JSAL

Volume 7 December 2015

Journal of South Asian Linguistics

CSLI PUBLICATIONS

Contents

Future Reference and Epistemic Modality in Hindi

Dave Kush, *Haskins Laboratories*

Received September 2014; Revised April 2015

Abstract

This paper presents an analysis of the Hindi morpheme *gaa* that is used in plain-future and epistemic ('presumptive', cf. Sharma 2008) statements. It is argued that *gaa* is a modal, but not a temporal, operator. It is also argued that apparent restrictions on the interpretation of *gaa* are due to independent restrictions on aspectual operators that *gaa* composes with.

1 Introduction

This paper is concerned with the interpretation of the Hindi morpheme *gaa*, which is often referred to as the marker of 'future tense'.[1] *Gaa* is commonly used in plain future assertions. For example, the sentences in (1)[2] can be felicitously uttered if the individuals in question are to arrive two days after the speech time.[3]

(1) a. amitaab^{h} do din=mẽ aa-e-gaa.
Amitabh two day=in come-SBJ.3.SG-MOD.M.SG
'Amitabh will come in two days.'

b. priti do din=mẽ aa-e-gii.
Priti two day=in come-SBJ.3.SG-MOD.F.SG
'Priti will come in two days.'

c. ve log do din=mẽ aa-ẽ-gee.
DEM.3.PL people two day=in come-SBJ.PL-MOD.M.PL
'They will come in two days.'

In addition to its use in plain future assertions, *gaa* is also used to make epistemic modal claims that lack future orientation.[4] (2a) is an epistemic modal claim with present temporal orientation. The indicative non-modal counterpart of (2a) is (2b). In (2a) the auxiliary bears subjunctive mood

[1] As evident in the glosses — and as discussed below — the final vowel of the morpheme changes under agreement with the number and gender of subject of the verb to which it attaches. Some authors use *-g-* as the exponent of this morpheme (e.g., Butt and Lahiri 2013), in order to reflect the variability of the final vowel. In this article I use *gaa*, the masculine singular form, to refer to the morpheme across all of its uses for the sake of convenience.

[2] Abbreviations used are as follows: 1 = first person, 2 = second person, 3 = third person, AUX = auxiliary, CORR = correlative pronoun, DEM = demonstrative, F = feminine, FAM = familiar, IMP = imperative, IMPF= imperfective, M = masculine, MOD = modal, NEG = negation, PFV = perfective, POSS = possessive pronoun, PL = plural, PROG = progressive, PRON = pronoun, PRS = present, PST = past, SG = singular, SBJ = subjunctive.

[3] I have chosen (somewhat tendentiously) to gloss *gaa* as a modal MOD, rather than as a future tense (e.g., FUT), which is more common.

[4] This use is often called the 'presumptive' (see, for example, Sharma 2008).

JSAL Volume 7, December 2015.

morphology and the morpheme *gaa*, whereas in (2b) the present indicative auxiliary *hai* 'be' is used. The felicity of the adverbial phrase *ab tak* 'by now' demonstrates the lack of future orientation.

(2) a. ve log ab=tak nahĩĩ aa-yee hõ-∅-gee.
DEM.3.PL people now=by NEG come-PFV.PL AUX-SBJ.PL-MOD.M.PL
≈'They must not have come by now.'[5]

b. ve log nahĩĩ aa-yee hãĩ.
DEM.3.PL people NEG come-PFV.PL AUX.PRS.3.PL
'They haven't come.'

The paper proceeds from the assumption that the *gaa* morphemes in (1) and (2a) are one and the same. Given that assumption, the primary focus of the paper is to offer a univocal analysis of *gaa* that specifies the semantic contribution of the morpheme in both instances. The study is partially informed by previous work on the semantics of English *will*, a morpheme that exhibits similar interpretive variability (i.e. it can be found in both plain future and epistemic utterances — see Jespersen 1924, Hornstein 1990, Enç 1996, Sarkar 1998).

(3) They *will* come in two days. PLAIN FUTURE

(4) They *will* not have come by now. ≈ They must not have come by now. EPISTEMIC

There are a number of possible analyses of *gaa*'s interpretive variability that would provide a univocal semantics of the morpheme, many of which have been explored for English *will*. The first possibility is that *gaa* is simply a future tense.[6] The second possibility is that *gaa* is both a future-shifting temporal and a modal operator. Prior work on English *will* has analyzed the morpheme as a portmanteau that supplies its complement with both a forward-shifted time argument and a quantified world argument (e.g., Abusch 1998, Sarkar 1998, Copley 2002, Condoravdi 2002). It is possible that a similar analysis would be appropriate for *gaa*. The third option is that *gaa* is solely a modal operator. It provides a mechanism to quantify over alternatives, but it has no effect on the temporal interpretation of its complement.

This paper advocates a variant of the third option. I analyze *gaa* as a Kratzerian modal that quantifies over a contextually-determined set of possible worlds. This account aligns itself with the tradition of treating apparent future markers as modals (e.g., Abusch 1998, Copley 2002, Condoravdi 2003, but differs from most previous accounts in one important regard. Despite *gaa*'s association with future orientation, I contend that it is not a tense. Instead, future-shift of *gaa*'s prejacent is contributed by a distinct temporal operator in the scope of the modal: the subjunctive. This analysis goes against a proposal due to Condoravdi (2002) that temporal semantics inhere in all modals, but is consistent with work by Matthewson (2011) that has argued for a separation of temporal and modal semantics in modal future constructions.

Finally, the paper catalogues and attempts to account for restrictions on the interpretation of *gaa*-marked sentences. Because modal flavor and temporal orientation are controlled by independent parameters, the account predicts a wide range of possible readings brought about by different combinations modal flavor, tense, and aspect. The paper explores the space of possible combinations to determine which readings are attested. It is shown that the account appears to over-generate some readings. In each case of over-generation, an attempt is made to explain the absence of the predicted reading on pragmatic or aspectual grounds.

The structure of the paper is as follows. Section 2 serves as a quick primer on relevant aspects of Hindi morpho-syntax and clause structure. Section 3 presents arguments for treating *gaa* as a modal

[5]Here and elsewhere my glosses commit to the presence of a subjunctive marker on the auxiliary. I represent this marker as ∅ because the marker is not realized as a separate vowel on the auxiliary as it is on main verbs. Despite the lack of the vowel, the presence of subjunctive agreement with the auxiliary can be seen in the nasalization of the preceding vowel with plural subjects.

[6]On this analysis future readings would be trivially derived, while epistemic readings would be arrived at pragmatically (see, e.g., Kissine's 2008 proposal for *will*).

operator. It also takes up the issue of *gaa*'s quantificational strength. Section 4 argues for divorcing forward-shifting temporal semantics from the denotation of *gaa*. Section 5 proposes a denotation for *gaa* and provides derivations of basic readings. Section 6 discusses restrictions on the interpretation of *gaa*-marked sentences and the origin of those restrictions. Section 7 discusses outstanding puzzles.

2 Morphological Preliminaries

Gaa is a morpheme that can attach to different verbal heads. The host verb can be a lexical verb that lacks overt aspectual marking as in (5).

(5) ve bacce do din=mẽ aa-ẽ-gee.
DEM.3.PL child.M.PL two day=in come-SBJ.PL-MOD.M.PL
'Those children will come in two days.'

The marker may also appear on an auxiliary verb *ho* when the main verb bears overt aspectual morphology. This is shown with the epistemic reading of *gaa* with present perfect, progressive, and imperfective aspect in (6a-c), respectively. The following sentences could be uttered in instances where I intend to make a statement about an event that I presume to have happened or one that I presume to be ongoing.

(6) a. ve log ab^{h}i=tak pahũc^{h}-ee hõ-∅-gee.
DEM.3.PL people now=by arrive-PFV.PL AUX-SBJ.PL-MOD.M.PL
'They must have arrived by now.'
b. ve log ab^{h}i naac rah-ee hõ-∅-gee.
DEM.3.PL people now dance PROG-M.PL AUX-SBJ.PL-MOD.M.PL
'They must be dancing now.'
c. ve log yahãã aksar aa-tee hõ-∅-gee.
DEM.3.PL people here often come-IMPF.M.PL AUX-SBJ.PL-MOD.M.PL
'They must come here often.'

Gaa can also attach to the auxiliary in copular constructions.

(7) ve bacce do saal=mẽ lambee hõ-∅-gee.
DEM.3.PL child.M.PL two year=in tall.M.PL be-SBJ.PL-MOD.M.PL
'Those kids will be tall in two years.'

When *gaa* attaches to a main verb (8a) or auxiliary (8b) in a standard assertion, it must be paired with subjunctive morphology. Subjunctive marking appears between the verb and *gaa* and agrees in person and number with the subject (see Butt and Rizvi 2010, Butt and Ramchand 2005).[7]

(8) a. ve log caaval banaa-*(ẽ)-gee.
DEM.3.PL people rice.M make-SBJ.M.PL-MOD.M.PL
'They will make rice.'
b. ve log caaval banaa rah-e hõ-∅-gee./*ho-gee
DEM.3.PL people rice.M make PROG-M.PL aux-SBJ.PL-MOD.M.PL/aux-MOD.M.PL
'They will/must be making rice.'

A compositional account of the semantics of *gaa*-marked constructions must determine *gaa*'s position vis-a-vis other morphemes in the clause in order to specify the order in which morphemes compose. A standard Hindi clause can contain, in the following order, a main verb, aspectual marker, auxiliary verb, and tense (agreement) morphology. These morphemes can be seen in the present progressive (9): the main verb *bol* 'speak' precedes the progressive aspectual morpheme *rah-aa*, which in turn precedes an auxiliary that agrees with the main subject in gender and number. The

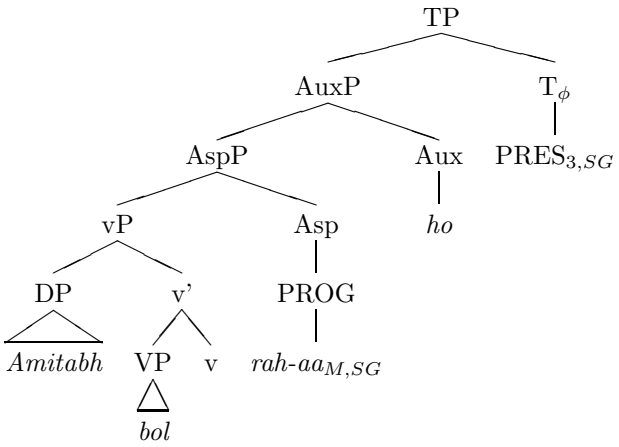
[7]Unlike the present and subjunctive markers, *gaa* does not inflect for person, only for number and gender.

auxiliary *hai* is a fusion of the auxiliary *ho* and the suffix *-ai*, which marks 3rd singular present indicative agreement.

(9) amitaab^{h} bol rah-aa hai.
Amitabh speak PROG-M.SG AUX.PRS.3.SG
'Amitabh is speaking.'

In accordance with the Mirror Principle (Baker 1985), the ordering of morpheme transparently reflects the commonly assumed underlying hierarchy of morphemes in the verbal domain on the assumption that phrases are head-final in Hindi. The *v*/VP is dominated by AspP (Ferreira 2005, Bhatt and Pancheva 2005), which is itself dominated by Tense. On the assumption that the auxiliary heads its own distinct phrase, it must sit between AspP and TP. Following Bhatt (2005), I assume that Tense (or T^0) is the locus of agreement in the Hindi clause: therefore, any heads or phrases that agree with the controller of agreement in the clause (in this case the subject), must originate in the scope/c-command domain of the local T^0. Taken together, these assumptions yield the structure (10) for (9).[8]

(10)



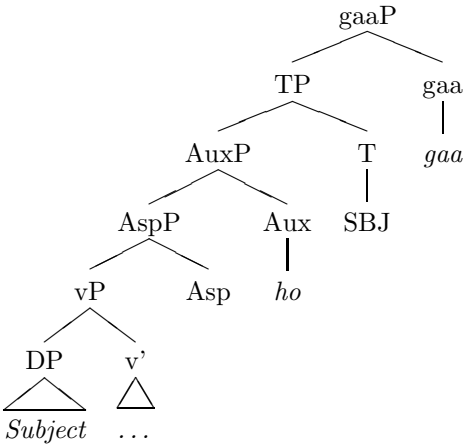
When specifying the structure of a *gaa*-marked clause, we must consider the position of two morphemes not present in (9): the subjunctive and *gaa* itself. The fact that the subjunctive agrees in person and number with the controller of agreement, on par with the present tense in (9), suggests that it should occupy Tense. The position of *gaa* within the clause is slightly more difficult to determine because different considerations point to distinct locations for the morpheme. On the one hand, *gaa* surfaces to the right of the subjunctive morpheme, which might be taken (all else equal) to indicate that the morpheme sits above TP. This interpretation, which hews to a very strict interpretation of the Mirror Principle equating surface linear order with syntactic scope, would provide the structure below for *gaa*-marked clauses.

Placing *gaa* above TP as in (11) would be odd for at least one reason: in (9) *gaa* agrees with the subject of the clause in gender and number, akin to an aspectual suffix. If *gaa* sat above T(P), it would not fall within the typical domain of agreement (again assuming that T is the head responsible for agreement in Hindi). If *gaa* is to agree, it should originate in the scope of T. However, *gaa* cannot be pronounced in this position, otherwise it would surface to the left of the subjunctive morpheme. I assume that if *gaa* originates below T, it head-moves so that it adjoins to the T head. This adjunction causes the morpheme to be spelled out to the right of T (either through direct right-adjunction, or via rules that determine the order of adjoined heads at a post-syntactic stage, see e.g., Embick

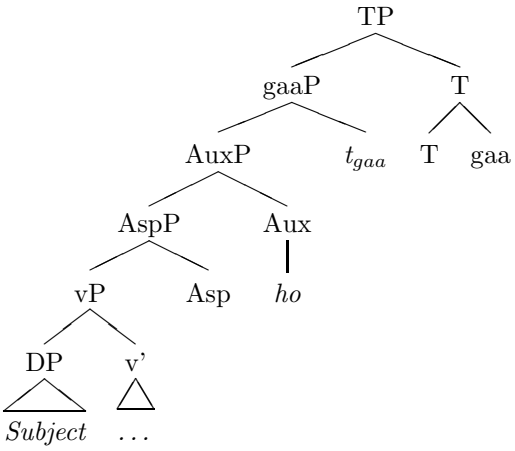
[8]In the diagrams below, the subject is shown in its interpreted (base) position (spec,*v*P) rather than its raised position in spec,TP. I do not intend this to suggest that the subject does not raise.

(2007).[9]

(11)



(12)



[9]One might ask what motivates the dislocation of *gaa*. Why should an affix ever surface to the right of T? The need to move *gaa* to this unusual position may be a diachronic remnant. A number of scholars hold that the present-day *gaa* can be traced back to a participle of the Sanskrit verb *ga* 'to go' (Kellogg 1893, 231, Beg 1988, 191, Butt and Lahiri 2013). Present-day *gaa*-marked constructions may descend from a periphrastic future construction, in which 'go' took a complement clause. Examples of such periphrastic future constructions are relatively common (see Fleischman 1982, Bybee et al. 1994), and can be observed, for example, in present-day Spanish (i) and English (ii).

(i) Yo voy a com-er.
I go.1.SG that eat-INF
'I'm going to eat./I will eat.'

(ii) I'm going to eat.

Over time, embedding verbs in periphrastic constructions may lose their independent status, becoming incorporated into what was once the embedded clause. Crucially, this process begins at the edge of the clause and moves inward: on their way to becoming true verbal affixes, erstwhile verbs first attach to the periphery of the clause, outside the domain of 'core' verbal morphology, and only later move inward Thus, if *gaa* began as a participle in a 'go' future that had a complement clause to its left, it would naturally cliticize to the right edge of the original clause (adjacent to T), as it lost its independence.

Before ending this digression, I note one interesting wrinkle in using a periphrastic 'go' future construction as the diachronic source for *gaa* futures, first remarked upon by Butt and Poudel (2012). In periphrastic futures crosslinguistically 'go' typically embeds non-finite complements. If subjunctive morphology is considered a kind of tense marking, the proposed source of the *gaa*-future would appear to violate a cross-linguistic norm.

For the purposes of this paper, I adopt the assumption that *gaa* does, in fact, originate below T, so that agreement is straightforwardly explained. I also assume that this head movement or dislocation does not have any interpretive consequences (Chomsky 2001); in all subsequent derivations, *gaa* is assumed to compose below T.[10]

3 Motivating Modality

Gaa is commonly glossed as a future tense, which suggests that the null hypothesis is that it is a simple tense operator, on par with the present or past tense. To argue against this simplistic view, I present two pieces of evidence that *gaa* exhibits properties of a modal operator. The first argument is morphological: *gaa* must obligatorily be used with subjunctive mood morphology, which arguably only appears in modal contexts (Portner 1998). The second argument in favor of analyzing *gaa* as a modal comes from the fact that appears to quantify over a domain of possibilities, as evidenced by its ability to license conditionals and participate in modal subordination. In this regard, *gaa* exhibits the behavior of other modal operators and contrasts simple tenses, which do not enable quantification over alternative states of affairs.

Gaa requires that its verb bear subjunctive marking. As in many Indo-European languages, the Hindi subjunctive appears in a limited variety of environments. The subjunctive primarily occurs in 'embedded' environments such as in: (i) the complement position of a bouletic verb (13); (ii) the antecedent of a conditional (14).[11]

(13) a. vo caah-taa hai ki mãĩ ja-ũũ.
PRON.3.SG want-IMPF.M.SG AUX.PRS.3.SG that PRON.1.SG go-SBJ.1.SG
'He wants me to go.' lit. 'He wants that I go.'

b. umiid hai ki vo ʈʰiik ho jaa-e
hope AUX.PRS.3.SG that PRON.3.SG okay be go-SBJ.3.SG
'(I) hope it gets better.' lit. 'Hope is that it gets better.'

(14) a. agar ve log duudʰ pi rah-e hẽ-∅, to un=ko
if DEM.3.PL people milk drink PROG.PL AUX-SBJ.3.PL then PRON.OBL=ACC
pareʃan mat kar-na.
bother NEG do-INF.M.SG
'If they're drinking milk, then don't bother them.'

b. agar us=ne duudʰ pi-yaa ho-∅ to
if PRON.OBL.3.SG=ERG milk drink-PFV.M.SG AUX-SBJ.3.SG then
us=ko pasifaiyer de-na.
PRON.OBL.3.SG=DAT pacifier give-INF.M.SG
'If he's drunk milk, give him the pacifier.'

The environments above can both be seen as containing a modal operator that scopes over the subjunctive (Portner 1998).

[10] See section 7 for some discussion that considers whether *gaa* is actually interpreted above T, as in (11).

[11] The morpheme can be added to the polite imperative (i) to produce (ii). An anonymous reviewer notes that the use of *gaa* in this construction is compatible with the generalization that *gaa* requires the subjunctive under the assumption that the ending *-iye* is an archaic subjunctive form.

i. abʰi caaval banaa-iye
now rice.M make-IMP.POLITE
'Please make rice now.'

ii (#abʰi) caaval banaa-iye-gaa
now rice.M make-IMP.POLITE-MOD.M.SG
'Please make rice (at some later point).' #'Please make rice now.'

(15) a. Complement clause of a bouletic
want/hope that X...≈
In all worlds consistent with speaker's desires/hopes that X ...
b. Antecedent of a conditional
If X ...≈
In all worlds in which X holds ...

In matrix environments, the distribution of the subjunctive is even more limited, but its presence is still conditioned by a modal. Matrix subjunctive is only licensed when there is an overt modal operator like *ʃaayad* 'maybe', or in deontic questions and in bouletic exhortations, two environments where a covert modal is present (Portner 1998).

(16) ʃaayad mãĩ kal kaam kar-ũũ.
maybe PRON.1.SG tomorrow work.M do-SBJ.1.SG
'Maybe I will work tomorrow.'[12]

(17) mãĩ kaam kar-ũũ?
PRON.1.SG work.M do-SBJ.1.SG
'Shall I work?'

(18) amitaabʰ zinda rah-e!
Amitabh alive stay-SBJ.1.SG
≈ 'Long live Amitabh!'

The subjunctive cannot typically be used in isolation to make plain declarative statements (19a). However, once *gaa* is added, a simple declarative is possible (19b).

(19) a. mãĩ kal kaam kar-ũũ.
PRON.1.SG tomorrow work.M do-SBJ.1.SG
#'I will work tomorrow.'
b. mãĩ kal kaam kar-ũũ-gaa.
PRON.1.SG tomorrow work.M do-SBJ.1.SG-MOD.M.SG
'I will work tomorrow.'

If the subjunctive requires a licensing modal then one must be present in (19b). The simplest assumption seems to be that *gaa* acts as the modal licensing the subjunctive.[13]

Gaa exhibits another characteristic behavior of a modal operator: its use seems dependent on a domain of quantification (e.g., a set of possible worlds). In this regard it contrasts with plain tenses which do not quantify over worlds. One way of showing that *gaa* has a domain of quantification is to observe how it interacts with conditional operators. According to one line of reasoning, conditional operators function to restrict a modal's domain of quantification (see Kratzer 1986, von Fintel 1994). Thus, if an item licenses the use of a conditional, it must be a modal.

The restrictive role that conditionals play is intuitively demonstrated with English *must*. In the sentence below, the modal claim is only evaluated with respect to the worlds (or situations) where there are mangoes on the table. The sentence can be interpreted in one of two ways. It can be taken as a statement of epistemic inference, or of obligation.

(20) If there are mangoes on the table, Amitabh must eat mangoes.
a. ≈ The presence of mangoes on the table indicates that Amitabh is a mango-eater.
b. ≈ The presence of mangoes on the table should cause Amitabh to eat mangoes.

[12]This and other examples gloss the frame adverbial *kal* as 'tomorrow' for the sake of simplicity. This is strictly speaking inaccurate because *kal* can mean either 'tomorrow' or 'yesterday' depending on context.
[13]I return to the syntax of subjunctive licensing in Section 7.

The sentence above shows that conditionals are licensed in the presence of *must*, which is analyzed as a modal. However, in order to motivate the use of conditionals as a diagnostic for modality, we must show that conditionals are not licensed with simple non-modal sentences. There is one small complication that bedevils our ability to do so: it appears that conditionals are sometimes licensed when no overt modal is present. For example, (21) is acceptable, but there is not a modal in sight.

(21) If Amitabh is quiet, he is angry.

(21) might seem to suggest that we cannot use the acceptability of a conditional to diagnose the presence of a modal. However, as Copley (2002) and Klecha (2014) have argued, this conclusion would be wrong. According to these authors, the interpretation of sentences such as (21) is limited in a way that attenuates the threat that they pose to the generalization that conditionals depend on modal licensors. The authors suggest that the conditional in (21) is, in fact, licensed by a modal, albeit a covert one. They note that in (21) and sentences like it, the antecedent cannot stand in a causation relation to the consequent: The meaning of (21) cannot be paraphrased as the causal (22a). Instead, an *indicational* inference must link the content of the two clauses. A suitable inferential paraphrase of is given in (22b).

(22) a. #Amitabh's quietude *causes* his being angry.
b. Amitabh's quietude implies/indicates that he is angry.

They argue that the inferential readings are tantamount to analyzing the original sentence as containing a covert epistemic modal in the consequent. This modal is responsible for licensing the conditional.

(23) If Amitabh is quiet, he *must* be angry.

Conditional sentences that contain an overt modal are not limited to inferential readings. A causal link can be established, for example, between the two clauses in (24). *Must* expresses modal obligation, which allows a causal paraphrase.

(24) If Amitabh is naughty, he must stand in the corner.
≈ Amitabh's naughtiness causes him to have to stand in the corner.

According to this diagnostic, then, the acceptability of a conditional can be used to infer the presence of an overt modal if the conditional can be given a causal reading. Inferential or epistemic readings of a conditional cannot be used to argue for the presence of an overt modal because these readings could, in principle, be achieved through a covert modal. Therefore, returning to examples like (20), the epistemic reading (20a) does not necessarily motivate the modality of *must*. It is the possibility of the causal reading (20b) and (24) that motivates analyzing *must* as a modal.

How does this influence our assessment of the modality of *gaa* in Hindi? First, we must determine if the distribution of indicational and causal readings is the same in Hindi as it is in English. The answer appears to be yes. In (25) there is is no overt modal; the consequent of the conditional contains a verb that bears perfective aspect and no other marking. There are no other operators in the clause.

(25) agar vo k^{h}ush t^{h}-aa, us=ne k^{h}aa-yaa.
if PRON.3.SG happy be-PST.M.SG, PRON.OBL.3.SG=ERG eat-PFV.M.SG
'If he was happy, he ate.'

The possible paraphrases of (25) track those of (21). (25) cannot mean the causal (26a), only the inferential (26b).

(26) a. #Amitabh's happiness *caused* his eating.
b. Amitabh's happiness implies/indicates that he ate.

It would seem that the ability to support causal readings of a conditional is an appropriate test for the presence of an overt modal in Hindi, just as in English. We can turn to using the test to probe *gaa*'s modal properties.

(27), the Hindi counterpart of (20), can receive an epistemic reading. On this reading, the conditional is unsurprisingly felicitous. This is consistent with *gaa* being a modal, but it could also indicate the presence of a covert modal. Unfortunately, (27) cannot receive a causal reading, so the acceptability of the conditional may not be very informative.

(27) agar aam mez=par rak^{h}-e hãĩ, amitaab^{h} aam k^{h}aa-taa
if mango.M table=on set-PPL.M.PL AUX.PST.3.PL, Amitabh mango.M eat-IMPF.M.SG
ho-∅-gaa.
AUX-SBJ.3.SG-MOD.M.SG

'If there are mangoes on the table, Amitabh must eat mangoes.'
≈ The presence of mangoes on the table indicates that Amitabh eats mangoes.
The presence of mangoes on the table should cause Amitabh to eat mangoes.

Although the causal readings cannot be found in (27), a causal reading of a conditional is possible with future-oriented *gaa*. In (28), where the verb in the consequent of the conditional is *gaa*-marked, Amitabh's future sadness is causally linked to his past quietude. We can conclude that this causal inference is supported by the presence of *gaa* because the same causal link cannot be established between an antecedent and a past-marked consequent (29).

(28) agar amitaab^{h} cup-caap rah-e, vo duk^{h}ii ho-∅-gaa.
if Amitabh quiet remain-SBJ.3.SG, PRON.3.SG sad be-SBJ.3.SG-MOD.M.SG

'If Amitabh remains quiet, he will be sad.'
≈ Amitabh's quietude causes future sadness.

(29) agar amitaab^{h} cup-caap rah-aa, vo duk^{h}ii t^{h}-aa.
if Amitabh quiet remain-PFV.M.SG, he sad be-PST.M.SG

'If Amitabh remained quiet, he was sad.'
≈ If Amitabh was quiet, he must have been sad.
≠ Amitabh's quietude cause his sadness.

Thus, we have our first piece of evidence in favor of analyzing *gaa* as a modal.

Another test for modality that relies on conditionals to restrict the domain of a modal can be used for future *gaa*. Some modals can undergo a kind of subordination that yields *Implicit Conditional* readings (Klecha 2009, Roberts 1989). (30a) shows that the modal *might* supports an implicit conditional reading; (30b) shows *gaa* permits similar subordination.

(30) a. Don't touch it! It might explode (if you do)!
b. us=ko mat cuu-na! vo ṭuuṭ ja-e-gaa (agar
that.OBL=ACC NEG touch-INF.M.SG! PRON.3.SG break go-SBJ.3.SG-MOD.M.SG (if
tum aisa kar-o-ge).
you that do-SBJ.2.PL-MOD.M.PL)
'Don't touch that. It'll break (if you do).'

Simple tenses cannot undergo similar subordination. Consider the availability of an implicit conditional reading in (31a) and the absence of the corresponding subordinated reading with the simple past tense (31b).[14] The same asymmetry is observed between *gaa*-marked clauses (31c) and perfective-marked clauses (31d).

[14]These sentences are adapted from examples in Klecha (2009).

