Experiencers vs. agents in Urdu/Hindi nominalized verbs of perception

Miriam Butt
University of Konstanz

Lucrezia Carnesale
University of Pavia-Bergamo

Tafseer Ahmed
University of Konstanz & Alexa Translations

Proceedings of the LFG'23 Conference

Miriam Butt, Jamie Y. Findlay and Ida Toivonen (Editors)

2023

PubliKon

lfg-proceedings.org

Abstract

Urdu/Hindi displays a curious construction in which a nominalized verb of perception combines with the verb 'give'. As an experiencer predicate, it takes a dative subject; however, there is no other instance in the language in which the subject of 'give' is a dative. Furthermore, the verb 'give' is a three-place predicate, but the N-V experiencer predicate is only two-place. We propose an analysis by which the construction originates in a ditransitive agentive N-V complex predicate whose goal argument is reanalyzed into an experiencer. We propose that the mechansim is similar to that posited by Schätzle (2018) for the rise of dative subjects in Icelandic, where an originally locative predication gave rise to experiencer predicates.

1 Introduction

In this paper, we propose an analysis for a puzzle we identified as part of a larger scale investigation of perception verbs.[†] The puzzle centers around nominalized versions of the verbs *dɪkʰ* 'appear to' and *sʊn* 'hear', illustrated in (1-a) and in (1-b), where they roughly mean 'seeing' and 'hearing', respectively.

(1) a. mujʰ-e is=ka koi lɑkṣɑn nahĩ
Pron.1.Sg-Dat this.Obl=Gen.M.Sg some sign.M.Sg.Nom not
dɪkʰ-a-i de-t-a
appear-Caus-F.Sg give-Impf-M.Sg
'I do not see any sign of it.'

b. un-hẽ gogi=ki mɑhin avaj
Pron.3-Pl.Dat Gogi=Gen.F.Sg sweet voice.F.Sg.Nom
sʊn-a-i d-i
hear-Caus-F.Sg give-Perf.F.Sg
'They heard Gogi's sweet voice.'

These examples are peculiar in many respects. While they do conform to an overall pattern in the language (and South Asian Languages more generally, e.g., see Verma and Mohanan 1990) by which experiencer predicates have a dative subject ('I' and 'we' in (1)), the nominalized experiencer verb combines with the verb *de* 'give' in (1). This verb is agentive and as such generally requires an ergative subject when in the perfective, as per the general ergative tense/aspect split (see section 2.1). As such, we would expect an ergative subject in (1-b).

We must thus ask how the combination of an agentive verb with a nominalized verb of perception gives rise to an experiencer predicate with a dative subject.

[†]In working on the material in this paper, we engaged in many stimulating discussions and would like to particularly thank Rajesh Bhatt for some valuable pointers, Shahina Butt for an interesting piece of data and observations and the members of the audiences at the Syntax Colloquium at the University of Konstanz, LFG23 in Rochester and SALA-37 in Venice for helpful comments. We would furthermore like to acknowledge funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) — Project-ID 251654672 — TRR 161 for Tafseer Ahmed.

Furthermore, the verb 'give' is generally a three-place predicate, in which an agent causes some entity to be transferred to a goal, but the predication formed by *dɪkʰai* 'seeing'/ *sʊnai* 'hearing' + *de* has only two arguments: an experiencer (the dative subject) and a stimulus (the nominative object). The combined argument structure of *dɪkʰai* 'seeing'/ *sʊnai* 'hearing' + *de* thus poses another puzzle.

In this paper, we take on the puzzle presented by examples as in (1) and propose that *dɪkʰai* 'seeing'/ *sʊnai* 'hearing' + *de* are instances of N-V complex predicates, and that the current structure emerged from the reanalysis of an original goal argument as an experiencer via a reinterpretation of the underlying mapping between event structure and event participants. In this we follow existing proposals by Schätzle (2018) for the rise of dative experiencer subjects in Icelandic via originally locational predications. For the analysis of complex predicates, we basically follow the existing proposals and machinery by Butt (2014) and extend it to cover the effects of nominalization at argument structure. We show that the overall combined complex argument structure then results naturally in just the two arguments found in (1) and in the dative-nominative configuration typical for experiencer verbs in South Asian languages (SALs).

Most of the examples we analyzed for this study were extracted from the three corpora listed below. We additionally also consulted previous literature and used Google search and our own native speaker intuitions to test our hypotheses.

1) Co-author Carnesale compiled a corpus consisting of **Hindi literary texts** from the 20th century for a total of 1,224,437 tokens.[1]

2) **hiTenTen21**: The corpus consists of texts collected from the Internet and belongs to the TenTen corpus family. It contains a total of 901,352,786 tokens and is available on SketchEngine.[2]

3) **UD Urdu UDTB:** The Universal Dependencies (UD) Treebank based on the Urdu Treebank from Bhat et al. (2017), consisting of 5,130 sentences.

The paper is structured as follows: in section 2 we provide the necessary background as to previous work and our assumptions with respect to how case and complex predicates work otherwise in the language. Urdu/Hindi makes extensive use of complex predication and two types in particular are pervasive in the language: complex predicates formed by the sequence of two verbs and complex predicates formed by the sequence of a noun and a verb. Since we argue that the sequence

[1]The corpus primarily consists of the works by Munshi Premchand and includes many of his novels (*Alãkar, Godan, Gaban, Karmabhūmi*) and short story collections, for a total of 888,322 tokens. The corpus also contains novels from other authors, in particular: Dharmavir Bharati's *Sūraj ka satvã ghoṛa* and *Gunahõ ka devta* (125,489), Mohan Rakesh's *andhere band kamre* (101,127), and Jaishankar Prasad's *Kankal* (67,742). Additionally, a minor section is also included collecting short stories by Mahadevi Varma (5,674), Amritlal Nagar (18,052) and Raghuveer Sahay (18,022). The corpus was interrogated through the Concordance tool on SketchEngine (https://www.sketchengine.eu/guide/concordance-a-tool-to-search-a-corpus/).

[2]https://www.sketchengine.eu/hitenten-hindi-corpus/

dɪkʰai/sʊnai + de functions as a complex predicate, section 2.2 is dedicated to the discussion of this phenomenon. In section 2.1 we briefly provide an overview of relevant information with respect to case alternations in Urdu/Hindi, focusing particularly on the nominative/dative/ergative alternation on subjects and the nominative/accusative alternation on objects.

Section 3 is dedicated to the functions of the verb *de* 'give' elsewhere in the language and shows that this verb otherwise consistently takes an ergative agentive subject and that the nominalized verbs of perception *dɪkʰai / sʊnai + de* constitute an exception to this pattern. Section 4 looks into the internal structure of the nominalized verbs of perception. Our discussions then lead us to our analysis in section 5, where we crucially rely on event-based linking as articulated in Schätzle (2018) and Beck and Butt (2023). Section 6 concludes.

2 Background: Case and Complex Predication

2.1 Case alternations in Urdu/Hindi

Case alternations are a typical feature of Urdu/Hindi (Davison 2004a,b, 1999; Montaut 2003, 2006, 2009; Ahmed Khan 2009; Ahmed and Butt 2011; Butt 2022). The alternations are motivated by a mix of semantic and morphosyntactic factors, making for a complex case system (see Butt and King 2004). Of particular relevance for this paper are the ergative/nominative alternation on subjects and the accusative/nominative alternation on direct objects.

Urdu/Hindi is split ergative so that (di)transitive predicates appear with either a nominative or an ergative subject. This alternation is based on the aspect of the verb: perfective (di)transitive predicates require the ergative, which is typically the case clitic *=ne*, as shown in (2-a). Any other tense/aspect type requires a nominative subject, as shown in (2-b) for a version in the future. Besides this morphosyntactic constraint, the ergative/nominative alternation also displays a semantic condition, as it can only appear with agentive arguments. On unergatives like the verb 'scream' in (3), the subject can optionally be marked with the ergative. When this happens, the ergative encodes an argument that is credited with more, rather than less, volition over the action, with animacy also playing a role (cf. Butt and King 1991; Mohanan 1994; de Hoop and Narasimhan 2005; Butt 2006; Ahmed 2010).

