Focus, Adjacency, and Nonspecificity

Miriam Butt and Tracy Holloway King
University of Konstanz, Xerox PARC

This paper investigates an intriguing interaction between preverbal structural focus and the nonspecific interpretation of preverbal objects in Urdu[1] and Turkish.[2] In both languages, accusative objects can scramble freely and are interpreted as specific. Nominative objects are semantically incorporating nonspecifics and cannot scramble, but they behave like accusative objects when they are coerced into a specific interpretation by the presence of preverbal focus. The apparently conflicting functions of the immediately preverbal position in Urdu and Turkish argue for two independent constraints: a differentiation between (non-)incorporating objects and the licensing of focus in a way similar to that of case, thus confirming the close association between focus and case licensing in two genetically unrelated but typologically similar languages.

1 Object interactions

Objects in Urdu and Turkish interact with morphology, syntax, and semantics. This section first examines the interaction of nominative objects with preverbal focus. Then the interaction of specificity with case marking and position are discussed.

1.1 Interaction between Focus and Unmarked Objects

Turkish and Urdu both allow objects to appear in the nominative case, i.e., without overt case marking or "bare". These bare objects are licensed in immediately preverbal position, which also licenses focus.

1.1.1 Turkish

Kornfilt (1995) describes an intriguing interaction between focus and unmarked objects in Turkish. Nominative objects, which are nonspecific, are normally restricted

[1]The South Asian language Urdu is closely related to Hindi, which is mostly spoken in India. In this paper, we primarily draw our data from the dialect of Urdu spoken in Lahore, Pakistan and the dialect of Hindi-speaking informants from New Delhi, India.

[2]The data and claims for Turkish are taken mainly from Hoffman (1995), who bases her findings on a corpus gleaned from the CHILDES database (Mac Whinney and Snow (1995)), transcribed colloquial speech, and contemporary novels. Further data is taken from Kornfilt (1995) and additional fieldwork conducted independently.

to the immediately preverbal position. (1a) shows a nonspecific nominal, *üç filim*, in preverbal position, while the minimally different (1b) shows that a nonspecific nominative cannot appear elsewhere in the clause.

(1) a. ben dün [üç filim] gör-dü-m
I.Nom yesterday three film.Nom see-Past.1Sg
'I saw three (nonspecific) films yesterday.' (Turkish)

b. *[üç filim] ben dün gör-dü-m
three film.Nom I.Nom yesterday see-Past.1Sg
'I saw three (nonspecific) films yesterday.' (Turkish)

However, foci are also licensed immediately preverbally. Exactly when a focus is licensed in this way, the nominative object can occur elsewhere in the clause in Turkish, as in (2). Note that the unmarked object in these cases cannot occur simply anywhere in the clause, but is restricted to the immediately postverbal position.

(2) dün [ben]$_F$ gör-dü-m [üç filim]
yesterday I.Nom see-Past.1Sg three film.Nom
'Yesterday I saw three films.' (Turkish)

1.1.2 Urdu

There is an intriguingly similar interaction found in Urdu. As with Turkish, focused phrases occur in immediately preverbal position in Urdu, as seen in (3).

(3) naadyaa=ne hassan=ko [xat]$_F$ di-yaa
Nadya.F=Erg Hassan.M=Dat letter.M.Nom give-Perf.M.Sg
'Nadya gave Hassan a particular letter$_S$.' (Urdu)

Also as in Turkish, nonspecific nominative objects are restricted to the immediately preverbal position, as in (4).

(4) a. naadyaa=ne hassan=ko [xat di-yaa]$_F$
Nadya.F=Erg Hassan.M=Dat letter.M.Nom give-Perf.M.Sg
'Nadya gave Hassan a letter$_{NS}$.' (Urdu)

b. * example

However, when there is a preverbal focus, nominative objects either receive a specific interpretation, as seen in (3). If the nominative object is itself the focus, it is either specific, as in (3), or results in predicate focus, as in (4a).

(5) naadyaa=ne [xat] [hassan=ko]$_F$ di-yaa
Nadya.F=Erg letter.M.Nom Hassan.M=Dat give-Perf.M.Sg
'Nadya gave Hassan a particular letter$_S$.' (Urdu)

Our goal in this paper is to provide a unified account of the interaction of the seemingly disparate phenomena of case marking, discourse function, specificity, and position. The basic approach of the analysis will be to draw parallels in the licensing requirements for case and discourse functions (Horvath 1995).