(31) a. If she goes to New York, she'll go to the wedding. She might enjoy herself (if she does).
b. If she went to New York, she went to the wedding. #She enjoyed herself (if she did).
c. agar vo nu york jaa-e-gii, to vo ʃaadi=mẽ
if PRON.3.SG New York go-SBJ.3.SG-MOD.F.SG, then PRON.3.SG wedding=in
jaa-e-gii. us=ko mazaa aa-e-gaa.
go-SBJ.3.SG-MOD.F.SG. PRON.OBL.3.SG=DAT enjoyment come-SBJ.3.SG-MOD.M.SG
'If she goes to New York, she'll go to the wedding. She will enjoy it.'
d. agar vo nu york ga-yii, to vo ʃaadi=mẽ ga-yii.
if PRON.3.SG New York go-PFV.F.SG, then PRON.3.SG wedding=in go-PFV.3.F.SG
#us=ko mazaa aa-yaa.
PRON.OBL.3.SG=DAT enjoyment come-PFV-M.SG
'If she went to New York, she went to the wedding. #She enjoyed it.'

In sum, its obligatory co-occurrence with the subjunctive and the readings it licenses with conditionals suggest that *gaa* is a modal.

3.1 Domain of Quantification

Following Kratzer (1977, 1991) I formalize a modal's domain of quantification using a modal base (MB), which provides a set of possible worlds over which the modal quantifies. MBs differ with respect to the accessibility relation that restricts the subset of quantified possible worlds. Kratzer (1977, 1991) identified two MBs that are relevant to our purposes: (i) the epistemic MB and (ii) the circumstantial MB. An epistemic modal base is the set of worlds that are consistent with some body of evidence available to the speaker, or the speaker's beliefs, at the evaluation time of the modal. The circumstantial, on the other hand, is the set of worlds compatible with (objective) facts or circumstances in the evaluation world (w) at a given time (t).

(32) $MB_{Epistemic}$(w,t) ={w': w' is a world consistent with a body of evidence/beliefs in w at t}

(33) $MB_{Circumstantial}$(w,t) ={w': w' is a world consistent with a relevant set of facts/circumstances in w at t}

Modal operators can be lexically underspecified for their domain of quantification. They can rely on the context to set that parameter. This underspecification has been used to account for the fact that the same modal can contribute different readings depending on its context of use (Kratzer 1977). For example, *must* receives a different interpretation in (30a,b) depending on the worlds it quantifies over. The differences in the set of worlds chosen in the sentences can be brought out by the addition of the conditional clauses.

(34) a. John *must* exercise (if he wants to lose weight).
b. John *must* be exercising now (according to his schedule).

The first *must*, which expresses a kind of teleological claim, makes use of the circumstantial MB. In order to achieve the goal of losing weight, the circumstances of the world dictate that John must exercise. The second claim is epistemic. The utterer is making the claim that it is consistent with her evidence (John's schedule) that he is currently working out.

In the same way that *must*'s interpretive variability can be attributed to differences in MB, I propose that the two readings of *gaa* exemplified in (1) and (2) come about as a result of keying the modal to the circumstantial and epistemic MB. Plain future readings come about when *gaa* quantifies over worlds that are circumstantially accessible (as proposed for other future markers by Copley 2002, Abusch 2007, Matthewson 2006, a.o.). For example, sentence (1), reprinted below, can be roughly paraphrased as a claim about worlds in the circumstantial MB.

(1) ve log do din=mẽ aa-ẽ-gee.
DEM.3.PL people two day=in come-SBJ.3.SG-MOD.3.PL
'They will come in two days.'
Paraphrase:
The people arrive two days after the present in the worlds that are consistent with the circumstances in the actual world.

Epistemic readings arise when the modal takes the epistemic MB. Speakers make a claim about worlds that are consistent with their current evidence. (2a) would receive the paraphrase:

(2a) ve log ab=tak nahĩ aa-yee hõ-∅-gee.
DEM.3.PL people now=by NEG come-PFV.PL AUX-SBJ.PL-MOD.M.PL
'They must not have come by now.'
Paraphrase:
The people's arrival did not occur at any point prior to the present in the worlds that are consistent with what is known/the evidence at the time of speech.

One might wonder why it is necessary to distinguish between the two MBs at all. For example why do we need to invoke the circumstantial modal base for plain future readings? Why not use the epistemic modal base for all readings of *gaa*?[15] In some sense all claims about the future are predictions, based on the set of facts at the speaker's disposal at utterance time. Speakers do not, in actuality, have direct knowledge of future events in the way they can of past and present events. Nor do they have access to an objective set of facts about the state of the world separate from their own beliefs. Given the objective uncertainty about the future, cautious language users would only be within their rights to make epistemic predictions about future events.

The question of whether speakers are ultimately justified in making definitive claims about the future is more the province of epistomology than language use. Yet, it appears that speakers do, in practice, distinguish between epistemic modal claims and plain future claims. Some evidence comes from the distribution of possible readings. If all future claims were simply epistemic modal claims, we would expect the distribution of future claims to be subject to the same constraints that govern the distribution of epistemic modals. Epistemic readings of *gaa* are not available under the propositional attitude verb *know* (35). Suppose that it is common knowledge that Amitabh is a big meat eater. Reporting this fact using an epistemic modal would be infelicitous. This is presumably because epistemic *gaa* carries with it some kind of indirect evidential semantics (see von Fintel and Gillies 2010 for discussion of evidential semantics of epistemic modals). Use of epistemic *gaa* implies some inference. In cases where a proposition is general knowledge, no inference is required.

(35) #sab log jaan-tee hãĩ ki amitaabʰ mãs kʰaa-taa
all people know-IMPF.M.PL AUX.PRS.3.SG that Amitabh meat eat-IMPF.M.SG
ho-∅-gaa.
AUX-SBJ.M.SG-MOD.M.SG
#'Everyone knows that Amitabh must eat meat.'

If all future claims were epistemic claims, we would expect future reference to be impossible under *know*, contrary to fact. Suppose Amitabh states in a press conference that he plans to eat meat at an upcoming gala. It would be felicitous to state (36).

(36) sab log jaan-tee hãĩ ki amitaabʰ mãs kʰaa-e-gaa.
all people know-IMPF.M.PL AUX.PRS.3.SG that Amitabh meat eat-SBJ.M.SG-MOD.M.SG
'Everyone knows that Amitabh will eat meat.'

If we associate evidential semantics with the epistemic MB, it would appear that future *gaa* is not an epistemic modal. A separate MB is required — the circumstantial.

[15] Such a proposal was put forward by Crouch (1993) for English *will*.

In addition to a modal base, Kratzer (1981) also argued that the meanings of modals are assessed relative to a second conversational background that provides an *ordering source* (OS) by which worlds in the base are ranked. Ordering sources are used to account for the range of interpretations associated with circumstantial (root) modals. For example, the modals below are analyzed by Kratzer as circumstantial modals, but they all take different ordering sources. Deontic modals result from pairing the circumstantial MB with a deontic ordering source. The modal *can* in (37a) is an example of this. A dispositional ordering source accounts for the ability reading of (37b), whereas a teleological ordering is required for (37c).

(37) a. John can work out here. (I give him permission.)
b. John can lift 500 lbs (in view of his monstrous lats).
c. John should/must exercise more (to get sculpted triceps).

The ordering sources above do not exhaust the list of possible ranking functions. Among others, a *stereotypical* ordering source was also hypothesized to exist, which ranks worlds according to how 'normal' they are. The OS is typically thought to be paired with the epistemic MB to ensure that only the most plausible, normal epistemic possibilities are considered.

The stereotypical OS can also be used with future circumstantial modals (Copley 2002, Matthewson 2006, Werner 2006, a.o.), in order to ensure that possible futures are only those that conform with plausible continuations of current states of affairs.

Apart from the stereotypical, circumstantial *gaa*'s repertoire of ordering sources contains the dispositional OS (38) and bouletic OS (39). In (38) the worlds in the circumstantial MB are ordered according to whether Amitabh acts according to his disposition to eat cake. In (39), the worlds in the MB are ranked according to whether Amitabh achieves his desire of obtaining a cake to eat.[16]

(38) kyũũki amitaab^{h} kek bahut pasand kar-taa hai, vo koi-b^{h}i kek
because Amitabh cake much like do-IMPF.M.SG AUX.PRS.3.SG, PRON.3.SG any cake
pyaar=se k^{h}aa le-∅-gaa.
love=with eat take-SBJ.M.SG-MOD.M.SG
'Because Amitabh likes cake a lot, he'll gladly eat any cake you give him.'

(39) kyũũki amitaab^{h} kek k^{h}aa-naa cahaa-taa hai, vo kek
because Amitabh cake eat-INF.M.SG want-IMPF.M.SG AUX.PRS.3.SG, PRON.3.SG cake
k^{h}arid-e-gaa.
buy-SBJ.M.SG-MOD.M.SG
'Amitabh will buy cake because he wants to eat cake.'

Gaa-marked constructions do not exhibit other readings that circumstantial modals take. The morpheme cannot be used to express deontic obligation. For instance, in (40) *gaa* cannot be paired with imperfective morphology to express regular obligation. Similarly *gaa* cannot be used to talk of obligations in the future (41). Nor can *gaa* be used teleologically. A translation of (37c) using *gaa* is not possible (42).[17] These readings are presumably blocked by lexical restriction.

[16]Thanks to an anonymous reviewer for pointing this reading out.

[17]An anonymous reviewer notes that *gaa* may appear in some constructions that express deontic obligation. For example, in the sentence below, the *gaa*-marked auxiliary composes with an infinitival and an oblique subject to make a statement about Amitabh's deontic obligation to clean his room. This statement could be read as expressing a future obligation or a presumed obligation at present.

i. amitaab^{h}=ko apn-aa kamraa saaf karnaa ho-∅-gaa.
Amitabh=DAT self's-M.SG room clean do.INF.M.SG AUX-SBJ.M.SG-MOD.M.SG
'Amitabh will/must have to clean his room.'

In spite of *gaa*'s presence, the deontic modality expressed by these sentences is independent of *gaa*. Instead, it is likely contributed by the infinitival (or via some constructional combination of the infinitival and the dative-marked subject, as suggested by Butt and King 2004). The same configuration expresses deontic modality when *gaa* is absent (for example, with a present indicative, instead of a *gaa*-marked auxiliary).

(40) amitaab^{h} roz apnaa kamraa saaf kar-taa ho-∅-gaa.
Amitabh daily self's-M.SG room clean do-IMPF.M.SG AUX-SBJ.M.SG-MOD.M.SG
'#Amitabh has to clean his room daily.'

(41) amitaab^{h} apnaa kamraa saaf kar-e-gaa.
Amitabh self's-M.SG room clean do-SBJ.M.SG-MOD.M.SG
#Amitabh will have to clean his room.'

(42) majbuut ban-ne=ke liye, jan vyaayaam kar-e-gaa.
Strong become-INF.OBL=GEN.OBL for, John exercise do-SBJ.M.SG-MOD.M.SG
#'To get strong, John should/must work out.'
'To get strong, John will work out.'

To sum up this subsection, we have seen that *gaa* can take either an epistemic or circumstantial MB. The use of two MBs was motivated intuitively by the need to account for two kinds of readings (plain future and epistemic uses), as well as by distributional evidence: plain future readings are not subject to restrictions on epistemic *gaa*. It was also argued, based on *gaa*'s inability to license the full range of root modal readings, that *gaa*'s ordering source parameter is lexically restricted to the stereotypical, dispositional, or bouletic OS.

3.2 Modal Strength

Modals can vary in quantificational strength, or force. On the traditional Kratzerian analysis, the strength of a modal is lexically fixed, in contrast to the MB or OS parameter. A modal can either be a *possibility* modal which quantifies over its domain of possible worlds existentially, or a *necessity* modal, which is a universal quantifier (see Kratzer 1977, among many others). An example of the former is *might*, the latter is *must*.

(43) John might play ball.
There exists some world provided by the MB in which John plays ball.

(44) John must play ball.
In every world provided by the MB John plays ball.

Prior glosses have made use of the English 'must' to translate the contribution of epistemic *gaa*, suggesting that *gaa* has universal force. Supposing universal force for *gaa* would also be consistent with prior work on future markers like English *will*, which are frequently analyzed as universal quantifiers over possible worlds (cf. Condoravdi 2002, Copley 2002, Sarkar 1998, a.o.). Moreover, this analysis comports with the basic facts. The *gaa* morpheme does, in many respects, display behavior consistent with universal force.

One test of a modal's strength involves how it interacts with conjunction. A claim of possibility ($\Diamond\phi$) can be conjoined with the possibility of its negation ($\Diamond\neg\phi$) without contradiction arising. On the other hand, conjoining the universal $\Box\phi$ with $\Box\neg\phi$ does result in a contradiction. This explains the acceptability of (46a), which uses the existential modal *might*, and the unacceptability of (46b), which uses the universal modal *must*.

(45) a. ✓ $\Diamond\phi \wedge \Diamond\neg\phi$ (No Contradiction)
b. # $\Box\phi \wedge \Box\neg\phi$ (Contradiction)

(46) a. John might come and John might not come.
b. # John must come and John must not come.

ii. amitaab^{h}=ko apn-aa kamraa saaf kar-naa hai.
Amitabh=DAT self's-M.SG room clean do-INF.M.SG AUX.PRS.3.SG
'Amitabh has to clean his room.'

Gaa-marked epistemic claims behave like *must* under conjunction. A scenario illustrates: Suppose my friend and I are planning a party and want to know what we should serve our guests. My friend asks whether the people we've invited eat mangoes. If I want to express that I consider it a possibility that they might or might not eat mangoes, I cannot utter (47).[18]

(47) #ve log aam kʰaa-tee hõ-∅-gee aur ve log
DEM.3.PL people mango.M eat-IMPF.M.PL AUX-SBJ.3.PL-MOD.M.PL and DEM.3.PL people
aam nahĩĩ kʰaa-tee hõ-∅-gee.
mango.M NEG eat-IMPF.PL AUX-SBJ.PL-MOD.M.PL
#'They must eat mangoes and they must not eat mangoes.'

The same behavior is observed with future *gaa*.

(48) #amitaabʰ kal aa-e-gaa aur amitaabʰ kal nahĩĩ
Amitabh tomorrow come-SBJ.3.SG-MOD.M.SG and Amitabh tomorrow NEG
aa-e-gaa.
come-SBJ.3.SG-MOD.M.SG
'Amitabh will come tomorrow and won't come tomorrow.'
(not 'Amitabh might come tomorrow and might not come tomorrow.')

Gaa also behaves like an obligatorily strong modal in another regard. von Fintel and Gillies (2008, 2010) note that modal claims of the form $\Diamond\phi$ allow a speaker to 'stick to her conversational guns' (i.e. maintain the validity of the modal statement) even after it has been shown that $\neg\phi$. Stated differently, if a speaker utters $\Diamond\phi$ and subsequently finds out that $\neg\phi$, the speaker can claim that she was nevertheless not in error when making the modal claim. If the speaker's original claim was $\Box\phi$, however, then such a conversational move is not possible. We can return to an instance where two friends are having a discussion about mango-eating. Suppose Saif and Amit enter Amitabh's kitchen and find a bag of mangoes that has been neglected on the table. Saif and Amit could have the exchange in (49) and no one could be accused of speaking infelicitously. However, the exchange in (50), where the strong *must* is employed instead of *might*, seems incoherent.[19]

(49)
Saif: Amitabh might not eat mangoes.
Amit: You're wrong, I've seen him eat mangoes.
Saif: Look, I wasn't wrong. I didn't say that he doesn't eat mangoes, I said that he *might* not eat mangoes.

[18]Curiously, despite its compatibility with strong readings, it is not unheard of for Hindi-speaking informants to offer English translations of *gaa*-marked clauses using the existential *might* instead of *must*. Although some instances appear to be cases of simple translational error, it appears that for some speakers, *gaa* displays the behavior of a weaker quantifier. For example, an anonymous reviewer notes that despite its behavior under conjunction, *gaa* behaves more like a possibility modal under disjunction. Imagine Saif and Amit are standing outside Amitabh's house, knocking on the door. The reviewer maintains that the translation in (ia) more appropriately captures the meaning of the *gaa*-marked construction in (i) than does (ib).

i. vo darvaazaa nahĩĩ kʰol rah-aa hai. vo bagiiche=mẽ ya bathrum=mẽ
PRON.3.SG door NEG open PROG-M.SG AUX.PRS.3.SG PRON.3.SG garden.OBL=in or bathroom=in
ho-∅-gaa.
AUX-SBJ.M.SG-MOD.M.SG
a. He's not opening the door. He might be in the garden or he might be in the bathroom.
b. He's not opening the door. He must be in the garden or he must be in the bathroom.

At the moment, I do not have an account of this phenomena, although I make two notes. First, modals under disjunction quite often display behavior that is puzzling from the perspective of well-motivated semantic analysis. Second, the apparent 'weakness' of the claim with respect to an indicative non-modal counterpart may be attributable to the evidential semantics of epistemic constructions and not modal strength *per se*.

[19]Importantly, one or two informants judged the dialogue below to be marginally coherent. When probed further, the informants seemed to be making a meta-linguistic assessment about the difference in form between the modal statement and the present indicative statement.

(50)

Saif: Amitabh must not eat mangoes.
Amit: You're wrong, I've seen him eat mangoes.
Saif: #Look, I wasn't wrong. I didn't say that he doesn't eat mangoes, I said that he *must* not eat mangoes.

If the dialog above is translated into Hindi, the *gaa*-marked verb patterns along the lines of English *must*, rather than *might*.[20]

(51)

Saif: amitaab^{h} aam nahĩĩ k^{h}aa-taa ho-∅-gaa.
Amitabh mango.M NEG eat-IMPF.M.SG AUX-SBJ.M.SG-MOD.M.SG
'Amitabh must not eat mangoes.'
Amit: tum=ne galat kah-aa. mãĩ=ne us=ko aam k^{h}aa-tee
you=ERG wrong say-PFV.M.SG I=ERG PRON.3.SG.OBL=ACC mango.M eat-IMPF.M.PL
hu-e dek^{h}-aa hai.
become-PFV.M.PL see-PFV.M.SG AUX.PRS.3.SG
'You're wrong. I've seen him eat mangoes.'
Saif: #mãĩ=ne galat nahĩĩ kah-aa. mãĩ=ne sirf yeh kah-aa ki vo
I=ERG wrong NEG say-PFV.M.SG I=ERG only this say-PFV.M.SG that PRON.3.SG
aam nahĩĩ k^{h}aa-taa ho-∅-gaa.
mango.M NEG eat-IMPF.M.SG AUX-SBJ.M.SG-MOD.M.SG
≈ 'I wasn't wrong. I only said that he *must* not eat mangoes.'

Finally, *gaa* patterns like a strong modal in downward-entailing (DE) environments such as the scope of negation. A possibility modal in a DE environment yields a stronger claim than a necessity modal in the same environment. For example, although necessity modals are stronger than possibility modals in upward-entailing contexts, in DE contexts this relation is reversed. A negated possibility modal is stronger (i.e. *not possible*) than a negated necessity.

[20]The same behavior is observed in the absence of negation, as in i..

i. Saif: amitaab^{h} aam k^{h}aa-taa ho-∅-gaa.
Amitabh mango.M eat-IMPF.M.SG AUX-SBJ.M.SG-MOD.M.SG
'Amitabh must eat mangoes.'
Amit: tum=ne galat kah-aa. us=ko aam pasand nahĩĩ hãĩ.
you=ERG wrong say-PFV.M.SG PRON.3.SG.OBL=DAT mango.M fond NEG be.PRS.PL
'You're wrong. He isn't fond mangoes.'
Saif: #mãĩ=ne galat nahĩĩ kah-aa. mãĩ=ne sirf yeh kah-aa ki vo aam
I=ERG wrong NEG say-PFV.M.SG I=ERG only this say-PFV.M.SG that PRON.3.SG mango.M
k^{h}aa-taa ho-∅-gaa.
eat-IMPF.M.SG AUX-SBJ.M.SG-MOD.M.SG
≈ 'I wasn't wrong. I only said that he *must* not eat mangoes.'

The same facts also obtain in different constructions, such as the copula.

ii. Saif: amitaab^{h}=ko aam pasand hõ-∅-gee.
Amitabh=DAT mango.M fond b-SBJ.PL-MOD.M.PL
'Amitabh must be fond of mangoes.'
Amit: tum=ne galat kah-aa. us=ko aam pasand nahĩĩ hãĩ.
you=ERG wrong say-PFV.M.SG PRON.3.SG.OBL=DAT mango.M fond NEG be.PRS.3.PL
'You're wrong. He doesn't like mangoes.'
Saif: #mãĩ=ne galat nahĩĩ kah-aa. mãĩ=ne sirf yeh kah-aa ki us=ko aam
I=ERG wrong NEG say-PFV.M.SG I=ERG only this say-PFV.M.SG that PRON.3.SG.OBL=DAT mango.M
pasand hõ-∅-gee.
fond be-SBJ.PL-MOD.M.PL
≈ 'I wasn't wrong. I only said that he *must* be fond of mangoes.'

(52) $\neg\Diamond > _{strength} \neg\Box$

(53) a. It is not the case that John might come.
b. It is not the case that John must come, but he might.

(53a) asserts that there is no possibility of John's coming, while (53b) states that it is not necessary (though it remains possible). In the example below, epistemic *gaa* is embedded under a negated propositional attitude verb.[21] In this context, it appears that *gaa* contributes a reading of necessity, rather than possibility.

(54) [Context: Someone accuses me of having claimed that it was an inescapable fact that the Agarwals live in Washington, DC.]

a. mãĩ=ne nahĩĩ kah-aa ki ve log DC=mẽ rah-tee
I=ERG NEG said-PFV.M.SG that DEM.3.PL people DC=in live-IMPF.M.PL
hõ-∅-gee.
AUX-SBJ.PL-MOD.M.PL
'I didn't say that they must live in DC.'

b. mãĩ soc-taa hũũ ki ve ʃaayad VA ya DC=mẽ
I think-IMPF.M.SG AUX.PRS.1.SG that PRON.3.PL maybe VA or DC=in
rah-tee hãĩ.
live-IMPF.M.PL AUX.PRS.3.PL
'I think that they might live in VA or DC.'

That (54a) can be followed by (54b) shows that a negated epistemic *gaa* follows the pattern $\neg\Box\phi$, which admits $\Diamond\phi$. Thus, epistemic *gaa* once again behaves like a necessity modal.[22]

To summarize this section briefly: It was shown that traditional tests to diagnose modal strength seem to show that *gaa* should be analyzed as a universal modal. I defer providing a denotation for *gaa* that accommodates the notion of force until section 5.

4 Locus of Forward-Shift

The previous section argued that *gaa* cannot be a simple tense, but the arguments did not establish that *gaa* was solely a modal operator. It is conceivable that *gaa* could be both a modal operator and a forward-shifter. There is a tradition, dating back to at least Abusch (1998), of encoding forward-shifting semantics directly into the denotation of particular modals that are used in constructions that make future reference. According to Abusch, English *will* not only quantifies over possible worlds, but also supplies its prejacent with the right-unbounded interval (t, ∞), where t is supplied by a commanding tense. A variant of Abusch's denotation for *woll*, the modal underlying *will*, is below.[23] M represents the set of worlds quantified over, evaluated with respect to the actual world (represented with the indexical w_o).

(55) $[\![woll]\!] = \lambda P \lambda t. \forall w' \in M(w_o)$: p(w')(t,∞)

In English, the conflation of modal and tense operators into a single lexical item like *will* does not constitute a marked departure from the general treatment of modals in the language. But one

[21] Embedding *gaa* in this fashion is necessary because *gaa* outscopes clausemate negation (see, e.g., (2a)).

[22] If one were to provide an account of *gaa* that took into account the fact that some speakers occasionally allow weaker-than-universal readings, one might adopt previous analyses of the apparent context-sensitivity of modal strength in languages like St'atimcets (Rullman et al. 2008), Nez Perce (Deal 2011), or Gitksan (Matthewson 2013) as a guide. These researchers have assumed that modals whose strength appears to depend on the context come lexically specified for strength, which additional mechanisms modulate in context. According to these analyses, a modal with variable force can be analyzed either as a universal quantifier whose meaning is contextually weakened (Rullman et al. 2008), or an existential quantifier whose meaning is pragmatically strengthened in certain contexts (Deal 2011). *Gaa*'s behavior under negation, displayed in (54), would seem to indicate that the former treatment would be more appropriate for *gaa*.

[23] Condoravdi (2002) has argued for extending this treatment. According to her analysis, all modals uniformly shift the time of evaluation forward.

might expect a more hygienic picture in a language that has more clearly demarcated mood, tense and aspect categories, such as Hindi does. I argue that the semantic labor is more evenly divided across different functional morphemes in Hindi than in English. *Gaa* need not act as a forward-shifter because all forward-shift is performed by the subjunctive in *gaa*'s scope. In order to establish this, it must be shown that the subjunctive (i) behaves like a Tense and (ii) makes its own temporal contribution independent of *gaa*.

To begin, note that the subjunctive distributes like a tense: it can attach to auxiliaries in the absence of any other verbal morphology (56a), like the present and past (56b,c).

(56) a. agar ve log bimaar hõ-∅, vo bimaar ho-∅-gaa.
if DEM.3.PL people sick be-SBJ.PL PRON.3.SG sick be-SBJ.PL-MOD.M.SG
'. . . They are/might be/will be sick.'

b. ve log bimaar hãĩ.
DEM.3.PL people sick be.PRS.3.PL
'They are sick.'

c. ve log bimaar t^{h}-ee.
DEM.3.PL people sick be.PST-3.PL
'They were sick.'

Note also that the subjunctive inflects for the person and number of the local subject in clauses without ergative agreement, as does the present Tense. Person and number agreement is seen on the present indicative auxiliary in (57a), and on the subjunctive-marked verb root in (57b).[24]

(57) a. mãĩ bimaar hũũ.
I sick be.PRS.1.SG
'I am sick.'

b. vo caah-taa hai ki mãĩ jaa-ũũ.
he want-IMPF.M.SG AUX-PRS.3.SG that I go-SBJ.1.SG
'He wants me to go.'

Under the assumption that the subjunctive makes the same temporal semantic contribution across all its uses, its interpretation in contexts where *gaa* is absent can inform our analysis of the interpretation of *gaa*-marked constructions. A subjunctive marked verb can have a present-oriented interpretation (58). In both of the sentences below, the subjunctive-marked predicate of the embedded clause ('eats mangoes') occurs or has occurred at the time of utterance.

(58) a. yeh samb^{h}av hai ki ve aam k^{h}aa-tee
this possible be.PRS.3.SG that PRON.3.PL mango.M eat-IMPF.M.PL
hõ-∅-gee.
AUX-SBJ.PL-MOD.M.PL
'It is possible that they eat mangoes.'

b. yeh samb^{h}av hai ki unhõ=ne ab=tak tiin aam
this possible be.PRS.3.SG that PRON.3.PL.OBL=ERG now=by three mangoes
k^{h}aa-yee hõ-∅.
eat-PFV.3.PL AUX-SBJ.PL-MOD.M.PL
'It is possible that they have eaten three mangoes by now.'

A subjunctive-marked verb can also receive a future-shifted interpretation in the absence of *gaa*. In (59a,b) the embedded predicate is read as occurring at a later point. The eating will be ongoing or will have happened by the day after the utterance time.

[24]Here, my discussion of the subjunctive and the present as distinct tenses is at odds with certain descriptions in the literature (e.g., Butt and Rizvi 2010 which collapse present and subjunctive into a single tense given the overwhelming similarity in their agreement paradigms).

(59) a. yeh sambʰav hai ki jab tuu kal aa-ee, tab
this possible be.PRS.3.SG that when you.FAM tomorrow come-SUBJ.2.sg, then
ve log kʰaa rah-e hõ-∅.
DEM.3.PL people eat PROG-M.PL AUX-SBJ.PL
'It is possible that when you come tomorrow they will/may be eating.'

b. yeh sambʰav hai ki jab=tak tu kal aa-ee, tab=tak
this possible be.PRS.3.SG that when=by you.FAM tomorrow come-SUBJ.2.sg, then=by
unhõ=ne tiin aam kʰaa-yee hõ-∅.
PRON.3.PL.OBL=ERG three mango eat-PFV.3.PL AUX-SBJ.PL
'It is possible that by the time you come tomorrow they will/may have eaten 3 mangoes.'