(2) a. nadya=ne bɑtʃtʃe=ko kɪtab d-i
Nadya.F=Erg child.M.Sg.Obl=Dat book.F.Sg.Nom give-Perf.F.Sg
'Nadya gave the child a/the book.'

b. nadya bɑtʃtʃe=ko kɪtab de-g-i
Nadya.F.Nom child.M.Sg.Obl=Dat book.F.Sg.Nom give-Fut-F.Sg
'Nadya will give the child a/the book.'

(3) a. lɑṛk-i tʃilla-i
girl-F.Sg.Nom scream-Perf.F.Sg
'The girl screamed.'

b. lɑṛk-i=ne tʃilla-ya
girl-F.Sg=Erg scream-Perf.M.Sg
'The girl screamed (purposefully).'

The agreement patterns of Urdu/Hindi have by now been researched in some detail (e.g., Mohanan 1994; Bhatt 2005; Butt 2014; Kulkarni 2011) and are complex. As illustrated by (2) and (3), the basic pattern is as follows. The verb agrees with the nominative subject if there is one. If the subject is overtly case marked, the verb agrees with the nominative object if there is one. If the object is also case marked, the verbal morphology defaults to masculine singular.

The dative is realized by the case clitic *=ko* (as in (2)) or the inflection *-e* (shown in (1)). The clitic *=ko* is form-identical with a Differential Object Marker (DOM; Bossong 1985; Comrie 1989; Aissen 2003) that signals animacy and/or specificity (Butt 1993; Bhatt and Anagnostopoulou 1996) and is generally glossed as accusative. This accusative alternates with an unmarked form of the noun that is generally glossed as nominative, as shown in (4).

(4) a. yasin=ne kamputar xarid-a
Yassin.M.Sg=Erg computer.M.Sg.Nom buy-Perf.M.Sg
'Yassin bought a/some computer.'
b. yasin=ne kamputar=ko xarid-a
Yassin.M.Sg=Erg computer.M.Sg=Acc buy-Perf.M.Sg
'Yassin bought a (certain)/the computer.'

The final alternation of relevance is a dative/ergative alternation on subjects which is conditioned by the type of the predication (see also Montaut 2016). In predications involving experiencer subjects as in (5-a), the subject is dative. In predications involving more agentive actions as in (5-b), the subject is ergative. The difference in (5) is effected by the use of the light verb 'come' (a typical unaccusative verb of motion) vs. 'do' (a typical agentive verb).

(5) a. nadya=ko kahani yad a-yi
Nadya.F.Sg=Dat story.F.Sg.Nom memory come-Perf.F.Sg
'Nadya remembered a/the story.'
(lit.: 'Memory of the story came to Nadya.')
b. nadya=ne kahani yad k-i
Nadya.F.Sg=Erg story.F.Sg.Nom memory do-Perf.F.Sg
'Nadya remembered a/the story.' (lit.: 'Nadya did memory of the story.')

Dative/ergative alternations are found in other parts of the grammar as well, not just with light verbs; see Butt (2006). However, all we need to know for the purposes of this paper is that these case alternations exist in Urdu/Hindi and that they are conditioned by a combination of verbal morphology and semantics.

2.2 Complex Predication

We have already introduced light verbs above in (5) as part of an N-V complex predicate. In this section, we provide more background on complex predication.

2.2.1 Basic Types

In South Asian linguistics, the term complex predicate covers many types, but two main categories can be singled out: aspectual V-V complex predicates (e.g., Hook 1974; Butt 1995; Butt and Geuder 2001) and N-V complex predicates (Mohanan 1994; Ahmed and Butt 2011). The V-V complex predicates consist of an uninflected main verb and an inflected light verb (also called vector or compound verb in the previous literature, see e.g., Pray 1970; Hook 1974, 1991). The light verbs interact with the semantics of the main verb, giving rise to specific semantic modulations which depend on the light verb. In particular, a light verb typically conveys meanings related to inception and completion, volitionality, benefaction and so on. An example is shown in (6), which also illustrates that the case of the subject is ultimately determined by the light verb (cf. Butt 1995). The same main verb *ga* 'sing' is used in both examples in (6). However, in (6-a) it is combined with the agentive light verb *ḍal* 'put' and the subject is consequentially required to be ergative. In contrast, in (6-b) the light verb is the unaccusative verb of motion *paṛ* 'fall' and the subject must be nominative (and the verb agrees with this nominative subject).[3]

(6) a. ʊs=ne gana ga ḍal-a
Pron.3.Sg=Erg song.M.Nom sing put-Perf.M.Sg
'He/she sang a song (completely, forcefully).'

b. vo gana ga pɑṛ-a
Pron.3.Sg.Nom song.M.Nom sing fall-Perf.M.Sg
'He fell to singing (burst out into song).'

The example in (5) has already shown that the case of the subject is also determined by the light verb in N-V complex predicates. N-V complex predicates consist of an uninflected noun that contributes the larger part of the contentful predication and an inflected light verb, which determines agreement and carries information about tense/aspect (Mohanan 1994; Ahmed and Butt 2011; Kulkarni 2011).

It is not always straightforward to determine whether an N+V sequence represents a complex predicate or is just, for example, an instance of an object followed by a verb in a typical SOV word order. There are two basic reliable cues which can be used to determine the status of an N+V sequence as a complex predicate. One involves agreement, the other the type and number of arguments licensed in the predication. Example (7), taken from Ahmed and Butt (2011), is clearly a complex

[3] A reviewer asks why the pronoun is interpreted as only masculine in (6-b). This is because the verb indicates this interpretation via its masculine agreement morphology. If the subject were feminine, the verb would show feminine agreement morphology. In (6-a) we cannot tell because the verb agrees with the object.

predicate with respect to these two characteristics. For one, the verb agrees with the noun 'house', but not with the noun 'construction', indicating that 'house' is the syntactic object of the overall predication, not 'construction'.

(7) bɪlal=ne mɑkan tɑmir ki-ya
Bilal.M.Sg=Erg house.M.Sg.Nom construction.F.Nom do-Perf.M.Sg
'Bilal built a/the house.'

For another, the verb 'do' by itself is transitive in that it takes two arguments: a doer (agent) and a done thing (patient/theme). The example in (8), however, contains three NPs of which only *Bilal* (the doer) and 'construction' (the done thing) are licensed by the verb 'do'. The extra noun 'house' is best analyzed as being contributed by 'construction' (the house is the thing that is constructed). However, 'house' is nominative in (7) and therefore must be licensed by a verbal predication. If it were licensed within a nominal domain, it would be expected to take the genitive case, as in (8). Note also the change in agreement, indicating that 'construction' is the syntactic object in (8).

(8) bɪlal=ne mɑkan=k-i tɑmir ki
Bilal.M.Sg=Erg house.M.Sg=Gen-F.Sg construction.F.Nom do.Perf.F.Sg
'Bilal did the construction of a/the house.'

Examples as in (7) thus provide evidence that a noun and a verb can combine their argument structures to result in a joint, complex predication in which the jointly determined arguments are licensed verbally by the complex predicate (*tɑmir+kiya*).

2.2.2 Argument Merger

A number of different types of approaches exist to account for complex predication. For within LFG alone, see the overview article by Andrews (2023). In this paper, we follow the theory developed originally in Butt (1995) and then refined and extended in Butt (1998, 2014). We furthermore combine Butt's basic theory of argument merger with the event-based version of mapping (or linking) theory as developed in Schätzle (2018) and Beck and Butt (2023).

Complex predication is taken to occur when two or more predicational elements enter into a relationship of co-predication: each element adds arguments to a joint monoclausal predication. Technically this is effected by representing the light verb as an incomplete predicate that needs to combine with another predicate (Alsina 1996). This is shown in (9), where the missing predicate is indicated by the variable %Pred.[4]

(9) GIVE < agent goal %Pred >

[4]This notation is taken directly from the grammar development platform XLE (Crouch et al. 2017). Note also that in order to allow for a simplified discussion of a(rgument)-structure and complex predication the representation in (i) uses standard thematic role names.