1.2 Case marking, position and specificity

Consider how object case marking, in particular the appearance of nominative versus accusative case, works in Turkish and Urdu. The choice of nominative or accusative case depends on the specificity of the object. In Turkish and Urdu, as well as in a large number of typologically similar languages, there are two ways of encoding information about specificity: case marking and word order. Specificity is encoded by case marking through the accusative case in Turkish and Urdu (Enç 1991; Butt 1993). In Urdu/Hindi a "non-canonical" positioning of the object, i.e., when it is not immediately preverbal, also forces a specific interpretation (T. Mohanan 1992).

First we need to define SPECIFIC and NONSPECIFIC. Enç (1991) argues that indefinite NPs can be either specific or nonspecific, but that definite NPs are always specific. Eskenazi (1996) elaborates on this to provide the definition in (6).

(6) ... *nonspecific* nominals include only those indefinites that are neither referential nor have wide scope;
specific nominals are simply the complement of this set, which includes nominals that either refer to or presuppose a familiar entity/entities.
(Eskenazi 1996:3)

In Turkish and Urdu there are two ways of encoding information about specificity: case marking and word order. Specificity is encoded by case marking through the accusative case in Turkish and Urdu (Enç 1991; Butt 1993). In Urdu/Hindi a "non-canonical" positioning of the object, i.e., when it is not immediately preverbal, also forces a specific interpretation (T. Mohanan 1992). These facts are demonstrated below.

1.2.1 Turkish

Enç (1991) argues that accusative objects are always specific, while nominative objects are nonspecific. This can be seen in the classic example below. Both (8a) and (8b) are possible continuations of the utterance in (7). However, if the object is marked with accusative case, as in (8a), the it must be specific and the girls must be part of the group of children who entered the room. In contrast, if the object is marked with genitive case, as in (8b), the it must be nonspecific and the girls must not be part of the group of children who entered the room.

(7) odam-a birkaç çocuk girdi
my room-Dat several child.Nom entered
'Several children entered my room.' (Turkish)

(8) a. [iki kız-ı] tanıyordum
two girl-Acc know-Prog.Past.1Sg
'I knew two (of the) girls$_S$.' (Turkish)

b. [iki kız] tanıyordum
two girl.Nom know-Prog.Past.1Sg
'I knew two girls$_{NS}$.' (Turkish)

These nominative objects are generally restricted to the immediately preverbal position (see also Hoffman 1995:50–51). So, (9a) is grammatical in which the nominative, nonspecific object is immediately preverbal; if it appears before the adverb, as in (9b), then the result is ungrammatical. In contrast, an accusatively marked object can precede the adverb and be nonadjacent to the verb, as in (9c), which differs from (9b) only in the case marking of the object (and hence also its specificity).

(9) a. Nadya Hasan'a çabucak [mektup] yaz-dı.
Nadya.Nom Hasan.Dat quickly letter.Nom write-Past.3Sg
'Nadya quickly wrote (a) letter(s) to Hasan.' (Turkish)

b. *Nadya Hasan'a [mektup] çabucak yaz-dı.

c. Nadya Hasan'a [mektub-u] çabucak yaz-dı.
Nadya.Nom Hasan.Dat letter-Acc quickly write-Past.3Sg
'Nadya quickly wrote the letter to Hasan.' (Turkish)

Thus, in Turkish nonspecific objects are marked with nominative case and in immediately preverbal position; specific objects are marked with accusative case and can scramble.

1.2.2 Urdu

Now consider the situation in Urdu which differs slightly from that in Turkish. As in Turkish, accusative objects are always specific. In the context of (10), only (11a) yields a felicitous interpretation, since in (11a) the nominative NP *murɣii* 'chicken' can be interpreted as nonspecific. The accusative NP in (11b), on the other hand, can only be interpreted as specific; since no *particular* chicken or set of chickens has been established as a referent in the preceding discourse, the sentence is infelicitous.

(10) adnaan aaj raat=kii salen ke-liye murɣii cah
Adnan.M.Nom today night.F=Gen.F curry for chicken.F.Nom want

rah-aa
Stat-M.Sg

t^{h}-aa
was-M.Sg
'Adnan wanted chicken$_{NS}$ for tonight's curry.' (Urdu)

(11) a. us=ke xansaame=ne bazaar=se [murɣii]
pro=Gen.Obl cook.M.Obl=Erg market.M=from chicken.F.Nom

xarid-ii
buy-Perf.F.Sg
'His cook bought a chicken$_{NS}$ from the market.' (Urdu)

b. #us=ke xansaame=ne bazaar=se [murɣii=ko]
pro=Gen.Obl cook.M.Obl=Erg market.M=from chicken.F=Acc

xarid-aa
buy-Perf.M.Sg
'His cook bought a particular/the chicken$_{S}$ from the market.' (Urdu)

However, in Urdu, a nominative NP in the immediately preverbal position *may* be interpreted as nonspecific, but it need not be, as in (12a). When a nominative object occurs in any position other than the immediately preverbal one, it must be interpreted as specific, as in (12b/c) (see also Singh (1994)).