The data above appear to suggest that future orientation can be achieved through use of the subjunctive in the absence of *gaa*. This could be taken to motivate analyzing the subjunctive as a kind of indefinite present that supplies a rightward open interval encompassing the present and any time thereafter.[25] This analysis would capture the temporal flexibility exhibited by the subjunctive in the preceding examples.

The subjunctive behaves somewhat differently from a simple indexical present. It appears that it can receive back-shifted interpretations under some conditions. Changing the matrix verb in (59) from present to past tense results in a past-reading of the subjunctive-marked complement clause. Similarly, in the when-clause in (61), the subjunctive-marked verb receives an interpretation that overlaps with the past tense of the matrix clause (seen on the auxiliary *tʰaa/tʰee*).[26] The complement clause of the past-tense *caah* 'want' in (62) receives a reading where the coming event follows the wanting but still occurs prior to the speech time.

(60) vo sambʰav tʰ-aa ki ve log us samay kʰaa rah-ee
it possible be.PST-M.SG that DEM.3.PL people DEM.3.SG time eat PROG-M.PL
hõ-∅.
AUX-SBJ.PL
'It was possible that they were eating at that time.'

(61) us zamaane=mẽ, agar koi garib darvaaze=pe aa-ee, tab ham
DEM.3.SG era.M.SG.OBL=in if some poor door.M.SG.OBL=on come-SBJ.3.SG, then we
us=ko kʰaanaa de-tee tʰ-ee.
PRON.3.SG.OBL=DAT food.M.SG give-IMPF.M.PL AUX.PST-3.PL
'In those days, if a poor person came to our door, then we would give him food.'

(62) mãĩ=ne caah-aa tʰ-aa ki vo picchʰle hafte
I=ERG want-PFV.PL AUX.PST-M.SG that PRON.3.SG last.M.SG.OBL week.M.SG.OBL
ja-ee.
go-SUBJ.3.SG
'I had wanted him to go last week.'

In all of the above examples, the interpretation of the subjunctive appears back-shifted by a past-tense operator that takes higher scope. This indicates that if we are to analyze the subjunctive as an interval whose right-hand is side open-ended, we must provide a way for the left-hand side of the interval to be bound by a commanding tense. To accommodate these facts I propose the following denotation for the subjunctive.

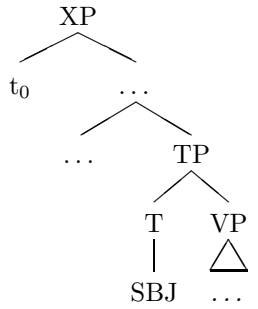
(63) $[\![SBJ_T]\!]^{g,c} = \lambda P \lambda t.\mathrm{P}(t, \infty)$

[25] See Butt and Rizvi (2010), Deo (2006), Montaut (2004) for discussion of the subjunctive in Hindi, as well as the diachronic development of the subjunctive from the present form in older Indo-Aryan languages.

[26] Thanks to a reviewer for bringing data point 61 to my attention and requesting further discussion of back-shifted readings of the subjunctive.

Under this analysis, the subjunctive passes the rightward open interval to its complement, licensing a forward-shifted interpretation. The forward-shift begins from t, whose value is supplied by the closest commanding tense. In embedded contexts t will be bound by a present or past tense in a higher clause. In matrix contexts no such tense is available to bind t. I therefore assume an indexical t_o sits atop matrix TP in order to bind any unbound time variables.

(64)



It is important to note that once we have analyzed the subjunctive as above, there is absolutely no need to suppose that *gaa* has forward-shifting semantics. The fact that the prejacent of *gaa* can be future-oriented can be attributed to the subjunctive in its complement.

5 Denotation of *gaa*

The foregoing sections have established that although it need not have any forward-shifting semantics, the denotation of *gaa* must have:

i) an underspecified MB parameter
ii) an underspecified OS parameter
iii) lexical specification for universal force

According to the Kratzerian tradition, the lexical entry of a modal must follow the abstract template:

(65) MODAL = λp. (Force $w' \in OS(MB(w_o))$:p(w')

The template above says that a modal (minimally) supplies its prejacent proposition (a property of worlds) with a world variable bound by some quantifier (represented as Force w' above). The set of worlds that the modal quantifies over is determined first by the parameter $MB(w_o)$, which stands for the contextually-supplied modal base, evaluated with respect to the real world (w_o). This initial set is passed to the ordering source OS which ranks and further restricts the worlds the modal quantifies over.

The abstract template can be fleshed out for *gaa* in the following manner. The MB parameter remains relatively unchanged, but for the fact that I propose to evaluate the MB with reference not just to a world, but also to a time (e.g., Condoravdi 2002, Hacquard 2010), which I assume to be hardwired as the speech time (the indexical t_0).[27] I make the function that applies the OS to the

[27] In the denotation the evaluation time is hardwired to t_0. If t_0 rigidly picks out the speech time back-shifted evaluation times of *gaa* embedded under a past tense propositional attitude verb, as in (27), could be considered problematic. This is observed with both epistemic *gaa*.

i. amitaab^h=ne soc-aa t^{h}aa ki saif aam pasand kar-taa ho-∅-gaa,
Amitabh=erg think-pfv.3.sg aux.pst-m.sg that Saif mango.m like do-impf.m.sg aux-sbj.3.sg-mod.m.sg,
par ab^{h}i vo jan-taa hai ki yeh galat t^{h}aa.
but now pron.3.sg know-impf.m.sg aux.prs.3.sg that this wrong be.pst-m.sg
'Amitabh thought that Saif must eat mangoes, but now he knows that's wrong.'

These results are not problematic if we adopt the assumption that the interpretation of t_0 in embedded contexts can be set to the internal now of the propositional attitude.

MB more precise by using the BEST_{OS} function from Portner (2009), which further picks out the subset of worlds that conform to the OS's ranking criteria.[28] The resultant denotation is (66).[29]

(66) $[\![gaa]\!]^c = \lambda \text{P} \lambda t. \forall \text{w'} \in \text{BEST}_{OS}(MB(\text{w}_0,\text{t}_0))$: P(w')(t)

With the denotation of *gaa* in hand, we can provide example derivations. Although the denotation provides a number of parameters that are open to contextual variation (i.e. the OS, the MB), for the purposes of illustration, the example derivations below hold the OS fixed as the stereotypical OS (indicated with BEST_{Stereo}), in order to more clearly investigate the space of possible readings created by manipulating the MB and the temporal orientation of the modal's prejacent (provided by the subjunctive).

Consider the derivation of the pair (67a,b), which are structurally identical. Both feature a verb bearing progressive aspect and a *gaa*-marked auxiliary. The two differ in temporal orientation, brought about by use of the future- and present-oriented adverbials *ab* 'now' and *kal* 'tomorrow'.

(67) a. ve log kal kʰaa rah-ee hõ-∅-gee.
DEM.3.PL people tomorrow eat PROG-M.PL AUX-SBJ.PL-MOD.M.PL
'They will be eating tomorrow.'
b. ve log ab kʰaa rah-ee hõ-∅-gee.
DEM.3.PL people now eat PROG-M.PL AUX-SBJ.PL-MOD.M.PL
'They must be eating now.'

To provide a compositional derivation of the truth conditions of these sentences we require a denotation for the progressive operator in addition to the denotation of *gaa*. I assume the denotation below, which makes use of the relation $\circ$ (Condoravdi 2002), to specify that the run-time of the event ($\tau(e)$) overlaps with the time provided by t. This is intended to capture that the progressive is viewed as contemporaneous with, or ongoing at, the reference time (Comrie 1976).

(68) $[\![PROG]\!] = \lambda P \lambda t. \exists \text{e}[\text{P(e)} \ \& \ \tau(e) \circ t]$

Under this analysis (67a) is assigned the structure in (70).[30] Composition proceeds bottom-up as specified in (69).

(69)
a. $[\![vP]\!]^c$ = $\lambda e \lambda w.$ eat(e, they)(w)
b. $[\![AspP_1]\!]^c$ = $\lambda t \lambda w. \exists e$[eat(e, they)(w) & τ(e) $\circ$ t]
c. $[\![AspP_2]\!]^c$ = $\lambda t \lambda w. \exists e$[eat(e, they)(w) & τ(e) $\circ$ t & t $\bigcap$ TOMORROW][31]
d. $[\![gaaP]\!]^c$ = $\lambda t. \forall \text{w'} \in \text{BEST}_{Stereo}(MB(\text{w}_0,\text{t}_0)): \exists e$[eat(e, they)(w') & τ(e) $\circ$ t & t $\bigcap$ TOMORROW]
e. $[\![TP]\!]^c$ = $\lambda t. \forall \text{w'} \in \text{BEST}_{Stereo}(MB(\text{w}_0,\text{t}_0)): \exists e$[eat(e, they)(w') & τ(e) $\circ$ (t_0,∞) & (t_0,∞) $\bigcap$ TOMORROW]
f. $[\![(67a)]\!]^c$ = $\forall \text{w'} \in \text{BEST}_{Stereo}(MB(\text{w}_0,\text{t}_0)): \exists e$[eat(e, they)(w') & τ(e) $\circ$ TOMORROW][32]

[28] If one were to attempt to formalize *gaa*'s occasional ability to receive weaker-than-universal interpretations, one could incorporate Rullman and colleagues' approach to analyzing modals as kinds of specific plural indefinites. Under their analysis, a choice function, represented below as f would serve as a final restrictor of the set of worlds that the modal quantifies over. Rullman et al. (2008) abstract away from the OS in their analysis of variable modal strength. It might be possible to incorporate the role of the choice function into the definition of the BEST function, perhaps by allowing context to supply further eligibility criteria to the OS, resulting in a narrower set of worlds.

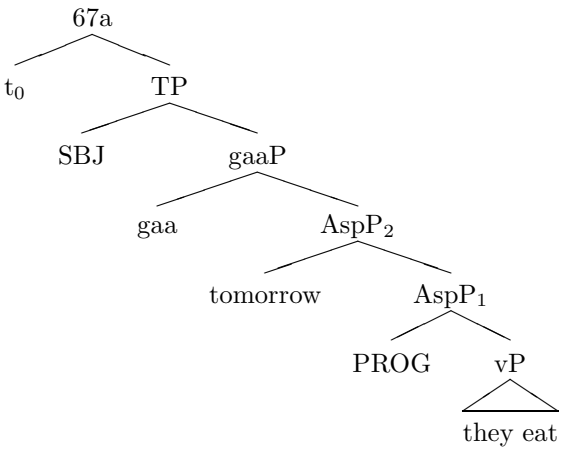
[29] Although it does not influence the temporal interpretation of its prejacent, *gaa* passes its prejacent the lambda-bound t as an argument in (66). This simply a book-keeping measure taken to permit the indexical t_0 to set the left-hand bound of the temporal interval provided by the subjunctive.

[30] I omit the auxiliary for the derivation because I assume it is semantically vacuous.

[31] I assume a simple intersective semantics for frame adverbials such as 'tomorrow' for simplicity: $[\![tomorrow]\!] = \lambda P \lambda t.[P(t) \ \& \ t \bigcap \text{TOMORROW}]$. See Condoravdi (2002) for a similar denotation.

[32] In this example τ(e) $\circ$ (t_0,∞) & (t_0,∞) $\bigcap$ TOMORROW has been reduced to the equivalent τ(e) $\circ$ TOMORROW.

(70)



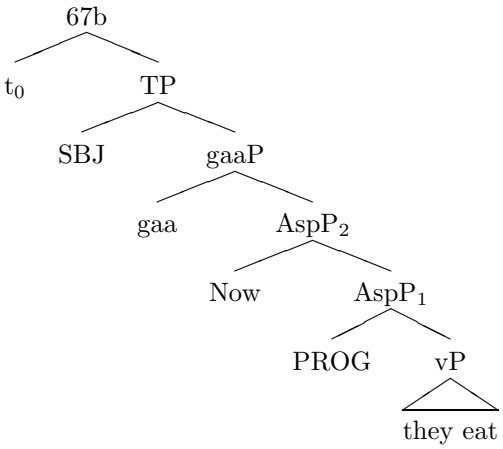
We arrive at the truth conditions in (69f), which state that the run-time of the event overlaps with the reference time. The reference time is the intersection of the temporal interval provided by the subjunctive (t_o, ∞) and the interval specified by the indexical 'tomorrow'. In the truth conditions the MB is still underdetermined. I assume that for future readings the circumstantial MB is the default MB provided by the context c.[33] Choice of the circumstantial MB produces the following truth conditions.

(71) $\forall$w' $\in$ $\textsc{Best}_{Stereo}(MB_{Circ}(w_0,t_0))$:$\exists e$[eat(e, they)(w') & τ(e) $\circ$ $\textsc{Tomorrow}$]

According to this denotation, in all of the worlds that are consistent with the circumstances in the actual world w_0 at the present (t_0), there is an event of eating that takes place tomorrow of which the people are the agents.

Turning to the derivation of (67b), in which the prejacent of the modal is present-oriented, composition proceeds as in (69) until the adverbial phrase. At this point, the temporal perspective of the prejacent is determined by the intersection of the present-oriented adverb 'now' and the interval (t_o, ∞) provided by he subjunctive.

(72)



[33]See the next section for discussion of other possible readings.

(73)
a. $[\![AspP_2]\!]^c$ = $\lambda t \lambda w.\exists e$[eat(e, they)(w) & τ(e) $\circ$ t & t $\bigcap$ NOW]
b. $[\![gaaP]\!]^c$ = $\lambda t.\forall$w' $\in$ BEST$_{Stereo}(MB(\text{w}_0,\text{t}_0))$:$\exists e$[eat(e, they)(w') & τ(e) $\circ$ t & t $\bigcap$ NOW]
c. $[\![TP]\!]^c$ = $\lambda t.\forall$w' $\in$ BEST$_{Stereo}(MB(\text{w}_0,\text{t}_0))$:$\exists e$[eat(e, they)(w') & τ(e) $\circ$ (t,∞) & (t,∞) $\bigcap$ NOW]
d. $[\![(67b)]\!]^c$ = $\forall$w' $\in$ BEST$_{Stereo}(MB(\text{w}_0,\text{t}_0))$:$\exists e$[eat(e, they)(w') & τ(e) $\circ$ (t_0,∞) & (t_0,∞) $\bigcap$ NOW]
= $\forall$w' $\in$ BEST$_{Stereo}(MB(\text{w}_0,\text{t}_0))$:$\exists e$[eat(e, they)(w') & τ(e) $\circ$ t_0][34]

Once again, the choice of MB is under-determined by the truth conditions and must be provided by the context. The easiest reading to illustrate is the familiar epistemic reading that arises when the epistemic MB is chosen. According to this reading, as stated in the truth conditions below, the people are eating at the present in all of the best epistemically accessible worlds.

(74) $\forall$w' $\in$ BEST$_{Stereo}(MB_{Epist}(\text{w}_0,\text{t}_0))$:$\exists e$[eat(e, they)(w') & τ(e) $\circ$ t_0]

The foregoing derivations illustrate how the account derives the default circumstantial future and present epistemic readings of a *gaa*-marked construction. These readings are not the only ones predicted by the account, however. The next section investigates other predicted readings.

6 Possible and Impossible Readings

6.1 Future Epistemics

Although it is commonly assumed that modals for the future take the circumstantial MB as a default (Abusch 2007, Copley 2002, Matthewson 2006), any account that allows the MB to vary independently of the temporal orientation of its prejacent predicts that other MBs may be selected with future orientation. In our case, there is nothing in the account that would bar selection of the epistemic MB for (5), resulting in the truth conditions in (75). The result would be an 'epistemic future' reading (see Condoravdi 2002, Matthewson 2013 for discussion of epistemic futures).

(5) ve log kal kʰaa rah-ee hõ-∅–gee.
DEM.3.PL people tomorrow eat PROG-PL AUX-SBJ.PL-MOD.M.PL
'They will be eating tomorrow.'

(75) $\forall$w' $\in$ BEST$_{Stereo}(MB_{Epist}(\text{w}_0,\text{t}_0))$:$\exists e$[eat(e, they)(w') & τ(e) $\circ$ TOMORROW]

According to these truth conditions, an event is ongoing at the reference time tomorrow in all epistemically accessible worlds. One might wonder how to distinguish epistemic futures from circumstantial futures. Though there may be a considerable amount of overlap between the worlds provided by the circumstantial MB, the set of worlds in the epistemic MB might be generated specifically from some body of evidence or (restricted) body of knowledge. (76) illustrates the ability to make future epistemic claims based on a body of evidence that may not be consistent with the larger set of circumstances.

(76) sabuut=ke hissab=se, ve log kal aa-ẽ-gee. par ham
evidence=of according=from DEM.3.PL people tomorrow come-SBJ.PL-MOD.M.PL but we
sab jaan-tee hãĩ ki ve nahĩĩ aa-ẽ-gee.
all know-IMPF.M.PL AUX.PRS.3.PL that PRON.3.PL NEG come-SBJ.PL-MOD.M.PL
'According to the evidence, they will come tomorrow. But we all know that they won't.'

If the statements above were both based on the same set of worlds the sentences above would seem as though they were in direct contradiction. However, the felicity of the discourse fragment above suggests that the first claim targets a set of epistemically accessible worlds based on the evidence, whereas the modal in the second sentence quantifies over worlds consistent with the objective facts

[34]The equivalent τ(e) $\circ$ t_0 has been introduced in place of the more complex τ(e) $\circ$ (t_0,∞) & (t_0,∞) $\bigcap$ NOW.

at present. Therefore, although both sentences are future-oriented, the first is an epistemic future and the second is a circumstantial future.

6.2 No Present Circumstantials

Because the analysis allows MB and temporal orientation to vary independently of one another, it permits both circumstantial and epistemic readings of *gaa* when the modal's prejacent is future-oriented. Both of these readings are attested. Circumstantial futures accounted for the default interpretation of future-oriented sentences, while epistemic futures could be used to reason about future events based on evidence or beliefs. Prima facie, we would expect the same type of interpretive flexibility with a present-oriented prejacent: *gaa*-marked claims should be interpretable as both present epistemics and present circumstantials. There is ample evidence of present epistemic readings. However, I show below that present circumstantial readings are not attested. Thus, it appears that the account, on its own, overgenerates. However, I propose that this overgeneration is taken care of by an independent pragmatic principle that governs the felicitous use of modal statements: Condoravdi's 2002 *Diversity Condition*.

In order to establish whether a present circumstantial reading is attested, we must be clear about its meaning. The truth conditions corresponding to a present circumstantial reading can be obtained by swapping the MB specification of a present epistemic. This is done below.

(77) $\forall$w' $\in$ $\textsc{Best}_{Stereo}(MB_{Circ}(\text{w}_0,\text{t}_0))$:$\exists e$[eat(e, they)(w') & τ(e) $\circ$ t_0]

The intended meaning can be drawn out if we consider the set of circumstantially accessible worlds that the modal would quantify over. The set would comprise all the worlds that are consistent with the facts/circumstances in the evaluation world at the time of evaluation (in this case, the present). All circumstantially accessible worlds must conform to the circumstances at the present moment, which entails that they are identical to the evaluation world at the present.[35] If ϕ holds at t_0 in the evaluation world, ϕ will hold at t_0 in all circumstantially accessible worlds. A universal present circumstantial modal statement $M(\phi)$ would therefore quantify over a set of worlds in which ϕ uniformly held at the present. Because the modal would be quantifying over counterparts of the evaluation world (w_0), the meaning of a present circumstantial modal claim would be, in principle, the same as its (non-modal) present indicative counterpart. Therefore, if a present circumstantial reading were possible, (67b) reprinted below, should also be able to be read as (78).

(67b) ve log ab kʰaa rah-ee hõ-∅-gee.
DEM.3.PL people now eat PROG-M.PL AUX-SBJ.PL-MOD.M.PL
'They must be eating now.'

(78) ve log ab kʰaa rah-ee hãĩ.
DEM.3.PL people now eat PROG-M.PL AUX.PRS.3.PL
'They are eating now.'

Contrary to the predictions of the relatively unconstrained theory, native speakers consistently reject such a reading of (67b). In light of this disparity between the predictions of the account and the range of attested meanings, an explanation must be provided. Why are present circumstantial readings unavailable? It is conceivable that one could cast the unavailability of the present circumstantial reading as the result of blocking or economy. The modal statement is 'more complex' than the non-modal statement, so under (semantic) equivalence, the latter should be preferred. Although this account has intuitive appeal, I do not pursue it here.

I follow a proposal due to Condoravdi (2002), which states that a general felicity condition on licit MB-time pairings blocks the present circumstantial reading. I believe this proposal is superior to a blocking explanation because it rules out use of the modal expression without direct reference

[35] Although they do not differ at the present, they may differ from one another in their future circumstances (see Condoravdi 2002, Matthewson 2013.

or comparison to other forms. The intuition behind Condoravdi's (2002) proposal is that modals can only be used felicitously to talk about live possibilities or alternatives. A proposition ϕ is a 'live' possibility if the conversational context or common ground does not conclusively determine whether ϕ or $\neg\phi$. Take epistemic *must* as an example. In the contexts of two detectives sorting through evidence at a murder scene, it would be felicitous to utter (79). However, it would be infelicitous to utter the same statement after Amitabh had been found guilty of the crime. This infelicity holds even though it is, in point of fact, still consistent with the evidence that Amitabh was the murderer. All that has changed between the two instances is that after the trial Amitabh's guilt is no longer a live issue.

(79) Amitabh must be the murderer.

Condoravdi (2002) formalizes the requirement that modals only discuss live alternatives with the *Diversity Condition* (paraphrased below)[36]. Under the analysis a context c, with common ground cg, can assign a modal base MB to a MODAL with temporal perspective t and applying to property P, only if cg and MB satisfy (80):

(80) DIVERSITY CONDITION
There is $w \in cg$ and any $w', w'' \in \text{MB}(w,t)$ such that:
$\text{P(t)}(w')$ and $\neg\ \text{P(t)}(w'')$

With this condition, we are in place to explain the absence of the present circumstantial reading of (8). According to Diversity, use of the circumstantial MB with (8) would only be felicitous if the issue of whether the people were eating was live. For the issue to be circumstantially live in the present moment, the following conditions would have to be met:

(i) there must be a circumstantially accessible world in which the people are eating at t_0, and
(ii) there must be a circumstantially accessible world in which the people are not eating at t_0.

Because all circumstantially accessible worlds are identical up to and including t_0, both conditions cannot be met simultaneously. Either all the worlds are those in which the people are eating or they are all worlds in which the people are not eating. Therefore, use of the modal is eschewed in favor of the simple present (78).[37]

6.3 Restrictions on Present Epistemic Readings

Diversity blocks all circumstantial present readings, but allows present epistemic readings. We therefore expect a present epistemic reading wherever present orientation is available. Such readings are attested across a variety of constructions. Present epistemic readings are available with the progressive, imperfective and perfect. A speaker could use each of the sentences below, to make a claim of epistemic certainty about an ongoing event of Amitabh dancing (81), Amitabh's habit of dancing (82), or Amitabh's having danced prior to the speech time (83).

(81) amitaab^{h} ab naac rah-aa ho-∅-gaa.
Amitabh now dance PROG-M.SG AUX-SBJ.3.SG-MOD.M.SG
'Amitabh must be dancing now.'

[36] See a very similar proposal in Werner (2006).

[37] Diversity may also be invoked to explain the lack of other present circumstantial readings cross-linguistically. For example, English *must* can be read circumstantially or epistemically. However, when the complement of the modal is unambiguously present-oriented, a circumstantial reading is not available. For example, a deontic reading is available for i., where the instantiation time of the modal is the future not available for ii.

i. When I arrive tomorrow, you must be eating!
ii. You must be eating now!
In all the best circumstantially accessible worlds ranked according to a deontic ordering source, you are eating now.

According to Diversity, present circumstantial readings are uniformly blocked. The upshot of this analysis is that all circumstantial modals are, in effect, future-oriented. A similar claim has been made recently by Matthewson (2011).

(82) amitaab^{h} bahut naac-taa ho-∅-gaa.
Amitabh much dance-IMPFV.M.SG AUX-SBJ.3.SG-MOD.M.SG
'Amitabh must dance a lot.'

(83) amitaab^{h} ab=tak naac-aa ho-∅-gaa.
Amitabh now=by dance-PFV.M.SG AUX-SBJ.3.SG-MOD.M.SG
'Amitabh must have danced by now.'

Present epistemic reference is also observed with the copula, as shown in the second sentence below.

(84) mãĩ=ne amitaab^{h}=ko daftar=mẽ nahĩĩ dek^{h}-aa. vo bimaar
I=ERG Amitabh=ACC office.M=in NEG see-PFV.M.SG PRON.3.SG sick
ho-∅-gaa.
be-SBJ.3.SG-MOD.M.SG
'I didn't see Amitabh at the office. He must be sick.'

In contrast to the examples above, when *gaa* attaches to a 'bare' main verb, a present epistemic reading is blocked. The sentence below must be interpreted as a claim about a future dancing event.

(85) amitaab^{h} naac-e-gaa.
Amitabh dance-SBJ.3.SG-MOD.M.SG
'Amitabh will dance.' #'Amitabh must dance/must be dancing.'

The principal difference between (85) and the acceptable present epistemic constructions is that (85) lacks the auxiliary *ho*. But it is not immediately obvious what it is about present epistemic readings that makes them incompatible with auxiliary-less constructions. There are three possible avenues of explanation: either (i) epistemic modal flavor requires the auxiliary, (ii) present orientation requires the auxiliary, or (iii) present orientation and epistemic flavor, when conjoined, require the auxiliary.

The first option can be easily dismissed. There is no general incompatibility between epistemic flavor and the absence of an auxiliary. As shown earlier with (76), a *gaa*-marked bare verb can receive epistemic interpretation when it has future orientation. Moreover although (86) is incompatible with a present epistemic reading, it can be read as a future epistemic claim (as evidenced by the adverbial which sets the MB).

(86) sabuut=ke hissab=se, Amitabh naac-e-gaa.
evidence=of according=from, Amitabh dance-SBJ.3.SG-MOD.M.SG
'According to the evidence, Amitabh will dance.'

Of the remaining options, (ii) should be preferred over (iii) on grounds of economy. An explanation that makes reference to temporal orientation alone is simpler than an explanation that posits an interaction of temporal orientation and modality. I contend that such a simple explanation is feasible. Below I provide evidence that present orientation is not possible in the absence of an auxiliary as a general rule. Moreover, I argue that this fact originates in the semantics of Aspect. In Hindi, the absence of an auxiliary entails that the verbal predicate is interpreted as a dynamic eventuality. Following previous authors, I assume that dynamic eventualities are not compatible with present orientation (Kamp and Reyle 1993, Partee 1984, Condoravdi 2002). When an auxiliary is present, on the other hand, the predicate is a (derived) state, which allows present reference.

A predicate's temporal orientation is determined by how aspect relates the run-time of the predicate-denoted event to the reference time set by Tense. If Tense provides an interval, aspect determines whether the run-time either (i) is included in the interval, (ii) overlaps with the interval, or (iii) does not overlap at all (e.g., if the event entirely precedes or follows the interval). Predicates that denote states (such as the progressive), permit temporal overlap (Kamp and Reyle 1993, Partee

1984) with the reference time set by Tense. The test below illustrates the possibility of overlap. When a matrix predicate is stative, it can be interpreted as simultaneous with an event in a modifying *when*-adverbial (Katz 1995).

(87) STATES

a. When I came home, Amitabh was eating mangoes. (mango-eating overlaps with arrival)
b. When I came home, Amitabh had eaten mangoes. (result of mango-eating overlaps with arrival)
c. When I came home, Amitabh was in the bathroom.

Dynamic predicates, on the other hand, do not permit simultaneous interpretation.

(88) DYNAMIC PREDICATES

a. When I came home, Amitabh ate mangoes. (mango-eating follows arrival)
b. When I come home, Amitabh will eat mangoes. (mango-eating follows arrival)

Auxiliary-marked predicates in Hindi behave like stative predicates in that they all allow simultaneous readings. This is not surprising when one considers that the constructions that require the auxiliary (i.e. the progressive, the imperfective, the perfect, and the copula), have been analyzed as denoting stative eventualities by a number of researchers (e.g., Dowty 1977, Ferreira 2005, Katz 1995, Husband 2012). The examples below illustrate the use of a past-marked auxiliary to create a past progressive, past perfect, and past copular construction, respectively.