The a-structure in (9) is what Butt (2014) posits for the permissive light verb 'give/let', illustrated in (10), where it combines with a verbal noun with invariant oblique infinitive inflection.

(10) nadya=ne yasin=ko kıtab pɑɽʰ-ne d-i
Nadya.F=Erg Yassin.M=Dat book.F.Sg.Nom read-Inf.Obl give-Perf.F.Sg
'Nadya let Yassin read a/the book.'

Butt (1995) shows that this permissve consists of a co-predication between the verbal noun and the light verb 'give' and that they together function as a monoclausal predication equivalent to that of a simplex verb. The two predicational elements work together and combine their individual arguments. The light verb *de* 'give' licenses an agent, a goal and an empty slot which must be filled by a predicate. This empty slot is filled by *pɑɽʰ* 'read', which contributes an agent and a patient. The composed a-structure is as shown in (11-a) and the monoclausal f-structure is as represented in (11-b).

(11) a. **Composed a-structure**:
GIVE/LET < agent goal$_i$ READ < agent$_i$ patient >>

b. **Monoclausal f-structure**

$$\begin{bmatrix} \text{PRED} & \text{'let-read} < \text{SUBJ}, \text{OBJ}_\theta, \text{OBJ} > \text{'} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'Nadya'} \\ \text{CASE} & \text{ERG} \end{bmatrix} \\ \text{OBJ}_\theta & \begin{bmatrix} \text{PRED} & \text{'Yassin'} \\ \text{CASE} & \text{DAT} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'book'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{TNS-ASP} & \begin{bmatrix} \text{TENSE} & \text{PAST} \\ \text{ASPECT} & \text{PERF} \end{bmatrix} \end{bmatrix}$$

Butt posits a theory of argument merger by which the highest embedded argument is merged with the lowest matrix argument. This is indicated by the subscript *i* on the arguments in (11-a). This argument merger has the effect that there is a total of three arguments to be expressed in the syntax (rather than the original five): the a-structure of the predicate 'read' is slotted into the %Pred argument of 'give' and the agent argument of 'read' merges with the goal argument of 'give'.

Butt (2014) motivates this theory of argument merger by pointing out that this is parallel to what has been established for syntactic control and raising whereby the lowest matrix grammatical function generally controls the highest embedded grammatical function (Bresnan 1982). Identification at the level of syntax results

in control/raising, while identification at the level of a-structure leads to complex predication. The full space of possibilities is illustrated in Table 1 (taken from Butt 2014) and accounts for various different types of argument mergers that are possible at a-structure.

	Control	**Raising**	**Complex Predicate**
syntax (f-structure)	PRO controlled	Exceptional Case Marking (ECM)	No
a-structure	argument controlled (fusion)	arguments unified (raising)	Yes

Table 1: Space of possibilities in argument combinations

The argument fusion possibility is what is illustrated in (11), an 'allow-to-do' reading. But as pointed out by Davison (2014) the permissive *de* 'give' can also be found in an 'allow-to-happen' reading seen in (12).

(12) ḍakṭar=ne mɑriz=ko buxar a-ne nahĩ̃ĩ̃
doctor.Sg=Erg patient.Sg=Dat fever.M.Sg.Nom come-Inf.Obl not
di-ya
give-Perf.M.Sg
'The doctor did not let the patient get a fever.'

Butt analyzes this as the a-structure version of syntactic raising and proposes a version of the light verb *de* 'give' that does not contain a goal argument as part of the account; see (13-a). This analysis is in analogy to that already proposed for causatives in Chichewa, Marathi, Malayalam and Urdu/Hindi (Alsina and Joshi 1991; Butt 1998), for which exactly two a-structure options have already been posited, one with a "causee" (or goal as this is the goal of the causation) argument in the a-structure of the causative and one without; see (13-b).

(13) a. GIVE/LET < agent %Pred >

b. CAUSE < agent goal %Pred >
CAUSE < agent %Pred >

If we deploy this goal-less option together with the idea of argument raising, we get the configuration in (14). The verb *a* 'come' is taken to consist of a goal (or location) and a theme that moves to this location. No arguments are identified/fused with one another; rather the two a-structures of the predicates are unified into a joint, complex predication. The "lower" arguments in the a-structure are "raised" to combine into a single a-structure and the individual arguments in this combined but also single a-structure are linked to grammatical functions as shown in (14), which assumes LFG's classic Mapping Theory for the moment (Bresnan and Zaenen 1990; Bresnan 2001) and provides exactly the right analysis for (12).

(14)

GIVE/LET <	agent	COME <	goal	theme	>>
	[−o]		[+o]	[−r]	
	|		|	|	
	SUBJ		OBJ_θ/~~OBJ~~	OBJ	
	doctor		patient	fever	
	Erg		Dat	Nom	

With this background information on case in Urdu/Hindi and our assumptions as to complex predication in place, we are now ready to examine the *dɪkʰai* and *sʊnai* + *de* construction in more detail. We begin with a closer look at *de* 'give'.

3 The verb *de* elsewhere in the language

As far as we have been able to ascertain, the verb *de* 'give' always licenses an ergative/nominative subject elsewhere in the language. We know of no other instance where it takes a dative subject. It can be used in Urdu/Hindi both as a main verb and as a light verb. When used as a main verb it is ditransitive and it licenses three arguments: an agent, a goal and a theme, encoded with nominative/ergative, dative, and a nominative/accusative, respectively, as illustrated in (15-a). It may also be used in idiomatic and metaphorical functions and as shown in (15-b), where it also licenses an ergative subject.

(15) a. nadya=ne bɑcce=ko kɪtab d-i
Nadya.F=Erg child.M.Sg.Obl=Dat book.F.Sg.Nom give-Perf.F.Sg
'Nadya gave the child a/the book.' (main verb)

b. protestar=ne ɪslamabad=mẽ dʰɑrna di-ya
protestor=Erg Islamabad=in sit-in.M.Sg.Nom give-Perf.M.Sg
'Protesters staged a sit-in in Islamabad.' (idiomatic use)

When *de* functions as a light verb it shows at least three different uses: as a light verb in V-V aspectual complex predicates and permissives and in N-V complex predicates. We have already seen its permissive use, where it always requires an ergative/nominative (even in the 'allow-to-happen' reading). When it occurs in V-V complex predicates, it tends to convey benefaction and completion and is associated with the responsibility for an action (Butt and Geuder 2001). As shown in (16), *de* also requires an ergative/nominative subject in this case.

(16) a. nadya=ne bɑṭua kʰo di-ya
Nadya.F=Erg wallet.M.Sg.Nom lose give-Perf.M.Sg
'Nadya lost a/the wallet (and it's her fault).' (based on Hook 1974: 310)

b. nadya bɑṭua kʰo de-ti
Nadya.F=Erg wallet.M.Sg.Nom lose give-Impf-F.Sg
'Nadya tends to lose a/the wallet (and it's her fault).'

The verb *de* is not used very often as a light verb in N-V complex predicates; however, examples as in (17) can be found. In (17) the argument *vividʰata=par* 'on diversity' is not contributed by the verb *de* 'give', which never licenses an oblique argument marked with the postposition *par* 'on'. The noun *vividʰata* is rather contributed by the noun *dʰyan* 'attention' and appears in the nominative rather than the genitive, thus indicating complex predication. Again, *de* 'give' requires an ergative subject, as is generally consistent with *de* 'give' as an agentive verb.

(17) bʰaṣa=ke vividʰɑta=pɑr hɑm=ne arɑmbʰ=se
language.F=Gen.Obl diversity.M.Sg=on 1.Pl=Erg beginning=from
dʰyan di-ya
attention.M.Sg give-Perf.M.Sg
'From the very beginning, we paid attention to the diversity of languages.'
(hiTenTen21)

We can thus conclude that in all its other uses in the language, the verb *de* 'give' licenses an ergative/nominative subject and never a dative one. We also have not found any ergative subject in combination with *dɪkʰai* / *sʊnai* + *de* in our corpora and native speakers judge the addition of an ergative argument to these constructions as severely ungrammatical. Finally, we have not found *de* 'give' in combination with any other experiencer predicate in the language, rendering the *dɪkʰai* / *sʊnai* + *de* instances an exceptional situation in the syntax of Urdu/Hindi.

