative phrase as the non-subject noun.

(4) a. duurbiin   banduuk=kii daayii taraf hai
         binoculars.F.Nom gun.F=F.Gen right   side   is
         'The binoculars are to the right of the gun.'

     b. banduuk=kii daayii taraf duurbiin   hai
         gun.F=F.Gen right   side   binoculars.F.Nom is
         'To the right of the gun are the binoculars.'

     c. daayii taraf banduuk=kii duurbiin   hai
         right   side   gun.F=F.Gen binoculars.F.Nom is
         'To the right of the gun are the binoculars.'

We describe next two eyetracking experiments using two different designs that address the major goals of this study.

2 Experiment 1

This experiment was an eyetracking study that followed the design of Hörnig et al. (2005): participants were asked to mentally visualize the relative positions of three objects as they read two successive sentences that described the relative position of these objects. Participants' fixations on the words in the sentences were recorded as they carried out the task. The details of the experiment are described next.

2.1 Participants

Thirty-two students at the Centre of Behavioural and Cognitive Sciences at Allahabad, India, participated for payment.

2.2 Design

The experiment design is summarized schematically in Table 1. As mentioned above, participants were shown two sentences successively and then shown a picture. The two sentences described the relative position of three objects; participants were asked to build a mental representation of the relative position of three objects as they read the sentences. After both sentences had been read, a picture showing two objects appeared on the screen after a latency of 1000 ms, and the participants had to decide, by pressing a button, whether the layout of these objects matched the description provided. An example picture is shown in Figure 1. The 1000 ms latency was used in order to stay as close to the original Hörnig et al. design as possible; it does not affect the main findings in our paper since we analyze the reading times preceding the picture verification task.

The first sentence was either in canonical order or non-canonical order (where non-canonical order refers to the locative inversion shown in (4b)), and the second sentence had the same word order as the first sentence. This parallelism in word order was maintained in order to prevent the task from becoming excessively difficult. Another common factor between the two sentences was that either the subject noun or the non-subject noun in the first sentence was repeated in the second sentence; this repetition of one of the nouns in the first sentence ensured that one noun (either the subject or the non-subject) in the second sentence was previously mentioned or given; the other noun in sentence 2 was therefore new, i.e., not previously mentioned.

We provide examples of the first sentence next. Consider the canonical order sentences first. As shown in (5), the canonical order sentences had two versions; in one version (5a) the target noun that would be repeated in the second sentence (in clefted or non-clefted form) was in subject position, and in the second version it was in non-subject position (5b). In the present example this word is *duurbiin*, 'binoculars'; hereafter, we will refer to this word in the first sentence as the antecedent of the given word in the second sentence.

Canonical Sentences

Sentence 1:

Subject Antecedent - *The binoculars are to the right of the gun.*

Non-Subject Antecedent - *The gun is to the left of the binoculars.*

Sentence 2:

	Given NP1	**New NP1**
Clefted	There is a chair but it is the binoculars that are to the left of the flag.	There is a chair but it is the flag that is to the right of the binoculars
Non-clefted	There is a chair and the binoculars are to the left of the flag.	There is a chair and the flag is to the right of the binoculars.

Non-canonical Sentences

Sentence 1:

Subject Antecedent - *To the right of the gun are the binoculars.*

Non-Subject Antecedent - *To the left of the binoculars is the gun.*

Sentence 2:

	Given NP1	**New NP1**
Clefted	There is a chair but it is the binoculars to whose right is the flag.	There is a chair but it is the flag to whose left are the binoculars.
Non-clefted	There is a chair and to the right of the binoculars is the flag.	There is a chair and to the left of the flag are the binoculars.

TABLE 1 Schematic design of Experiment 1.

The antecedent was systematically controlled to be the grammatical subject or object in order to balance the design; i.e., we wanted to avoid the asymmetry whereby only subject antecedents or only object antecedents would be used, since such an asymmetry could lead to participants developing a strategy for processing the sentences. There was no theoretical question associated with this manipulation.

FIGURE 1 Example of picture shown on screen in Experiment 1.

(5) a. Canonical order, Subject antecedent
   duurbiin   banduuk=kii daayii taraf hai
   binoculars.F.Nom gun.F=F.Gen right   side is
   'The binoculars are to the right of the gun.'

   b. Canonical order, Non-Subject antecedent
   banduuk   duurbiin=kii   baayii taraf hai
   gun.F.Nom binoculars.F=F.Gen left   side is
   'The gun is to the left of the binoculars.'

After participants finished reading the first sentence (they indicated this by means of a button press), they were shown the second sentence, which was either clefted or non-clefted, and the first noun phrase was either given or new. The second sentence was the target sentence, i.e., we analyzed reading times only for this sentence.

Each target sentence began by introducing a distractor item (in the example below, this phrase was *There is a chair*); the purpose of this distractor was to set up a contrast that made the cleft sound more natural. Thus, in the cleft condition the distractor phrase was *There is a chair but...*, and in the non-cleft condition (which did not require an explicit contrast set) the distractor phrase was *There is a chair and....* The purpose of providing the distractor phrase in the non-cleft sentence was simply to create as similar a sentence as possible to the clefted sentence. The noun mentioned in the distractor phrase never appeared in the picture shown at the end of a trial.

Shown below are example target sentences for the canonical word order case; the first NP (NP1) following the distractor phrase (*There is an X and/but...*) is either clefted or non-clefted, and is either given or new. The front-slashes (/) indicate the partitioning of regions of interest for purposes of data analysis (the sentences were presented to participants without such a marker).

(6) a. Canonical order, Non-Clefted, NP1 given
   [ek kursii   bhii hai aur] **/ duurbiin /**   jhanDee=/kii baayii taraf / hai
   one chair.F.Nom also is   and   binoculars.F.Nom flag.M=F.Gen left   side   is
   'There is a chair and the binoculars are to the left of the flag.'

   b. Canonical order, Non-Clefted, NP1 new
   [ek kursii   bhii hai aur] **/ jhanDaa /** duurbiin=/kii   daayii taraf / hai
   one chair.F.Nom also is   and   flag.M.Nom binoculars=F.Gen right   side   is
   'There is a chair and the flag is to the right of the binoculars.'

    c. Canonical order, Cleft, NP1 given
        [ek kursii   bhii hai lekin] / **duurbiin** /   hai jo / jhanDe=/kii   baayii taraf / hai
        one chair.F.Nom also is but   binoculars.F.Nom is that flag.M=F.Gen left   side   is
        'There is a chair but it is the binoculars that are to the left of the flag.'

    d. Canonical order, Cleft, NP1 new
        [ek kursii   bhii hai lekin] / **jhanDaa** /   hai jo / duurbiin=/kii   daayii taraf / hai
        one chair.F.Nom also is but   flag.M.Nom is that binoculars.F=F.Gen right side is
        'There is a chair but it is the flag that is to the right of the binoculars.'

    The non-canonical counterparts of the first sentence are shown in (7), and the corresponding target sentences are shown in (8).