(12) a. naadyaa=ne hassan=ko [xat] di-yaa
Nadya.F=Erg Hassan.M=Dat letter.M.Nom give-Perf.M.Sg
'Nadya gave Hassan a (particular) letter$_{S/NS}$.' (Urdu)

b. naadyaa=ne [xat] hassan=ko di-yaa
Nadya.F=Erg letter.M.Nom Hassan.M=Dat give-Perf.M.Sg
'Nadya gave Hassan a particular letter$_{S}$.' (Urdu)

c. naadyaa=ne hassan=ko [xat] jaldii=se di-yaa
Nadya.F=Erg Hassan.M=Dat letter.M.Nom quickness=Inst give-Perf.M.Sg
'Nadya gave Hassan a particular letter$_S$ quickly.' (Urdu)

Thus, in Urdu nonspecific objects are marked with nominative case and must be in immediately preverbal position; specific objects are either marked with accusative or nominative case and in either case can scramble (see XXX on when the accusative and when the nominative case is chosen).

2 Theoretical background

Having seen the interaction between case marking, object specificity, and focus, we now introduce the technical machinery necessary to formulate the analysis of these data. First, we discuss the encoding of discourse functions in Urdu and Turkish. Then, we discuss the syntactic framework which we assume. Finally, we formalize the different object position interpretations found in these languages.

2.1 Information structure

Vallduví (1992, 1993) for Catalan and English and King (1995) for Russian argue that the traditional bipartite divisions of a sentence drawn in terms of *topic-focus*, *theme-rheme*, *old information-new information* are best understood in terms of a tripartite distinction. Vallduví views the information structure of a sentence as instructions to the hearer on how to update his/her current knowledge store. He couches the idea of a knowledge store in terms of a Heimian collection of file-cards (Heim (1982)). He defines the relevant notions in new terms in order to avoid potential confusion with existing terminologies (S = focus, ground; ground = link, tail). The FOCUS part of a sentence can be seen as an instruction to update a given file-card or to add an entirely new one. The GROUND represents the information that is already known. However, a distinction is made between the kind of information that represents a LINK, and the kind that is contained in the TAIL. The link points the hearer to the file card that is to be updated, while the tail further specifies how the new information fits onto the given file card. Here, we will refer to the link as TOPIC and the tail as BACKGROUND.

Choi (1996) proposes an extension of Vallduví's system whereby there is a four-way distinction based on two features: [±New] and [±Prominent].[3] Following Choi,

[3]Choi's use of these features is different than ours in that her division of focus differentiates

we assume that the information structure must be further divided and make the following proposal based on her two features. Topic and focus share the discourse function feature [+Prominent] which differentiate them from their relatively less prominent pairs, background and completive information respectively. Completive information and focus share the feature [+New] since they both introduce new information into the discourse. While focus's major function is to fill the informational gap between the speaker and the hearer, i.e., packaging and marking information as "new",[4] completive information provides information which is new to the hearer, and hence [+New], but which is not of primary importance to the information structure of the discourse at hand, and hence [−Prom].

(13) [+New] = FOCUS [+Prom]
COMPLETIVE INFORMATION [−Prom]
[−New] = TOPIC [+Prom]
BACKGROUND INFORMATION [−Prom]

The remainder of this section discusses how Turkish and Urdu encode these discourse functions. We connect the extension of Vallduví's proposal with the structural approach to topic and focus (É. Kiss (1995), King (1995)). The syntactic encoding of discourse functions in Turkish and Urdu (see Hoffman 1995 for Turkish; Verma (1970), Gambhir (1981), Dwivedi (1994), and Kidwai (1997) for Hindi/Urdu) can be summarized as below. In this paper we are primarily interested in the syntactic encoding of focus in preverbal position.

between CONTRASTIVE FOCUS and COMPLETIVE FOCUS. We do not discuss contrastive focus in this paper.

[4]This is what is often refered to as new-information and presentational focus (Dik et al. (1981); Rochemont (1986); Rochemont and Culicover (1990)).