4 The nominalized verbs of perception and *de* 'give'

In this section, we turn to the exact make up of the nominalized verbs of perception, *dɪkʰai* 'seeing' and *sʊnai* 'hearing'. These each consist of a verb stem (*dɪkʰ* 'appear to' and *sʊn* 'hear'), which is causativized via the addition of the causative morpheme *-a* and is further nominalized via the feminine nominalization affix *-i* (Chatterji 1926). Since the nominalized verbs of perception contain a causative, one would expect an agent/causer argument even more strongly, either from *de* 'give' or from the causativization. Furthermore, while the verbs *dɪkʰ* 'appear to' and *sʊn* 'hear' are both verbs of perception, the language treats one as more agentive than the other. The verb *dɪkʰ* 'appear to' is an experiencer predicate with a dative subject, as shown in (18-a), while *sʊn* 'hear' takes an ergative/nominative subject, illustrated in (18-b), adding yet another source of agentivity to the mix.

(18) a. mujʰ-e cand dɪkʰ-a
Pron.1.Sg-Dat moon.M.Sg.Nom appear.to-Perf.M.Sg
'I saw the moon (the moon appeared to me).'
b. hɑm=ne avaz sʊn-i
Pron.3-Pl.Dat voice.F.Sg.Nom hear-Perf.F.Sg
'We heard a voice/sound.'

In what follows, we attempt to derive the number and type of arguments in the

dɪkʰai / *sʊnai* + *de* examples in (1) via what is known about complex predicate formation in Urdu/Hindi (Butt 1995, 1998, 2014; Mohanan 1994) and, in particular, with respect to the theory of complex predication sketched in section 2.2.2.

The overall core arguments of *dɪkʰai* and *sʊnai* + *de* work out to be a dative subject and a nominative object. In this, they exactly parallel other N-V experiencer predicates in the language, as shown in (19), where the noun *bʰuk* 'hunger' and the verb *lag* 'be attached to' together mean 'feel hungry'.

(19) a. mʊjʰ-e bʰuk lɑg-i
Pron.1.Sg-Dat hunger.F.Sg be.attach-Perf.F.Sg
'I felt hunger (lit. hunger is attached to me).' (N-V experiencer)

b. mʊjʰ-e jahaz dɪkʰ-a-i
Pron.1.Sg-Dat plane.M.Sg.Nom appear.to-Caus-F.Sg
di-ya
give-Perf-M.Sg
'I saw a plane' (seeing+give)

We know that the dative argument must be a subject via subjecthood tests we can apply (Mohanan 1994). This is demonstrated with respect to anaphora resolution and control for *dɪkʰai* + *de* in (20). Reflexives in Urdu/Hindi are subject-oriented and, as can be seen in (20-a), the reflexive pronoun *apne* is coreferential with the dative pronoun *mujʰe* 'to me'. This shows that the dative argument is functioning as a subject (Gurtu 1985; Mohanan 1994).

(20) a. mʊjʰ-e ɑpn-e gʰɑr=mẽ=se ek buṛʰ-i
Pron.1.Sg-Dat self-Obl house=in=Abl one old-F.Sg
aurɑt nikɑl-t-i hu-i
woman.F.Sg.Nom emerge-Impf-F.Sg become-Perf.F.Sg
dikʰ-ɑ-i d-i
appear-Caus-F.Sg give-Perf.F.Sg
'I saw an old woman coming out of my house.'

b. [age ja=kɑr] un-hẽ ran=ke pas ek kʰubsurɑt
ahead go=having 3.Pl.Obl-Dat Ran=Gen.Obl near one beautiful
bag dɪkʰ-a-i de-t-a hɛ
garden.M.Nom appear-Caus-F.Sg give-Impf-M.Sg be.Prs.3.Sg
'They continue forward and they see a beautiful garden next to Ran.'
(hiTenTen21)

Further evidence comes from control: an unexpressed embedded (PRO) subject in Urdu/Hindi can generally only be controlled by a matrix subject. In (20-b), the unexpressed subject of the adverbial clause 'having gone ahead' is coreferential with the dative pronoun *mʊjʰe*, thus providing another piece of evidence that the dative argument functions as a subject in these constructions.

We thus have a complex predication consisting of many individual parts that need to be combined so that it ends up with only two core arguments: a dative

subject and a nominative object. However, given the individual pieces, one would expect to have at least four arguments, as shown in (21), which is one of several possible ways of putting together the argument structures. We assume the classic version of LFG's Mapping Theory (Bresnan and Zaenen 1990; Bresnan 2001) for ease of exposition and also use standard thematic role labels for the same reason. We also assume the Urdu/Hindi specific mappings worked out by Butt (1998) for causatives, which posit that goals are always intrinsically assigned a [+o].

(21)

GIVE <	agent	goal_i CAUSE <	agent_i	goal_j APPEAR.TO <	exp.$_j$	theme >>>
	[−o]	[+o]		[+o]		[−r]
	|	|		|		|
	SUBJ	OBJ_θ		OBJ_θ		OBJ
	Erg/Nom	Dat		Dat		Nom

The possibility in (21) uses the verb *dɪkʰ* 'appear to' as the base verb. As seen in (18), this is an experiencer predicate with two arguments, an experiencer and a theme (stimulus). This verb is causativized. Given our existing analyses for causatives, we have two possibilities, shown in (22-a). For purposes of illustration, we choose the one with a goal argument and make the same choice for 'give', which embeds the causativized verb. The complex predication in (21) works with argument fusion, with fused arguments indicated via subscripts. Only the higher argument of a fused pair is linked to a grammatical function. As can be seen, we end up with four arguments that should be linked rather than two. Furthermore, while we do have a nominative object, we cannot arrive at a dative subject.

(22) a. CAUSE < causer/agent causee/goal %Pred >
CAUSE < causer/agent %Pred >

b. GIVE/LET < agent %Pred >
GIVE/LET < agent goal %Pred >

We can try various different ways of putting the individual pieces together; one other possibility is shown in (23), where we have used the version of the causative without a goal/causee argument and work with argument raising instead of argument fusion to model the relationship between the causative and the embedded verb of perception. But the result is still no better.

(23)

GIVE <	agent	goal_i CAUSE <	agent_i	APPEAR.TO <	experiencer	theme >>>
	[−o]	[+o]			[−o]	[−r]
	|	|			|	|
	SUBJ	OBJ_θ			OBL	OBJ
	Erg/Nom	Dat			Loc	Nom

We invite our readers to try out all the other possible combinations of the predicational parts, utilizing either version of 'give' and the causative or varying the application of argument fusion vs. argument raising in putting the pieces together. In all cases, we are stuck with a predicted nominative/ergative subject and in most

cases have too many predicted arguments. The same applies to any and all tinkering around with the verb *sʊn* 'hear'. Indeed, getting to the right number and type of arguments is even more difficult with this verb, since it contributes an agent rather than an experiencer to the overall predication.

One thing we have not as yet factored in is the nominalization. As shown in (24) for the linking configuration in (23), this should serve to prevent at least one argument from being realized overtly in the syntax, thus cutting down the number of overall arguments to be linked. However, as the nominalization is of the causativized verb stem and not of the complex predication as a whole (the N+V is verbal, not nominal), this does not help with problem of how to arrive at a dative subject, rather than the ergative/nominative agentive subject predicted by (24).

(24)

GIVE <	agent	goal$_i$	CAUSE <	agent$_i$	APPEAR.TO	< experiencer	theme >>>
	[−o]					[−o]	[−r]
	|			|		|	|
	SUBJ			Ø		OBL	OBJ
	Erg/Nom					Loc	Nom

The same overall problem will persist, no matter which version of linking/-mapping we assume. For the current space of possibilities, see the overview in Findlay et al. (2023).