(7) a. Non-canonical order, Subject antecedent
        banduuk=kii daayii taraf duurbiin   hai
        gun.F=F.Gen right side binoculars.F.Nom is
        'To the right of the gun are the binoculars.'

    b. Non-canonical order, Non-Subject antecedent
        duurbiin=kii   baayii taraf banduuk   hai
        binoculars.F=F.Gen left   side gun.F.Nom is
        'To the left of the binoculars is the gun.'

(8) a. Non-canonical order, Non-Clefted, NP1 given
        [ek kursii   bhii hai aur] / duurbiin=/kii   daayii taraf / jhanDaa /   hai
        one chair.F.Nom also is and   binoculars.F=F.Gen right side flag.M.Nom is
        'There is a chair and to the right of the binoculars is the flag.'

    b. Non-canonical order, Non-Clefted, NP1 new
        [ek kursii bhii hai aur] / jhanDee=/kii baayii taraf / duurbiin /   hai
        one chair also is   and flag.M=F.Gen left   side binoculars.F.Nom is
        'There is a chair and to the left of the flag are the binoculars.'

    c. Non-canonical order, Cleft, NP1 given
        [ek kursii   bhii hai lekin] / duurbiin /   hai jis=kii /   daayii taraf / jhanDaa /   hai
        one chair.F.Nom also is but   binoculars.F.Nom is that=F.Gen right side flag.M.Nom is
        'There is a chair but it is the binoculars to whose right is the flag.'

    d. Non-canonical order, Cleft, NP1 new
        [ek kursii   bhii hai lekin] / jhanDaa /   hai jis=kii /   baayii taraf / duurbiin /   hai
        one chair.F.Nom also is but   flag.M.Nom is that=F.Gen left   side binoculars.F.Nom is
        'There is a chair but it is the flag to whose left are the binoculars.'

    After participants had indicated that they had finished reading the second sentence, the picture shown in Figure 1 was presented after an interval of one second. Once the participant had indicated that they had finished viewing the picture, the picture disappeared and a question such as (9) appeared on the screen.

(9) dikhaye gaye chitron=kii   aapsii   sthitii   varnan=ke   anusaar hai
     shown were images.M=F.Gen respective position.F.Nom description.M=M.Obl.Gen according is
     ya nahiiN?
     or not
     'Does the layout provided here correspond to the description provided?'

    If the layout shown was incorrect given the sentences, the participant was expected to respond to the question with a *No*. There were an equal number of correct no and yes responses.

    The noun phrases used in the experiment were common nouns such as names of fruits, animals, musical instruments and everyday objects. The positions of these objects were described in terms of prepositions such as 'to the left of', 'to the right of', 'above', 'below' and combinations of these. The prepositions were balanced across items. The full list of items is available from the authors.

To summarize, the experiment had three factors, each with two levels: (a) Clefting of first NP; (b) Givenness of first NP (given/new); and (c) Word order (canonical or non-canonical). In addition, although this was not theoretically interesting, in order to maintain balance in the items, the antecedent of the given NP was either the subject or non-subject of the second sentence. 48 sets of items were constructed. There were no fillers because (a) Hörnig et al. (2005) also used no fillers and we wanted to stay close to their design, and (b) pre-testing showed that participants found the task quite difficult, so including fillers would have made the task intractably hard for them (it was probably for this reason that that Hörnig and colleagues also did not use fillers in their study).

2.3 Method

Participants were seated 55 cm from a 17" color monitor with 1024×768 pixel resolution. The eyetracker used was an IView-X (SensoMotoric Instruments) running at 500 Hz sampling rate, 0.025 degree tracking resolution, < 0.5 degree gaze position accuracy. Participants were asked to place their head in a frame and to position their chin on a chin-rest for stability. The angle per character was 0.26 degrees (3.84 characters per degree of visual angle).

Participants were asked to avoid large head movements throughout the experiment. A keyboard was used to record button responses. The presentation of the materials and the recording of responses was controlled by two PCs running proprietary software (the software used was Presentation, and SensoMotoric Instruments' own software for eyetracker control).

At the start of the experiment the experimenter performed a standard calibration procedure, which involves participants looking at a grid of thirteen fixation targets in random succession in order to validate their gazes. Calibration and validation were repeated after every 10–15 trials throughout the experiment, or if the experimenter noticed that measurement accuracy was poor (e.g., after large head movements or a change in the participant's posture).

2.4 Dependent measures

In reading studies, several dependent measures are usually presented. Some of the commonly discussed dependent measures and their interpretation as regards word and sentence comprehension processes are as follows. *First fixation duration* (FFD) is the time elapsed during the first fixation during first pass (the first encounter with a region of interest as the eye traverses the screen from left to right), and has been argued to reflect lexical access costs (Inhoff 1984). *Gaze duration* or *first pass reading time* (FPRT) is the summed duration of all the contiguous fixations in a region before it is exited to a preceding or subsequent word; Inhoff (1984) has suggested that FPRT reflects text integration processes, although Rayner and Pollatsek (1987) argue that both FFD and FPRT may reflect similar processes and could depend on the speed of the cognitive process. *First-pass regression probability*, the probability of the eye making a leftward saccade to a previous word during first pass, is another measure that is sometimes used as an index of processing difficulty. *Right-bounded reading time* (RBRT) is the summed duration of all the fixations that fall within a region of interest before it is exited to a word downstream; it includes fixations occurring after regressive eye movements from the region, but does not include any regressive fixations on regions outside the region of interest. RBRT may reflect a mix of late and early processes, since it subsumes first-fixation durations. *Re-reading time* (RRT) is the sum of all fixations at a word that occurred after first pass; RRT has been assumed to possibly reflect the costs of late-stage processes (Clifton et al. 2007, 349), and recent work suggests that it may be informative about retrieval costs (Gordon et al. 2006, 1308, Vasishth et al. 2008, Vasishth et al. 2010). Re-reading time is usually computed including zero re-reading times, an issue we return to later. Another measure that has been invoked in connection with late-stage processes is *regression path duration*, which is the sum of all fixations from the first fixation on the region of interest up to, but excluding, the first fixation downstream from the region of interest. Finally, *total reading time* (TRT) is the sum of all fixations on a word.

In Experiment 1, we present results based on regression probabilities and re-reading times because none of the early measures showed any effect; note that previous work (Birch and Rayner 1997) also found effects due to clefting in these measures. We follow the literature in presenting re-reading times including zeros, although the distinction (including or excluding zeros) is moot in the present case because the proportion of

re-reading was very high; the results do not change substantially regardless of whether we include zeros or not.[6]

2.5   Statistical analysis

A linear mixed model (LMM) was fitted to the data, with crossed random intercepts for participants and items, and with the factors as orthogonally coded predictors (fixed factors). LMMs have several advantages over repeated-measures ANOVA, one of them being that they allow us to take by-item and by-participant variance into account simultaneously, which is an improvement over separate analyses or the calculation of min-F (Clark 1973, Raaijmakers et al. 1999); Baayen (2008) presents further discussion of this issue. Throughout this paper we present coefficient estimates, their standard errors, t- or z-scores (depending on the dependent measure). An absolute t-score of 2 or greater indicates significance at the α level 0.05. Note also that the t-score is not accompanied by degrees of freedom or p-values. This is because degrees of freedom cannot be computed precisely in LMMs (Baayen 2008).