Topic:	
Interpretation:	what the sentence is about ([−New,+Prom])
Position:	occurs in clause initial position: SpecIP
Focus:	
Interpretation:	provides new information relevant to the discourse ([+New,+Prom])
Position:	must appear immediately preverbally: SpecVP
Background:	
Interpretation:	provides more detailed information as to how the new focus information fits in with the already known information ([−New,−Prom]
Position:	postverbal: right adjoined to IP
Completive Information :	
Interpretation:	preverbal nontopicalized nonfocused information which is new, but is backgrounded relative to the focus ([+New,−Prom])
Position:	preverbal elsewhere case: generated in S

2.2 Syntactic structure

Out analysis exploits the architecture provided by Lexical-Functional Grammar (LFG). In the projection-based architecture of LFG a grammar is encoded as several (mathematically defined) projections which represent mutually constraining but essentially independent levels of linguistic representation. The core levels of representation, or projections, in "classic" LFG have been c(onstituent)-structure, which encodes linear word order and constituency, and f(unctional)-structure, which primarily encodes predicate-argument relations in terms of grammatical functions (SUBJ, OBJ, OBL, etc.) and head-modifier relationships. In addition to these two, a s(emantic)-structure and an a(rgument)-structure form the core projections within most current LFG analyses. Here we also exploit i(nformation)-structure which encodes the discourse functions of the sentence in context (Kaplan 1987, Choi 1996, King 1997).[5]

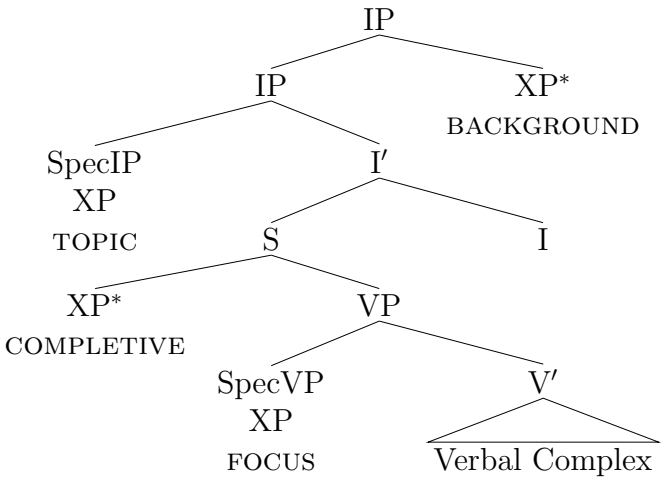
For the purposes of this paper, we assume the treatment of phrase structure presented in Bresnan (1995,1997) and King (1995). There is a basic X′ syntax with a specifier, head, and complement structure, as well as adjunction to maximal projections. Specifiers are filled either by traditional grammatical functions, e.g.,

[5]In many LFG treatments of discourse, the discourse functions are represented as part of f-structure (e.g., Bresnan 1995, King 1995) as there are some phenomena which necessitate the interaction of these two. However, we believe that information structure should be accorded its own independent status within the theory (see Choi 1996 and King 1997 for a similar conclusion).

subject or object, or by the prominent discourse functions, i.e., topic or focus. In addition, there is a lexocentric category S which behaves like a small clause in that it does not project according to the X′ schema. This S category captures the non-configurational portion of the phrase structure.[6] Positions which are not lexically realised in a given sentence are not projected in the c-structure.

For Urdu/Hindi and Turkish we posit the structure in (14). Arguments are taken to be generated under S where they receive default discourse function interpretation, and the specifiers are associated with topic (SpecIP) and focus (SpecVP). Right-adjoined to the IP is backgrounded information. Just as syntactic positions are associated with grammatical functions via functional equations, so are these syntactic positions associated with the appropriate discourse functions.[7]

(14)



2.3 Incorporating vs. non-incorporating objects

The basic pattern of object interpretation introduced in section 1 is summarized in (15). As this is a recurring pattern across languages (with slight variations), there have been a number of analyses trying to capture it. Here we briefly outline proposals by de Hoop (1992) and Van Geenhoven (1996).

[6]We adopt a version of the VP-internal subject hypothesis (Fukui and Speas 1986) in that all arguments can be generated internally, in this case internal to S.

[7]The use of functional equations on the c-structure (examples shown below in the sample analyses) in conjunction with functional uncertainty paths (Kaplan and Zaenen 1989) allows for a great deal of flexibility. In the examples in this paper we have not spelled out the functional uncertainty equations; rather, for the sake of simplicity, we have treated them as "resolved" and have simply provided the appropriate grammatical function.

(15)

Object Case	Position	Specific	Nonspecific	Language
Accusative	Anywhere	√	*	Turkish, Urdu
Nominative	Anywhere	√	*	Urdu
	Postverbal (+F)	√?	*	Turkish
Nominative	Immed. Preverbal	√	√	Urdu
	Immed. Preverbal	*	√	Turkish

De Hoop (1992) sees Turkish as an instance of a more general crosslinguistic pattern and proposes that there are two types of object case assigned in the VP: STRONG Case, which gives rise to specific interpretations, is licensed at S-structure, and allows movement; and WEAK Case, which is associated with nonspecific interpretations, is licensed at D-structure, and does not allow movement. We do not follow the particulars of de Hoop's approach here, but do base our analysis on her fundamental insight that there are two types of object position which give rise to differing interpretations.