(89) AUXILIARY-MARKED CONSTRUCTIONS

a. jab mãĩ gʰar=pe aa ga-yaa, tab amitaabʰ aam kʰaa rah-aa
when I home=on come go-PFV.M.SG, then Amitabh mango.M eat PROG-M.SG
tʰ-aa.
AUX.PST-M.SG
'When I came home, Amitabh was eating mangoes.' (mango-eating overlaps with arrival)

b. jab mãĩ ghar=pe aa ga-yaa, tab amitaabʰ=ne aam kʰaa-yaa
when I home=on come go-PFV.M.SG, then Amitabh=ERG mango.M eat-PFV
tʰ-aa.
AUX.PST-M.SG
'When I came home, Amitabh was eating mangoes.' (result of mango-eating overlaps with arrival)

c. jab mãĩ ghar=pe aa ga-yaa, tab amitaabʰ bathroom=mẽ **tʰ-aa**.
when I home=on come go-PFV.M.SG, then Amitabh bathroom=in be.PST-M.SG
'When I came home, Amitabh was in the bathroom.'

In Hindi, there are only two constructions that are formed without the auxiliary: the simple perfective (90) and the familiar bare 'future' construction (91). In the perfective, a suffix attaches directly to the verb. In the bare future construction, there is no overt marker of aspect. Both of these auxiliary-less constructions pattern together in that they disallow simultaneous readings with *when*-adverbials, as shown in (92).

(90) amitaabʰ=ne kaam ki-yaa.
Amitabh=ERG work.M do-PFV.M.SG
'Amitabh worked.'

(91) amitaabʰ kaam kar-e-gaa.
Amitabh work.M do-SBJ.3.SG-MOD.M.SG
'Amitabh will work.'

(92) Auxiliary-less Constructions

a. jab mãĩ ghar=pe aa ga-yaa, tab amitaab^{h}=ne aam k^{h}aa-yaa.
when I home=on come go-PFV.M.SG, then Amitabh=ERG mango.M eat-PFV.M.SG
'When I came home, Amitabh ate mangoes.' (mango-eating follows arrival)

b. jab mãĩ ghar=pe aa-ũũ-gaa, tab amitaab^{h} aam
when I home=on come-SBJ.1.SG-gaa, then Amitabh mango.M
k^{h}aa-e-gaa.
eat-SBJ.3.SG-MOD.M.SG
'When I come home, Amitabh will eat mangoes.' (mango-eating follows arrival)

It thus appears that the presence of the auxiliary marks that a predicate is stative (i.e., allows temporal overlap with the reference time), while the absence of an auxiliary marks that a predicate is dynamically interpreted (i.e., it does not allow temporal overlap). Although the auxiliary covaries with stative interpretations, in most cases, it does not appear that auxiliary itself controls the stativity of the predicate. In many cases, the stativity of the auxiliary's complement is determined by Asp^{0}. For example, the denotation for the progressive used in the previous section specified the relevant temporal overlap. By the same token, aspectual operators such as the perfective (PFV below) are often thought to encode non-overlap with the reference time directly (often through use of the operator $<$ which denotes non-overlapping temporal precedence — Singh 1998).

(68) $[\![PROG]\!] = \lambda P \lambda t \exists e.[\mathrm{P(e)} \ \& \ \tau(e) \circ t]$

(93) $[\![PFV]\!] = \lambda P \lambda t \exists e.[\mathrm{P(e)} \ \& \ \tau(e) < t]$

If aspect is responsible for determining whether predicates are stative or dynamic in the constructions above, it stands to reason that aspectual operators perform this function uniformly across all constructions in the language. This entails that aspect is responsible for the dynamic interpretation of bare main verbs. Accordingly, I propose that a covert *prospective* Aspect operator obligatorily occurs with a 'bare' lexical verb (Matthewson 2013). The operator forces an eventive, future-oriented reading of the verb by specifying that the run-time of the event is properly contained within the interval provided by Tense ($\subseteq$). Future orientation is forced because total inclusion within the interval entails that $\tau(e)$ does not overlap with either side of the interval.[38]

(94) $[\![PROSP]\!] = \lambda P \lambda t.\exists e[\mathrm{P(e)} \ \& \ \tau(e) \subseteq t]$

According to this analysis, the prospective operator turns its complement VP into a 'derived' dynamic predicate. Some evidence in favor of treating these constructions as 'derived' dynamic predications, as opposed to having aspectual properties determined by the lexical aspect of the main verb, comes from the fact that lexically stative verbs in this construction must also be interpreted as non-overlapping with the reference time. Lexically stative predicates such as 'stay', 'know French', and 'think' (95a-c) cannot receive a present (epistemic) reading when they are bare. They are obligatorily future-oriented (indicated in the case of 'know French' and 'think' by the necessity of an inceptive reading of the predicate).

(95) a. amitaab^{h} {tab | #ab | ab=se} vahãã rah-e-gaa.
Amitabh then now now=from there stay-SBJ.3.SG-MOD.M.SG
'Amitabh will stay there then/#now/starting now.'
'Amitabh must be staying there now.'

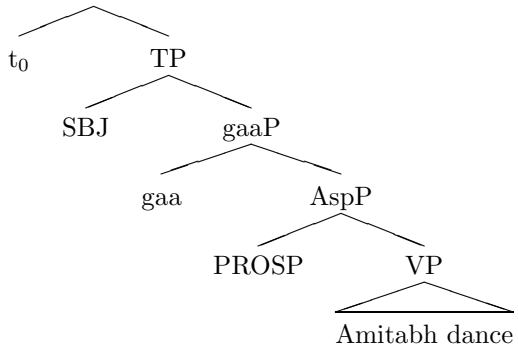
[38] See Condoravdi (2002) for use of the inclusion operator to achieve future-shifted readings. It is necessary to state the meaning of the prospective in terms of inclusion ($\subseteq$) rather than simple temporal subsequence ($>$) because we use the open-ended interval (t_0,∞) for the indefinite present tense. If subsequence were used instead of inclusion composing a prospective-marked VP with the open-ended interval (t_0,∞) would result in the following impossible statement $\tau(e) > (t_o, \infty)$, which states that the run-time of the event in question occured after the infinite interval.

b. amitaab^{h} frenc jaan-e-gaa.
Amitabh French know-SBJ.3.SG-MOD.M.SG
'Amitabh will (come to) know French.'

c. ve log soc-ẽ-gee ki amitaab^{h} aam k^{h}aa-taa
DEM.3.PL people think-SBJ.PL-MOD.M.PL that Amitabh mango.M eat-IMPF.M.SG
hai.
AUX.PRS.3.SG.
'They will (come to) think that Amitabh eats mangoes.'
'They must think that Amitabh eats mangoes'

Under this analysis, the structure assigned to the obligatorily future-oriented (85) is (96), and the truth conditions are (97).

(96)



(97) $\forall$w' $\in$ BEST$_{Stereo}(MB(\text{w}_0,\text{t}_0))$:$\exists e$[dance(e, Amitabh)(w') & τ(e) $\subseteq (t_o, \infty)$]

The truth conditions state that the run-time of the event of people dancing is properly contained within the open-ended interval (t_o, ∞) provided by the subjunctive. This forces a future orientation on the assumption that the present moment is not fully contained within this interval. Importantly, the truth conditions enable either an epistemic or circumstantial reading.

Positing a null prospective aspect engenders one minor complication. Copular constructions appear superficially similar to plain-future constructions like those in (95) in that they lack overt aspectual marking. In (98), there is no aspectual operator between the verb and the subjunctive.

(98) ve log k^{h}uʃ hõ-∅-gee.
DEM.3.PL people happy be-SBJ.PL-MOD.M.PL
'Those people will become/must be happy.'

Prima facie, we might suppose that this configuration would involve use of the prospective aspect, on analogy with the analysis of (95). This analysis would explain the future-oriented reading of (98): with the prospective aspect, the copular construction would be read as a kind of dynamic predicate (consistent with the reading 'Those people will become happy.'). Although this analysis explains the future-shifted reading of (98), it cannot handle the equally possible present epistemic reading of (98). For the present epistemic reading, the copular construction would need to be interpreted as a state, which the prospective aspect marker would block. The question that arises is thus: how could the copular construction receive a stative interpretation? The auxiliary itself cannot act as a stativizer, so the interpretation must arise elsewhere.

There are at least two ways to explain how the bare copula gets interpreted as a stative predicate. One may either assume that a separate (covert) aspectual operator occurs in copular constructions, a default operator that permits overlapping reference with the commanding tense. Such an operator might be required to perform other tasks such as 'stativizing' the copula's predicate (adjective,

noun, etc.) for composition (see Husband 2012. Alternatively, if the predicates selected by copulas are already states (Schwarzschild 2011) that provide a time variable (Tonhauser 2006), it might be possible to assume that bare copular constructions need not have any aspectual marking whatsoever.

To recapitulate the proposal: the absence of a present epistemic reading with superficially 'bare' verbs arises because bare verbs bear a (covert) prospective aspectual operator. This operator makes present temporal reference impossible.[39] Present epistemic readings are possible when *gaa*'s prejacent is a (derived) state because stative predicates permit present temporal reference. The possibility of a present epistemic reading seems to track use of the auxiliary *ho* because the auxiliary's presence correlates with the stativity of the predicate.

7 Remaining Puzzles

Though the analysis covers a wide range of possible readings, there are still a number of empirical and theoretical puzzles that deserve mention. I discuss a few below.

7.1 Matrix Back-shifted Readings

The data discussed in the foregoing text all deal with present and future-oriented readings. Yet it has also been observed that past-oriented epistemic readings are also possible with *gaa* (see Kush 2011 and Sharma 2008). This can occur with a variety of different aspectual operators. In the example below, a perfective-marked verb, in conjunction with a *gaa*-marked auxiliary is used to convey that the time of the comet fall occurred prior to speech time. In previous sections, this construction was referred to as the 'perfect'.

(99) jis d^{h}umketu=ke vajah=se sab dinosaurs mar-e, vo
CORR.OBL comet=of cause=from all dinosaurs die-PFV.M.PL, PRON.3.SG
mesozoic=ke dauran gir ga-yaa ho-∅-gaa.
mesozoic=of during fall go-PFV.3.SG AUX-SBJ.3.SG-MOD.M.SG
'The comet that killed the dinosaurs must have fallen during the mesozoic.'

Somewhat surprisingly, back-shifted readings are also possible with morphological configurations that provide less evidence for a 'perfect'. For example, (100) illustrates a back-shifted progressive construction. Imagine Amitabh awakens from his nap to find food laid out for him. He can say:

(100) jab mãĩ so rah-aa t^{h}aa, mer-e naukar k^{h}aanaa banaa
when I sleep PROG-M.SG AUX.PST-M.SG, 1.SG.POSS-PL servant.M food.M make
rah-ee hõ-∅–gee.
PROG-M.PL AUX-SBJ.PL-MOD.M.PL
'While I was sleeping, my servants must have been making food.'

Back-shifted readings are also possible with imperfective morphology. Suppose that it is believed that the only way to become a successful actor is to have eaten bananas during childhood. If we know that Amitabh is a successful actor, we can utter the following epistemic claim about his banana-eating activities as a child:

(101) bacpan=mẽ amitaab^{h} bahut kele k^{h}aa-taa ho-∅-gaa.
childhood=in Amitabh many bananas eat-IMPF.M.SG AUX-SBJ.3.SG-MOD.M.SG
'Amitabh must have eaten many bananas as a child.'

These sentences are form identical to their present counterparts, save for the temporal adverbials that indicate past orientation: *gaa* attaches to a subjunctive-bearing auxiliary, which scopes above an aspectual operator. The sentences pose a compositional challenge for our account because the

[39] Under this analysis, dispositional readings of bare verbs marked with *gaa*, such as (38), are obligatorily future-oriented. I follow Kissine (2008) in assuming that although statements of this sort express present dispositional modality, they quantify over possible behaviors in the future.

run-time of the VP-denoted event does not overlap within the indefinite present interval provided by the subjunctive. Moreover, these interpretations cannot be due to a back-shifting of the subjunctive. Back-shifted subjunctives were only possible when a past-tense scoped over an embedded subjunctive. The subjunctive is not embedded in either of the examples above.

I tentatively propose that a covert existential perfect operator is responsible for the back-shift in these instances, a move in line with Condoravdi's (2002) decompositional account of back-shifted epistemic modals in English. The exact analysis of the perfect in Hindi is beyond the scope of this paper, but it should be noted that there is independent evidence for a covert perfect operator in Hindi. The forms of non-perfect and perfect sentences are often indistinguishable. The plain progressive is compatible with the non-perfect frame-adverbial *abʰi* 'now', as well as the *se*-adverbial (comparable to English *since* adverbials). On the assumption that *se*-adverbials require a perfect context, a covert perfect operator is required to accommodate (102b).

(102) a. mãĩ abʰi kaam kar rah-aa hũũ.
I now work.M do PROG-M.SG AUX.PRS.1.SG
'I am working now.'
b. mãĩ dopahar=se kaam kar rah-aa hũũ.
I noon=from work.M do PROG-M.SG AUX.PRS.1.SG
'I have been working since noon.'

There is further suggestive, though by no means conclusive, evidence in favor of a covert perfect operator. In non-modal contexts, the perfect readings require the presence of the auxiliary *ho*. For example, when paired with an auxiliary, perfective morphology can yield a perfect interpretation, as evidenced by the felicity of the since-adverbial. However, without the auxiliary, the perfect reading is unavailable.

(103) amitaabʰ=ne dopahar=se kʰaa-yaa hai.
Amitabh=Erg noon=from eat-PFV.M.SG AUX.PRS.3.SG
'Amitabh has eaten since noon.'

(104) #amitaabʰ=ne dopahar=se kʰaa-yaa.
Amitabh=Erg noon=from eat-PFV.M.SG
#'Amitabh ate since noon.'

If, for whatever reason, the covert perfect requires the presence of an auxiliary and back-shifted *gaa*-marked constructions are only possible with the covert perfect, we would expect that *gaa*-marked constructions that lack the auxiliary should not be able to be back-shifted. This is what we find. Bare verbs marked with *gaa* cannot have back-shifted interpretations in matrix contexts.

(105) #amitaabʰ do din pehele kʰaa-e-gaa.
Amitabh two days before eat-SBJ.3.SG-MOD.M.SG
#'Amitabh was going to eat two days ago.'

7.2 Future Imperfectives

In previous sections, we saw that both present and future orientation were generally possible when there was overt aspect on a verb. For example, progressive and perfect constructions could receive a present (epistemic) reading or one of two kinds of future readings. Somewhat surprisingly, there is one construction where only present (epistemic) orientation is possible. It appears that when *gaa* attaches to an auxiliary and the main verb bears imperfective aspect all future orientation is blocked. (106a) shows that we can use an imperfective-marked verb to make a presumptive statement about a group of people's daily rice-making duties, if the rice-making is supposed to be currently ongoing. However, (106b) shows that the analogous future statement is unacceptable. Even if it is assured that daily rice-making is a part of a Japanese chef's job, we cannot use the imperfective-marked

verb and a *gaa*-marked auxiliary to describe the future event. This holds true even though there is nothing *conceptually* incoherent in talking about future imperfective events like habits, dispositions or states which the imperfective is usually used to express.

(106) a. aajkal, japan=mẽ chef ban kar, ve log roz caaval
nowadays, Japan=in chef become having, DEM.3.PL people daily rice.M
banaa-tee hõ-∅-gee
make-IMPF.M.PL AUX-SBJ.3.SG-MOD.M.SG
'Nowadays, after having become chefs in Japan, they must make rice on a daily basis.'

b. #do saal=mẽ, japan=mẽ chef ban kar, ve log roz caaval
two year=in, Japan=in chef become having, DEM.3.PL people daily rice.M
banaa-tee hõ-∅-gee
make-IMPF.M.PL AUX-SBJ.PL-MOD.M.PL
'In two years, after having become chefs in Japan, they will make rice on a daily basis.'

The restriction on future imperfectives is observed in *gaa*'s absence. Bare subjunctive imperfectives display the same resistance to future reference in the absence of *gaa*. When embedded under a statement of epistemic possibility, a subjunctive auxiliary without *gaa* can pair with an imperfective verb with present orientation. But, as above, any attempt to shift the temporal orientation of the imperfective verb forward results in unacceptability (107b).

(107) a. ho sak-taa hai ki ve roz caaval banaa-tee
be can-IMPF.M.SG AUX.PRS.3.SG that PRON.3.PL daily rice.M make-IMPF.M.PL
hõ-∅
AUX-SBJ.3.PL
'It's possible that they make rice on a daily basis (nowadays).'

b. *ho saktaa hai ki agle saal ve roz caaval banaa-tee
be can-IMPF AUX.PRS.3.SG that next year PRON.3.PL daily rice.M make-IMPF.M.PL
hõ-∅
AUX-SBJ.3.PL
#'It's possible that they will make rice on a daily basis next year.'

It therefore appears that imperfectives disallow future reference *as a general rule.*[40]

[40] An anonymous reviewer comments that there may be sentences in which an imperfective-marked verb is future oriented. The reviewer offers the example i. as evidence. In the antecedent of the conditional, the event of going to school daily must occur in the future.

i. (?*)agar agle saal=tak ye bacca roz skul jaa-taa ho, to
if next.OBL yearM.SG=by this child.M.SG daily school go-IMPF.M.SG AUX.SBJ.3.SG then
us=ko das rupya inaam mil-ẽ-gee.
PRON.3.SG.OBL=DAT ten rupees reward get-SBJ.PL-MOD.M.PL
'If the kid goes/has gone to school every day till next year, then he will get a 10 rupee reward.'

It is unclear what to make of this data point because I have had difficulty confirming the reviewer's intuition with other native speakers. The reviewer presents the data as acceptable, but most native speakers I have consulted reject the sentence outright (hence the equivocal '?*' in parentheses). Insofar as some speakers understand the intended meaning of the sentence, marginal acceptability appears highly dependent on the use of a *by*-adverbial 'by next year' above. Changing the by-adverbial to a simple adverbial that does not specify a right boundary in the way a by-adverbial does, results in complete unacceptability.

ii. *agar agle saal ye bacca roz skul jaa-taa ho, to us=ko das
if next.OBL year.M.SG this kid daily school go-IMPF.M.SG AUX.SBJ.3.SG then PRON.3.SG.OBL=DAT ten
rupya inaam mil-ẽ-gee.
rupees reward get-SBJ.PL-MOD.M.PL
#'If the kid goes to school every day next year, then he will get a 10 rupee reward.'

Use of a by-adverbial also results in a marginal improvement of (107b), as shown below.

While I do not offer a concrete proposal for this restriction, it suffices to show that this restriction is independent of *gaa*. We may therefore assume that the restriction stems from a restriction imposed by the semantics of the imperfective operator itself.[41]

7.3 Attachment and Interpretation Height

In section 2, I rejected the idea that *gaa* is base generated above Tense on the grounds of agreement. I reasoned that *gaa* must originate lower so that it fell within the scope of the controller of agreement in the clause (the local T head). *Gaa*'s surface position to the right of T was explained with appeal to head-movement or local-dislocation. This movement was assumed not to have any semantic consequences: I provided derivations of *gaa*-marked sentences that interpreted *gaa* in its base position.

There may be reason to suppose, however, that *gaa* should be interpreted above T — or that at least the head movement proposed has semantic consequences. In section 3, following Portner (1998), I argued that the subjunctive must be licensed by a modal and concluded that *gaa* served as the licensor. However, *gaa* should not be able to license the subjunctive in my structures under standard assumptions. It is typically assumed that licensing occurs under c-command/LF-scope, but *gaa* does not c-command the T head in its base or head-moved position. This problem would be solved if *gaa*'s base position were situated above TP, as in (11). However, this clause structure would require additional stipulation to explain how *gaa* agrees with the subject of the clause. Under such an account, *gaa* might act as its own agreement probe, separate from T.

If we do not wish to situate *gaa* above T, two analytical options that present themselves. The first is that head-movement and adjunction of *gaa* to the subjunctive is sufficient to license the morpheme — such a proposal was made for certain cases of NPI licensing in Japanese by Nakao and Obata (2007). The second is that *gaa* must (covertly) raise above T so that it can license the subjunctive. If such movement were indeed to occur, I assume that it would have to originate from *gaa*'s base position because movement of *gaa* out of the complex T head created via adjunction would violate the commonly presumed ban on excorporation (though see Roberts 1991 for arguments that excorporation may be possible). Movement of *gaa* above TP in one fell swoop would run afoul of the head movement constraint (HMC, Travis 1984; see also Matushansky 2006), which prohibits head movement to skip intervening heads as landing sites.

I leave exploration of these options to future research, but note that the denotation of *gaa* provided above will work equally well regardless of whether the morpheme sits above or below TP.

8 Conclusion

This paper has defended a univocal analysis of the Hindi morpheme *gaa* on its uses in plain future and epistemic modal constructions. It was argued that *gaa* is a necessity modal that quantifies over worlds in either a circumstantial or epistemic modal base. It was also argued that *gaa* does not shift the temporal orientation of its prejacent forward. The temporal orientation of *gaa*-marked clauses

iii. #?ho sak-taa hai ki agle saal=tak ve roz caaval banaa-tee
be can-IMPF.M.SG AUX.PRS.3.SG that next.OBL year.M.SG=by PRON.M.PL daily rice.M make-IMPF.M.PL
hõ-∅
AUX-SBJ.PL
#'It's possible that they will have been making rice on a daily basis by next year.'

The use of a by-adverbial indicates that these sentences are interpreted as a species of perfect construction. Thus, these data may indicate that future orientation with imperfective marking is only acceptable when the imperfective is embedded underneath another aspectual operator such as the covert perfect. I leave exploring this possibility to future research.

[41] It is possible that the imperfective in Hindi encodes some degree of anteriority, thus making it incompatible with future reference. Rajesh Bhatt (p.c.) suggests support for this idea might come from auxiliary drop patterns with the imperfective. Though the imperfective construction usually requires a tense-marked auxiliary to accompany the verb, this auxiliary can be dropped with past imperfective constructions. This is not the case with present imperfectives. If the imperfective encodes anteriority, deletion of a past auxiliary is recoverable, but deletion of a present auxiliary is not.

is determined by the subjunctive morpheme in the scope of the modal.

Because the account allows modal flavor and temporal orientation and vary as independent parameters, it predicts a wide number of readings. The paper considered the full range of possible interpretations of *gaa*-marked constructions that result from possible combinations of MB, temporal orientation, and aspectual marking. Certain combinations were unattested. The empirical landscape is summarized below.

	Circumstantial	Epistemic	Circumstantial	Epistemic
	Present		Future	
Impf	*Diversity	✓	*?	*?
Prog	*Diversity	✓	✓	✓
Perfect	*Diversity	✓	✓	✓
Copula	*Diversity	✓	✓	✓
Bare	*Diversity/*Aspect	*Aspect	✓	✓

For the most part, the absence of certain readings was argued to arise from independent pragmatic or semantic principles. For example, all present circumstantial readings were argued to be blocked by Condoravdi's (2002) *Diversity Condition*. The obligatory future orientation of superficially bare main verbs was attributed to a covert prospective aspect. Finally, it was observed that the gaa-marked imperfectives could not have future orientation. Although a complete account of this fact was not given, it was suggested that it might reflect idiosyncratic semantic restrictions imposed by the imperfective operator.

Acknowledgments

I am grateful, first and foremost, to Valentine Hacquard for extensive comments on many previous versions of this paper. Thanks are also due to Rajesh Bhatt, Miriam Butt, Ashwini Deo, Ayesha Kidwai, Angelika Kratzer, Alexander Williams, and audiences at Jawaharlal Nehru University and Formal Approaches to South Asian Languages 1 at UMASS for helpful comments. My primary informants for this paper were Ashok, Manu, Tanu, and Ambrish Kush, native speakers of Hindi from northwestern Uttar Pradesh. I also received judgments from Rajesh Bhatt and Anurag Rimzham. All of my informants deserve thanks. None of those mentioned in these acknowledgments bear any responsibility for errors in this paper.

References

Abusch, Dorit. 1998. Generalizing tense semantics for future contexts. In S. Rothstein, ed., *Events and Grammar*, pages 13–33. Dordrecht: Kluwer Academic Publishers.

Abusch, Dorit. 2007. Temporal and Circumstantial Dependence in Counterfactual Modals. In *Proceedings of the Sixteenth Amsterdam Colloquium*, pages 37–42. ILLC, University of Amsterdam.

Baker, Mark. 1985. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16:373–415.

Beg, Mirza K. A. 1988. *Urdu Grammar: History and Structure*. New Delhi: Bahri Publications.

Bhatt, Rajesh. 2005. Long distance agreement in Hindi-Urdu. *Natural Language & Linguistic Theory* 23(4):757–807.

Bhatt, Rajesh. and Roumyana Pancheva. 2005. Aspect. LSA Summer Institute, Class Notes.

Butt, Miriam and Tracy Halloway King. 2004. The status of case. In V. Dayal and A. Mahajan, eds., *Clause Structure in South Asian Languages*, pages 153–198. Dordrecht: Kluwer Academic Publishers.

Butt, Miriam and Aditi Lahiri. 2013. Diachronic pertinacity of light verbs. *Lingua* 135:7–29.

Butt, Miriam. and Tikaram Poudel. 2012. A comparative look at the future in Indo-Aryan. Talk given at *South Asian Languages: Theory, Typology, and Diachrony*, Yale University, New Haven.

Butt, Miriam and Gillian Ramchand. 2005. Complex Aspectual Structure in Hindi/Urdu. In N. Ertischik-Shir and T. Rapoport, eds., *The Syntax of Aspect*, pages 117–153. Oxford: Oxford University Press.

Butt, Miriam and Jafar Rizvi. 2010. Tense and aspect in Urdu. In P. Cabredo-Hofherr and B. Laca, eds., *Layers of Aspect*, pages 43–66. Stanford: CSLI Publications.

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

Chomsky, Noam. 2001. Derivation by phase. In M. Kenstowicz, ed., *Ken Hale: A life in language*, pages 1–52. Cambridge, MA: The MIT Press.

Comrie, Bernard. 1976. *Aspect: An introduction to the study of verbal aspect and related problems*, vol. 2. Cambridge University Press.

Condoravdi, Cleo. 2002. Temporal interpretation of modals: modals for the present and for the past. In D. Beaver, S. Kaufmann, B. Clark, and L. Casillas, eds., *The Construction of Meaning*, pages 59–88. Stanford: CSLI Publications.

Condoravdi, Cleo. 2003. Moods and modalities for *will* and *would*. In *The Proceedings of the Amsterdam Colloquium — Workshop on Mood and Modality*.

Copley, Bridget L. 2002. *The Semantics of the Future*. Ph.D. thesis, MIT.

Crouch, Richard. 1993. *The temporal properties of English conditionals and modals*. Ph.D. thesis, University of Cambridge.

Deal, Amy Rose. 2011. Modals without scales. *Language* 87:559–585.

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

Dowty, David. 1977. Towards a semantic analysis of verb aspect and the english 'imperfective progressive'. *Linguistics and Philosophy* 1:45–77.

Embick, David. 2007. Linearization and local dislocation: Derivational mechanics and interactions. *Linguistic Analysis* 33:2–35.

Enç, Mürvet. 1996. Tense and modality. In S. Lappin, ed., *The Handbook of Contemporary Semantic Theory*, pages 345–358. Oxford: Blackwell.

Ferreira, Marcelo Barra. 2005. *Event Quantification and Plurality*. Ph.D. thesis, MIT.

Fleischman, Suzanne. 1982. *The Future in Thought and Language: Diachronic evidence from Romance*. Cambridge: Cambridge University Press.

Hacquard, Valentine. 2010. On the event relativity of modal auxiliaries. *Natural Language Semantics* 32:279–315.

Hornstein, Norbert. 1990. *As Time Goes By*. Cambridge, MA: The MIT Press.

Husband, E. Matthew. 2012. *On the Compositional Nature of States*. Amsterdam: John Benjamins. Linguistcs Today 188.

Jespersen, Otto. 1924. *The Philosophy of Grammar*. London: George Allen.

Kamp, Hans and Uwe Reyle. 1993. *From Discourse to Logic; Introduction to the Model-theoretic Semantics of natural language*. Berlin: Springer.