The statistical analyses on reading times were carried out on log-transformed values. We carried out log-transformed analyses in order to achieve approximately normal residuals.

2.6   Results

Using multiple regression models and orthogonal contrast coding, we examined the following effects: word order, clefting, givenness status of NP1, the interaction of NP1's givenness status with clefting, the interaction of word order with givenness, and the interaction of clefting with word order.

The relevant regions of interest were the first noun phrase in the second sentence (abbreviated below as NP1), and the integration region. Note that the first noun phrase is the one that occurs after the distractor phrase (e.g., after the phrase *There is a chair. . .* in Table 1); this would be the subject noun in canonical order and the non-subject noun in non-canonical order. The first noun phrase region was chosen for analysis because, as described above, previous work (e.g., Birch and Rayner 1997) has shown that the effect of clefting a noun phrase appears at the NP itself.

The integration region is the locative phrase (e.g., *kii daayii taraf*, 'to the right side of') where the relationship between the two referents becomes clear for the first time. This region is theoretically interesting because it is here that the first noun phrase would need to be retrieved in order to construct the spatial relationship between the two noun phrases.

We present the results by region of interest.

2.6.1   Region of interest: NP1

First-pass regression probability

First-pass regression probability was investigated using a generalized linear mixed model with a binomial link function; participants and items were the crossed random factors. The reading time measure (re-reading time) was also evaluated using linear mixed models, also with participants and items as crossed random factors.

	Estimate	Std. Error	z value	p-value
WO	-0.05	0.06	-0.81	0.41
Cleft	0.20	0.06	3.43	<0.01 *
Given$_{N1}$	0.05	0.06	0.84	0.40
Cleft $\times$ WO	-0.11	0.06	-1.88	0.06
Cleft $\times$ Given$_{N1}$	0.14	0.06	2.35	0.02 *
WO $\times$ Given$_{N1}$	-0.03	0.06	-0.46	0.65

TABLE 2   Summary of statistical analyses for first-pass regression probability at the first NP. The asterisk (*) signifies statistical significance at $\alpha = 0.05$.

---

[6]In Experiment 2, by contrast, the proportion of re-reading times is relatively low (approximately 20%); there, it is important to consider the consequences of including versus excluding zero re-reading times. We discuss this in the Results section of Experiment 2.

The results are summarized in Table 2. Regression probability at the first noun phrase was higher in the clefted versus the non-clefted condition (32.33% vs. 24.9%); an interaction was found between clefting and NP1's givenness status, such that when NP1 was given, clefting NP1 resulted in higher regression probability (34.12%) compared to the non-clefted case (21.67%), while when NP1 was new, the clefted noun had a regression probability (30.55%) that was comparable to the non-clefted case (28.12%). In addition, a marginal interaction was seen between clefting and word order such that the difference between clefted and non-clefted conditions was smaller in canonical order (clefted 31.07%, non-clefted 27.6%), than in non-canonical order (clefted 33.6%, non-clefted 22.19%).

Re-reading time The mean re-reading times are shown in Table 3 and the corresponding statistical analyses are in Table 4. We found:

1. *a word order effect*: re-reading time at the first noun phrase was faster in non-canonical word order than canonical;
2. *an effect of clefting*: the first noun phrase was re-read faster when it was clefted versus non-clefted, consistent with the advantage due to clefting found in previous work on clefting;
3. *no effect of givenness status of the first NP*;
4. *an interaction between clefting and word order*: a cross-over interaction was found between the clefting status of NP1 and word order, such that in canonical word order the clefted noun phrase was read faster than its non-clefted counterpart, whereas in non-canonical word order the clefted noun phrase was read slower than its non-clefted counterpart;
5. *an interaction between givenness status of NP1 and clefting*: a cross-over interaction was found between the givenness status of NP1 and clefting, such that when NP1 was given, clefting resulted in slower reading time, and when NP1 was new, clefting resulted in faster reading time;
6. *an interaction between givenness status of NP1 and word order*: a cross-over interaction was found between the givenness status of NP1 and word order, such that a given NP1 was read faster in non-canonical order (compared to canonical order), and a new NP1 was read slower in non-canonical order.

Main effects of Word Order and Clefting

Word order		Clefting	
Canonical	Non-canonical	Non-clefted	Clefted
715	685	711	670

Clefting and Word Order interaction

	Non-clefted	Clefted
Canon.	777	658
Non-canon.	644	725

N1 Givenness status, Clefting and WO

	Given	New
Cleft	704	676
Non-cleft	682	739
Canon.	758	672
Non-canon.	621	744

TABLE 3 Experiment 1. Summary of effects at NP1 (re-reading times in msecs).

2.6.2 Region of interest: Integration region

In the integration region, we looked at the dependent variable re-reading time.

As summarized in Tables 5 and 6, we found (i) a main effect of clefting, such that the integration region was read faster in the clefted condition; (ii) a main effect of word order, such that non-canonical order was

	Estimate	Std. Error	t value
WO	-0.14	0.05	-3.04 *
Cleft	-0.20	0.05	-4.27 *
Given$_{N1}$	-0.02	0.05	-0.37
Cleft $\times$ WO	-0.13	0.05	-2.80 *
Cleft $\times$ Given$_{N1}$	0.12	0.05	2.67 *
WO $\times$ Given$_{N1}$	-0.12	0.05	-2.64 *

TABLE 4 Summary of statistical analyses for re-reading time at the first NP. The asterisk (*) signifies statistical significance at $\alpha = 0.05$.

read slower than canonical; and (iii) an interaction between word order and givenness, such that when NP1 was given the integration region was read equally quickly in canonical and non-canonical order, but when NP1 was new, the integration region was read faster in canonical order than non-canonical.

Main effects of word order and clefting			
Word order		Clefting	
Canonical	Non-canonical	Non-clefted	Clefted
759	962	903	854
Clefting and Word Order interaction			
	Non-clefted	Clefted	
Canon.	654	875	
Non-canon.	1077	840	
N1 Givenness status, Clefting and WO			
	Given	New	
Cleft	743	964	
Non-cleft	762	1041	
Canon.	757	761	
Non-canon.	750	1164	

TABLE 5 Summary of re-reading time results at the integration region.

	Estimate	Std. Error	t value
WO	0.90	0.06	14.60 *
Cleft	-0.14	0.06	-2.31 *
Given$_{N1}$	-0.08	0.06	-1.33
Cleft$\times$ Given$_{N1}$	0.02	0.06	0.38
WO$\times$ Given$_{N1}$	-0.16	0.06	-2.59 *
Cleft$\times$WO	0.08	0.06	1.26

TABLE 6 Re-reading time (log ms) in integration region the effect of word order, clefting, clefting-word order interaction, and clefting-NP1-givenness status interaction.