One option at this point is to throw up one's hands and look to analyze *dɪkʰai* and *sʊnai* + *de* as fixed constructions which are the product of some form of grammaticalization. Indeed, an investigation of the morphology involved shows that while both the causative and the nominalization morphemes are each separately fully productive in Urdu/Hindi, the nominalized causative is not. Some fixed examples still exist and are in everyday use, but their number is very small. We have found, for example: *caṛʰ-ai* 'climb, ascent', *lɪp-ai* 'painting', *laṛ-ai* 'fight', *luṭ-ai* 'plundering', *paṛʰ-ai* (Kachru 1980; Saksena 1982).

We could thus hypothesize that, although the compositional nature of the nominalization is still transparent, *dɪkʰai* and *sʊnai* have been lexicalized to be nouns of perception with an Experiencer-Stimulus argument structure and that the verb *de* has completely lost its original semantics of 'give' and consequently its capacity to license an argument structure and with that the bothersome agentive argument. For *dɪkʰai*, this would leave us with the arguments in (25-a) and we could assume that the fused argument is simply reinterpreted as an experiencer as part of the lexicalization, thus giving us the overall argument structure in (25-b).

(25) a. **Originally:**
cause < causer/agent$_i$ appear.to < experiencer$_i$ theme >>

b. **After Lexicalization:**
seeing < experiencer theme >

However, this explanation does not feel quite as natural for *sʊnai*, whose trajectory would need to be as in (26). We would have to assume that the original agentive

semantics became reinterpreted as experiencer semantics in analogy to *dɪkʰai*, for which there is no real good evidence.

(26) a. **Originally:**
cause < causer/agent$_i$ hear < agent$_i$ theme >>
b. **After Lexicalization:**
hearing < experiencer theme >

Furthermore, if we assume that *de* 'give' is syntactically and semantically empty, then it is not clear why it is exactly this verb that is used in the construction, rather than 'go' or 'come', for example. In particular, the verb *a* 'come' is frequently used as a light verb in complex predicates expressing experiential semantics. Consider (27), in which the visual perception is expressed by the complex predicate *nazar a*, consisting of the noun *nazar* 'seeing' and the light verb *a* 'come'.

(27) bɑhʊt tɑlaʃ=ke bad use ɑpn-a puran-a
much search=Gen.Obl after Pron.3.Sg.Obl self-M.Sg old-M.Sg
ghɑr nɑzɑr a-ya
house.M.Sg.Nom seeing.F come-Perf.M.Sg
'After much searching, he saw his old house.'

If *de* were completely empty, it is also not clear what its syntactic status or function would be. It is not a general purpose verbalizer in the language—that function is taken up by *kar* 'do'. Nor can *de* in this construction be analyzed as an auxiliary (situating the event in time) or a modal (situating the event in terms of possible worlds) either semantically or syntactically (cf. Butt 2010).

Given the syntactic (and semantic) parallels with other N-V experiencer complex predicates, we see the most likely successful analysis as being one in which *de* 'give' functions as a light verb when combining with 'seeing/hearing'. In what follows, we pursue an analysis with respect to an event-based version of linking that we have been developing within LFG and suggest that this provides us with the right tools to understand *dɪkʰai* and *sʊnai* + *de*.

5 Our Proposal: An Event-based Approach

Unlike many other proposals for relating argument structure to syntactic roles, standard LFG does not assume an event-based representation. However, Schätzle (2018) and Beck and Butt (2023) propose to relate argument structure to syntactic roles by integrating Ramchand's (2008) tripartite organization of subevental structure. They further integrate this with ideas in Kibort's (2014) version of LFG's Mapping Theory, which posits four abstract argument positions as an independent tier of representation ('argument slots'), eschewing thematic role labels. The latter does justice to the many critical discussions of thematic role labels (cf. for example Dowty 1991; Grimshaw 1990; see also approaches to a-structure that experiment with glue semantic representations such as Findlay 2016; Findlay et al. 2023).

Ramchand (2008) decomposes an event into three subevents, each of which causes/initiates the other: 1) a causing or initiating subevent (init); results in 2) a process subevent (proc), which results in 3) a result state (res). In addition, rhemes (rh) are taken to be in a static relationship with one of the three subevents of a predicate, like a static spatial Figure/Ground relationship.[5] We let each of the three event slots plus the rheme license an argument participant, corresponding to Kibort's four basic a-structure slots.

In (28) we illustrate our general linking schema. A maximum of four abstract argument slots are licensed per predication by the subevents init, proc, res as well as the rheme (rh). These can be linked to the standard LFG grammatical functions. We further associate argument slots with Figure/Ground relations (Talmy 1975).[6] These relations serve to "promote" or "demote" individual arguments, for example, backgrounding an agent (initiator) argument in passivization or participle formation. They also simultaneously exert pressure on the configuration to prefer sentient, initiator arguments as subjects (Grimm 2005).

(28) General Linking Schema

		init	proc	res	rh	
PREDICATE	<	x	x	x	x	>
		FIGURE	GROUND			
Grammatical Functions		SUBJ	OBJ	OBJ_θ	OBL	

We further integrate insights from Zaenen (1993), who proposes to inform the linking of the arguments to grammatical functions via the Proto-Agent and Proto-Patient entailments defined by Dowty (1991) within LFG's Mapping Theory. We use a combination of the Proto-Agent and Proto-Patient entailments, the Figure/Ground relation and the semantics of the subevent/rheme an argument is licensed by to determine the final linking of argument slots to grammatical functions. Generally, the argument with the most Proto-Agent properties will be linked to the subject, while the argument with the most Proto-Patient properties should work out to be the object, but as is well known, the interface between semantic arguments and syntactic grammatical relations is more complex than that, which is why we posit all the various moving parts in (28).

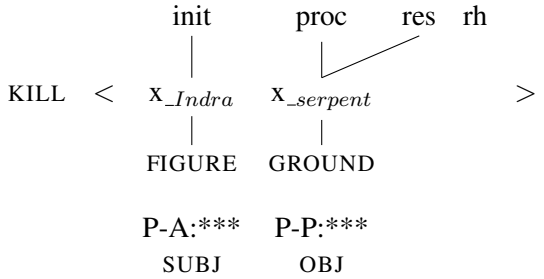
To illustrate the basic system we demonstrate how both an active agentive clause and an experiencer predicate work. We begin with the active transitive exam-

[5]Note that Ramchand's proposals differ significantly from other event-based approaches such as that by Rappaport Hovav and Levin (1998), Van Valin and Polla (1997) or Croft (2012) in that she posits a tripartite rather than a binary Cause-Result distinction, works with the concept of rhemes and embeds her analysis within a formal semantic approach to events.

[6]A reviewer would like to know more about our assumptions as to Figure/Ground, but this would take us too far afield. We instead refer the interested reader to Schätzle (2018), where the general linking schema and its theoretical motivations have been worked out in some detail.

ple in (29), which consists of three subevents in Ramchand's system: an init(iation) event that leads to a proc(ess) of killing, which in turn leads to a res(ult) in which the serpent is dead. The initiator of the event sequence is Indra, licensed by init. The serpent is involved in both the proc (getting killed) and the res (being dead). Note that we use the underscore notation in (29) to indicate the actual arguments for ease of exposition. They have no theoretical status.

(29) Indra killed the serpent.



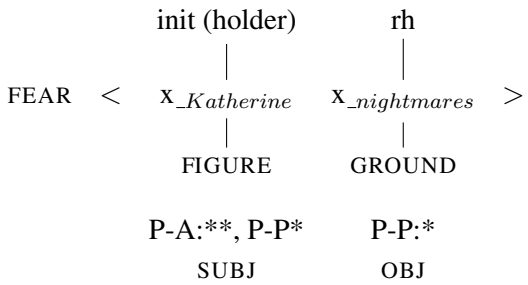
Barring any other information coming from morphosyntax (such as passive morphology, for example; see Beck and Butt 2023), Indra as the initiator is the Figure, which leaves the serpent as the Ground. If one goes through the set of Dowty's Proto-Role entailments, one will find that Indra accumulates two Proto-Agent properties because he is the sentient initiator of an action. A further Proto-Agent property is accrued by his status as the Figure, yielding a total of three Proto-Agent (P-A) properties. This is marked by the three '*' for P-A in (29).