Katz, Graham. 1995. *Stativity, Genericity, and Temporal Reference*. Ph.D. thesis, University of Rochester.

Kellogg, Samuel H. 1893. *Grammar of the Hindi Language*. Delhi: Munshiram Manoharlal Publishers Pvt. Ltd., reprinted 1990, 2nd edn.

Kissine, Mikhail. 2008. Why *will* is not a modal. *Natural Language Semantics* 16:129–155.

Klecha, Peter. 2009. The modality of English futures. In *Proceedings of Chronos*, vol. 9. Paris.

Klecha, Peter. 2014. Diagnosing modality in predictive expressions. *Journal of Semantics* 31(3):443–455.

Kratzer, Angelika. 1977. What 'must' and 'can' must and can mean. *Linguistics & Philosophy* 1:337–355.

Kratzer, Angelika. 1981. The notional category of modality. In H.-J. Eikmeyer and H. Rieser, eds., *Words, worlds, and contexts. New approaches in word semantics*. Berlin: De Gruyter.

Kratzer, Angelika. 1986. Conditionals. In *The Proceedings of the Chicago Linguistics Society*, vol. 22, pages 1–15. Chicago.

Kratzer, Angelika. 1991. Modality. In A. von Stechow and D. Wunderlich, eds., *Semantik: Ein internationales Handbuch zeitgenössischer Forschung*, pages 639–650. Berlin: De Gruyter.

Kush, Dave. 2011. Height-Relative Determination of (Non-Root) Modal Flavor: Evidence from Hindi. In *Proceedings of SALT*, vol. 21, pages 413–425.

Matthewson, Lisa. 2006. Temporal semantics in a supposedly tenseless language. *Linguistics & Philosophy* 29:673–713.

Matthewson, Lisa. 2011. On the (non)-future orientation of modals. Paper presented at *Sinn und Bedeutung 16*, Utrecht.

Matthewson, Lisa. 2013. Gitksan Modals 1. *International Journal of American Linguistics* 79(3):349–394.

Matushansky, Ora. 2006. Head movement in linguistic theory. *Linguistic Inquiry* 37(1):69–109.

Montaut, Annie. 2004. *Hindi Grammar*. München: Lincom-Europa.

Nakao, Chizuru and Miki Obata. 2007. Parametric Variations in NPI-Licensing and the Role of LF X0-Movement. In D.-W. Lee, ed., *Proceedings of the 9th Seoul International Conference on Generative Grammar*, vol. 9, pages 135–152.

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

Portner, Paul. 1998. The progressive in modal semantics. *Language* 74:760–787.

Portner, Paul. 2009. *Modality*. Oxford: Oxford University Press.

Roberts, Craige. 1989. Modal Subordination and Pronominal Anaphora in Discourse. *Linguistics & Philosophy* 12:683–721.

Roberts, Ian. 1991. Excorporation and minimality. *Linguistic Inquiry* 22:209–218.

Rullman, Hotze, Lisa Matthewson, and Henry Davis. 2008. Modals as distributive indefinites. *Natural Language Semantics* 16(16):317–357.

Sarkar, Anoop. 1998. The conflict between future tense and modality: the case of *will* in English. *Penn Working Papers in Linguistics* 5:91–117.

Schwarzschild, Roger. 2011. Stubborn distributivity, multipaticipant nouns and the count/mass distinction. In *Proceedings of NELS*, vol. 39.

Sharma, Ghanshyam. 2008. A pragmatic account of the Hindi presumptive. In *Annual Review of South Asian Languages and Linguistics*, pages 83–116.

Singh, Mona. 1998. On the semantics of the perfective aspect. *Natural Language Semantics* 6(2):171–199.

Tonhauser, Judith. 2006. *The temporal semantics of noun phrases: Evidence from Guarani*. Ph.D. thesis, Stanford University.

Travis, Lisa. 1984. *Parameters and effects of word order variation*. Ph.D. thesis, MIT.

von Fintel, Kai. 1994. *Restrictions on Quantifier Domains*. Ph.D. thesis, University of Massachusetts, Amherst.

von Fintel, Kai and Anthony S. Gillies. 2008. CIA leaks. *Philosophical Review* 117(1):77–98.

von Fintel, Kai and Anthony S. Gillies. 2010. Must...stay...strong! *Natural Language Semantics* 18:351–383.

Werner, Tim. 2006. Future and non-future modal sentences. *Natural Language Semantics* 14:235–255.

Urdu/Hindi Complex Predicates of Motion and the Manner/Result Complementarity

Annette Hautli-Janisz, *University of Konstanz*

Received March 2015; Revised December 2015

Abstract

This paper surveys complex predicates of motion in Urdu/Hindi (Hook 1974, Hautli-Janisz 2013), a spatial resultative construction that denotes manner of motion along a path. In particular, I show that the combinatorial possibilities between main and light verb are driven by the principles of manner/result complementarity set forth by Levin and Rappaport Hovav (2008, 2013). In order to identify these meaning components in Urdu/Hindi verbs, I propose a set of truth-conditional and syntactic tests that identify manner versus result components in the class of Urdu/Hindi motion verbs. Moreover, I provide an analytical framework that shows how the dichotomy drives the patterns found in this type of complex predicate.

1 Introduction

Urdu/Hindi has been shown to exhibit various types of complex verbal constructions, including N+V, ADJ+V and V+V complex predicates (CPs) (e.g., Mohanan (1994), Butt (1995), inter alia). A lesser known construction, but one that occurs fairly frequently is that of complex predicates of motion (Hook 1974, Hautli-Janisz 2013): Here, two motion verbs combine to express manner of motion and direction in one complex event, as shown in (1).[1]

(1) cor mɑkan=se bahɑr **kud nɪkl-a**
thief.M.Sg.Nom house.M.Sg=Source outside **jump emerge-Perf.M.Sg**
'The thief jumped out of the house.' (Hook 1974, p. 69)

From a surface point of view, CPs of motion are similar to aspectual complex predicates (Butt 1995) in that a verb in the root form is followed by a finite verb. However, it is not aspectual information that is contributed by the finite verb, but rather motional information. Taking together the events denoted by both verbs renders a complex motion event, i.e. the manner of motion along a path. In (1), the thief undergoes the path of emerging by way of jumping out of the house.

In this paper I show that Urdu/Hindi motion verbs adhere to the dichotomy of manner and result proposed by Levin and Rappaport Hovav (2008, 2013), with CPs being a way of expressing a combination of both components in a monoclausal structure. After an introduction of the data in Section 2 and a comparison of the construction with other spatial resultatives (Section 3), I establish a set

[1]The Hindi judgements on the data come from Hook (1974), the Urdu judgements come from two informants, Asad Mustafa (from Karachi) and Ghulam Raza (from Punjab).

JSAL Volume 7, December 2015.

of truth-conditional and syntactic criteria that manifest manner and result meaning components for Urdu/Hindi, taking into account language-specific properties (Section 4). The analytical framework in Section 5 shows that the complementarity of manner and result drives the combinatorial possibilities between main verbs and light verbs in CPs of motion as in (1): Here, manner is contributed by the main verb *kud-na* 'to jump' and the result component is contributed by the light verb *nɪkal-na* 'to emerge'. In Section 6, I show that these verb complexes complement the ways of telic path formation in Urdu/Hindi described in Narasimhan (2003). Section 7 concludes the paper.

2 Data

The expression of manner of motion together with the direction of motion is a phenomenon which, from a syntactic point of view, is realized very differently across languages (Talmy 1991). One possibility is the usage of a sequence of motion verbs, where each verb contributes its share in the overall interpretation. Li (1993) shows that this strategy is employed in a wide range of languages, in particular in many West African and South-East Asian languages, but also in languages of Papua New Guinea and East Asia. A cross-linguistic overview of the patterns is shown in examples (2)–(4), with (2) for Korean (Zubizarreta and Oh 2007), (3) for Edo (Baker and Stewart 1999, Ogie 2003) and (4) for Thai (Wechsler 2003).

(2) John-i kongwen-ey **kel-e ka-ss-ta**
John-Nom park-Loc **walk-L go-Past-Decl**
'John walked to the park.' Korean (Zubizarreta and Oh 2007, (7))

(3) Òzó **rhùlé-rè làá** òwá
zo **run-Past enter** house
'Ozo ran into the house.' Edo (Ogie 2003, (19))

(4) Piti **den khâw** roo nrian
Piti **walk enter** school
'Piti entered the school walking.' Thai (Wechsler 2003, (2))

Urdu/Hindi also employs this strategy of expressing complex motion events, with the construction first noted by Hook (1974). The verb sequence consists of two motion verbs, where the first motion verb is in the root form and the second verb is finite and inflects according to the common agreement and tense/aspect patterns in the language. These syntactic properties are shared with aspectual CPs in Urdu/Hindi, an example of which is shown in (5): Here, the main verb *gɪr-na* 'to fall' combines with the light verb *paṛ-na* 'to fall', which contributes a sense of suddenness to the event of falling denoted by the main verb.

(5) am **gɪr pɑṛ-a**
mango.M.Sg.Nom **fall fall-Perf.M.Sg**
'The mango fell (suddenly).'

For CPs of motion, Hautli-Janisz (2013) shows that according to the principles of Butt (1995), the construction is an instance of a complex predicate (CP) in Urdu/Hindi, where the root verb is the main verb of the CP and the finite verb serves as the light verb. This analysis is based on several facts: Firstly, the light verbs in these motion verb sequences do not have a systematic contribution, secondly they contribute a bleached version of their lexical semantics and thirdly, only a restricted set of verbs can serve as light verbs. Moreover, the verbs in these CPs share their arguments in a similar way as aspectual CPs in the language.

As with the cross-linguistic examples from (2)–(4), the Urdu/Hindi constructions are used to convey the manner of the motion which is executed along a path. The following examples illustrate

the phenomenon: In (6), the verb *bʰag* 'run' in the root form combines with the finite verb *nɪkla* 'emerged' which inflects for tense and number and agrees with the masculine subject *sanp* 'snake'. Combining the two motion verbs results in an interpretation similar to 'shoot out of'. Example (7) shows a construction with the two motion verbs *baṛʰna* 'to advance' (root verb) and *doṛna* 'to run' (finite verb), which together denote the event of 'charge into'.

(6) sanp bɪl=se **bʰag nɪkl-a**
snake.M.Sg.Nom snake-pit.M.Sg.Obl=Source **run emerge-Perf.M.Sg**
'A snake shot out of the pipe.'

(7) sanḍ gayõ=ke ṛevɑṛ=ke tɑrɑf
buffalo.M.Sg.Nom cow.F.Pl.Obl=Gen herd.M.Sg=Gen.Obl direction

bɑṛʰ doṛ-a
advance run-Perf.M.Sg
'The buffalo charged into a herd of cows.'

In similar constructions in Yoruba (Ekundayo and Akinnaso 1983) and Sranan (Sebba 1987), the manner of motion verb always comes before the directional motion verb, a pattern that is not found in Urdu/Hindi. Instead, the order is more flexible: Whereas in (6), the manner of motion verb *bʰagna* 'to run' precedes the directional verb *nɪkalna* 'to emerge', the order in (7) is reversed with the directional verb *baṛʰna* 'to advance' preceding the manner of motion verb *doṛna* 'to run'.

However, the combinations are not restricted to cases where manner of motion combines with directional motion. Example (9) shows a construction where two directional motion verbs, *gʰʊsa* 'enter" and *ja* 'go', combine to form an increased directional reading towards a location.[2]

(10) cʰʊri ʊs=ke peṭ=mẽ **ja gʰʊs-i**
knife.F.Sg Pron.3.Sg=Gen stomach.M.Sg=in **go enter-Perf.F.Sg**
'The knife sank into his stomach.'

An interesting property of CPs of motion is that some combinations facilitate the swapping of their motion verbs, while retaining the overall interpretation of the sentence. The nature of the reversal is not one where root and finite verb change their order in the CP, but instead, the verb that is light becomes the main verb, whereas the main verb turns into the light verb. This is illustrated in (11), with the verbs *calna* 'to walk' and *uṛna* 'to fly': Whereas in (11a), the main verb *uṛ* 'fly' in the root

[2]The combinatorial possibilities go further, as exemplified in (8): Here, the sequence comprises two nearly synonymous verbs, *doṛna* 'to run' and *bʰagna* 'to run', with the interpretation of 'running away'. In the sequence in (8), both verbs have a deverbal noun that shares its root with the verb root, namely *doṛ* and *bʰag*, both meaning 'run'. As a nominal compound, *doṛ bʰag* (and its reversed version *bʰag doṛ*) means 'much running around' and is commonly used in the language. Therefore, the verbal sequence in (7) could be interpreted as a deverbal noun of the compound as a whole. As this pattern is exceptional in CPs of motion in Urdu/Hindi, I refrain from drawing any generalizations and merely note that the construction in (8) is uncommon in that sense.

(8) gʰoṛa **doṛ bʰag-a**
horse.M.Sg.Nom **run run-Perf.M.Sg**
'The horse ran away.'

Another exceptionality of the construction is that *bʰag doṛana* (and its inverse *doṛ bʰagana*) 'to run away' are the sole instances of Urdu/Hindi MVSs where causativization only applies to the finite verb, as shown in (9) with *doṛana* 'to cause to run'.

V_1.base + V_2-Caus

(9) malɪk=ne gʰore=ko **bʰag doṛ-a-ya**
owner.M.Sg=Erg horse.M.Sg.Obl=Acc **run run-Caus-Perf.M.Sg**
'The owner made the horse run away.'

form precedes the finite light verb *cala* 'walked'; in the alternative realization in (11b), *cal* 'walk' serves as the main verb and *uṛa* 'flew' is the light verb. According to native speaker judgement, the construction in (11a) is preferred.[3]

(11) a. hɑva=ke jʰonke=ke satʰ pɑtɑng **uṛ cɑl-i**
wind.M.Sg=Gen gust.M.Obl=Gen with kite.F.Sg **fly move-Perf.F.Sg**
'The kite flew up with a gust of wind.' (Hook 1974, p. 57)

b. hɑva=ke jʰonke=ke satʰ pɑtɑng **cɑl uṛ-i**
wind.M.Sg=Gen gust.M.Obl=Gen with kite.F.Sg **move fly-Perf.F.Sg**
'The kite flew up with a gust of wind.' (Hook 1974, p. 57)

The ability to swap verbs in CPs of motion is not dependent on the particular lexical semantics of the combined verbs, i.e. it is not only CPs with two manner of motion verbs that allow for the alternation. The trigger of this alternation, which is not allowed for all CPs of motion and is uncommon for aspectual and permissive complex predicates in the language, still needs to be investigated in further detail.

An interesting property of CPs of motion, which distinguishes them from other CPs in the language, is the possibility of causative formation of (at least one of) the verbs in the CP. The examples in (12) and (13) show the causativized versions of the constructions in (6) and (10), respectively. While in (12), both verbs, *bʰagana* 'to cause to run' and *nɪkalna* 'to cause to emerge' are in the causative, (14) shows that only the verb in the root form, *ʊṛana* 'to cause to fly', causativizes.

V₁-Caus + V₂-Caus

(12) malɪk=ne sanp=ko bɪl=se
owner.M.Sg=Erg snake.M.Sg=Acc snake-pit.M.Sg.Obl=Source

bʰɑg-a nɪkal-a
run-Caus emerge.Caus-Perf.M.Sg
'The owner made the snake shoot out of the snake pit.'

V₁-Caus + V₂.base

(13) hɑva pɑtɑng=ko **ʊṛ-a cɑl-i**
air.F.Sg.Nom kite.M.Sg=Acc **fly-Caus move-Perf.F.Sg**
'The gust made the kite fly up.'

In both cases, an external argument, the causer, is added to the overall event. If, as in (11), both verbs are in the causative, the external argument is shared between the two verbs. Verbs that do not causativize as simple verbs also do not allow for causativization in CPs of motion (e.g. *cʰorna* 'to leave', *ḍagmagana* 'to toddle', *rɪpaṭna* 'to slip'). In turn, verbs that can causativize as simple verbs do not necessarily allow for causativization in these CPs, where certain constraints seem to hold between the two motion verbs.

Despite the combinatorial freedom between motion verbs, native speakers have a clear intuition with respect to the grammaticality or ungrammaticality of certain CPs. For instance, the verb *nɪkalna* 'to emerge' can form a CP with the manner of motion verb *bʰagna* 'to run' as in (11), but the CP with *ḍagmaga* 'to stagger' in (14) is ungrammatical, despite 'to run' and 'to stagger' both being manner of motion verbs. Similarly, *rengna* 'to crawl' can appear in a CP with *nɪkalna* 'to emerge', but is ungrammatical in combination with *gʰʊsna* 'to enter', as shown in (15).

[3]The acceptability of (11b) varies between native speakers, but the majority of the consulted Urdu/Hindi informants judged the swapped version as being grammatical.

(14) * ʃɑrabi kɑmre=se bahɑr ḍɑgmɑga nɪkl-a
drunkard.M.Sg room.M.Sg=Source outside stagger emerge-Perf.M.Sg
'The drunkard staggered out of the room.'

(15) * bɑcca kɑmre=mẽ reng gʰʊs-a
child.M.Sg.Nom room.M.Sg=in crawl enter-Perf.M.Sg
'The child crawled into the room.'

Based on a quantitative investigation of the construction in three corpora, Hautli-Janisz (2013) shows that a number of cps are in fact preferred and used across different genres, for instance *bʰag nɪkalna* 'to run out of (lit. to run emerge)', *baṛʰ caṛʰna* 'to climb up (lit. to advance climb)' and *ʊtar caṛʰna* 'to climb down (lit. to descend climb)'. Moreover, the verbs *bʰagna* 'to run', *doṛna* 'to run' and *calna* 'to move/walk' are often used as light verbs in cps of motion and combine with a range of different main verbs. The most flexible motion verb is *nɪkalna* 'to emerge', which can be used both as a main and a light verb in various combinations. A comparatively large number of motion verbs does not appear in cps of motion at all, particularly very special motion concepts such as *langaṛana* 'to limp' or *matakna* 'to dance (style often found in Bollywood movies)'. The data also shows that causative cps of motion are less frequent than their base counterparts, whereas cps with indirect causatives do not occur at all.

In the following section I show that Urdu/Hindi cps of motion are instances of *spatial resultatives* with telic paths, paving the way for analyzing the construction along the lines of manner and result complementarity.

3 cps of motion as spatial resultatives

Resultative constructions, in particular the group of spatial resultatives, can appear in various guises across languages: For instance, English allows for the usage of an adjectival phrase as in (16a) or a prepositional phrase as in (16b) to denote the resultativity of a path (Goldberg and Jackendoff 2004, inter alia). In both cases, the paths denoted by the constructions are telic, i.e. the motion event is completed once the path denoted by the ap or the pp has been traversed.

(16) a. He jumped *clear* of the traffic.
b. John ran *out of the room*.

Goldberg and Jackendoff (2004) also show that path resultatives can be atelic as in (17), or can have a stative interpretation as shown in (18) (examples (68) and (69) in Goldberg and Jackendoff (2004), respectively): In (17), the path of floating is unbounded, but still expresses resultativity in that the moving entity ends up in a different location. In (18), the road occupies the entire length of the path and is homomorphic to the structure of the path, which, according to Jackendoff (1996), renders a resultative construction.

(17) The boat floated *down the river*.

(18) The road zigzagged *down the hill*.

Other languages employ complex verbal constructions to denote spatial resultativity, as shown in (19) for Sranan (Sebba 1987) and (20) for Ịjọ (Williamson 1965): In (19a), the construction with *waka* 'walk' has a locative reading, whereas with the addition of the verb *go* 'go' in (19b), the interpretation is one of walking along a bounded path. A similar pattern emerges in (20), where the telic path reading with *bómi* 'come' in (20b) is only available when the verb *pá* 'come/go-out' is added.[4]

[4]SPa in the glosses stands for simple past.

(19) a. A **waka** na wowoyo
Pron.3.Sg.M **walk** Loc market
'He walks (about) on the market.'
b. A **waka go** na wowoyo
Pron.3.Sg.M **walk go** Loc market
'He walks to the market.' Sranan (Sebba 1987, (15))

(20) a. toboụ-bì **bó**-mi
child-T **come**-SPa
'The child came.'
b. toboụ-bì **pá** **bó**-mi
child-T **come/go-out** **come**-SPa
'The child came out.' Ịjọ (Williamson 1965, p. 48)

In Urdu/Hindi, simple motion verbs and CPs of motion also differ in their event structure, similar to the difference in (19) and (20) for Sranan and Ịjọ. This is illustrated by way of the verbs *uṛna* 'to fly' and *bʰagna* 'to run' in (21) and (22), respectively: In (21a), the simple verb usage denotes the process of flying, whereas in combination with the light verb *calna* 'to walk' in (21b), the construction turns into a spatial resultative. Here, as a consequence of flying, the kite ends up in a different location, namely away. The same holds for (22), where the activity reading of *bʰagna* 'to run' in (22a) is turned into a resultative event when the verb is used in a CP with the light verb *nɪkalna* 'to emerge', as shown in (22b).

(21) a. pɑtɑng **ʊṛ-i**
kite.F.Sg.Nom **fly-Perf.F.Sg**
'The kite flew.'
b. %pɑtɑng **ʊṛ cɑl-i**
kite.F.Sg.Nom **fly** walk-Perf.F.Sg
'The kite flew up.'

(22) a. lɑṛki bʰag-i
girl.F.Sg.Nom run-Perf.F.Sg
'The girl ran.'
b. lɑṛki bʰag nɪkl-i
girl.F.Sg.Nom run emerge-Perf.F.Sg
'The girl ran out.'

Urdu/Hindi CPs of motion only denote resultative events that are telic, in contrast to the examples shown in (17) and (18) for English. This is illustrated in (23) for the CP *ʊṛ cal-na* 'to fly away (lit. to fly walk)' from (21b), where the telic modifier *palak jʰapakne=mẽ* 'in the blink of an eye (lit. in the blinking of an eyelash)' in (23a) is grammatical. The unbounded modifier *kayi gʰanṭõ=se* 'for many hours' in (23b) is ungrammatical in the atelic reading with the CP, however it is valid as a temporal modifier ('a few hours ago').[5]

(23) a. pɑtɑng pɑlɑk jʰɑpɑkne=mẽ **ʊṛ cɑl-i**
kite.F.Sg eyelash.M.Sg.Nom blink.Inf.Obl=in **fly walk-Perf.F.Sg**
'The kite flew away in the blink of an eye.'

[5]The *for* test, paralleled here for Urdu/Hindi, is complicated insofar as resultative events are accomplishments with a durative event structure, using an atelic modifier sometimes yields acceptable results by highlighting the process that leads to the result state (cf. Beavers 2006).

b. pɑtɑng kɑyi g^{h}ɑnṭõ=se **ʊṛ cɑl-i**
kite.F.Sg many hour.M.Pl.Obl=Source **fly walk-Perf.F.Sg**
*'The kite flew away for many hours.'
'The kite flew away a few hours ago.'

In sum, Urdu/Hindi cps of motion parallel the pattern of resultative formation in languages like Sranan or Ịjọ, where a sequence of motion verbs renders manner of motion along a bounded path.

In the remainder of the paper, I show that the claim of manner-result complementarity by Levin and Rappaport Hovav (2008, 2013) is the crucial factor in the formation of motion cps. In order to manifest the complementarity in Urdu/Hindi, I propose a set of syntactic diagnostics that consistently identifies these meaning components for Urdu/Hindi motion verbs in Section 4. With this distinction at our hands, we can then go on to resolve the combinatorial complexities of cps of motion in Section 5.

4 The complementarity of manner and result

4.1 In general

With the goal of filtering out the *lexicalized meaning* of verbs, i.e. the meaning that each verb contributes irrespective of its context, Levin and Rappaport Hovav (2008, 2013) introduce the notion of the *manner/result complementarity*. According to this concept, a verb either instantiates the manner with which an action is carried out or it denotes the result of an action. This means that manner and result meaning components are in complementary distribution, i.e. in a particular construction, a verb can express only one.

The crucial factor that distinguishes result from manner is the concept of the *scalar event structure* (Rappaport Hovav 2008): The traversal of a path can be considered a scalar change when the moving entity changes its location on the path in a monotonic relation to the event that progresses (Krifka 1998). Scalar changes are the basis for results, in that at the endpoint of the path, the entity is in a state (location) as a result of moving along the path. This reading is entailed in a sentence such as 'The water rose to the top of the dike', where the water level rises on a path that is mappable onto the progress of the event (scalar change). In contrast, manner verbs entail nonscalar change, i.e. dancing involves infinitely many small movements that cannot be mapped on a temporal or spatial path.

For English, Beavers and Koontz-Garboden (2012) establish a set of truth-conditional diagnostics that extract scalar result and nonscalar manner meaning components (for a summary see Table 1). For instance, the *denial of the result* is acceptable for nonscalar manner verbs like 'sweep' in (24) because they do not entail a result state: The floor does not necessarily become cleaner as a consequence of sweeping it. In contrast, the modifier phrase like 'nothing is different about it' is ungrammatical for scalar result verbs like 'break' in (25): The property of the vase changes from being whole to being broken and the result state cannot be reversed.

(24) Shawn swept the floor, but nothing is different about it. (✓manner)

(25) # Shawn broke the vase, but nothing is different about it. (✓result)

In turn, the *denial of action* is only grammatical with scalar result verbs, where actual motion of "various parts of the human body" (Beavers and Koontz-Garboden 2012, p. 345) is not necessarily implied. The test confirms the nonscalar manner nature of 'sweep' (26) and the scalar result structure of 'break' (27).

(26) # Shawn swept the floor, but didn't move a muscle. (✓manner)

(27) Shawn broke the vase, but didn't move a muscle. (✓result)

This classification is further underpinned by the *object deletion* diagnostic: For nonscalar manner verbs in English, the object can be dropped (see (28) for 'sweep'), a pattern that is impossible for

scalar result verbs (see (29) for 'break').

(28) Shawn swept the floor. All last night, Shawn swept. (✓manner)

(29) Shawn broke the vase. # All last night, Shawn broke. (✓result)

A somewhat more vague diagnostic is the *selectional restrictions* test: manner verbs generally impose more selectional restrictions on the subject, because they reject inanimates and natural forces, whereas result verbs accept these. The examples in (30) and (31) illustrate the pattern for 'sweep' and 'break', respectively.

(30) Shawn/#The stiff brush swept the floor. (✓manner)

(31) Shawn/The earthquake/The hammer broke the vase. (✓result)

Another test examines the set of accepted *resultative phrases*: Manner verbs are assumed to have fewer restrictions on the kinds of result they appear with than is the case with result verbs.

(32) Shawn swept the floor clean/shiny/bare. (✓manner)

(33) Shawn broke the vase in half/#purple/#into the ground. (✓result)

Regarding the *complexity of the action*, manner verbs consist of a series of non-trivial intervals of change and are durative (e.g. dancing involves complex bodily motion), whereas result verbs denote scalar change and can either be durative if they denote a multipoint scale (e.g. 'the temperature rises') or are punctual if they denote a two-point scale (e.g. the change from non-broken to broken).

MANNER	DIAGNOSTICS	RESULT
✓	Denial of result	#
#	Denial of action	✓
✓	Object deletion	*
✓	Selectional restrictions	—
✓	Complexity of action (Durative event)	multipoint scale: ✓ two-point scale: —
—	Restricted resultatives	✓

TABLE 1 Diagnostics for manner versus result in English (Beavers and Koontz-Garboden 2012)

Now the question is to what extent the tests employed for English can also be used to detect manner and result meaning components in Urdu/Hindi and the data at hand. Due to language-specific properties of Urdu/Hindi, a number of tests put forth by Beavers and Koontz-Garboden (2012) cannot be applied: For instance the criterion of object deletion does not yield consistent results for a pro-drop language like Urdu/Hindi and for the mostly intransitive verbs of motion. Moreover, I refrain from using diagnostics like the selectional restrictions test, because the patterns are generally difficult to quantify and motion verbs form a class that tends to be more acceptable with animates than inanimates and natural forces.

Taking this into account, §4.2 illustrates how manner versus result meaning components can be identified in Urdu/Hindi motion verbs by applying a set of truth-conditional and syntactic tests. In addition I show that in the class of Urdu/Hindi motion verbs, some verbs express both manner and result – a pattern that is confirmed in Section 5 when analyzing the behavior of these verbs in complex predicates of motion.