2.7 Discussion

To summarize the main results at the first noun phrase, (i) regression probability was higher in the clefted versus non-clefted condition; (ii) regression probability was higher in the clefted versus non-clefted condition when the first NP was given but not when it was new; (iii) re-reading time was faster in the clefted condition compared to the non-clefted one; (iv) re-reading time at the first noun phrase was faster in non-canonical versus canonical order; (v) faster re-reading time was seen in the clefting condition in canonical order but not in non-canonical order (where clefted noun phrases were read *slower* than non-clefted ones); (vi) clefted

nouns were read faster when the first noun phrase was new, not when it was given (when the noun phrase was given, clefts were read slower than non-clefts).

The fact that regression probability increases at NP1 when the noun is clefted suggests increased encoding cost, similar to that found by Birch and Rayner (1997) and Experiment 3B of Birch and Rayner (2010).[7] As discussed earlier, the computationally expensive encoding process includes steps such as building the exhaustiveness interpretation associated with clefts. One important consequence of this encoding process is that the clefted noun receives a higher net activation, which renders it easier to process during second-pass. We see this in the faster re-reading time at the clefted noun. A similar result was also found by Birch and Rayner (2010); in their Experiment 3B, they also find higher regression probability at the clefted noun (which we interpret as encoding cost that results from computing the exhaustiveness interpretation) and easier processing of the clefted noun. Note, however, that Birch and Rayner (2010) found a processing advantage due to clefting in first-fixation durations, gaze duration, total number of fixations, and total reading time; no effect was seen in second-pass (re-reading) time. It is worth pointing out here that their data had relatively few cases of re-reading; their re-reading times are quite low on average (23 and 25 ms for non-clefted and clefted nouns, respectively). Since the vast majority of rereading times consists of zeroes, the statistical model is questionable, as discussed in the appendix.

The interaction between clefting of NP1 and givenness status of NP1 is potentially interesting. When NP1 was new, clefting the noun did not result in higher regression probability, whereas when NP1 was given, clefting resulted in higher regression probability. Assuming that processing a given noun induces a reactivation of the previous mention of that noun in the preceding sentence, clefting a given noun might additionally involve accessing the contents of the preceding sentence simultaneously while building the exhaustiveness interpretation. This could explain the greater processing difficulty at the clefted NP1 when NP1 was given. By contrast, when NP1 is new, since the meaning of the preceding sentence need not be accessed, a floor effect might result, i.e., there may be no significant cost to encoding a clefted new NP compared to a non-clefted new NP. An alternative explanation that we cannot rule out is that the interaction is due to the region preceding NP1; here, the clefted noun is preceded by *lekin*, 'but', and the non-clefted noun by *aur*, 'and'. It could well be the case that this difference in the preceding region is responsible for the interaction. We cannot answer this question definitively in the present work.

The re-reading time at NP1 is also revealing. The faster re-reading time at the clefted NP1 versus its non-clefted counterpart suggests easier processing of NP1 during second-pass. Taken together, the regression probability and re-reading time differences at NP1 show clear evidence for an early encoding cost and subsequent retrieval advantage due to clefting that was hypothesized at the beginning of this paper. An encouraging fact is that these findings are consistent with the existing literature on the processing of clefts, as discussed earlier.

An important finding is that NP1 is read faster in non-canonical order than canonical order; this is the reverse of the usual pattern found in the literature on word-order variation. However, this finding is related to that suggested by Hörnig et al. (2005). As discussed earlier, Hörnig and colleagues provide evidence that non-canonical order sentences are read faster than canonical order sentences when the initial part of a non-canonical order sentence provides important linking information to the preceding sentence; in their case this was so when the initial part of the sentence was given. Hörnig and colleagues refer to this as the *relatum=given* principle. Our results suggest an additional constraint independent of givenness: relatum-first. When the relatum occurs first, as it does in non-canonical word order, processing is easier, regardless of whether the relatum contains a given NP or not. In addition to this finding, the word order and givenness interaction provides support for the *relatum=given* principle. At NP1, a given NP1 was read faster in non-canonical order than canonical, as predicted by *relatum=given*. We should point out here that these "principles" are not intended to have the same level of generality as general principles of grammar; they may well be specific to the processing of relational expressions. Another important point, noticed by a reviewer, is that it is quite possible that the faster re-reading time at NP1 in non-canonical order could be due to increased

---

[7]As a reviewer correctly points out, we cannot rule out the possibility that the increased regression probability at NP1 may be due to differences arising from the preceding word (*lekin* 'but' versus *aur* 'and'). However, even if this finding turns out to be confounded, our main claim is about facilitation due to clefting at the point of retrieval of the clefted noun.

re-reading of the integration region (which showed longer re-reading times in non-canonical order); if that is the reason for the effect, the relatum-first effect would be simply epiphenomenal.

An unexpected finding is that the clefting advantage is reversed in non-canonical order — clefted nouns are harder to read than non-clefted ones. We had expected the advantage due to clefting to apply uniformly in both canonical and non-canonical word order. A plausible explanation for the adverse effect of word order is that in canonical order fewer processing steps are involved than in non-canonical order, and clefting the noun in non-canonical order may overload the processing system. In canonical order, only the exhaustiveness interpretation needs to be computed when a noun is clefted. But in non-canonical order the presence of the relational phrase at the beginning of the sentence also reactivates the memory trace of the preceding sentence; although non-canonical order makes processing easier (as discussed above), clefting the noun could disrupt the attempt to connect the content of the current sentence with the previous one, resulting in an interaction between clefting and word order. This is admittedly a post-hoc and speculative explanation and needs to be followed up with planned experiments.

Finally, clefted nouns are read faster than non-clefted nouns when the noun is new, and the opposite pattern is seen when the noun is given. This finding is consistent with the observation, discussed by Delin and Oberlander (1995), that in focus-background clefts the clefted noun is new. When the clefted noun is given, processing is expected to slow down because clefting a given noun is not the normal case in focus-background clefts. If this interpretation is correct, a clear prediction is that in topic-comment clefts the reverse preference should be seen — given nouns should be processed faster when clefted versus non-clefted, and new nouns should be processed slower when clefted. We intend to test this prediction in future work.

Turning next to the results at the integration region, we find (i) faster re-reading times in the clefted versus non-clefted conditions; (ii) faster re-reading time in canonical versus non-canonical order; and (iii) an interaction between word order and givenness, such that when NP1 was given the integration region was read equally fast in canonical and non-canonical order, but when NP1 was new the integration region was read faster in canonical than non-canonical order.

It is interesting that the integration region is read faster in canonical order compared to non-canonical order, especially when one considers the fact that the opposite pattern was found at NP1 (the *relatum-first* effect). In other words, although non-canonical order can facilitate processing at NP1 (due the *relatum-first* effect), processing the locative phrase (which occurs immediately after NP1) is more difficult in non-canonical than canonical order. This pattern is hard to reconcile with the relatum-first pattern we saw at NP1.

The interaction between word order and givenness status of NP1 at the integration region is also quite revealing. The well-known preference for canonical word order is seen when we consider sentences with a new NP1: the integration region is read faster in canonical versus non-canonical order. However, when NP1 is given, the canonical order preference is neutralized. This seems to be at least consistent with the *relatum=given* principle that Hörnig and colleagues propose: in non-canonical order the relational information comes first, and when the relational expression includes a given NP, this may render non-canonical order just as easy to process as canonical order.