The serpent, on the other hand, is causally affected and undergoes a change of state (getting killed), which results in two Proto-Patient (P-P) properties. It accrues another P-P property through its status as Ground, yielding a total of three P-P properties and no P-A properties. In this configuration, the linking is quite clear: as the argument with the most P-A properties, Indra is linked to the SUBJ, and the serpent with its accumulation of only P-P properties is linked to the OBJ.

We now turn to an example of an experiencer predicate. In Ramchand's system, these involve a holder (an experiencer) of a state. The holder is identified with the init subevent, the state with a rh(eme). In our system, this results in the linking configuration in (30). Katherine accrues a P-A property because of her sentience and another P-A property because as the sentient argument she is the Figure. However, as she is not an agentive initiator, but a holder of a state, we also assign a P-P property. The nightmares are the Ground and thus acquire a P-P property. Since rhemes are not event participants, they do not contribute any P-A or P-P properties.

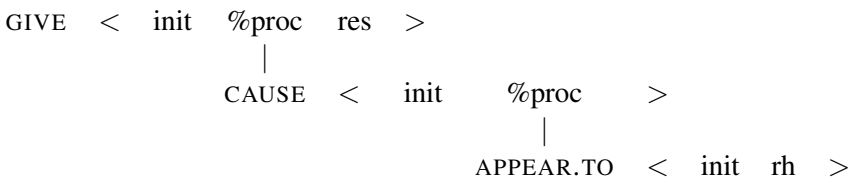
One can see that this configuration is not as clear-cut a case as the transitive active example. Basically Katherine ends up being linked to the SUBJ because of her property as a sentient being. If sentience were not involved, the linking could easily go the other way and this is indeed what we do find with experiencer predicates that are derived from former locative predications (Schätzle 2018; Beck and Butt 2023).

(30) Katherine fears nightmares.



We propose that the adoption of this event-based perspective provides exactly the right kind of perspective on the argument composition of the complex predicates under investigation here. We begin by considering the combination of *dɪkʰai* + *de* 'seeing give'. As we saw above, the verb *dɪkʰ* 'appear to' is an experiencer predicate and we analyze it analogously to 'fear' in (30) as involving an init holder of a state (rheme). This basic verb is then causativized. Here we assume Ramchand's approach, which sees causation as the addition of a init (causing) event (Ramchand 2008; Butt et al. 2010). This complex predicate is in turn embedded by the light verb *de* 'give', which contributes an init and expects some event being engaged in (the *%proc*) that will yield a result. The expected event predication is furnished by the causativized verb of perception, resulting in the full basic predication in (31).

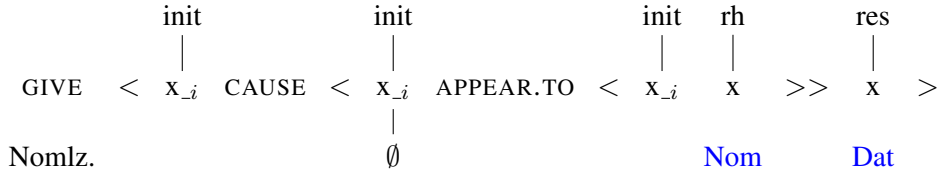
(31)



We now apply Butt's ideas on Argument Merger and we also need to factor in the effect of the nominalizing morpheme *-i*. This is done in (32), where coindexation again indicates argument merger.[7] The nominalization (Nomlz.) suppresses the highest argument contributed by the causative, thus effectively also suppressing all the argument slots merged with it. This has the effect that only the desired number of arguments remains: two. These two arguments can also be assigned the right case marking: rhemes are generally nominative and the result, as a natural goal or location, can be marked with a dative.

[7]Note that Butt's rule of complex predicate formation, which sees the lowest matrix argument being merged with the highest embedded one, assumes that this applies to the lowest argument ordered *before* the embedded a-structure. That is the two init arguments merge, rather than the res argument of GIVE with the init argument of CAUSE. We thank a reviewer for asking us to clarify this.

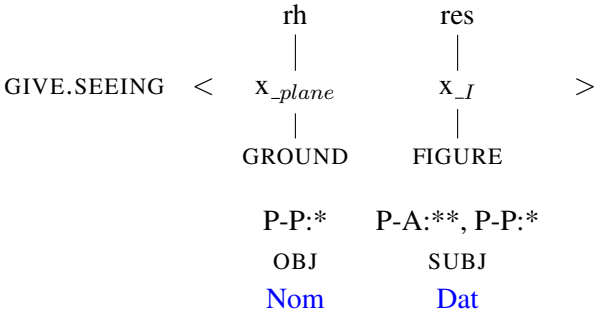
(32) Full predication of *dɪkʰai de* ‘seeing give’



In (34), we take just the two arguments of the combined predication in (32) that are available for linking to the syntax to show how this works out for an example as in (33) (repeated from above). In the configuration in (34), we have an argument licensed by the rheme and one by a res. The res is effectively what ends up as a dative subject. The rheme is a Ground by definition (following Ramchand’s system) and therefore receives one P-P property. The res argument accumulates one P-P property because of its status as a result. But it is also the Figure and it is also sentient, resulting in two P-A properties and thus leading to the linking to a subject.

(33) mʊjʰ-e jahaz dɪkʰ-a-i di-ya
Pron.1.Sg-Dat plane.M.Sg.Nom appear.to-Caus-F.Sg give-Perf-M.Sg
‘I saw a plane’

(34)



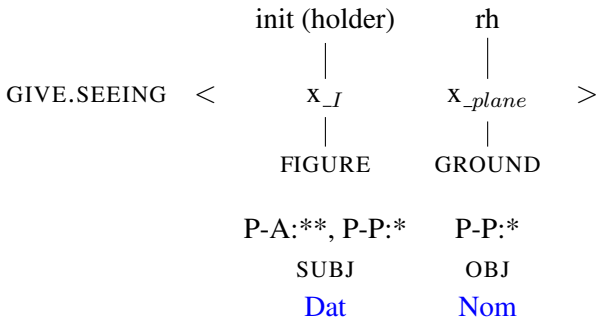
The dative case marking is due to the original goal (result) semantics (cf. Butt and King 1991, 2003; Butt and Ahmed 2011) this argument is associated with and is thus similar to the overall pattern found with N-V experiencer predicates in the language, which arguably involve an originally locative predication that has been reanalyzed as an experiencer predicate; see for example, (35), repeated from above.

(35) mʊjʰ-e bʰuk lɑg-i
Pron.1.Sg-Dat hunger.F.Sg be.attach-Perf.F.Sg
‘I felt hunger (lit. hunger is attached to me).’ (N-V experiencer)

We assume that a similar reanalysis took place with respect to the nominalized verbs of perception, so that the configuration in (34) was ultimately reinterpreted as representing a configuration in which there is a holder of a state, as in (36), rather

than something "arriving" at a destination as in (34). The configuration in (36) is thus then exactly parallel to that posited by Ramchand for experiencer predicates.

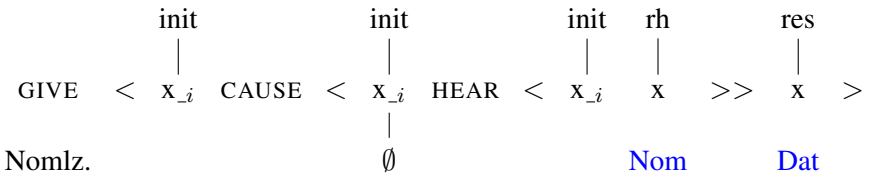
(36)



Although case is an important part of our analysis, the determination of case marking is not part of the linking/mapping between arguments and grammatical functions per se. We essentially adopt a *Constructive Case* (Nordlinger 1998) approach and follow Butt and King (1991, 2004), who posit lexical entries for individual case markers. These lexical entries encode constraints such as that the ergative is associated with perfective morphology, but also provide semantic information such as the specificity interpretation associated with accusative DOM. In our case, datives specify that they can only occur with grammatical functions that broadly express goal semantics: indirect goal objects, locatives (obliques) and experiencer objects (cf. Ahmed 2006). Ergatives require initiator semantics, but are constrained by morphosyntactic restrictions as to aspect. Nominative is assigned as a default case when no other semantic information is specified. Urdu/Hindi case marking is thus dependent on the right configuration of a-structure to f-structure mapping.