4.2 Diagnosing manner and result in Urdu/Hindi motion verbs

For Urdu/Hindi, three tests, namely the denial of the result (§4.2.1), the test with directional versus locational *mẽ* 'in' (§4.2.2) and the telic path alternation (§4.2.3), allow for a reliable distinction

between manner and result meaning components. §4.3 summarizes the results and provides a list of motion verbs and their classification.

4.2.1 Denial of the result

If a verb denotes scalar change (result verbs), at least one property of the entity that undergoes the event is necessarily different than before the event. As a consequence, constructions with scalar verbs are ungrammatical when this change is being contradicted, for instance with the phrase 'but nothing is different about it'. For the class of motion verbs, the denial of the result can be attested with the phrase 'but X is not somewhere else', showing that as a result of the motion, the entity ends up in a different location. Similar to 'to break', a small number of scalar motion verbs inherently license the result state of the event. In the class of Urdu/Hindi motion verbs, these are *gʰʊsna* 'to enter', *nɪkalna* 'to emerge' and *pahʊncna* 'to arrive'. As exemplified in (34a) with *nɪkalna* 'to emerge', the girl gradually emerges from the room and as a result of emerging, she ends up in a final location which lies outside of the room. The resultativity is attested with the phrase *lekɪn vuh kahĩ aur jɑgɑh nahĩ gayi* 'but she did not go to some (other) place' in (34b). This test shows that the verb itself entails a result state which cannot be reversed by the modifier phrase.[6]

(34) a. laṛki kɑmre=se nɪkl-i
girl.F.Sg.Nom room.M.Sg=Source emerge-Perf.F.Sg
'The girl emerged from the room.'

b. # laṛki kɑmre=se nɪkl-i
girl.F.Sg.Nom room.M.Sg=Source emerge-Perf.F.Sg

lekɪn vuh kɑhĩ aur jɑgɑh nɑhĩ ja-ti hɛ
but Pron.3.Sg somewhere and place.M.Sg not go-Impf.F.Sg be.Pres.3.Sg
#'The girl emerged from the room, but she isn't somewhere else.'

Other scalar motion verbs do not entail a result, as shown in (35) with the verb *ʊtar-na* 'to descend', where the boat moves from a higher position to a lower position. Despite the lack of an explicit endpoint with *nɪkal-na* 'to emerge', the denial of the result test still fails due to the scalar structure of the event: The path of motion is monotonic to the event structure, i.e. as the event progresses, the boat moves lower and the location at the end of the event is different than at the beginning.

(35) # kɑʃti ʊtr-i
boat.F.Sg.Nom descend-Perf.F.Sg

lekɪn vuh kɑhĩ aur jɑgɑh nɑhĩ ja-ti hɛ
but Pron.3.Sg somewhere and place.M.Sg not go-Impf.F.Sg be.Pres.3.Sg
#'The boat descended, but it isn't somewhere else.'

In contrast, a large number of motion verbs in Urdu/Hindi express the manner of motion and are nonscalar, i.e. the motion they denote is not mappable on a temporal or spatial path. For instance, motion concepts like *lɑngaṛa-na* 'to limp', *maṭak-na* 'to dance (style often found in Bollywood movies)' and *ḍagmaga-na* 'to wobble' involve infinitely many small parts of motion that combine to form the overall motion but do not necessarily entail a change in location. Therefore, these verbs allow for the denial of the result, as shown in (36) for *ter-na* 'to float': Here, the boat moves in(side) the water, but does not necessarily change its position in the course of the event.

[6] (34b) is felicitous in a context where the girl emerges from the room, but then stays right outside the room and does not go somewhere else, however this still entails that she left her initial starting position inside the room.

(36) kɑʃti ter-i
boat.F.Sg.Nom float-Perf.F.Sg

lekɪn vuh kɑhĩ aur jɑgɑh nɑhĩ ja-ti hɛ
but Pron.3.Sg somewhere and place.M.Sg not go-Impf.F.Sg be.Pres.3.Sg
'The boat floated, but it isn't somewhere else.'

In the following, the established distinction is confirmed by the results of a test that is based on language-specific properties of Urdu/Hindi, namely the test for locative versus directional *mẽ*.

4.2.2 Locative versus directional mẽ

Similar to languages like German (Gehrke (2007), inter alia), some postpositions in Urdu/Hindi either have a locative or directional interpretation when they appear with motion verbs. In particular the postposition *mẽ* 'in' can have a telic reading of directed motion ('into') or an atelic reading of locational movement ('in(side)'), depending on the motion verb it appears with. The contrast is shown in (37) and (38): With *ʊtarna* 'to descend' in (37b), *mẽ* 'in' explicitly records the end point of the path with the postpositional phrase *pani=mẽ* and describes the state of the boat as a result of its descent ('into the water'). This means that, together with the denial of result test for *ʊtarma* 'to descend' in (35), the verb can be unambiguously classified as a result verb — together with all other result verbs in Table 2 (§4.3) share.

(37) a. kɑʃti ʊtr-i
boat.F.Sg.Nom descend-Perf.F.Sg
'The boat descended.'

b. kɑʃti pani=mẽ ʊtr-i
boat.F.Sg.Nom water.M=in descend-Perf.F.Sg
'The boat descended into the water.'

In contrast, *mẽ* can also have a locative usage, exemplified in (39b) for the verb *bʰagna* 'to run': Here, *kɑmre=mẽ* 'room.Loc' denotes the location in which the motion happened.

(38) a. lɑṛki bʰag-i
girl.F.Sg=Nom run-Perf.M.Sg
'The girl ran.'

b. lɑṛki kɑmre=mẽ bʰag-i
girl.F.Sg.Nom room.M.Sg.Obl=in run-Perf.M.Sg
'The girl ran in(side) the room.'

The directional interpretation of *mẽ* 'in' as in (37b) correlates with the ungrammaticality of the denial of the result, i.e. verbs like *pahʊncna* 'to arrive', *ʈapakna* 'to drop' and *nɪkalna* 'to emerge' license a telic, directional interpretation with *mẽ* 'in' and do not allow for the denial of the result. These verbs are therefore result verbs, because the motion event can be mapped onto a spatial path which implies motion from one location to another. On the other hand, the atelic, locational interpretation of *mẽ* with verbs like *bʰagna* 'to run' as in (38b) concurs with the grammaticality of the denial of the result, i.e. the motion is restricted to a specific location and is not mapped onto a spatial path. These verbs, among them *ɪʈʰlana* 'to strut', *pʰʊdakna* 'to hop' and *cakarana* 'to stagger', therefore have a nonscalar event structure and are manner verbs.

4.2.3 Exceptional cases

Two verbs in the class of motion verbs in Urdu/Hindi, namely *bʰagna* 'to run' and *calna* 'to walk', exhibit an exceptional pattern in that they can express either manner or result meaning, depending

on the context they appear in. This can be tested with an alternation that is grammatical only with those two manner verbs, namely the *telic path alternation* (Hautli-Janisz 2014): Similar to languages like English, the intransitive verb frame alternates with a frame that takes an oblique denoting the bounded path of motion. Example (39) shows the grammaticality of the alternation with the verb *bʰagna* 'to run', (40) employs *nacna* 'to dance' to illustrate the ungrammaticality with other manner verbs.

(39) a. lɑṛki bʰag-i
girl.F.Sg=Nom run-Perf.M.Sg
'The girl ran.'

b. lɑṛki=ne lɑmbi bʰag=ko bʰag-a
girl.F.Sg=Erg long.F.Sg run.F.Sg=Acc run-Perf.M.Sg
'The girl ran the marathon.'

(40) a. lɑṛki nac-i
girl.F.Sg=Nom dance-Perf.M.Sg
'The girl danced.'

b. * lɑṛki=ne lɑmbi nac=ko nac-a
girl.F.Sg=Erg long.F.Sg dance.F.Sg=Acc dance-Perf.M.Sg
'The girl danced the long dance.'

Although both *bʰagna* 'to run' and *calna* 'to walk' have been identified as manner verbs according to the denial of the result test in §4.2.1 and the locative interpretation of *mẽ* 'in' in §4.2.2, they behave like result verbs in the telic path alternant — a property that other manner verbs do not have (for instance see the manner verb *nac-na* 'to dance' in (40)). The confirmation that they are solely resultative when they lexicalize their result meaning component and do not also lexicalize manner at the same time is shown in (41): Here, the telic path alternant of *bʰagna* 'to run' is modified with 'but she isn't somewhere else', a phrase that is grammatical only with manner verbs. In (41), this modification renders the construction ungrammatical, showing that in this context, the verb expresses solely a resultative meaning. This means that *bʰagna* 'to run' and *calna* 'to walk' are two verbs in Urdu that can lexicalize either a manner or a result meaning component depending on the context they appear in.

(41) * lɑṛki=ne lɑmbi bʰag=ko bʰag-a
girl.F.Sg=Erg long.F.Sg run.F.Sg=Acc run-Perf.M.Sg

lekɪn vuh kɑhĩ aur jɑgɑh nɑhĩ gɑ-yi
but Pron.3.Sg somewhere and place.M.Sg not go-Perf.F.Sg
'The girl ran the marathon, but she isn't somewhere else.'

This alternation shows that, similar to 'to climb' in English (Levin and Rappaport Hovav (2013) along with Kiparsky (1997)), *bʰagna* 'to run' and *calna* 'to walk' can denote nonscalar motion at a specific location as well as motion that is mappable onto a spatial path, for instance *lambi bʰag* 'marathon (lit. long run)' as in (40b). The exceptional behavior of *bʰagna* 'to run' and *calna* 'to walk' coincides with a cross-linguistic pattern: Levin et al. (2009) show that the Spanish *corer* 'to run' and *caminar* 'to walk' as well as the Italian *correre* 'to run' can express manner and result, depending on the context they occur in. This, together with the telic path alternation in (41), provides an explanation as to why these two verbs stand out in Urdu/Hindi cps of motion, in a way I present later on.

4.3 Summary

The syntactic diagnostics above suggest that three groups of motion verbs in Urdu/Hindi exist: Those that have a nonscalar event structure (manner verbs), those that denote scalar motion (result verbs) and those that can express either one depending on the context they appear in. The members of each class are listed in Table 2.

Manner		Result		Either one	
tʰarak-na	'to stomp'	*ʤʰul-na*	'to swing'	*cal-na*	'to walk'
ʈɛhɪl-na	'to lollop'	*gʰʊs-na*	'to enter'	*bʰag-na*	'to run'
ʈʰʊmak-na	'to strut'	*mʊɽ-na*	'to turn'		
lapak-na	'to dash'	*pʰalang-na*	'to leap over'		
kud-na	'to jump'	*ʊbʰar-na*	'to rise'		
sɛrak-na	'to slither'	*pahʊnc-na*	'to arrive'		
reng-na	'to crawl'	*pʰand-na*	'to leap over'		
rapaʈ-na	'to slip'	*ʊtar-na*	'to descend'		
pʰɪsal-na	'to slip'	*nɪkal-na*	'to emerge'		
laɽkʰara-na	'to stumble'	*cʰor-na*	'to leave'		
kʰɪsak-na	'to slide'	*gɪr-na*	'to fall'		
ɪʈʰla-na	'to strut'	*palaʈ-na*	'to turn'		
matak-na	'to sashay'	*ʈapak-na*	'to drop'		
pʰʊdak-na	'to hop'	*gʊzar-na*	'to cross'		
ter-na	'to float'	*a-na*	'to come'		
langara-na	'to hobble'	*ja-na*	'to go'		
caɽʰ-na	'to climb'	*bɛhɛ-na*	'to run (water)'		
cakara-na	'to stagger'	*baɽʰ-na*	'to advance'		
mandɛla-na	'to wander'				
ʤʰapaʈ-na	'to scram'				
lapeʈ-na	'to roll'				
doɽ-na	'to run'				
gʰum-na	'to roll/rotate'				
lʊɽʰak-na	'to tumble'				
nac-na	'to dance'				

TABLE 2 Manner and result in Urdu/Hindi motion verbs

Now that the manner-result complementarity for Urdu/Hindi motion verbs is established, the next question concerns the kinds of mechanisms that govern the compositionality between the verbs in CPs of motion. To that end, I propose a general schema in the following section, based on the patterns elicited above.

5 Manner and result in CPs of motion

This section shows how the complementarity of manner and result plays a crucial role in unraveling the mechanisms that govern the combinatorial possibilities of motion verbs in the CP. The data in Table 3 summarizes the different patterns of combining manner and result in CPs of motion in Urdu/Hindi and is drawn from Hook (1974), my own field research and a corpus investigation (Hautli-Janisz 2013). They are the basis for the following discussion.

5.1 Scalar + nonscalar motion

The basic pattern in CPs of motion is that an unambiguous result verb combines with an unambiguous manner verb. As shown in Section 2, the syntactic function of the verb in the CP does not correlate with a particular lexical semantic function, i.e. neither is the light verb restricted to express only

		CPs of motion		
Basic	(a)	MANNER + RESULT	*kud nɪkal-na*	'to jump out of (lit. to jump emerge)
	(b)	RESULT + MANNER	*baɽʰ doɽ-na*	'to charge into' (lit. to advance run)
	(c)	RESULT + MANNER	*gʰʊs cal-na*	'to move into (lit. to enter walk)'
Disjunctive	(d)	MANNER + RESULT	*ʊɽ cal-na*	'to fly away' (lit. to fly walk)
	(e)	MANNER + RESULT	*bʰag ja-na*	'to run away (lit. to run go)
	(f)	MANNER + RESULT	*doɽ bʰag-na*	'to run away' (lit. to run run)
Deixis	(g)	DEIC + RESULT	*a nɪkal-na*	'to come out' (lit. to come emerge)
	(h)	DEIC + RESULT	*ja gʰʊs-na*	'to go into' (lit. to go enter)

TABLE 3 Urdu/Hindi CPs of motion

result, nor does the main verb obligatorily express manner, or vice versa. As shown with the CP in (a) in Table 3, the manner verb *kudna* 'to jump' is the main verb and combines with the result light verb *nɪkalna* 'to emerge'. Conversely, in (b) in Table 3, the result main verb *baɽʰna* 'to advance' combines with the manner light verb *doɽna* 'to run'. Consequently, the light verb contributes the information that the main verb in the CP does not express, in Table 3 (a) it is the resultative information, in (b) it is the manner information.

This leads to our first conclusion, which is captured in Figure 1: If an unambiguous result verb serves as the light verb (Vlight) in a CP with an unambiguous manner verb, it contributes resultative information to the main verb meaning (Vmain) — represented by the `+Result` arrow. If an unambiguous manner verb serves as the light verb and has an unambiguous result main verb, it contributes manner information to the overall CP, represented by the `+Manner` arrow from right to left.

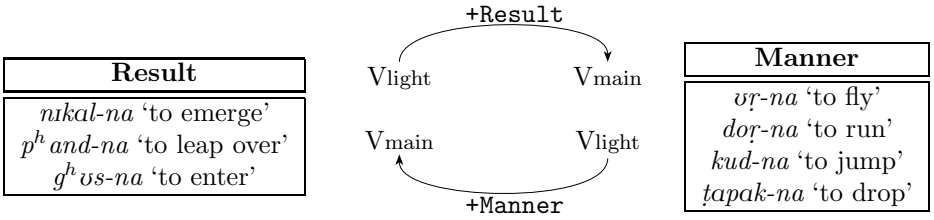


FIGURE 1 Combining result with manner of motion

5.2 Verbs with disjunctive behavior

As established in §4.2, two Urdu/Hindi motion verbs, namely *calna* 'to walk' and *bʰagna* 'to run', exhibit a disjunctive behavior with respect to the manner/result complementarity. In particular, these verbs can express either a scalar or a nonscalar event structure, but realize only one depending on the companion verb in the motion CP. The CPs in (c) and (d) in Table 3 and the examples in (42) illustrate the pattern for the verb *calna* 'to walk': If it combines with a scalar verb like *gʰʊsna* 'to enter' in (42a), it denotes the manner of continuous movement, expressing its manner interpretation. In contrast, when combined with a nonscalar verb like *ʊɽna* 'to fly' as in (42b), it adds a scalar path interpretation to the nonscalar event of flying.

(42) a. sanḍ mɑkan=mẽ **gʰʊs cɑl-a**
ox.M.Sg.Nom house.M.Sg=in **enter move-Perf.M.Sg**
'An ox got into the house.'

b. pɑtɑng **ʊɽ cɑl-i**
kite.F.Sg **fly walk-Perf.F.Sg**
'The kite flew away.'

This multifunctionality also pertains to *bʰagna* 'to run', illustrated by the constructions in (43): In combination with the light verb *jana* 'to go' in (43a), *bʰagna* 'to run' expresses its manner meaning and *jana* 'to go' contributes a scalar meaning in laying out the path of motion to the final location. In the CP in (43b) with the nonscalar verb *doṛna* 'to run', however, *bʰagna* 'to run' expresses a scalar meaning and lays out the path that the child traverses by the manner of motion *doṛna* 'to run'.

(43) a. bɑcca **bʰag gɑ-ya**
child.M.Sg.Nom **run go-Perf.M.Sg**
'The child ran away.'

b. bɑcca **doṛ bʰag-a**
child.M.Sg.Nom **run run-Perf.M.Sg**
'The child ran away.'

These patterns are shown in Figure 2: If *bʰagna* 'to run' and *calna* 'to walk' serve as light verbs (Vlight) in a CP, their contribution depends on the main verb in the CP. Combined with an unambiguously scalar motion verb like *nɪkalna* 'to emerge' as the main verb (Vmain), *bʰagna* 'to run' and *calna* 'to walk' realize their `Manner` meaning component ('to run out of' and 'to walk out of', respectively). If combined with the nonscalar main verb *doṛna* 'to run', the light verb realizes its `Result` component and contributes the path of motion ('to run away').

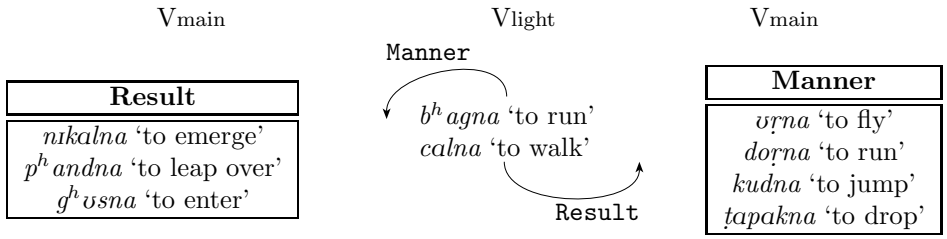


FIGURE 2 Disjunctive verbs *bʰagna* 'to run' and *calna* 'to walk'

The patterns indicate that the main verb in the CP is the discriminant factor that decides what meaning component the light verb realizes. The main verb unifies its meaning with the meaning of the light verb — an assumption that is also made for aspectual and permissive complex predicates in Urdu (c.f. Butt (1995)). For *bʰag-na* 'to run' and *cal-na* 'to walk', I conclude that if they are used as light verbs, they are underspecified – they can either lexicalize manner or result, depending on the main verb they combine with. The view of the "semantic primacy" of the main verb for selecting light verbs and their meaning contribution is further corroborated by the patterns found for the verbs *ana* 'to come' and *jana* 'to go', illustrated in the following.

5.3 The verbs *ana* 'to come' and *jana* 'to go'

The verbs *jana* 'to go' and *ana* 'to come' are two more motion verbs in Urdu/Hindi that exhibit multifunctional properties in CPs of motion (also see (g) and (h) in Table 3). Especially *jana* 'to go' is flexible in its event modulation, depending on the main verb it appears with in the CP. Whereas the contribution in a CP with *rapaṭna* 'to slip' is one of sudden inception ('to slip suddenly'), in combination with *nacna* 'to dance', *jana* 'to go' triggers a serial interpretation of the event ('to dance and then go'). Another usage of *jana* 'to go' entails telicity: In a CP with a verb that has a scalar interpretation like *ʊbʰarna* 'to rise' in (44), *jana* delimits the path of rising by contributing the end point of the motion ('to rise completely'), i.e. the atelic event structure of *ʊbʰarna* 'to rise' turns telic by way of adding the light verb *jana* 'to go'. This is the aspectual light verb usage of *jana* 'to go' as discussed in Butt (1995).

(44) surɑj **ʊbʰɑr gɑ-ya**
sun.M.Sg.Nom **rise go-Perf.M.Sg**
'The sun rose (completely).'

This function of *jana* 'to go', I claim, is independent of the principles of the manner/result dichotomy that drive the combinatorial possibilities between main and light in motion CPs. In these constructions no manner of motion is entailed: In a CP with an atelic scalar motion verb like *ʊbʰarna* 'to rise', the telic scalar nature of *jana* 'to go' merges its path with the main verb and adds the end point of the path that is not entailed by the main verb. This means that the verb is so light that it does not contribute an event in itself, it merely functions as a telic aspectual modifier of the main verb, in accordance with the assumptions in Butt (1995).

However, in combination with nonscalar verbs like *doṛna* 'to run', *jana* 'to go' fulfills the function that is predicted by the manner/result complementarity: While the nonscalar motion denoted by *doṛna* 'to run' is not mappable on a path that the entity traverses, the scalar path interpretation is contributed by *jana* 'to go', rendering a meaning of 'running away' as exemplified in (45). This differentiates the contribution of *jana* 'to go' from the construction in (44): Whereas in (44), it merely marks the end of the path because the main verb already denotes the path, the nonscalar motion of *doṛna* 'to run' in (45) does not conflict with the scalarity of *jana* 'to go' and the light verb can contribute its full path meaning. Therefore, the manner/result complementarity makes the right predictions as to the interpretation of *jana* 'to go': Together with a scalar motion verb, *jana* 'to go' unifies its scalar nature and serves as a telic aspectual modifier, in combination with a nonscalar motion verb, *jana* 'to go' expresses its full result meaning.

(45) bɑcca **doṛ gɑ-ya**
child.M.Sg.Nom **run go-Perf.M.Sg**
'The child ran away.'

However, *jana* 'to go' can also have a different spatial contribution in motion CPs that is independent from the manner/result complementarity, namely the notion of deixis. Following Fillmore (1966) for English, 'to go' entails that "the place to which one goes is the place where [the speaker] is not" (p. 223). The deictic nature of *jana* 'to go' explains the function of the verb when it combines with verbs that are scalar and have an end point already inherent in their event structure, for instance *nɪkalna* 'to emerge' in (46), where *jana* 'to go' cannot contribute any scalar path information. I claim that in these constructions, *jana* 'to go' adds a purely deictic meaning in the sense that the moves out of the room and to a place, which does not concur with the location of the speaker. The CP in (46) with *nɪkalna* 'to emerge' therefore only denotes result, the principles of the manner/result complementarity do not apply.

(46) lɑṛki kɑmre=se nɪkɑl gɑ-yi
girl.F.Sg.Nom room.Sg.Obl=Source emerge go-Perf.F.Sg
'The girl emerged from the room (and went away).'

The behavior of *jana* 'to go' is mirrored by the verb *ana* 'to come' in Urdu/Hindi: In (47), the main verb of the CP, *gɪrna* 'to drop', is a scalar verb with an event structure that is mappable on a path. In combination with *ana* 'to come', the event structure becomes bounded, yielding a telic construction with the final location *asman=mẽ* 'into the sky'. This construction, I claim, is in fact an aspectual complex predicate similar to the construction with *jana* 'to go' in (44), in that *ana* 'to come' solely contributes an endpoint to the event, but no event on its own.

(47) asman=mẽ badɑl **gɪr a-e**
sky.M.Sg=in cloud.M.Pl **drop come-Perf.M.Pl**
'Clouds flooded into the sky.' (Hook 1974, p. 79)

Following Fillmore (1966) in his deictic interpretation of the English 'to come', the contribution of *ana* 'to come' in CPs of motion is that the path of motion is directed towards the location of the speaker. This pattern is found in example (48), the equivalent of the construction in (46) with *jana* 'to go'. Instead of the interpretation of 'run away' with a path directed away from the speaker, *ana* 'to come' denotes the motion of running towards the speaker.

(48) bɑcca **doṛ a-ya**
child.M.Sg.Nom **run come-Perf.M.Sg**
'The child came running.'

As with *jana* 'to go', *ana* 'to come' can also form CPs of motion with verbs that are scalar and have an inherent endpoint, illustrated in (49) with the verb *pʰandna* 'to leap over'. Parallel to *jana* 'to go', *ana* 'to come' is solely used to mark the deictic structure of the event. Here, Ali leaps over the wall, coming towards the speaker. As above, I argue that these combinations do not violate the manner/result complementarity, because *ana* 'to come' situates the motion event in relation to a reference point and does not influence the nature of the motion event itself, i.e. the light verb does not act upon the manner with which the event is carried out nor does it affect the resultative nature of the event. Again, I claim that deixis is a factor outside of the manner/result dichotomy, with the CP in (49) only realizing result, but no manner component.

(49) ɑli devar=ko **pʰand a-ya**
Ali.M.Sg.Nom wall.F.Sg=Acc **leap-over come-Perf.M.Sg**
'Ali came leaping over the wall.'

In sum, the light verbs *ana* 'to come' and *jana* 'to go' fulfill a set of functions depending on the main verb they combine with, illustrated in Figure 3. If they combine with a scalar motion verb that licenses an inherent endpoint to the event, for instance *nɪkalna* 'to emerge', the light verbs contribute a sense of `Deixis` to the event structure in that the event is located with relation to a reference point. Here, the light verbs function outside of manner/result dichotomy and are not complex events in that they only have a resultative path interpretation without any manner of motion entailed. If they combine with a scalar verb that is atelic, e.g. *ʊbʰarna* 'to rise', then they function as aspectual light verbs and only contribute the endpoint of the motion (`Telicity`). In combination with nonscalar motion verbs like *doṛna* 'to run' and *kudna* 'to jump', they consistently realize their scalar component, which triggers the `Result` interpretation of the construction, rendering "true" CPs of motion that denote manner of motion along a path.

5.4 Interim summary

This section has shown that manner and result are combined in CPs of motion and are the principles that govern the combinatorial possibilities between main verbs and light verbs. In general, the meaning between the verbs in the CP is unified — the main verb is the driving force that selects the meaning component of the light verb, which results in a varied contribution of a subset of light verbs, in particular *bʰag-na* 'to run', *cal-na* 'to walk', *ja-na* 'to go' and *a-na* 'to come'. The principle of unification is mirrored on the level of syntax, where Hautli-Janisz (2013) shows that the arguments of main verb and light verb are merged and form a joint predication. An interesting theoretical question is how to combine and formalize the semantic and syntactic processes of unification: Instead of considering them as subsequent modules where first the semantics decides on the combinatorial possibilities and then the syntax combines the arguments, it would be preferable to have a joint analysis — an issue that I leave for further research. Before I conclude, I briefly discuss how Urdu/Hindi CPs of motion fit in the typology of constructions denoting manner of motion along a path.

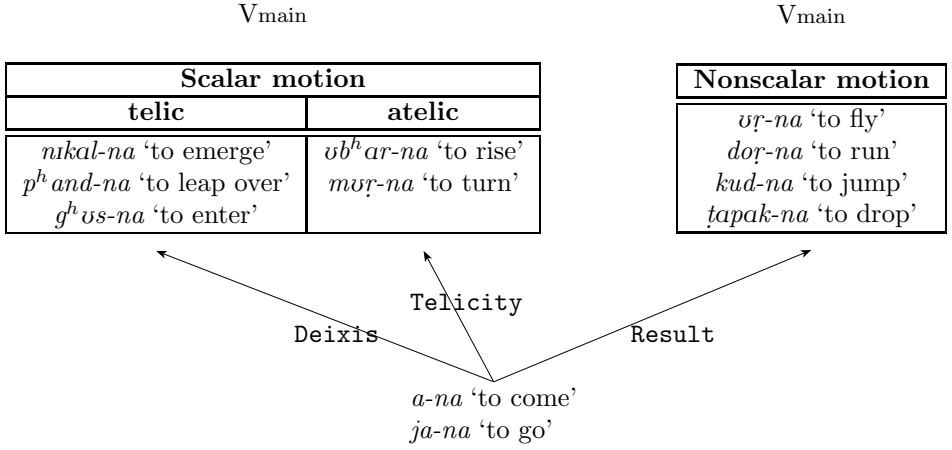


FIGURE 3 Disjunctive verbs *a-na* 'to come' and *ja-na* 'to go'

6 Urdu/Hindi CPs of motion and the typology

Manner of motion along a path is expressed very differently across languages. In verb-framed languages like Spanish, French, Korean, Japanese (Talmy 1985, 1991), a construction parallel to the Spanish example in (50) is employed: Here, the direction of motion is expressed with the main verb *subir* 'to rise', whereas the manner of motion is an adjunct to the matrix clause, here *flotando* 'floating'.