Our next goal was to attempt to replicate the two effects we found for Hindi clefts: the increased encoding cost followed by facilitation during later processing. One aim of the second experiment was to determine whether the effect could be replicated in a simpler reading task; a second aim was to determine whether the clefting results in a short-lived activation-increase of the clefted noun — i.e., is the increased activation of the noun limited to the current sentence being processed — or does the clefted noun remain in a permanently high-activation state even when the clefted noun is retrieved in a subsequent sentence. A short-lived activation is predicted by decay-based models such as Dependency Locality Theory (Gibson 2000), and decay- and interference-based models of sentence comprehension such as the cue-based retrieval model of Lewis and Vasishth (2005). In both these classes of models, the memory trace for the prominent (clefted) element would degrade over time, either due to decay or due to interference, or perhaps a combination of both. An alternative possibility is that the noun made more prominent by clefting may remain active in memory, perhaps because it is in a focus buffer that is not subject to decay. Of course, evidence for a short-lived effect of clefting would necessarily be a null result, i.e., not much can be concluded from it. The more conclusive finding would be a

long-lasting facilitation due to clefting.

In addition to the above goals, we were also interested in the interaction of clefting with topic status, specifically whether the topic was a continuation or a shift (these terms are defined below) relative to the preceding sentence. It is well-known that topic continuations are easier to process than topic shifts (Grosz et al. 1995); we were interested in determining whether the processing advantage due to clefting is independent of this information-structural factor or whether there is an interaction. If the clefting advantage occurs independently of topic status, then this suggests that there is something about the syntactic structure itself that controls the activation level of the clefted noun, independent of other information structure factors. If an interaction is found, this would suggest that the activation-boost afforded by clefting is not an automatic reflex of the syntactic structure, but rather can be counteracted by information-structural (extra-syntactic) constraints like topic status (because topic status is dependent on preceding context). A priori, we would expect that clefting would facilitate processing in both types of topic, but that processing topic shifts would be more difficult overall; in other words, we expect a main effect of clefting, and a main effect of topic status, but no interaction. We describe the details of this experiment next.

3 Experiment 2

3.1 Participants

Thirty two native speakers of Hindi participated for payment at the Centre of Behavioural and Cognitive Sciences, Allahabad, India.

3.2 Design

Each trial consisted of three sentences. The first was a context sentence that introduced two referents; this context sentence always had a masculine and a feminine noun (the order of occurrence of masculine and feminine gender was counterbalanced across items). An example is shown in (10).

(10) ek   aadmii   saṛak=ke   kinaare ek aurat=ke   saath kharaa   thaa
    one man.M.Nom road.F=M.Obl.Gen near   one woman.F=M.Obl.Gen with standing was
    'A man was standing by the street corner with a woman.'

Following standard definitions for Hindi and Urdu (e.g., Kidwai (2000), Dwivedi (1994), Butt and King (1997)), the first noun of the context sentence was considered the topic. The context sentence was followed by a second sentence, which was either in canonical (SOV) or non-canonical (OSV) word order, and the first noun phrase in this sentence was either clefted or non-clefted. When the second sentence was in canonical order, the topic was a continuation of the preceding context sentence (a topic continuation); and when the sentence was in non-canonical order, the topic was a considered to have been shifted (a topic shift).

An example of the second sentence in canonical order is shown in (11). The clefting manipulation in the present experiment differs from that in the first experiment in two important respects. First, the noun phrase (which is either clefted or non-clefted) is not preceded by a distractor phrase; this has the consequence that the only way that the reader can infer that a contrast is intended for the first noun is by fixating on the first noun and then parafoveally processing the two words following the noun, or by reading the two words following the clefted noun. Second, in the canonical order sentence shown in (11), the topic of the sentence is the same as the topic of the first context sentence. We will refer to this as a topic continuation (Grosz et al. 1995).

(11) {aadmi   hai jo   / aadmi} aurat=ko   dekh rahaa thaa
     man.M.Nom is who man   woman.F=Acc see PROG was
     'It is the man who / The man was looking at the woman.'

By contrast, the non-canonical word order version of the second sentence, shown in (12), has a topic shift: the new topic is the object of the preceding sentence.

(12) {aaurat   hai jisko   / aurat=ko}   aadmi dekh rahaa thaa
     woman.F.Nom is who.Acc woman.F=Acc man   see   PROG was
     (Approximate translation) '(It is) the woman (whom) the man was looking at.'

The second sentence is followed by a third sentence; an example is shown in (13). This sentence begins with a pronoun *vo*, 'he/she'; in Hindi, such pronouns are unspecified for gender. Further on in the sentence, an adjective is introduced that disambiguates the pronoun (adjectives agree with the subject in gender marking); the disambiguation is either to the subject or the object of the second sentence. For example, in (13), masculine marking *-aa* on the adjective *moṭ-aa*, 'fat-masc' and the auxiliary *th-aa*, 'was-masc' signals that the person who being described as fat is the man. The corresponding feminine marker is *-ii*.

(13) vo   mere vicaar=me bahut=hii   moṭ-**aa/-ii** th-**aa/-ii**
       he/she my   opinion=in much=EMPH fat-**M/-F**   was-**M/-F**
       (Lit.) 'He/she, in my opinion, is very fat'.

After the triplet of sentences was read, a yes/no question was presented to probe the participants' comprehension of the sentences. There were 48 items, which are available from the authors. Here, too, the experiment items were presented without fillers; although this is not the standard practice in psycholinguistic research, logistical constraints associated with running this experiment in India forced us to keep the experiment as short as possible.

3.3   Method

The same method was used as in Experiment 1: fixations were recorded as participants read the sentences. Each sentence was presented separately; i.e., each trial began when the first sentence was presented, and then the participant signaled with a button-press that he/she has finished reading it; the button-press resulted in the first sentence being removed from the screen and the second sentence being presented; and so on. Participants had to fixate on a circle on the left edge of the screen to trigger the display of a sentence on the screen; this ensured that they began reading from the left edge of the sentence.

3.4   Predictions

The predictions in this experiment are as follows. First, clefting the noun in the second sentence should result in a processing advantage later in the sentence, i.e., at the verb, when the clefted noun is to be retrieved. This follows from the previous psycholinguistic research on clefting, and the results of experiment 1. Second, if the processing advantage due to clefting interacts with discourse-level factors like topic status, the clefted noun should be easier to process (should be fixated for a shorter duration) when it is a topic continuation than when it is a topic shift; alternatively, if the syntactic marker for clefting induces a processing advantage (by increasing activation of the clefted noun) regardless of discourse factors, no interaction should be seen between topic status and clefting. Third, if clefting status results in permanently prominent status of the clefted noun, fixation durations at the disambiguating region should be shorter when the antecedent is clefted rather than non-clefted; if topic status has a similar effect, we should see a similar facilitation at the disambiguating region when the antecedent is a topic continuation rather than a topic shift (main effects of clefting and topic status when the ambiguous region resolves to the first noun in the preceding sentence). Alternatively, if the advantage due to clefting and topic status is short-lived (i.e., is subject to decay and/or interference processes as some sentence comprehension theories assume), no advantage due to clefting and topic status should be seen in the third sentence.