Work done by van Geenhoven (1995, 1996), among others (e.g., Farkas (1995), McNally (1992, 1995)), shows quite clearly that a more precise semantics than that assumed by Enç is needed to provide a satisfactory analysis of the distribution of bare NPs in a variety of languages. In particular, van Geenhoven draws parallels between bare plurals in West Germanic and incorporation structures in West Greenlandic, formulating a unifying semantics for these syntactically disparate structures. We sketch van Geenhoven's proposal as far as it pertains to the Urdu and Turkish data.

Van Geenhoven (1996:137) assumes that semantically incorporating predicates absorb one (or more) predicates as restrictions on the variable representing its internal argument, as in (16a). Nonincorporating predicates are represented as ordinary n-place predicates, as in (16b).

(16) a. Incorporating: $\lambda P \lambda x \exists y$ [eat(x,y) $\wedge$ P(y)]

b. Nonincorporating: $\lambda y \lambda x$ [eat(x,y)]

The ability of bare objects to take wide scope can be tested through negation. A bare object may often initially appear to be a candidate for a specific interpretation; however, if it cannot take wide scope with respect to negation, a possible analysis involving a specific interpretation must be discarded. For example, West Greenlandic noun incorporation and West Germanic bare plurals take narrow scope with respect to negation, as illustrated in (17).

(17) a. juuna kaali-mit allagar-si-nngi-l-a-q
Junna.Abs Kaali-Abl letter-get-Neg-Ind-[−tr].3Sg
'It is not the case that Juuna got a letter/letters from Kaali'

'There is/are a letter/letters from Kaali that Juuna did not get.'

(West Greenlandic)

b. John didn't see spots on the floor.

'It is not the case that John saw spots on the floor.'

'There were spots on the floor that John didn't see.'

In addition, neither can be used to pick up a salient referent that was established in previous discourse, and neither can be used partitively, as shown in (18b) in the context of (18a) for West Greenlandic.

(18) a. nillataartisivim-mi tallima-nik manne-qar-p-u-q
fridge-Loc five-Inst.Pl egg-have-Ind-[−tr]-3Sg
'There are five eggs in the fridge.' (West Greenlandic)

b. jensi marlunnik manni-tu-saa-a-q
Jensi-Abs two egg-eat-Fut-[−tr]-3Sg
'Jensi will eat two eggs.'

'Jensi will eat two of the eggs.' (West Greenlandic)

These nouns are still discourse transparent, as is well-known from the literature on noun incorporation. However, when the incorporated nouns or bare plurals are embedded under negation, as in (17), they become opaque for discourse purposes, indicating that these nominals do not introduce a referent of their own. Rather, as argued by van Geenhoven (1996), they must be interpreted as properties, i.e., predicatively (for similar argumentation see also de Hoop (1992), Ramchand (1993, forthcoming) for Scottish Gaelic and Bengali, and McNally (1996) for Spanish).

There is a three-way syntactic distinction between these two types of NPs. NPs corresponding to arguments of nonincorporating predicates are essentially strong case NPs which are not restricted in position. In contrast, nonspecific arguments of incorporating predicates may or may not be syntactically incorporated. The difference between the syntactically incorporating and syntactically nonincorporating languages is motivated by language-dependent factors. West Germanic, for example, merely requires syntactic adjacency for semantic incorporation. West Greenlandic, however, requires morphological adjacency. In sum, van Geenhoven proposes a three-way syntactic distinction and a two-way semantic distinction, which gives rise to the distribution shown in (19).

(19)

Semantics	**Case**	**Syntactic Position**
incorporated	weak	adjacent to verb morphologically incorporated
nonincorporated	strong	free

The fact that objects in the preverbal weak case position are assigned a predicative, and hence nonspecific, interpretation is exactly in line with the data seen for Urdu and Turkish. In LFG, we encode the above generalizations via Lexical Mapping Theory (LMT). In particular, following XXX we assume that grammatical functions are decomposable into the features:

(20) [± r(estricted)] [± o(bject)]

Following Ramchand 1993, we additionally assume that the two types of object (those grammatical functions which are [+o]) are interpretated as follows. Restricted objects, OBJ_θ ([+r]), are semantically *enriched*; this corresponds to strong case. Non-restricted objects, OBJ ([−r)], are only associated with a nonspecific semantic interpretation; this corresponds to weak case.