Now let us turn to *sʊnai* + *de*, which we saw works as a more agentive predication in its base use. However, when viewed through Ramchand's event-based approach, we actually end up with an identical configuration to *dɪkʰai* + *de*. This is because the base predicate 'hear' also involves an init event with a rheme: whatever is heard does not undergo any process and it does not change as a result of the process. So the stimulus argument of 'hear' must also be classified as a rheme. The overall argument merger, nominalization and linking thus proceed just as we saw with SEEING+GIVE, providing an explanation as to why the resulting complex predication works similarly despite the difference in the base verbs.

(37) Full predication of *sʊnai de* 'hearing give'



Our corpus study also yielded examples in which *dɪkʰai* and *sʊnai* can combine with *paṛ* 'fall', illustrated in (38). The dative experiencer argument would here originally be derived from the locative argument contributed by the verb 'fall'. In our current approach, the linking configuration with 'fall' would work out to be quite similar to that of (31), with the difference that 'fall' does not contribute an init argument to the complex predication. We do not provide the full analysis here due to lack of space and because the linking falls out quite straightforwardly.

(38) a. ɑcanɑk (mʊjʰ-e) ek hirɑn dɪkʰ-a-i
suddenly Pron.1.Sg-Dat one deer.M.Sg.Nom see-Caus-F.Sg
paṛ-a
fall-Perf.M.Sg
'Suddenly a deer appeared (to me).'

b. un-hẽ kʊc ʃor sʊn-a-i
Pron.Pl-Dat some loud.noise.M.Sg.Nom hear-Caus-F.Sg
paṛ-a
fall-Perf.M.Sg
'He (polite) heard some loud noise.'

The event-based approach to linking thus has not only been able to make sense of an otherwise extremely puzzling phenomenon, but also allows a straightforward analysis of further data as in (38).

6 Conclusion

We began this paper with a puzzle that involved the hitherto unexpected licensing of a dative subject in conjunction with the verb *de* 'give'. This verb otherwise requires an ergative/nominate subject in all its other uses. Furthermore, the puzzling construction with *de* involves a combination with just two nominalized causativized verbs of perception (*dɪkʰ* 'appear to' and *sʊn* 'hear'), which come with highly complex argument structures of their own, but the overall joint predication has just two core arguments: a dative subject and a nominative object. This is in line with experiencer predicates in the language more generally.

We delved into the morphosyntax of this construction and concluded that it must be the result of a complex predication that is best understood in terms of the event-based approach to linking articulated by Schätzle (2018) and Beck and Butt (2023) in combination with Butt's theory of complex predication. When viewed through the lens of this approach, the data and argument patterns fall out naturally, making exactly the right predictions and providing a natural connection to other experiencer predicates in the language.

References

Ahmed, Tafseer. 2006. Spatial, Temporal and Structural Usages of Urdu *ko*. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG06 Conference*, pages 1–13, Stanford: CSLI Publications.

Ahmed, Tafseer. 2010. The Unaccusativity/Unergativity Distinction in Urdu. *Journal of South Asian Linguistics* 3, 3–22.

Ahmed, Tafseer and Butt, Miriam. 2011. Discovering Semantic Classes for Urdu N-V Complex Predicates. In *Proceedings of the Ninth International Conference on Computational Semantics*, IWCS '11, pages 305–309, Oxford, United Kingdom.

Ahmed Khan, Tafseer. 2009. *Spatial Expressions and Case in South Asian Languages*. Ph.D. thesis, University of Konstanz.

Aissen, Judith. 2003. Differential Object Marking: Iconicity vs. Economy. *Natural Language and Linguistic Theory* 21, 435–483.

Alsina, Alex. 1996. *The Role of Argument Structure in Grammar*. Stanford: CSLI Publications.

Alsina, Alex and Joshi, Smita. 1991. Parameters in causative constructions. In L. Dobrin, L. Nichols and R.M. Rodriguez (eds.), *Papers from the 27th Regional Meeting of the Chicago Linguistic Society (CLS)*, pages 1–16.

Andrews, Avery D. 2023. Complex Predicates. In Mary Dalrymple (ed.), *The Handbook of Lexical Functional Grammar*, Empirically Oriented Theoretical Morphology and Syntax, Berlin: Language Science Press.

Beck, Christin and Butt, Miriam. 2023. The Rise of Dative Subjects: Relative Prominence in Event Structure. In Christina Sevdali, Dionysios Mertyris and Elena Anagnostopoulou (eds.), *The place of Case in grammar*, Oxford: Oxford University Press, in Press.

Bhat, Riyaz Ahmad, Bhatt, Rajesh, Farudi, Annahita, Klassen, Prescott, Narasimhan, Bhuvana, Palmer, Martha, Rambow, Owen, Sharma, Dipti Misra, Vaidya, Ashwini, Ramagurumurthy Vishnu, Sri and Xia, Fei. 2017. The Hindi/Urdu Treebank Project. In Nancy Ide and James Pustejovsky (eds.), *Handbook of Linguistic Annotation*, pages 659–697, Dordrecht: Springer Netherlands.

Bhatt, Rajesh. 2005. Long distance agreemnet in Hindi-Urdu. *Natural Language and Linguistic Theory* 23, 757–807.

Bhatt, Rajesh and Anagnostopoulou, Elena. 1996. Object shift and specificity: Evidence from ko-phrases in Hindi. In Lise Dobrin, Kora Singer and Lisa McNair (eds.), *Papers from the main session of the Chicago Linguistics Society (CLS) 32*, pages 11–22, Chicago: University of Chicago Linguistics.

Bossong, Georg. 1985. *Differentielle Objektmarkierung in den Neuiranischen Sprachen*. Tübingen: Gunter Narr Verlag.

Bresnan, Joan. 1982. Control and Complementation. *Linguistic Inquiry* 13(3), 343–434.

Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Oxford: Blackwell.

Bresnan, Joan and Zaenen, Annie. 1990. Deep Unaccusativity in LFG. In Katazyna Dziwirek, Patrick Farrell and Errapel Mejías-Bikandi (eds.), *Grammatical Relations: A Cross-Theoretical Perspective*, pages 45–57, Stanford: CSLI Publications.

Butt, Miriam. 1993. Object specificity and agreement in Hindi/Urdu. In *Papers from the 29th Regional Meeting of the Chicago Linguistic Society (CLS)*, volume 1, pages 80–103.

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

Butt, Miriam. 1998. Constraining Argument Merger through Aspect. In Erhard Hinrichs, Andreas Kathol and Tsuneko Nakazawa (eds.), *Complex Predicates in Nonderivational Syntax*, pages 73–113, New York: Academic Press.

Butt, Miriam. 2006. The Dative-Ergative Connection. In Patricia Cabredo-Hofherr and Oliver Bonami (eds.), *Empirical Issues in Formal Syntax and Semantics 10*, pages 69–92, The Hague: Thesus.

Butt, Miriam. 2010. The Light Verb Jungle: Still Hacking Away. In Mengistu Amberber, Brett Baker and Mark Harvey (eds.), *Complex Predicates: Cross-linguistic Perspectives on Event Structure*, pages 48–78, Cambridge: Cambridge University Press.

Butt, Miriam. 2014. Control vs. complex predication: Identifying non-fnite complements. *Natural Language and Linguistic Theory* 32, 165–190.

Butt, Miriam. 2022. Case Markers in Indo-Aryan. In Mark Aronoff (ed.), *Oxford Research Encyclopedia of Linguistics*, Oxford: Oxford University Press.

Butt, Miriam and Ahmed, Tafseer. 2011. The Redevelopment of Indo-Aryan Case Systems from a Lexical Semantic Perspective. *Morphology* 21(3), 545–572.