3.5   Results

We present the results for sentence 2, where the regions of interest were the first NP in the sentence and the integration region (the verb); and for sentence 3, where the region of interest was the adjective where the antecedent was disambiguated. There was no effect of word order and no cleft × word order interaction; in other words, we failed to find any advantage due to topic continuation at the first noun in sentence 2.

3.5.1   Sentence 2: first NP (re-reading times)

As in Experiment 1, the clefted NP had faster re-reading times (480 ms) than the non-clefted one (630 ms), replicating the processing advantage due to clefting that we describe in Experiment 1. Re-reading times here exclude zero re-reading times (see Appendix) and all reading times greater than 2000 ms; this resulted in removal of 80% of the data (i.e., re-reading probability was only about 20%). The results of the linear mixed model are presented in Table 7.

No difference was found between the clefted and non-clefted conditions when zero re-reading times were included, and an analysis of re-reading probability also showed no effects. However, note that when re-reading times include zeros, the residuals are remarkably non-normal; this is regardless of whether one log-transforms re-reading time (see Figure 2 in Appendix).

	Estimate	Std. Error	t value
Cleft	-0.14322	0.06883	-2.08 *
WO	0.03537	0.06324	0.56
Cleft×WO	-0.04421	0.06129	-0.72

TABLE 7 Experiment 2. Effect of clefting, word order, and the interaction of the two factors on log re-reading times at the first NP. Crossed random intercepts were fit for items and participants in a linear mixed model.

3.5.2 Sentence 2, integration region (first-fixation durations)

Unlike Experiment 1, where only re-reading time showed any effect in the integration region, in Experiment 2 only first-fixation durations showed significant effects. The results of the analyses are summarized in Tables 8 and 9. Overall, the integration region was read faster when the first NP was clefted, replicating the effect found in Experiment 1. Interestingly, however, although we found no main effect of word order, an interaction was found between clefting and word order, such that the facilitation due to clefting was seen only in the non-canonical word order condition (the topic-shift condition); in the canonical word order condition (topic continuation) there was no advantage due to clefting in the integration region.

	Canon.	Non-canon.
Non-cleft	526	573
Cleft	532	496

TABLE 8 Experiment 2. Mean first-fixation durations by word order and clefting conditions at the integration region.

	Estimate	Std. Error	t value
Cleft	-0.046132	0.021565	-2.14 *
WO	0.002713	0.021645	0.13
Cleft×WO	-0.050441	0.021617	-2.33 *

TABLE 9 Experiment 2. Effect of clefting, word order, and the interaction of the two factors on log first fixation duration at the integration region. Crossed random intercepts were fit for items and participants in a linear mixed model.

3.5.3 Sentence 3, integration region (re-reading times and re-reading probability)

We analyzed the re-reading times at the adjective separately for canonical and non-canonical order because this allows us to investigate the effect of antecedent resolution while holding word order constant. In the non-canonical word order condition no effect was statistically significant. In canonical order, however, a main effect was found of antecedent resolution, such that the integration region was processed faster when the antecedent was the object rather than the subject. No effect of clefting was found, and no interaction was found between clefting and antecedent.

Here, re-reading times exclude zero re-reading times, that is, we are looking at pure second-pass re-reading time. Excluding zeros and all re-reading times greater than 2000 ms resulted in the removal of 85% of the data (i.e., re-reading occurred only in 15% of trials). Including zeros in re-reading times showed no effects, but, as discussed earlier, the residuals were remarkably non-normal, leading us to question whether the coefficients are interpretable. Re-reading probability at the disambiguating region for canonical and non-canonical word order (analyzed separately) showed no effects of clefting or antecedent type.

	Subj	Obj
Non-cleft	892	677
Cleft	735	565

TABLE 10 Experiment 2. Mean re-reading time by antecedent type and clefting conditions in the integration region (canonical word order).

	Estimate	Std. Error	t value
Cleft	-0.071756	0.065306	-1.10
Ant	0.144902	0.062597	2.31 *
Cleft×Ant	-0.004072	0.062283	-0.07

TABLE 11 Experiment 2. Effect of clefting, antecedent, and the interaction of the two factors on log re-reading time at the integration region in sentence 3, canonical word order. Crossed random intercepts were fit for items and participants in a linear mixed model.

3.6 Discussion

We present the experiment design below to save the reader the effort of revisiting the earlier sections of the paper:

(14)   Sentence 1:
        ek aadmii   saṛak=ke   kinaare ek aurat=ke   saath kharaa   thaa
        one man.M.Nom road.F=M.Obl.Gen near   one woman.F=M.Obl.Gen with standing was
        'A man was standing by the street corner with a woman.'
        Sentence 2 (canonical):
        {aadmi   hai jo   / aadmi} aurat=ko   dekh rahaa thaa
         man.M.Nom is   who   man   woman.F=Acc see   PROG was
        'It is the man who / The man was looking at the woman.'
        Sentence 2 (non-canonical):
        {aaurat   hai jisko   / aurat=ko}   aadmi dekh rahaa thaa
         woman.F.Nom is   who.Acc   woman.F=Acc man   see   PROG was
        (Approximate translation) '(It is) the woman (whom) the man was looking at.'
        Sentence 3:
        vo   mere vicaar=me bahut=hii   moṭ-**aa/-ii** th-**aa/-ii**
        he/she my   opinion=in much=EMPH fat-**M/-F**   was-**M/-F**
        (Lit.) 'He/she, in my opinion, is very fat.'

To summarize the results of Experiment 2, we find (i) a facilitatory effect of clefting on the first noun phrase in sentence 2, replicating the results of Experiment 1; (ii) an early facilitatory effect of clefting in the integration region; (iii) an interaction between word order and clefting in the integration region, such that the advantage due to clefting was only seen in non-canonical word order (the topic-shift condition); and (iv) in sentence 3, it is easier to resolve the antecedent to the object in sentence 2 than to the subject, when sentence 2 is in canonical order.

This experiment thus replicates our finding in Experiment 1 that clefting yields a processing advantage when the clefted noun needs to be retrieved later in the sentence: the re-reading time on the clefted noun is faster (result (i)) and the processing time at the integration region in sentence 2 is faster when the first NP is clefted (result (ii)). The interaction between clefting and word order (result (iii) above) suggests that in topic continuations clefting yields no processing advantage but in topic shifts, where processing is more difficult, clefting facilitates comprehension. Topic-continuation sentences may be relatively easy to process compared to topic shifts (Grosz et al. 1995); this could have the consequence that any processing advantage due to clefting may be masked the low overall processing difficulty.

The finding that the antecedent is preferentially resolved to the object in canonical word order (result (iv)) suggests a recency preference. This finding runs counter to the claims of Prasad and Strube (2000),

who found an overwhelming preference for resolution to subjects. However, their claims were based on production data (corpus data), and it is quite possible that the constraints on antecedent resolution differ in real-time comprehension.