(21)

	Position	**Role**	**Specific**	**Case**
Strong	Anywhere	OBJ_θ	+	Acc/Nom
Weak	Verb Adjacent	OBJ	−	Nom

3 Modeling the interaction

Our basic analysis is as follows.[8] Focus and OBJ are licensed in the c-structure (see Horvath 1995 on the similarity of focus and case licensing mechanisms). In particular, they are both licensed in SpecVP. As seen in the VP rules in (22), the constituent in SpecVP is always the focus and is optionally assigned the role of OBJ. In Urdu and Turkish, this is the only way in which the OBJ grammatical function is assigned; as such an unrestricted object can only occur in immediately preverbal

[8]Kidwai (1997) independently notices the same problem and formulates essentially the same generalizations in a minimalist framework. Her approach differs from ours in that she proposes that there is scrambling via movement which is driven by feature-checking and that agreement is done via Spec-head relations in AgrP both VP-internally and VP-externally.

Reinhart (1996) and Neeleman and Reinhart (in press) argue that Case Checking is preferentially done within a Prosodic Phrase (in VO languages), which leads to an adjacency requirement on case assignment. We are primarily concerned with understanding the interaction between nonspecificity and focus, rather than the interaction between definiteness and scrambling, and hence their approach is not immediately comparable to ours.

position. This contrasts with other grammatical functions in these languages which are freely assigned and are constrained via the case markers (REFs: Butt, Butt and King, Nordlinger). As such, restricted objects, OBJ_θ, are base generated in any position in the c-structure.

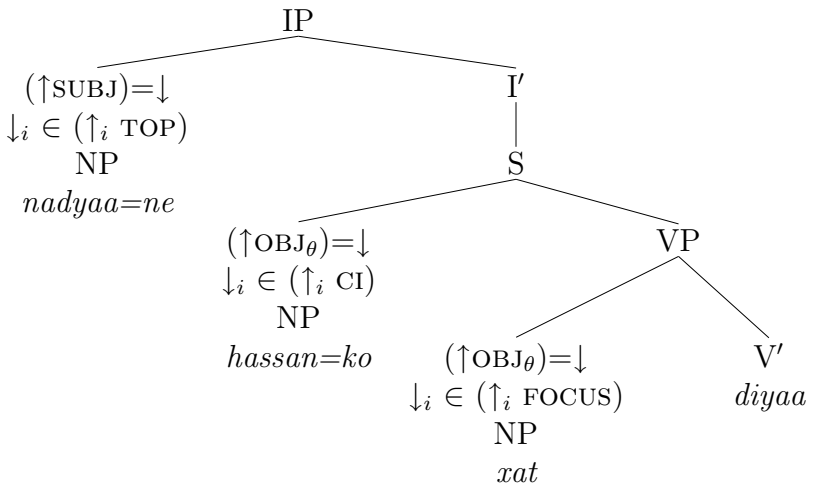
(22) VP $\longrightarrow$ XP (↑FOCUS)=↓ ((↑OBJ)=↓) V′ ↑=↓

A constituent in SpecVP is obligatorily interpreted as focused. If a restricted object is generated in SpecVP, it is the focus of the clause and is interpreted as a focused argument. If a non-restricted (and hence semantically incorporated) object is in SpecVP, the result is predicate focus since the object is semantically part of the verbal predicate. These two situations are illustrated below.

First consider argument focus in which a specific restricted object, an OBJ_θ, is in SpecVP, as in (23a). The annotated c-structure corresponding to this sentence is shown in (23b). Note that arrows without an subscript describe projections to the f-structure and ones with subcript i to i-structure.

(23) a. naadyaa=ne hassan=ko $[\text{xat}]_F$ di-yaa
Nadya.F=Erg Hassan.M=Dat letter.M.Nom give-Perf.M.Sg
'Nadya gave Hassan a particular letter$_S$.'

b.



The f- and i-structures corresponding to the c-structure in (23b) are shown in (24a) and (24b) respectively.[9] The f-structure shows the main predicate and its arguments,

[9]Note that the f- and i-structures of a sentence are not isomorphic. For example, predicates which are in a set in i-structure need not be in f-structure. As multiple topics and foci are in principle possible, we consider the discourse functions to have sets as a value.

the subject and two OBJ_θs. In the i-structure, the focus corresponds to the OBJ_θ in preverbal position in the c-structure.