(50) El globo subió por la chimenea (flotando).
the balloon moved-up through the chimney (floating)
'The balloon rose up the chimney, floating.' (Talmy 1985, (15e))

In satellite-framed languages, in contrast, illustrated in (51) for English and (51) for Dutch, the manner of motion verb is the main predicate of the sentence, with the path of motion contributed by the prepositions 'up' and *binnen* 'in', respectively. This way of combining manner of motion with the path of motion is not available in verb-framed languages.

(51) The ballon floated up the chimney.

(52) Jan rende de kamer binnen.
John ran Art room in
'John ran into the room.' (Slobin 2005, (2))

For Urdu/Hindi, Narasimhan (2003) claims that the language falls into the class of verb-framed languages, based on examples like (53): The path of motion to the target location *kamre=mẽ* 'into the room' with the manner of *langaɽana* 'to limp' requires the manner verb to appear in the subordinate clause *langaɽate hue* 'limping', with the path contributed by the matrix verb *ana* 'to come'.

(53) lɑṛka lɑngɑṛa-te hu-e
boy.M.Sg.Nom limp.Impf.M.Sg.Obl become-Perf.M.Sg.Obl

kɑmre=mẽ a-ya
room.M.Sg.Obl=in come-Perf.M.Sg
'The boy limped into the room (came into the room, limping).'

With the CPs of motion investigated in this paper, it becomes clear that a simple classification along the lines of verb-framed versus satellite-framed languages is problematic for Urdu/Hindi, because the language features other ways of expressing complex motion in a monoclausal construction. In fact, the CPs of motion show that Urdu/Hindi shares crucial properties with *equipollently-framed languages*, a typological category put forth by Slobin (2004), complementing the typology established by Talmy (1985). Languages of this type mark the manner and the path of motion with elements "that are equal in formal linguistic terms and appear to be equal in force and significance" (p. 9). Across languages, sequences of verbs, for instance [manner verb + path verb] in Niger-Congo languages, [manner + path]$_{\text{Verb}}$ constructions in Algonquian languages and [manner preverb + path preverb + verb] constructions in Jaminjungan languages express the manner of motion along a designated path.

In equipollently-framed languages, the participating verbs are equal, with none being subordinate to the other. This, I claim, is what is found in Urdu/Hindi: Although the status of the finite verb as a light verb entails that main verb and light verb are not completely equal concerning their syntactic and semantic "weight", they are equal in the sense that from a syntactic point of view, neither the main verb nor the light verb is subordinated (Hautli-Janisz 2013). From a semantic point of view, they are equal from a manner/result point of view in that one verb contributes the manner of the motion, while the other verb contributes the path (or 'result' in terms of Levin and Rappaport Hovav). I therefore argue that Urdu/Hindi cannot be clearly allocated to the group of verb-framed languages, because the language features constructions that are typical for equipollently-framed languages. Due to the fact that manner of motion verbs in Urdu/Hindi cannot realize telic paths by way of using postpositions, I claim that complex predicates of motion are a method for compensating for this gap in that they avoid a complicated subordinated structure similar to the one shown in (53). Complex predicates of motion are therefore the only way of syntactically realizing manner of motion verbs and telic paths in a monoclausal construction in Urdu/Hindi.

7 Discussion and conclusion

This paper shows that Urdu/Hindi CPs of motion are telic spatial resultatives, where the dichotomy between manner and result explains the various combinatorial possibilities between main verbs and light verbs. In order to support this claim, a set of syntactic tests was presented that establish result and manner components in Urdu/Hindi motion verbs. The resulting classification shows that the dichotomy of manner and result introduced by Levin and Rappaport Hovav (2008, 2013) holds, namely that a verb can express only one meaning component per construction, in contrast to the patterns that Beavers and Koontz-Garboden (2012) find for English manner-of-killing verbs, which lexicalize both manner and result in a single construction.

An important conclusion is that light verbs of motion, which can license both components, adjust their contribution to the structure of the main motion verb in the CP. Consequently, each CP has only one motion verb which contributes the manner and the result meaning component. This explains why exactly two motion verbs combine in CPs of motion and do not allow for further augmentation, as for example possible in languages like Dagaare, with four different motion verbs. Urdu/Hindi seems to strictly adhere to the principle that each meaning component can only be expressed by one verb, i.e. nonscalar motion cannot be contributed by more than one verb.

This principle also explains the unavailability of aspectual modification by using CPs of motion in aspectual complex predicates, in particular with *jana* 'to go': The lexical semantic "slots" of

result are already filled by the combining motion verbs in the CP and it is therefore impossible to merge another event modifier into the construction. The parameters of result and manner and the stipulation of their complementary distribution therefore do not only make the right predictions for the combinatorial possibilities of motion verbs in these CPs, they also explain other properties of the construction.

Across languages, the same verbs seem to be used in complex motion events to denote either path or manner of motion, independent of the exact syntactic status of the verb complex. For instance, the Sranan verbs *komoto* 'to come out', *komopo* 'to come from', *fadon* 'to fall', *opo* 'to arise', *gwe* 'to go away', *go* 'to go', *kon* 'to come', *waka* 'to walk' are commonly used to modify motion events (Sebba 1987). Chinese encodes spatial resultativity using a set of directionals such as *jin* 'enter', *chu* 'exit', *duo* 'cross', *lai* 'come' or *qu* 'go' (Scott 1996). These combine with other motion verbs in complex predicates, for example *pao jin* 'to enter running (lit. run enter)', *tong guo* 'to cross (lit. to traverse cross)' (taken from Butt (2010)). Therefore, Chinese seems to feature the same underlying principles of encoding complex motion than Urdu/Hindi, both from a syntactic as well as a lexical semantic point of view.

An interesting area for further research concerns an Urdu/Hindi-wide investigation of manner/result complementarity. This includes an extension of the truth-conditional and syntactic tests that diagnose manner and result in order to identify the two meaning components independent of the verb class. Resolving these issues will show whether manner and result meaning components are truly complementary or whether there are cases like 'to guillotine' in English, where the dichotomy is violated, as shown by Beavers and Koontz-Garboden (2012). This will also pave the way for a more general analysis of the effects of manner and result in the language.

Acknowledgments

I am greatly indebted to Asad Mustafa from KICS Lahore, Pakistan, for his support in data collection and Ghulam Raza for cross-checking the syntactic patterns. I also thank Rajesh Bhatt, Miriam Butt and Gillian Ramchand for valuable discussions and the two anonymous reviewers for their helpful comments. All remaining errors are my own.

References

Baker, Mark and Osamuyimen T. Stewart. 1999. On Double-Headedness and the Anatomy of the Clause. Manuscript.

Beavers, John. 2006. *Argument/Oblique Alternations and the Structure of Lexical Meaning*. Ph.D. thesis, Stanford University.

Beavers, John and Andrew Koontz-Garboden. 2012. Manner and result in the roots of verbal meaning. *Linguistic Inquiry* 43(3):331–369.

Butt, Miriam. 1995. *The Structure of Complex Predicates in Urdu*. Stanford: CSLI Publications.

Butt, Miriam. 2010. The light verb jungle: still hacking away. In M. Amberber, B. Baker, and M. Harvey, eds., *Complex Predicates*, pages 48–78. Cambridge: Cambridge University Press.

Ekundayo, S. Ayotunde and F. Niyi Akinnaso. 1983. Yoruba serial verb string commutability constraints. *Lingua* 60(2-3):115–133.

Fillmore, Charles J. 1966. Deictic categories in the semantics of 'come'. *Foundations of Language* 2(3):219–227.

Gehrke, Berit. 2007. Putting path in place. In E. Puig-Waldmüller, ed., *Proceedings of Sinn und Bedeutung*, vol. 11, pages 244–260.

Goldberg, Adele E. and Ray Jackendoff. 2004. The English resultative as a family of constructions. *Language* 80(3):532–568.

Hautli-Janisz, Annette. 2013. Moving Right Along: Motion Verb Sequences in Urdu. In M. Butt and T. H. King, eds., *In Proceedings of the LFG13 Conference*, pages 295–215.

Hautli-Janisz, Annette. 2014. *Urdu/Hindi Motion Verbs and Their Implementation in a Computational Lexical Resource*. Ph.D. thesis, Universität Konstanz.

Hook, Peter Edwin. 1974. *The Compound Verb in Hindi*. The University of Michigan: Center for South and Southeast Asian Studies.

Jackendoff, Ray. 1996. The proper treatment of measuring out, telicity, and perhaps even quantification in English. *Natural Language and Linguistic Theory* 14:305–354.

Kiparsky, Paul. 1997. Remarks on denominal verbs. In A. Alsina, J. Bresnan, and P. Sells, eds., *Complex Predicates*, pages 473–499. Stanford: CSLI Publications.

Krifka, Manfred. 1998. The origins of telicity. In S. Rothstein, ed., *Events and Grammar*, pages 197–235. Dordrecht: Kluwer Academic Publishers.

Levin, Beth, John Beavers, and Shiao Wei Tham. 2009. Manner of motion roots across languages: Same or different? In *Roots Workshop, University of Stuttgart*.

Levin, Beth and Malka Rappaport Hovav. 2008. Lexicalized manner and result are in complementary distribution. Talk given at the 24th Meeting of the Israeli Association for Theoretical Linguistics, The Hebrew University of Jerusalem.

Levin, Beth and Malka Rappaport Hovav. 2013. Lexicalized meaning and manner/result complementarity. In B. Arsenijević, ed., *Studies in the Composition and Decomposition of Event Predicates*, vol. 93 of *Studies in Linguistics and Philosophy*. Dordrecht: Springer.

Li, Yafei. 1993. Structural head and aspectuality. *Language* 69(3):480–504.

Mohanan, Tara. 1994. *Argument Structure in Hindi*. Dissertations in Linguistics. Stanford: CSLI Publications.

Narasimhan, Bhuvana. 2003. Motion events and the lexicon: a case study of Hindi. *Lingua* 113:123–160.

Ogie, Ota. 2003. About multi-verb constructions in edo. In D. Beermann and L. Hellan, eds., *In Proceedings of the Workshop on Multi-Verb Constructions, Trondheim Summer School 2003*.

Rappaport Hovav, Malka. 2008. Lexical meaning and the internal temporal structure of events. In S. Rothstein, ed., *Theoretical and crosslinguistic approaches to the semantics of aspect*, pages 13–42. Amsterdam: John Benjamins.

Scott, Biljana. 1996. *Aspectogenesis and the Categorisation of Directionals in Chinese*. Ph.D. thesis, University of Oxford.

Sebba, Mark. 1987. *The Syntax of Serial Verbs*. Amsterdam: John Benjamins.

Slobin, Dan. I. 2004. The many ways to search for a frog: Linguistic typology and the expression of motion events. In S. Strömqvist and L. Verhoeven, eds., *Relating Events in Narrative: Vol. 2. Typological and Contextual Perspectives*, pages 219–257. Mahwah: Lawrence Erlbaum Associates.

Slobin, Dan. I. 2005. Linguistic representations of motion events: What is signifier and what is signified? In C. Maeder, O. Fischer, and W. Herlofksy, eds., *Iconicity Inside Out: Iconicity in Language and Literature*, vol. 4, pages 307–322. Amsterdam: John Benjamins.

Talmy, Leonard. 1985. Lexicalization patterns: semantic structure in lexical forms. In T. Shopen, ed., *Language Typology and Syntactic Description: Grammatical Categories and the Lexicon*, vol. 3, pages 57–149. Cambridge: Cambridge University Press.

Talmy, Leonard. 1991. Path to realization – via aspect and result. In *Berkeley Linguistics Society (BLS)*, vol. 17, pages 480–519.

Wechsler, Stephen M. 2003. Serial verbs and serial motion. In D. Beermann and L. Hellan, eds., *Proceedings of the Workshop on Multi-Verb constructions*, pages 1–27.

Williamson, Kay. 1965. *A Grammar of the Kolokuma Dialect of Ịjọ*. Cambridge: Cambridge University Press.

Zubizarreta, Maria Luisa and Eunjeong Oh. 2007. *On the Syntactic Composition of Manner and Motion*. Cambridge: The MIT Press.

Review of *Morphosyntactic Categories and the Expression of Possession*

Sebastian Sulger, *University of Konstanz*

Received Date September 2015; Revised December 2015

Bibliographic information:
Morphosyntactic Categories and the Expression of Possession. Edited by Kersti Börjars, David Denison and Alan K. Scott. Amsterdam: John Benjamins Publishing Company. 2013.

1 Introduction

The book *Morphosyntactic Categories and the Expression of Possession* collects 11 papers that were originally presented at the workshop *Morphosyntactic Categories and the Expression of Possession*, which took place in Manchester on the 3rd and 4th of April 2009. The theme of the workshop, namely the realization of the concept of *possession* (across languages, with a focus on English) through various morphosyntactic constructions, has long been a challenge to linguistic theories. The papers all investigate aspects of the morphosyntactic marking of possession from the perspective of a variety of linguistic theories. Since in all of the surveyed languages there are different constructions available for realizing possession, particular attention is paid to the distribution of the relevant constructions, using corpus data and statistical analysis.

2 Summary

In the *Introduction*, **Kersti Börjars, David Denison and Alan Scott** set the stage for the volume by touching briefly upon the historical and theoretical implications of the construction that takes centre stage throughout the volume: the English genitive *s*. They also mention its usage in different constructions, all of which to be discussed in the volume (e.g., the group genitive vs. split genitive). They further note that this marker has received much attention in the literature, as it provides a window on a range of issues that influence the way we think about the architecture of grammar. They sketch the development of the marker as well as its theoretical treatments over time — as regards the latter, they draw attention to the literature on clitics vs. phrasal (edge) affixes. Moreover, the introduction mentions that most languages referred to in the volume have more than one way of expressing possession, noting in particular West Flemish and Urdu; here, it becomes clear that while the focus of the book is clearly on English (including its earlier stages), other languages are taken into consideration as well.

The first paper in the volume, by **Cynthia L. Allen**, *Dealing with postmodified possessors in early English — Split and group genitives*, takes a diachronic look at two post-modified possessor constructions in the development of written English: the group genitive construction, as in (1a), on the one hand, and the split genitive construction, as in (1b), on the other hand. Allen states that

in general, postmodified possessors bring about two conflicts with principles of English grammar (and other Germanic languages) — the first principle states that it is the head (i.e., not some complement) of the possessor phrase which should receive the possessive marking, while the second principle states that the possessive marking should be adjacent to the possessum (i.e., at the right edge of the possessor phrase). Group genitives such as (1a) violate the first principle; split genitives such as (1b) violate the second principle.

(1) a. the king of France's daughter
b. the king's daughter of France

Through careful analysis of a host of corpora, Allen documents the rise of the group genitive in the late Middle English period, and further the quite sudden favouring of the group genitive over the split genitive, which took place near the start of the Early Modern English period (but see below for a different angle on the grammatical availability of the split genitive to speakers of English). For the period where both strategies were available (Allen refers to that period as *e2, 1570-1639*), she provides evidence for the claim that the complexity of the involved possessor phrases plays a role: it turns out that the group genitive has always been used predominantly where possessor phrases were maximally simple and involved only the possessor N premodified by a determiner or a possessive and postmodified by the simplest possible PP. The split genitive, on the other hand, was found where possessor phrases had more premodifying material than just a determiner (i.e., titles, adjectives, etc.). Allen plausibly ties this to processing factors: the group genitive was favoured when the possessor was short and simple and thus did not involve much effort to create or parse the resulting structure.

The second paper, *Variation in the form and function of the possessive morpheme in Late Middle and Early Modern English* by **Teo Juvonen**, in a way supplements the preceding paper by Allen as it surveys the use of different strategies in the morphological marking of the possessive in a corpus of about 900.000 words of Late Middle and Early Modern English. The corpus comprises sermons, historical writings and letters. The morphological strategies discussed by Juvonen are: *s*-ending as in (2a), *s*-less ending as in (2b) and separated genitive as in (2c). Unfortunately, the examples in the paper do not include full glosses (i.e., the morphosyntactic glossing as well as the translations are missing); for most of the Early Modern English data this is not a severe shortcoming, but some of the Middle English data are in part hardly intelligible. It is for this reason that the examples below do not contain glosses and translations.

(2) a. the Kynges brother
b. wyth onye of my maister councell.
c. to yower worly worschyppe and herte ys desyre.

Juvonen also discusses the different morphosyntactic contexts that the genitive occurs in, e.g., specifying genitives as in (2a), descriptive genitives as in (3a), locative genitives as in (3b), among others. He then examines the different text types (sermons, historical writings, letters) with respect to these morphosyntactic contexts and finds that the possessive appears in the same structures as in present-day English; he also notes that the three mentioned genres show significant differences in the distribution and that private letters make the most varied use of the possessive. Naturally, the genres also show differences with respect to the possessor type, which Juvonen classifies into the categories *Human*, *Your*, *King*, *God*, *Collective*, *Animal* and *Inanimate*. The increased use of the categories *Your* in letters and *God* in sermons is expected, for example.

(3) a. an yryn rammys horne
b. shee should com home, eyther hither or to her fathers

The rest of the paper focuses on the different possessive encoding strategies, and how they were used across the different genres. One important finding from this section of the paper forms a counter-argument against the claim that the separated genitive constitutes an intermediate stage between the Old and Middle English inflectional genitive and the present-day possessive. The claim entails that the separated genitive marker is in fact associated with the third person possessive pronoun;

however, as Juvonen shows, most writers used two different words for the separated genitive marker (*ys*) and the third person possessive pronoun (*his*), thus presenting evidence against the original claim. Juvonen also includes a discussion of the placement of the possessive morpheme in complex possessors, which is central in determining its morphosyntax; unfortunately, the discussion suffers from the data sparsity (as complex possessors are infrequent in the material used) and stays inconclusive as individual writers exhibit much variation in the morphological marking of split vs. group genitives. What is affirmed by Juvonen is the fact that the group genitive had become the dominant form by the end of the 15th century (as discussed by Allen); he gives the split genitive example in (4), and mentions that by the latter half of the 17th century there seems to be increasing uncertainty about the status of the *of*-phrase: does it modify the possessor *a Bishops* or the possessum *son*?

(4) it was Gerelius a Bishops son of Suedeland

Based on this observation, Juvonen suggests that the group genitive replaced the split genitive because of grammaticalization and semantic bleaching of the preposition *of*: he claims that this is what caused the *of*-phrase to be linked more closely to its head, thus disabling the split genitive.

The third paper in the collection titled *The great regression* by **Benedikt Szmrecsanyi** again looks at possessive encoding strategies in English, this time focusing on the *s*-genitive/*of* -genitive alternation in Late Modern English news texts as in (5a) vs. (5b). By using the ARCHER corpus (more specifically: the British English news texts section of the ARCHER corpus), the paper first documents the collapse in the frequencies of the *s*-genitive in the early 19th century, and its subsequent recovery. The goal of the paper is to explain the v-shaped pattern in the variability of the strategies by taking into account different conditioning factors.

(5) a. the president's speech
    b. the speech of the president

Szmrecsanyi views any subtle changes in the conditioning factors as evidence of a change in the genitive choice grammar of English. The study thus explores the degree to which syntactic change manifests by changing the weights of the following language-internal factors: genitive relation, possessor animacy, possessor length, possessum length, possessor phonology, possessor givenness, possessor thematicity, lexical density. In addition, three language-external factors are considered: corpus file, ARCHER time slice and year. The considered portion of ARCHER was annotated (in part automatically using Perl scripts, in part manually) for these factors; the paper discusses the factors' relevance in some detail. Some of this annotation is done in a rather simplistic way; e.g., when annotating possessor givenness, the annotation considers a possessor instance as given if the same word had already occurred in a local discourse context of 50 words prior to the instance. This excludes, e.g., synonyms, renaming, etc.

Szmrecsanyi mentions the problem that by only looking at the frequency distributions of said factors, nothing can be said about the actual changes in the genitive grammar of English, since the input frequencies could also have changed; e.g., there are more institutional (i.e., inanimate) possessors in Late Modern English texts than in Early Modern English ones. The study thus explores whether the genitive frequency changes are merely a function of variable input frequencies (and these are in fact responsible for the v-shaped pattern observed). To do so, a linear regression model is fit that tries to predict a dependent variable (the *s*-genitive proportions) from several independent variables (e.g., *possessor animacy*). The minimal model contains three independent variables: mean possessum length, percentage of human possessors, percentage of ownership relations; Szmrecsanyi explains that the model accurately predicts the decrease in *s*-genitive frequencies in the first half of the 19th century and their recovery, but fails to explain the extent of the decrease as well as the extent of the recovery (in fact surpassing the *of*-genitive).

Therefore, a second model was built using mixed-effects binary logistic regression analysis. The aim here was to bring time into the equation — i.e., to examine whether human possessors favoured the *s*-genitive more strongly in the 17th century than in the 20th century. The dependent variables no longer are the aggregated *s*-genitive frequencies, but individual coding decisions (*s*-genitive vs. *of*-

genitive) for a given gentitive instance in a given corpus text. The minimal model has a predictive accuracy of 96% (i.e., 96% of the genitive occurrences' coding strategies can be correctly predicted by the model) and shows that interaction between the variables *ARCHER time slice* on the one hand and *genitive relation*, *possessor animacy*, *possessum length* and *possessor thematicity* on the other hand accounts for most of the variability. This allows Szmrecsanyi to draw a number of interesting conclusions with respect to genitive grammar change over time. For example, in the first ARCHER period, if there was an *ownership* genitive relation to be encoded, the odds in favour of an *s*-genitive increased by a factor of 19, while in the last period, the corresponding value is 243. Another conclusion drawn by Szmrecsanyi explains why the slump in the *s*-genitive frequencies around the first half of the 19th century was so much more severe than expected: this has to do with a change in the status of the factor *possessor animacy*. While the sheer input frequencies of human possessors dropped, the animacy constraint (due to which human/animate possessors favour the *s*-genitive) was relaxed — an actual change in the grammar of genitive choice speeding up the decrease in *s*-genitives.

How does all of this help in understanding the morphosyntactic status of the *s*-genitive across the periods? Szmrecsanyi, by employing Lehmann's *paradigmatic variability* parameter (Lehmann 1995), notes that in grammaticalization processes, it is common that "expressions for human concepts come to be used also for concepts that are inanimate", citing Heine (1997). Thus, dropping the selectional restriction concerning animacy in 19th century news texts can be seen as a sign of grammaticalization. Conversely, the increasing sensitivity of the *s*-genitive towards the factors *possessor thematicity* and *possessum length* can be seen as a development towards a freer "choice of items according to communicative intents" (Lehmann 1995, 164) and thus as degrammaticalization. The fact that the *s*-genitive increasingly attracted ownership relations can be explained using the *paradigmatic integrity* parameter, again by Lehmann (1995): grammaticalization involves semantic bleaching until only grammatical features remain; here, however, we find the opposite process, where lexical-semantic features (i.e., ownership) are in fact added to the *s*-genitive. Thus we again find degrammaticalization.

The fourth paper by **Catherine O'Connor, Joan Maling and Barbora Skarabela** titled *Nominal categories and the expression of possession — A cross-linguistic study of probabilistic tendencies and categorical constraints* is the first paper in the collection that moves beyond English in presenting a cross-linguistic study of what the authors call the Monolexemic Possessor Construction (MLP). The aim of the paper is to compare the stochastic patterns of prenominal possessives in English to the MLP found in a variety of languages (Germanic, Slavic, Romance). The choice of a prenominal possessive over a postnominal one in English correlates strongly with the features animacy, weight and discourse status, but is not categorical in nature (i.e., in a given context, a certain coding choice may be more probable than the other). This is not the case with the MLP; in this construction, the possessor occurs immediately left to the possessum, and the possessor may not be longer than a single word, as in the Czech examples in (6a–a'). If the possessor is to be expressed as a full phrase, the MLP is ungrammatical (6b), and the full phrase adnominal genitive has to be used as in (6c).

(6) a. Milan-ova kniha
Milan-POSS.ADJ book
'Milan's book'

a'. Kunder-ova kniha
Kundera-POSS.ADJ book
'Kundera's book'

b. * Koupila jsem Milan-ovu Kunder-ovu
buy.PAST.1.SG.FEM be.PRES.1SG Milan-POSS.ADJ.ACC Kundera-POSS.ADJ.ACC
knih-u
book-ACC
'I bought Milan Kundera's book.'

c. Koupila jsem knih-u Milan-a Kunder-y
buy.past.1.sg.fem be.pres.1sg book-acc Milan-gen Kundera-gen
'I bought (a/the) book of Milan Kundera.'

The paper disusses the stochastic generalization put forward by e.g., Bresnan et al. (2001) which asserts that statistically noticeable but noncategorical patterns found in one language (e.g., the genitive alternation in English) often correlate with categorical and inviolable patterns in other languages (e.g., the MLP). The focus of the paper is to show how the very same values of the animacy, weight and discourse status features involved in the English pattern also account for the variation in the MLP/non-MLP pattern, but in a categorical manner. Evidence from Czech, Russian, Icelandic, German etc. is adduced showing that optimal weight, discourse status and (less strictly) animacy are all grammaticalized in the MLP. With respect to discourse status, a cross-linguistically valid accessibility hierarchy ermerges as in (7). If a language has an MLP, it will (always) allow it with pronouns (which are the most accessible elements in discourse); if a language allows it with e.g., kinship terms, it will also allow it with any element occurring to the left of the kinship terms in (7). It would be interesting to see how this typological prediction holds up when looking at an even larger language sample.

(7) *Monolexemic Possessor Accessibility Hierarchy:*
Pronoun ⟩⟩ Proper Noun ⟩⟩ Kinship Term ⟩⟩ Common Noun
Most accessible ⟨——————————⟩ Least accessible

The scale in (7) implies that in a given context, pragmatic decisions must take place to resolve the possessor in an MLP; in particular, the question arises how a possessor is resolved if there are e.g., multiple pronouns available. To address this question, the paper further includes a discussion of whether the categorical restriction is at work in terms of pragmatic communicative decisions, or whether it just constitutes a frozen remnant of the stochastic tendencies observed e.g., in English; by citing elicitation experiments with native speakers of Czech, the authors confirm that the categorical restrictions reflect an active discourse pragmatic requirement.

In the fifth paper, *Expression of possession in English — The significance of the right edge* by **Kersti Börjars, David Denison, Grzegorz Krajewski and Alan Scott**, returns to the topic of *s*-genitive/*of*-genitive alternation in English. The focus of the paper is on the categorization of the *s*-genitive as a clitic or an affix. Analyzing spoken data from the British National Corpus, the authors are especially interested in the right edge criterion, which is key evidence for the *s*-genitive's status as a clitic: the item's ability to occur at the right edge even in cases where the possessor is postmodified, as in e.g., (8).

(8) the man in the car's wallet

The authors conduct a regression analysis based on non-structural factors (animacy, discourse structure, phonology etc.) again confirming the findings of the papers discussed above and of others such as Rosenbach (2002): the animacy of the possessor is the dominant factor in the choice of the possessive. The study then turns to structural factors and discusses a regression analysis examining the influence of weight in the construction choice; the authors mention that while the factor weight is discussed in several studies (cf. the paper by O'Connor, Maling and Skarabela discussed above), fewer studies distinguish as to how the weight is distributed in the phrase. Weight is expected to play a role in choosing the preferred construction since language has the general tendency to put longer constituents at the end of the phrase.