Butt, Miriam and Geuder, Wilhelm. 2001. On the (Semi)Lexical Status of Light Verbs. In Norbert Corver and Henk van Riemsdijk (eds.), *Semi-lexical Categories: On the content of function words and the function of content words*, pages 323–370, Berlin: Mouton de Gruyter.

Butt, Miriam and King, Tracy Holloway. 1991. Semantic Case in Urdu. In L. Dobrin, L. Nichols and R.M. Rodriguez (eds.), *Papers from the 27th Regional Meeting of the Chicago Linguistic Society*, pages 31–45.

Butt, Miriam and King, Tracy Holloway. 2003. Case Systems: Beyond Structural Distinctions. In Ellen Brandner and Heike Zinsmeister (eds.), *New Perspectives on Case Theory*, pages 53–87, Stanford: CSLI Publications.

Butt, Miriam and King, Tracy Holloway. 2004. The Status of Case. In Veneeta Dayal and Anoop Mahajan (eds.), *Clause Structure in South Asian Languages*, pages 153–198, Berlin: Kluwer Academic Publishers.

Butt, Miriam, King, Tracy Holloway and Ramchand, Gillian. 2010. Complex Predication: Who Made The Child Pinch The Elephant? In Linda Ann Uyechi and Lian-Hee Wee (eds.), *Reality Exploration and Discovery: Pattern Interaction in Language and Life*, pages 231–256, Stanford: CSLI Publications.

Chatterji, Suniti Kumar. 1926. *The Origin and Development of the Bengali Language, Volume II*. Calcutta: D. Mehra, Rupa & Co, 1975 edition.

Comrie, Bernhard. 1989. *Language universals and linguistic typology*. Oxford: Blackwell.

Croft, William. 2012. *Verbs: Aspect and Causal Structure*. Oxford: Oxford University Press, https://doi.org/10.1093/acprof:oso/9780199248582.001.0001.

Crouch, Richard, Dalrymple, Mary, Kaplan, Ronald M., King, Tracy Holloway, Maxwell III, John T. and Newman, Paula. 2017. XLE Documentation. https://ling.sprachwiss.uni-konstanz.de/pages/xle/doc/xle_toc.html.

Davison, Alice. 1999. Ergativity: Functional and Formal Issues. In Michael Darnell, Edith Moravcsik, Frederick Newmeyer, Michael Noonan and Kathleen Wheatley (eds.), *Functionalism and Formalism in Linguistics, Volume I: General Papers*, Amsterdam: John Benjamins.

Davison, Alice. 2004a. Non-nominative subjects in Hindi-Urdu: VP structure and case parameters. In Peri Bhaskararao and Karumuri Venkata Subbarao (eds.), *Non-nominative subjects*, pages 141–168, Amsterdam: John Benjamins.

Davison, Alice. 2004b. Structural Case, Lexical Case and the Verbal Projection. In Veneeta Dayal and Anoop Mahajan (eds.), *Clause Structure in South Asian Languages*, pages

199–225, Berlin: Kluwer Academic Publishers.
Davison, Alice. 2014. Non-finite complements and modality in dee-naa 'allow' in Hindi-Urdu. *Natural Language and Linguistic Theory* 32, 137–164.
de Hoop, Helen and Narasimhan, Bhuvana. 2005. Differential case-marking in Hindi. In Mengistu Amberber and Helen de Hoop (eds.), *Competition and variation in natural languages: The case for case*, pages 321–346, Amsterdam: Elsevier.
Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67(3), 547–619.
Findlay, Jamie Y. 2016. Mapping theory without argument structure. *Journal of Language Modelling* 4(2), 293–338.
Findlay, Jamie Y., Taylor, Roxanne and Kibort, Anna. 2023. Argument structure and mapping theory. In Mary Dalrymple (ed.), *The Handbook of Lexical Functional Grammar*, Empirically Oriented Theoretical Morphology and Syntax, Berlin: Language Science Press, 10.5281/zenodo.10037797.
Grimm, Scott. 2005. *The Lattice of Case and Agentivity*. Master's Thesis, Universiteit van Amsterdam, iLLC Master of Logic Series MoL 2005-9.
Grimshaw, Jane. 1990. *Argument Structure*. Cambridge, MA: The MIT Press.
Gurtu, Madhu. 1985. *Anaphoric Relations in Hindi and English*. Ph.D. thesis, Central Institute of English and Foreign Languages.
Hook, Peter Edwin. 1974. *The Compound Verb in Hindi*. Center for South and Southeast Asian Studies: The University of Michigan.
Hook, Peter Edwin. 1991. The Emergence of Perfective Aspect in Indo-Aryan Languages. In Elizabeth Traugott and Bernd Heine (eds.), *Approaches to Grammaticalization*, volume II: Types of grammatical markers, pages 59–89, Amsterdam: John Benjamins.
Kachru, Yamuna. 1980. *Aspects of Hindi Grammar*. New Delhi: Manohar Publications.
Kibort, Anna. 2014. Mapping out a construction inventory with (Lexical) Mapping Theory. In M. Butt and T. H. King (eds.), *Proceedings of the LFG14 Conference*, pages 262–282, Stanford: CSLI Publications.
Kulkarni, Amba. 2011. Agreement in Hindi Conjunct Verbs. In *Proceedings the 9th International Conference on Natural Language Processing (ICON-2011)*, Macmillan Publishers.
Mohanan, Tara. 1994. *Argument Structure in Hindi*. Stanford: CSLI Publications.
Montaut, Annie. 2003. Oblique Main Arguments in Hindi/Urdu as localizing predications: Questioning the Category of Subject. In Peri Bhaskarao and K.V. Subbarao (eds.), *Non Nominative Subjects*, pages 33–56, Amsterdam: John Benjamins.
Montaut, Annie. 2006. The Evolution of the Tense-Aspect System in Hindi/Urdu: The Status of the Ergative Alignment. In Miriam Butt and Tracy Hollowy King (eds.), *Proceedings of the LFG06 Conference*, pages 365–385, Stanford: CSLI Publications.
Montaut, Annie. 2009. Ergative and Pre-ergative Patterns in Indo-Aryan as Predications of Localization: A Diachronic View of Past and Future Systems. In A.R. Faithi (ed.), *Language Vitality in South Asia*, pages 295–325, Aligarh Muslim University.
Montaut, Annie. 2016. Noun-Verb Complex Predicates in Hindi and the rise of non-canonical subjects. In Pollet Samvelian and Lea Nash (eds.), *Approaches to Complex Predicates*, pages 142–174, Amsterdam: John Benjamins.
Nordlinger, Rachel. 1998. *Constructive Case: Evidence from Australian Languages*. Stanford: CSLI Publications.
Pray, Bruce. 1970. *Topics in Hindi-Urdu Grammar*. Berkeley, California: Center for South and Southeast Asia Studies.
Ramchand, Gillian. 2008. *Verb Meaning and the Lexicon: A First Phase Syntax*. Cam-

bridge: Cambridge University Press.

Rappaport Hovav, Malka and Levin, Beth. 1998. Building verb meanings. In Miriam Butt and Wilhelm Geuder (eds.), *The Projection of Arguments: Lexical and Compositional Factors*, pages 97–134, Stanford: CSLI Publications.

Saksena, Anuradha. 1982. *Topics in the Analysis of Causatives with an Account of Hindi Paradigms*. University of California Press.

Schätzle, Christin. 2018. *Dative Subjects: Historical Change Visualized*. Ph.D. thesis, University of Konstanz.

Talmy, Leonard. 1975. Figure and Ground in Complex Sentences. In *Proceedings of the First Annual Meeting of the Berkeley Linguistics Society*, pages 419–430.

Van Valin, Robert D. and Polla, Randy J. La. 1997. *Syntax: Structure, Meaning and Form*. Cambridge: Cambridge University Press.

Verma, M. K. and Mohanan, K.P. (eds.). 1990. *Experiencer Subjects in South Asian Languages*. Stanford: CSLI Publications.

Zaenen, Annie. 1993. Unaccusativity in Dutch: Integrating Syntax and Lexical Semantics. In James Pustejovsky (ed.), *Semantics and the Lexicon*, pages 129–161, Dordrecht: Kluwer Academic Publishers.