(24) a. **Functional structure:**

$$\begin{bmatrix} \text{PRED} & \text{'give<SUBJ,OBJ}_\theta\text{, OBJ}_\theta\text{>'} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'NADYA'} \\ \text{CASE} & \text{ERG} \end{bmatrix} \\ \text{OBJ}_\theta & \begin{bmatrix} \text{PRED} & \text{'HASSAN'} \\ \text{CASE} & \text{DAT} \end{bmatrix} \\ \text{OBJ}_\theta & \begin{bmatrix} \text{PRED} & \text{'LETTER'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \end{bmatrix}$$

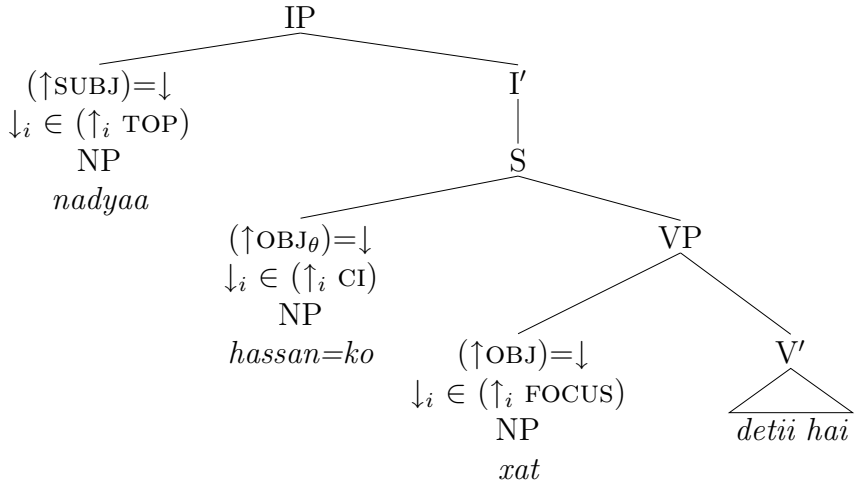
b. **Information structure:**

$$\begin{bmatrix} \text{TOPIC} & \left\{\begin{bmatrix} \text{PRED} & \text{'NADYA'} \end{bmatrix}\right\} \\ \text{FOCUS} & \left\{\begin{bmatrix} \text{PRED} & \text{'LETTER'} \end{bmatrix}\right\} \\ \text{COMP.INF} & \left\{\begin{bmatrix} \text{PRED} & \text{'HASSAN'} \end{bmatrix}\right\} \end{bmatrix}$$

Next consider predicate focus in which a nonspecific unrestricted object, an OBJ, is in SpecVP, as in (25a). The annotated c-structure corresponding to this sentence is shown in (25b).

(25) a. naadyaa hassan=ko [xat de-tii hai]$_F$
Nadya.F.Nom Hassan.M=Dat letter.M.Nom give-Impf.F.Sg be.Pres.3Sg
'Nadya gives Hassan a letter$_{NS}$/letters$_{NS}$.'

b.



The f- and i-structures corresponding to the c-structure in (25b) are shown in (26a) and (26b) respectively. The f-structure shows the main predicate and its arguments, the subject, OBJ_θ, and OBJ. In this case, the focus in the i-structure is the predicate *write* and its OBJ *letter*, shown here as *letter-write* reflecting the semantic incorporation which occurs with OBJ.

(26) a. **Functional structure:**

$$\begin{bmatrix} \text{PRED} & \text{'give<SUBJ,OBJ}_\theta\text{, OBJ>'} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'NADYA'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{OBJ}_\theta & \begin{bmatrix} \text{PRED} & \text{'HASSAN'} \\ \text{CASE} & \text{DAT} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'LETTER'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \end{bmatrix}$$

b. **Information structure:**

$$\begin{bmatrix} \text{TOPIC} & \left\{\begin{bmatrix} \text{PRED} & \text{'NADYA'} \end{bmatrix}\right\} \\ \text{FOCUS} & \left\{\begin{bmatrix} \text{PRED} & \text{'LETTER-WRITE'} \end{bmatrix}\right\} \\ \text{COMP.INF} & \left\{\begin{bmatrix} \text{PRED} & \text{'HASSAN'} \end{bmatrix}\right\} \end{bmatrix}$$

Finally, consider what happens if a nominative object is interpreted as nonspecific, as in (27a). The result is ungrammatical. This falls out of our analysis because

a non-specific object must be an OBJ, not an OBJ_θ. However, there is no way for it to receive this grammatical function except in SpecVP.[10]

(27) a. # naadyaa=ne [xat] [hassan=ko]$_F$ di-yaa
Nadya.F=Erg letter.M.Nom Hassan.M=Dat give-Perf.M.Sg
'Nadya gave Hassan a letter$_{NS}$.'

b. NO WELL-FORMED C- AND F-STRUCTURE

4 Further predictions

This section examines two predictions of this analysis: the distribution of nominative objects in Turkish when there is a preverbal focus; and the interpretation of N-V complex predicates in Urdu.