Finally, since in sentence 3 clefting and antecedent-type did not interact, and since no effect of clefting was found, it appears that the facilitatory effect of clefting may not extend beyond the sentence in which the cleft appears. The facilitatory effect due to clefting appears to be short-lived, as predicted by decay and/or interference accounts of parsing. Note, however, that there was a numerical tendency towards faster reading time at the disambiguating region when the subject or object was clefted (see Table 10). This is suggestive of a slight facilitation in resolving to the subject or object antecedent when the antecedent is clefted. A cautionary note is in order here: since the findings in sentence 3 are null results, we cannot conclude much from them.

4   General Discussion

The two experiments presented here provide several new results regarding the effects of clefting on sentence comprehension. Encoding a clefted noun consumes processing resources, but such deeper encoding renders subsequent processing of the clefted noun easier, e.g., when it is retrieved later on in the sentence. The increased encoding cost due to clefting, as expressed by the higher regression probability in clefted versus non-clefted nouns in Experiment 1, appears to be restricted to cases where the NP is given, not when it is new.[8] We believe that this difference between given and new NPs with respect to clefting may be due to the particulars of the experiment design — when the NP is given, extra processing load is presumably incurred because the relational information described in the preceding sentence must be recalled, and it is in this high-load condition that we observe the greater processing cost associated with clefting an NP. Evidence for this conclusion comes from other work by Birch and Rayner (2010), who have shown similar increases in processing difficulty at the clefted noun (increased regression probability) even when the clefted noun was not explicitly mentioned in the preceding context (their Experiment 3B; although in their experiment, the clefted noun was contextually implied, e.g., referring to a waitress in a context where a restaurant is being discussed).

In addition to encoding cost, clefting also delivers a processing advantage. This is clear from the faster re-reading times at the clefted versus non-clefted noun[9] and the faster re-reading time at the integration region, where the clefted noun must be retrieved to construct relational information (i.e., what is to the left/right of what). Experiment 2 replicated our main findings for the clefting manipulation and demonstrated that even in simpler reading tasks the processing advantage due to clefting can be observed.

Experiment 1 also provides evidence consistent with the *relatum=given* principle that Hörnig and colleagues propose for German: the first NP is read faster in non-canonical order when it is given than when it is new. In addition, at NP1 we also find evidence for a new principle that we call *relatum-first*: NP1 is read faster in non-canonical order versus canonical order (however, as mentioned earlier, this result may be just epiphenomenal). These results suggest that the *relatum=given* principle may indeed be cross-linguistically applicable, and that the default canonical word order preference can be counteracted and even reversed given appropriate context.

Interestingly, we did not find any evidence for the given-new principle in any region of interest (the first NP and the integration region). This may be the case because the effect of given-new ordering emerges only after the entire sentence has been processed. In fact, all the studies mentioned in the introduction that found evidence for given-new ordering looked at total sentence reading times. A clear prediction is that we should also find evidence for a given-new ordering preference if we look at whole sentence reading times in Experiment 1. Indeed, when we considered total reading times on the whole second sentence in Experiment 1, we found a main effect of givenness, such that sentences with given-new ordering were read faster than new-given (t=2.14, p<0.05). As a reviewer notes, however, it may well be that the effect of given-new ordering occurs as the sentence is being read but is spread out over several regions, with the result that a region-by-

---

[8]Note that we found no effect of encoding cost in Experiment 2, but this is a null result and as a consequence we cannot conclude anything from it, especially given the relatively small sample size of the study.

[9]A reviewer insightfully notes that we cannot be sure that the re-reading of the noun is a consequence of having started integration; this is a reasonable objection but cannot be investigated with our current data.

region analysis cannot detect it.

Experiment 2 served to replicate our main findings for the effect of clefting in sentence comprehension.

To conclude, the two experiments presented here provide new evidence regarding the role of clefting, word order and given-new order in online sentence comprehension. The present work contributes toward our understanding of how exactly information structure serves to facilitate comprehension in language. Regarding given-new ordering, the present work suggests that the given-new preference is a rather late effect in processing that comes into play after the whole sentence has been processed; variation in given-new ordering does not seem to make a difference during incremental processing. Regarding word order, the experiments show that despite the fact that non-canonical word order is more complex syntactically, it can be easier to comprehend than canonical order if the non-canonical order makes it easier to connect sentences in context. Finally, regarding clefting, although clefting initially involves a greater processing cost, this has the important consequence that the clefted noun is easier to retrieve from memory at a later stage due to its heightened activation.

Acknowledgments

5   Appendix: A note on re-reading time

Most eyetracking studies present re-reading times (e.g., Sturt (2003), Birch and Rayner (2010)) including zero re-reading time. However, when (as in the present experiment) re-reading occurs in a low proportion of the trials, it is worth considering what it means to include zero re-reading time. One problem that arises in such cases is that if the dependent measure contains a large proportion of zeros, the statistical model will have extremely non-normal residuals in the model fit, rendering model assumptions invalid. We illustrate this problem in Figure 2.

A second reason to take re-reading time excluding zeros seriously is that it may reflect later integration processes. Evidence consistent with this suggestion comes from other work (Vasishth et al. 2010) where we have systematically compared re-reading time excluding zeros with self-paced reading data and found that this measure yields similar results to the self-paced reading data.

In cases (such as Experiment 2) where a low proportion of trials contains re-reading events, one issue is the loss of statistical power: with fewer data points it become less likely that a true effect would be discovered. Here, re-reading probability could be a useful measure. Although not used standardly, this could be a very informative measure. We therefore include this measure in our analyses where we look at re-reading time.
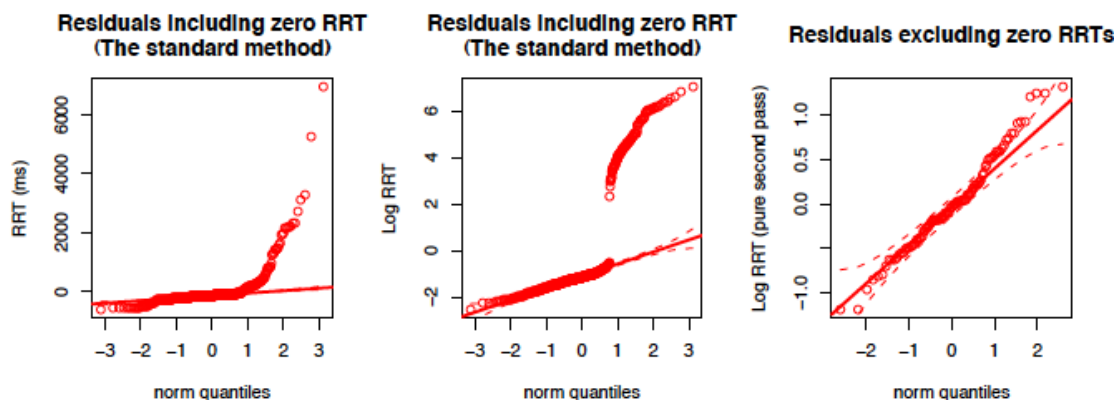
FIGURE 2 Residuals in the linear mixed model when the dependent measure is the standardly used re-reading time with zero re-reading times (raw re-reading time or log-transformed), versus pure second-pass time (log re-reading time excluding zeros).