The crucial question for the authors in deciding about the morphosyntactic status of the *s*-genitive then is: what happens if the possessor is postmodified; i.e. if the *s*-genitive does not appear on the head of the possessor phrase, as in (8)? The authors discuss two new variables, length of premodifying sequence as well as length of postmodifying sequence, for checking whether it makes a difference where the weight of the possessor is located: before or after the head. It turns out that the effect of premodification is weaker than of postmodification, so that the latter decreases the odds of the

s-genitive more strongly (unfortunately, the data the authors work with is too sparse to examine any further the effects of the actual length of the postmodification). The so-called split possessive (also discussed by Allen in the first paper of the volume) then is argued to be a strategy for avoiding standard *s*-genitives where the possessor contains postmodification, and the data shows a clear correlation between the presence of a split and the length of the postmodification. What is less clear from their account is how the choice of the *escape hatch* (*of*-genitive vs. split genitive) is driven: is this also weight-based, and does the distribution of weight matter in the choice?

For Börjars, Denison, Krajewski and Scott, the evidence adduced is not in favor of a simple dichotomy *affix* vs. *clitic*; in fact, they argue that this is an oversimplification that does not do justice to the mixed properties of the English *s*-genitive and the corresponding Swedish construction examined by Börjars (2003), and propose a spectrum of categories with an *affix end* and a *clitic end*. It sounds reasonable to do away with notions such as *clitic-like affix* and instead recognize the fact that in morphosyntactic reality, things tend to get more complicated than that and certain items display non-uniform behavior.

The sixth paper, *A cognitive analysis of John's hat* by **Richard Hudson**, takes an entirely different approach and presents a cognitive analysis of the English *s*-genitive, couched within Hudson's Word Grammar framework (Hudson 2010). The general assumption here is that a string such as *John's hat* in fact spawns two different syntactic analyses in the mind of a speaker of English. Under the first analysis, the morpheme {*z*} (*morph* in Hudson's terminology) behaves like a suffix and is a direct descendant of the Old English inflected genitive case; here, the string *John's* behaves like a single word which doubles in function as a determiner. Under the second analysis, the same morpheme {*z*} behaves like a clitic (a separate syntactic word POSS, which in Hudsons's view is a special type of determiner and is realized as a suffix) giving rise to the group genitive.

Hudson claims that each of these analyses has advantages and disadvantages for a learner of English. The suffix analysis involves a straightforward morphology/syntax mapping, but the possessor phrase must receive a complex analysis, doubling in syntactic classification as a (possessive) pronoun and a noun (common or proper); it also involves a separate referent in the semantics. Hudson acknowledges that this analysis seems more intuitive in cases where we have simple (e.g., proper noun/one word) possessors, like *John's hat*. On the other hand, when the possessor is complex, and the *s*-genitive is not adjacent to the possessor phrase head (e.g., the possessor involves postmodification), the group genitive is the only analysis available. Here, Hudson states a simple mapping at the syntax-semantics and morphology-syntax interfaces, at the cost of the special morphology involving a clitic.

Hudson thus assumes that in any example, {*z*} will spawn one of the analyses above; everywhere it occurs, it will be either an affix or a clitic. Compare this to the analysis by Börjars, Denison, Krajewski and Scott where the *s*-genitive will only receive a single analysis, retaining all of its morphosyntactic properties in a single entry. Hudson further discusses the construction which is in competition with the *s*-genitive, the *of*-genitive, and argues (like others above) that the variation is due to a processing effect: people prefer the *s*-genitive with short possessors, since they put the *landmark* relation (a semantic relation in Word Grammar terminology) first. If the distance between the head of the possessor and the possessum becomes too large, processing benefits dictate the *of*-genitive. Hudson's account lacks a discussion of the split genitive; it is not evident how this construction can be handled under his account, since the *s*-genitive in such cases seems to display mixed properties (adjacent to possessor head *and* taking scope over the whole possessor phrase).

John Payne's paper *The oblique genitive in English* deals with the construction in (9), which is known as the *oblique genitive* (OG below) and also commonly referred to as the *double genitive*. Payne notes that the construction has previously been analyzed as a variant of the *s*-genitive (e.g., *the Prime Minister's friend*; Payne refers to this as the subject-determiner construction), as a variant of the *of*-genitive (e.g., *a friend of the Prime Minister's*; Payne calls this the *of*-oblique) as well as an equivalent of the partitive (e.g., *one of the Prime Minister's friends*).

(9) a friend of the Prime Minister's

Payne proceeds to compare the OG to all of these correspondents in turn. In short, the OG is much more semantically restricted than the *s*-genitive and involves a quite different pattern in the selection of determiners; the *of*-genitive does not quite stand in complementary distribution with the OG either, and patterns differently with respect to weight; and finally, the partitive always involves anti-uniqueness, while the OG does not always do so, contrary to what is claimed by Barker (1998); see example (10).

(10) that nose of his

The paper comes up with a separate analysis for the OG, carried out on data from the British National Corpus. The analysis presents the insight that the semantic relations realized by the OG in fact constitute a subset of the relations that are possible with the *s*-genitive (as noted by other researchers, such as Barker (1998) and Chomsky (1970), and provides attested examples for 13 different thematic relations occurring with the OG. A further insight from Payne's data analysis is that there are many cases where the OG occurs with a definite article, some of which present clear evidence against the anti-uniqueness claim by Barker (1998).

The question then is brought up what drives the choice between the OG and the *s*-genitive in these cases, and Payne suggests that this choice is largely a matter of information structure: in the *s*-genitive, the referent is identified by first processing the genitive NP, which provides an *anchor* (a term coined by Fraurud 1990) for the identification, while in the OG, the function of that genitive NP anchor is reduced, and processing happens largely by contextual anchors. An analysis along the lines of Abel (2006) involving focus is challenged by Payne; however, since his counterexample involves contrastive stress on the possessor, it does not necessarily speak against Abel's account.

In the eighth paper, *The marker of the English "Group Genitive" is a special clitic, not an inflection* by **Stephen R. Anderson**, the author develops a formal account of the possessive marker s in English. Basically, the paper paints the picture already presented in Anderson (2005), and makes a bigger theoretical point in that Anderson claims that accounting correctly for the English group genitive (EGG below) can tell a general story about various grammatical categories.

Anderson establishes the feature [POSS] (realized by the *s*-genitive) as a feature which is marked on the phrasal level (in his view: on a possessor DP residing in the specifier position of a higher DP), then discusses two quite different accounts of phrasal properties. One is the account that he is in favor of, namely to treat the group genitive as a *special clitic* (Zwicky 1977). Under this account, rules modify the phonological makeup of phrases by introducing what Anderson refers to as "affix-like phonological content" (i.e., clitics or particles) at a certain point within the phrase. These rules are technically not different from morphological rules at the word level. The EGG marker, then, is such a special clitic, introduced at the right edge of a (possessor) DP in a certain structural configuration. The clitic bears the feature [POSS] which is percolated up to the possessor DP via one of these rules.

The other account called "EDGE inflection", as put forward by e.g., Nevis (1986) and Zwicky (1987), treats the group genitive as a special inflectional pattern applied at the edges of words. Here, certain features are identified as EDGE features of a phrase, and transmitted via the daughters of that phrase until they land on a terminal node, where they are realized through inflection. To explain the EGG, this account would involve a feature [POSS] which is identified as of type EDGE: LAST. The feature would then successively walk through the rightmost daughters of the DP tree, and being an EDGE: LAST feature, it is realized only when it lands on the rightmost terminal node.

Anderson mentions that both accounts produce the right facts for the EGG, but the theoretical implications and mechanisms are different: One involves a clitic as a single marker of the [POSS] feature at the edge of the phrase, the other realizes the feature (through intermediate constituents) on a single grammatical word, as an affix. The author then proceeds to discuss some general properties of clitics and affixes, then demonstrates that there are rather clear cases where one analysis is favorable over the other: Heiltsuk clearly makes use of independent clitics for marking determiners, while it seems clear that Australian Kuuk Thaayorre uses edge inflection affixes to mark the ergative. From these examples, Anderson establishes three diagnostics for distinguishing clitics from edge

inflection: selection of certain parts of speech is more likely to apply to affixes; lexical gaps as well as idiosyncratic shapes are more likely to occur with affixes. The EGG shows prototypical properties of clitics, with the exception of the idiosyncratic shapes of possessive pronouns (*my*, *your*, *mine*, *yours* etc.). Anderson accounts for this by assuming that pronouns are special determiners that do not lexicalize just a single terminal node, but an entire DP; with possessive pronouns, this lexicalized DP is assigned the feature [POSS].

Phonologically, Anderson argues that the possessive */z/* is in fact adjoined to the final syllable (instead of being incorporated into it). Here, possessive */z/* is no different from plural */z/*, which is also adjoined. This way, Anderson can nicely account for the data in (11a) vs. (11b), e.g., by saying that two instances of adjoined */z/* are collapsed into one in (11a) (thus in fact modeling the *Haplology Criterion* discussed by e.g., Zwicky 1987), while in the examples in (11b) we only have a single instance of adjoined */z/*; the other one is in fact part of the base syllable.

(11) a. anyone who likes kids' (*kids's) ideas
b. the fuzz's old cars; at Buzz's

Liliane Haegeman discusses two kinds of prenominal possessor patterns in West Flemish, a dialect of Dutch, in the twelfth article titled *Two prenominal possessors in West Flemish*. The paper is of a descriptive nature, and the main focus is to show that, while several other papers propose a unitary account for the two patterns, they actually show different syntactic features and thus cannot have an identical syntax. The first pattern is shown in (12a), referred to by Haegeman as the doubling construction (DC below); the second pattern, called the *sen construction* (SC), is shown in (12b). In (12a), the DP possessor *Valère* is doubled by the possessive pronoun *zenen*, and the latter can also occur on its own; Haegeman mentions that when this is the case, as in (12c), the properties of the pronoun are the same as in the DC.

(12) a. (Valère) zen-en hoed
(Valère) his-MSG hat
'Valère's hat'
b. Valère sen hoed
Valère sen hoed
'Valère's hat'
c. zen-en hoed
his-MSG hat
'his hat'

It is important to note that the possessive pronoun displays double agreement, matching both the possessor (person, gender in the singular, number) as well as the possessum (gender, number). In the SC construction, *sen* does not agree with either; the only restriction imposed by *sen* is that the possessor DP's head be singular. The DC and SC pattern alike with respect to constituency, thematic relations, possessor properties (semantic and syntactic), recursion as well as definiteness. However, the author presents abundant evidence against a unifying approach to the two constructions, including the mentioned agreement patterns, reciprocal possessors, as well as adjacency effects. The latter are discussed by Haegeman at some length: while in the SC, the prenominal possessor DP always occurs adjacent to *sen*, in the DC the possessor may occur in at least two different positions in the containing DP. These patterns can be observed with quantifier phrases, appositives, discourse particles, and others, and Haegeman gives examples of all of these. (13) illustrates the obligatory adjacency for the SC, while (14) shows the possible patterns for the DC; both of these examples involve quantifier phrases.

(13) a. K'ee-n [al [Marie sen boek-en]] gezien.
I have-1SG all Marie sen book-PL see-PTCP
'I have seen all Marie's books.'
b. * K'ee-n [Marie al sen boek-en] gezien.

(14) a. K'ee-n [al [Marie eur boek-en]] gezien.
I have-1SG all Marie her book-PL see-PTCP
'I have seen all Marie's books.'

b. K'ee-n [Marie al eur boek-en] gezien.
I have-1SG Marie all her book-PL see-PTCP
'I have seen all Marie's books.'

Haegeman mentions that the possibility of the possessor to move to the left of the quantifier may be influenced by information structure, although she does not develop this point any further. To account for the adjacency effects, she proposes that there are three different prenominal positions for possessor DPs: in the SC, the possessor occupies a specifier position of an IP in the nominal domain (the details of this particular assumption are not evident), while in the DC, the possessor may either occupy the specifier of DP or the specifier of a higher projection (e.g., the specifier of the quantifier phrase in (14b)). As a more general point, Haegeman suggests that it is because of the agreement pattern in the DC that the possessor may be non-adjacent, which seems like a reasonable assumption.

In a descriptive paper, entitled *A Mozart sonata and the Palme murder — The structure and uses of proper-name compounds in Swedish*, **Maria Koptjevskaja-Tamm** describes Swedish nominal compounds where the first nominal constitutes a personal proper name (proper name compounds — PNC below). The author asserts that there has been little work on such compounds in the past, but that they are nonetheless an important means of constructing possession (in the broadest sense) across Germanic languages. This is illustrated by the fact that the said compounds are (almost) synonymous with other possessive nominals (using the Swedish equivalents of either the *s*-genitive or the *of*-genitive). The natural question brought up by Koptjevskaja-Tamm is what influences the choice between the constructions, but the author also discusses the similarities and differences between PNCs and common noun compounds (CNCs). An empirical/statistical analysis is missing from Koptjevskaja-Tamm's account, as she focuses on an initial description of the compounds, using mainly Google searches and informants.

Discussing the formal properties of PNCs, Koptjevskaja-Tamm notes that they are structurally more simpler and more complex than CNCs — simpler because morphological processes (such as the German *Fugenelemente* ('linking elements')) are rare inside PNCs, but more complex since the proper noun part may itself be morphologically complex. However, this complexity is argued to be deceiving, as proper nouns are lexically fixed expressions; inflectional morphemes, attributive adjectives and other modification are generally integrated parts of the proper nouns involved. This limitation in the formal makeup constitutes a disadvantage when compared to *s*-genitives and postnominal possessives, where such restrictions do not apply, and plays an important role in the choice between the three constructions, according to Koptjevskaja-Tamm.

In what follows, Koptjevskaja-Tamm provides a detailed discussion of the uses of PNCs in Swedish. She establishes that PNCs may be used as proper names (15a–15b) and common nouns (15c); Koptjevskaja-Tamm calls cases such as (15b) and (15c) "commemorative compounds" as they involve no inherent relation between the referent of the proper noun and the common noun head, while (15a), on the other hand, is non-commemorative as there is a clear or inherent relation between the two (note that the *s* morpheme in (15a) is a *Fugenmorphem* ('linking morpheme') rather than an *s*-genitive). Koptjevskaja-Tamm does not provide clearcut characteristics for distinguishing between the two (probably as there aren't any), and acknowledges that commemorativeness must be a tendency at best.

(15) a. Strindberg-s + muse-et
Strindberg-LNK + museum-DEF.SG.N
'the Strindberg museum' non-commemorative, proper name

b. Kristoffer + skola-n
Christopher + school-DEF.SG.COM
'the Christopher school' commemorative, proper name

c. en Gustav Adolf + bakelse
a Gustav Adolf + cake
'a Gustav Adolf cake' commemorative, common noun

PNCs compete with possessive NPs in Swedish for naming streets, churches and other entities; this even happens for one and the same entity: the hospital that goes by the official name *Astrid Lindgren-s sjukhus* is sometimes also named *Astrid Lindgren+sjukhus-et*, which is a PNC. The author notes that the heaviness of the proper noun might play a role, so that longer proper nouns appear mostly with genitives, while shorter ones appear mostly within PNCs. The connection to processing seems obvious (see also the papers by Allen, Hudson, and Payne), but as the exceptions from the rule are numerous, Koptjevskaja-Tamm notes that this must be regarded as a tendency only. The difference between (16a) and (16b), on the other hand, is rather clear: (16a) refers to the particular moustache that was part of Hitler's face, while the PNC in (16b) specifies a typified moustache of the kind Hitler wore. The difference between identifying particular instances and typified instances is general and productive in Swedish (and other Germanic languages, one might add).

(16) a. Hitler-s mustasch
Hitler-GEN moustache
'Hitler's moustache'

b. Hitler + mustasch
Hitler + moustache
'Hitler moustache'

Koptjevskaja-Tamm further finds that PNCs may be created on the spot for anaphoric reference within a discourse, where the function of the compound is thus purely anaphoric and does not reflect an independent category outside of the particular discourse. Here, PNCs also compete with adnominal possessors, and the author sees the alternation in line with Ariel's accessibility theory for noun phrase antecedents (Ariel 1990). Regarding structural simplification, the author explains that a large proportion of PNCs provide concise and salient labels for entities that would otherwise need long descriptions, and therefore lack any direct correspondence with adnominal possessives. Morphosyntactically, PNCs are often chosen over their possessive counterparts where the referent is indefinite (since the possessive NPs normally have a definite interpretation).

Koptjevskaja-Tamm tries to approach two theoretical questions at the end of the paper: 1) whether the issue of (non-)referentiality is relevant for the occurrence of proper names within compounds, and 2) whether the distinction between instance specification and type specification is relevant for choosing between PNCs and the corresponding *s*-genitives. Regarding the first question, the author notes that referentiality is not a fully testable notion with respect to PNCs, and that anaphor to parts of words (i.e., parts of compounds) is fully grammatical, but governed by independent pragmatic principles. In addition, the accessibility of a proper noun referent inside a PNC will depend also on the type of compound; e.g., in commemorative PNCs, pronouns will probably not be able to refer to back to the proper noun of a PNC. Regarding the second question, Koptjevskaja-Tamm again argues for a differentiated approach, as the distinction between instances and types is not always clearcut either (e.g., the English *a Picasso picture* may describe both). She adds that proper names may conflate types and instances, referring to Langacker (1991), and concludes that instead of treating PNCs as a single construction, it might in fact be more fruitful to split the construction apart into several distinct patterns.

The last paper, *Possessive clitics and ezafe in Urdu* by **Tina Bögel and Miriam Butt**, compares two ways of expressing possession in the Indo-Aryan language Urdu. The first pattern exemplified in (17a) uses the genitive case marker *k-* which, according to the authors, can be analyzed in a straightforward way as a case clitic (in the fashion of Butt and King 2004). The second pattern,

shown in (17b), uses the *ezafe*, a loan construction from Persian. The paper discusses the formal properties and syntactic distributions of both constructions and provides analyses couched within Lexical-Functional Grammar (LFG). In the discussion of the grammatical framework, the authors pay special attention to the projection architecture of LFG, where different modules of grammar are interrelated through well-defined projections. This point is of central importance for the alignment of prosody and syntax with respect to clitics.

(17) a. yasin=ki gari
Yassin.M.SG=GEN.F.SG car.F.SG
'Yassin's car'

b. sahib=e takht
owner.M.SG=EZ throne.M.SG
'the owner of the throne'

Focusing on the genitive case marker, the authors note that it may be used for several different functions: kinship, subjects of verbal nouns, objects (of actions, feelings, notions), as well as possession and the description of properties. They establish that the range of uses of the genitive does not vary much from languages such as English; all the uses reflect an abstract sense of possession and express a more or less clearly-defined relationship between two entities. Regarding the genitive's morphosyntax, they find that it agrees in gender and number with the head noun, a unique property among the case markers of Urdu. Otherwise, they state that it adheres to the general head-final syntax of Urdu, and present conclusive evidence for its status as a clitic stemming from coordination and clitic inclusion. In the LFG analysis, the genitive case clitic is the functional head of a possessive case phrase, which ends up as a possessive specifier in the functional structure. The prosodic alignment of the genitive is straightforward: the case clitic integrates with the possessor (*yasin* in (17a)) under a single prosodic node, as this is its prosodic host; syntax and prosody are thus aligned.

The authors then turn to the *ezafe* construction, mentioning that it has been extensively discussed by other papers (e.g., Samvelian 2007), some of which identify it as a clitic, while others see it as an affix and a part of nominal morphology. In Persian as well as in Urdu, the *ezafe* construction displays a head-initial pattern, and modifiers appear to the right, which is exceptional in both languages. Moreover, the *ezafe* always prosodically forms a unit with the head noun to its left, while at the same time licensing modifiers to the right; syntactic function and prosodic realization thus differ. Bögel and Butt discuss the account of Samvelian (2007) in some detail, who argues that the Persian *ezafe* is a phrasal affix. Unlike Anderson (this volume, 2005), however, who refers to phrasal affixes as *special clitics* using the terminology of Zwicky (1977), Samvelian analyzes *ezafe* as part of word-level morphology, and not as being introduced post-lexically; Samvelian's main evidence comes from other phrasal affixes which seem to be in complementary distribution with the *ezafe*, thus they must be generated on the same level by the *Haplology Criterion* (see, e.g., Zwicky 1987). Bögel and Butt challenge Samvelian's morphological account and argue that different groups of phrasal affixes might belong to different classes, and that the *Haplology Criterion* must not hold in the morphological component, but may in fact apply in the phonological/prosodic part of the grammar (see Anderson's paper in this volume for an example involving syllables).

Bögel and Butt's own account of the (Urdu) *ezafe* involves ample evidence that the marker behaves like a clitic in many respects: e.g., it is separable from its host using parentheticals, it can take scope over noun conjunction (see (18) for an example), it does not display morphophonological idiosyncracies. On the other hand, it also has some non-clitic-like properties; for example, it displays a high degree of lexical selection, as it only occurs with nouns of Persian origin as its head. All in all though, the authors conclude that Urdu *ezafe* should be analyzed as a clitic (a phrasal affix in the original, post-lexical sense of e.g., Anderson 2005). With respect to the semantics, *ezafe*, just like the genitive, is not restricted to possession but shows a wider range of usage; although the authors mention that Urdu *ezafe* is mainly found in high literature and attaches to Persian nouns only, a more detailed discussion of the choice between the genitive and *ezafe* is not included in the paper.

(18) [ye [mal o daulat]=e dunya
this material and wealth=EZ world
'this material and wealth of the world'

Their own LFG analysis involves separate modules of grammar, taking into account the misalignment within the *ezafe* construction: while it is a functional head selecting a modifier to its right, prosodically it attaches to the word on its left. The first (syntactic) property is modeled via LFG's standard c(onstituent)-structure and f(unctional)-structure, while the second (prosodic) property is modeled on a separate p(rosodic)-structure; nothing in the LFG architecture requires these structures to align. The final prosodic analysis by Bögel and Butt for examples such as the one in (19 is as given in (20); the authors emphasize that the *ezafe* is analyzed as an independent functional item in the lexicon, whose prosodic properties are special and must be dealt with post-lexically. Their account is thus remarkably similar to Anderson's *special clitic* account of the English group genitive in this volume.

(19) hukumat=e pakistan
government.M.SG=EZ pakistan
'the government of Pakistan'

(20) ((hukumat)ω e)ω (pakistan)ω)ϕ

3 Evaluation

Shortcomings of the volume as a whole are of a technical nature. Some examples in some of the papers lack glosses. The numbering of the examples is also off in some cases. In addition, some cited references are not included in the bibliography at the end of the volume (I have found at least four such instances across all the papers).

There is also some unclarity in the volume regarding the terminology of clitics. Anderson uses the term *special clitic* for the English group genitive in the sense coined by Zwicky (1977). Special clitics are clitics which are special in their morphosyntax (when compared to the regular syntax of the particular language); they are morphemes bound to and existing as part of their host, and thus show similarities with morphological affixes. This distinguishes them from other, *simple clitics* in the sense that these are free and independent morphemes. Both simple and special clitics are, however, prosodically dependent to adjacent material. Anderson in his earlier work has coined the term *phrasal affix* (Anderson 1992) which turns out to be equivalent to Zwicky's *special clitics*. The intention behind this term was to point out the parallelism of distribution and function between morphological affixes and clitics. Anderson seems, however, to use Zwicky's term presently, which is why the term *phrasal affixes* does not feature in his paper in this volume. Bögel & Butt in their paper, referring to Anderson's work, use his earlier terminology and talk about *phrasal affixes*. To complete the confusing picture, Samvelian (2007) in her discussion of the Persian *ezafe* considers phrasal affixes to be part of the morphology and treats them essentially as EDGE affixes in the sense of Zwicky (1987) and Miller (1992).

The paper by Börjars et al. may in part provide the answer for such issues. Anderson as well as Bögel and Butt in their papers acknowledge that the markers they analyze (English group genitive and Urdu *ezafe*) display mixed properties of affixes and clitics, but both papers proceed to analyze the markers as clitics. A dichotomy *affix* vs. *clitic* may in fact turn out to be an oversimplification that does not do justice to the mixed properties of such items, and Börjars et al. instead suggest a scale of grammatical categories with a *clitic end* and an *affix end*. While this is an interesting proposal, the exact makeup of the proposed scale is left for further research.

The volume provides an interesting perspective on possessive alternations, which is the key theme of several papers (Börjars et al., Bögel and Butt, Szmrecsanyi, Haegeman, O'Connor et al., Allen, Juvonen). Throughout the paper, the features animacy, weight and topicality/discourse status crop up, and clear correlations are established between these features and the choice of a particular possessive encoding strategy. The question arises whether these tendencies form part of the grammar,

or whether they belong in a separate component capturing language use. O'Connor et al. answer this by looking at languages where those factors are implicated in categorical distinctions between separate constructions, and thus clearly form part of the grammar. It can therefore be concluded that the statistical patterns displayed e.g., by English actually form part of the grammar and need to be represented in a model of grammar. A question I would add is in how far the features animacy, weight and topicality hold up in a cross-linguistic study of possessive patterns, or whether there are more features that involve categorical distinctions and/or statistical preferences in other languages; this is, however, a question that must be resolved in a typological study.

All in all, the volume is essential reading for any linguist interested in the morphosyntactic realization of possession, and in fact presents the state of the art with respect to possession marking in Germanic languages and beyond. While the overall focus is clearly on the English language, this is not necessarily a hurtful restriction: it enables the volume to approach the various issues in English from several distinct angles, while maintaining a manageable set of data. Empirical-statistical (Allen, Juvonen, Szmrecsanyi), cognitive (Hudson) and theoretical-explanatory (Börjars et al., Anderson, Payne) accounts add up to render a rather complete picture of the English possessive constructions from a synchronic as well as from a diachronic perspective. In addition, descriptive (O'Connor et al., Haegeman, Koptjevskaja-Tamm) and analytical (Bögel & Butt, O'Connor et al.) papers team up to provide insights into other languages' possessive structures as well.

References

Abel, Jennifer. 2006. That crazy idea of hers: The English double genitive as focus construction. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique* 51(1):1–14.

Anderson, Stephen R. 1992. *A-Morphous Morphology*. Cambridge: Cambridge University Press.

Anderson, Stephen R. 2005. *Aspects of the Theory of Clitics*. Oxford: Oxford University Press.

Ariel, Mira. 1990. *Accessing Noun-Phrase Antecedents*. London: Routledge.

Barker, Chris. 1998. Partitives, double genitives and anti-uniqueness. *Natural Language and Linguistic Theory* 16(4):679–719.

Börjars, Kersti. 2003. Morphological status and (de)grammaticalisation: The Swedish possessive. *Nordic Journal of Linguistics* 26:133–136.

Bresnan, Joan, Shipra Dingare, and Christopher D. Manning. 2001. Soft constraints mirror hard constraints: Voice and person in English and Lummi. In M. Butt and T. H. King, eds., *Proceedings of the LFG01 Conference*, pages 13–32. Stanford, CA: CSLI Publications.

Butt, Miriam and Tracy Holloway King. 2004. The Status of Case. In V. Dayal and A. K. Mahajan, eds., *Clause Structure in South Asian Languages*, pages 153–198. Dordrecht: Kluwer.

Chomsky, Noam. 1970. Remarks on nominalization. In R. Jacobs and P. Rosenbaum, eds., *Readings in English Transformational Grammar*. Waltham: Blaisdell.

Fraurud, Kari. 1990. Definiteness and the processing of noun phrases in natural discourse. *Journal of Semantics* 7:395–433.

Heine, Bernd. 1997. *Possession: Cognitive Sources, Forces, and Grammaticalization*. Cambridge: Cambridge University Press.

Hudson, Richard. 2010. *Word Grammar and Cognition*. Cambridge: Cambridge University Press.

Langacker, Ronald. 1991. *Foundations of Cognitive Grammar*, vol. 2. Stanford, CA: Stanford University Press.

Lehmann, Christian. 1995. *Thoughts on Grammaticalization*. München: Lincom.

Miller, Philip. 1992. *Clitics and Constituents in Phrase Structure Grammar*. New York: Garland.

Nevis, Joel A. 1986. *Finnish Particle Clitics and General Clitic Theory*, vol. 33 of *Working Papers in Linguistics*. Department of Linguistics.

Rosenbach, Anette. 2002. *Genitive Variation in English: Conceptual Factors in Synchronic and Diachronic Studies*. Berlin: Mouton de Gruyter.

Samvelian, Pollet. 2007. A (phrasal) affix analysis of the Persian ezafe. *Journal of Linguistics* 43:605–645.

Zwicky, Arnold M. 1977. *On Clitics*. Bloomington: Indiana University Linguistics Club.

Zwicky, Arnold M. 1987. Suppressing the z's. *Journal of Linguistics* 23:133–148.