References

Baayen, R. Harald. 2008. *Practical data analysis for the language sciences*. Cambridge University Press.

Baguley, Thom and Stephen J. Payne. 1998. Given-new versus new-given? An analysis of comprehension times for spatial descriptions. In S. O'Nualláin, ed., *Spatial cognition: Foundations and applications*, pages 317–328. Amsterdam: Benjamins.

Birch, Stacy and Keith Rayner. 2010. Effects of syntactic prominence on eye movements during reading. *Memory & Cognition* 38(6):740.

Birch, Stacy L. and Susan M. Garnsey. 1995. The effect of focus on memory for words in sentences. *Journal of Memory and Language* 34(2):232–267.

Birch, Stacy L. and Keith Rayner. 1997. Linguistic focus affects eye movements during reading. *Memory and Cognition* 25(5):653–60.

Bredart, Serge and Frédéric Docquier. 1989. The Moses illusion: A follow-up on the focalization effect. *Cahiers de Psychologie Cognitive, European Bulletin of Cognitive Psychology* 9:357–362.

Bredart, Serge and Karin Modolo. 1988. Moses strikes again: Focalization effects on a semantic illusion. *Acta Psychologica* 67:135–144.

Butt, Miriam and Tracy H. King. 1997. Structural topic and focus without movement. In *Proceedings of the LFG97 Conference. Stanford: CSLI Publications*.

Clark, Herbert H. 1973. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior* 12(4):335–59.

Clark, Herbert H. and Susan E. Haviland. 1977. Comprehension and the given-new contract. In *Discourse processes: Advances in research and theory*, vol. I. Norwood, NJ.

Clifton, Charles, Adrian Staub, and Keith Rayner. 2007. Eye Movements in Reading Words and Sentences. In R. V. Gompel, M. Fisher, W. Murray, and R. L. Hill, eds., *Eye movements: A window on mind and brain*, chap. 15. Elsevier.

Cutler, E. Anne and Jerry A. Fodor. 1979. Semantic focus and sentence comprehension. *Cognition* 7:49.

Delin, Judy and Jon Oberlander. 1995. Syntactic contraints on discourse structure: The case of it-clefts. *Linguistics* 33(4):465–500.

Drenhaus, Heiner, Malte Zimmermann, and Shravan Vasishth. 2011. Exhaustiveness effects in clefts are not truth-functional. *Journal of Neurolinguistics* 24:320–337.

Dwivedi, Veena. 1994. *Syntactic Dependencies and Relative Phrases in Hindi*. Ph.D. thesis, University of Massachusetts, Amherst, MA.

Ehrlich, Kate and Phillip N. Johnson-Laird. 1982. Spatial descriptions and referential continuity. *Journal of Verbal Learning and Verbal Behavior* 21:296–306.

Foraker, Stephani. and Brian McElree. 2007. The role of prominence in pronoun resolution: Active versus passive representations. *Journal of Memory and Language* 56(3):357–383.

Frazier, Lyn and Charles Clifton. 2004. Should given information appear before new? Yes and no. *Memory and Cognition* 32:886–895.

Frazier, L. and K. Rayner. 1982. Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology* 14:178–210.

Gambhir, Vijay. 1981. *Syntactic Restrictions and Discourse Functions of Word Order in Standard Hindi*. Ph.D. thesis, University of Pennsylvania, Philadelphia.

Gibson, Edward. 2000. Dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita, and W. O'Neil, eds., *Image, Language, brain: Papers from the First Mind Articulation Project Symposium*. Cambridge, MA: MIT Press.

Gordon, Peter C., Randall Hendrick, Marcus Johnson, and Yoonhyoung Lee. 2006. Similarity-based interference during language comprehension: Evidence from eye tracking during reading. *Journal of Experimental Psychology: Learning Memory and Cognition* 32(6):1304–1321.

Grosz, Barbara J., Aravind K. Joshi, and Scott Weinstein. 1995. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics* 21:203–225.

Hofmeister, Phillip. 2009. Encoding effects on memory retrieval in language comprehension. In *Proceedings of CUNY conference*. Davis, CA: University of Davis.

Hörnig, R., K. Oberauer, and A. Weidenfeld. 2005. Two principles of premise integration in spatial reasoning. *Memory and Cognition* 33(1):131–139.

Hörnig, R., T. Weskott, R. Kliegl, and G. Fanselow. 2006. Word order variation in spatial description with adverbs. *Memory & Cognition* 34(5):1183–1192.

Inhoff, Albrecht W. 1984. Two stages of word processing during eye fixations in the reading of prose. *Journal of verbal learning and verbal behavior* 23(5):612–624.

Johnson-Laird, P. N. 1983. *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge, MA: Harvard University Press.

Kidwai, Aeysha. 2000. *XP-Adjunction in Universal Grammar: Scrambling and Binding in Hindi-Urdu*. New York: OUP.

Lewis, Richard L. and Shravan Vasishth. 2005. An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science* 29:1–45.

Morris, Robin K. and Jocelyn R. Folk. 1998. Focus as a contextual priming mechanism in reading. *Memory and Cognition* 26(6):1313–22.

Prasad, Rashmi and Michael Strube. 2000. Discourse salience and pronoun resolution in Hindi. In A. Williams and E. Kaiser, eds., *Penn Working Papers in Linguistics, Volume 6.3*, pages 189–208. University of Pennsylvania.

Raaijmakers, Jeroen G.W., Joseph M.C. Schrijnemakers, and Frans Gremmen. 1999. How to deal with the "language-as-fixed-effect fallacy": Common misconceptions and alternative solutions. *Journal of Memory and Language* 41(3):416–426.

Rayner, Keith and Alexander Pollatsek. 1987. Eye movements in reading: A tutorial review. *Attention and performance XII: The psychology of reading* pages 327–362.

Singer, Murray. 1976. Thematic structure and the integration of linguistic information. *Journal of Verbal Learning and Verbal Behavior* 15(5):549–558.

Srivastav, Veneeta. 1991. The syntax and semantics of correlatives. *Natural Language & Linguistic Theory* 9(4):637–686.

Sturt, Patrick. 2003. The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language* 48:542–562.

Vasishth, Shravan, Sven Bruessow, Richard L. Lewis, and Heiner Drenhaus. 2008. Processing polarity: How the ungrammatical intrudes on the grammatical. *Cognitive Science* 32(4).

Vasishth, Shravan, Katja Suckow, Richard L. Lewis, and Sophie Kern. 2010. Short-term forgetting in sentence comprehension: Crosslinguistic evidence from head-final structures. *Language and Cognitive Processes* 25(4):533–567.

Ward, Peter and Patrick Sturt. 2007. Linguistic focus and memory: An eye movement study. *Memory and Cognition* 35(1):73–86.

Yekovich, Frank R, Carol H Walker, and Harold S Blackman. 1979. The role of presupposed and focal information in integrating sentences. *Journal of Verbal Learning and Verbal Behavior* 18(5):535–548.