4.1 The Turkish "fleeing" effect

This analysis predicts that Turkish should exhibit a situation similar to that of Urdu. That is, when there is a focus phrase in SpecVP, it should be possible to have a nominative object elsewhere in the clause as long as is is not interpreted nonspecifically. This is usually claimed to be limited to the situation in which the nominative object appears post-verbally in the presence of a a preverbal focus, as seen in (28). In (28a), the previous discourse may have been talking about films, some films. The postverbal *üc filim* 'three films' here could be picking up on the general categories of "films", but need not introduce three particular films as discourse referents.

(28) a. dün [ben]$_F$ gör-dü-m [üç filim]$_{Back}$
yesterday I.Nom see-Past.1Sg three film.Nom
'Yesterday I saw three films.' (Turkish)

b. dün [kim]$_F$ gör-dü [üç filim]$_{Back}$
yesterday who.Nom see-Past.3Sg three film.Nom
'Who saw three films yesterday?' (Turkish)

However, the situation in Turkish is more complex (Hoffman p.c.) and we will explore this in more detail.

[10]Note that the inverse case, where a nonspecific interpretation is assigned to an accusative NP is out because the accusative case marker constrains its NP to be an OBJ_θ.

As noted by Kornfilt (1995), Turkish nominative objects cannot generally appear in a position other than the immediately preverbal one. Kornfilt argues that, in fact, nominative objects can only ever appear in a different position when the immediately preverbal position is focused, and then only in the immediately postverbal position. She proposes an explanation in which the focus feature allows Case assignment to the right, thus licensing immediately postverbal nominative objects. However, this proposal faces empirical difficulties: nominative NP objects do not in fact have to appear immediately postverbally, as (29) illustrates. In addition, postverbal nominative NPs are referential, not nonspecific, in contrast to their preverbal cousins, as predicted by our analysis.

(29) a. gönderdim Ayşe'ye [üc kitap]
send.Past Aeysha-Dat three book.Nom
'I sent Aeysha three books.' (Turkish)

b. Ayşe'ye gönderdim Ali [üc kitap]
Aeysha-Dat send.Past Ali three book.Nom
'Ali sent Aeysha three books.' (Turkish)

(30) dün [ben]$_F$ gör-düm [bir filim]$_{Back}$
yesterday I.Nom see-Past.1Sg one film.Nom
'Yesterday I saw one (particular) film$_S$.' (Turkish)

This obligatory difference in interpretation in non-immediately preverbal position is supported by the fact that truly bare NPs like *filim* do not lend themselves to a specific interpretation, as in (31),and hence only felicitously appear in immediately preverbal position. This is because postverbal (background) position requires referentiality ([− New]) and in Turkish bare NPs resist referentiality.

(31) ??dün [ben]$_F$ gör-düm [filim]$_{Back}$
yesterday I.Nom see-Past.1Sg film.Nom
'Yesterday I saw some film(s)$_{NS}$.' (Turkish)

Nominative NPs are also possible in preverbal contexts other than immediately preverbal, as in (32). In these cases, as predicted, the NPs receive a specific interpretation, as evidenced by the contrast between (32) and (33). In (33) a habitual reading is forced by the quantifier *her yıl* 'every year' and the habitual marking on the verb: the resulting sentence is unacceptable since the habitual reading requires a predicative, nonspecific interpretation of the object.

(32) ?Ayşe'ye, [üc kitap] ben, [beş kitap] da kardeşim
Aeysha-Dat three book.Nom I.Nom five book.Nom too my sister.Nom

yolladı
send.Past
'I sent Aeysha three books, and my sister sent her five books.' (Turkish)

(33) *her yıl Ayşe'ye [kitap] ben [mecmua] da kardeşim
every year Aeysha-Dat book.Nom I.Nom magazine.Nom too my sister.Nom

yollar
send.Hab
'Every year, I send Aeysha books, and my sister sends her magazines.'(Turkish)

As seen above, there is an asymmetry in acceptability between postverbal and not immediately preverbal nominative NPs in Turkish, unlike in Urdu. We suggest that this follows from the difference in discourse requirements between preverbal material and the postverbal backgrounded material. The semantic burden on the preverbal material is heavier than on the postverbal background information: the preverbal NPs must either introduce a new referent into the discourse, or refer to a definite or specific entity (topicalization and strong case, respectively). The postverbal NP merely has to pick up on another referent already mentioned previously, but not necessarily a specific entity (i.e., it may be partitive without it being clear *which* of the set of items is being referred to). So, in (34), taken from Eskenazi (1996), where a context is set up in which there is a specific book the speaker has in mind, the nominative NP is good in a position that is not immediately preverbal.

(34) a. Ali says no one ever sent him any books, but I know he was sent a book because ...

b. ona (bir) kitap [ben]$_F$ yolladım
he.Dat one book.Nom I.Nom send.Past.1.Sg
'I sent him (a) book$_S$.' (Turkish)

Thus, Turkish is very much like Urdu after all. Nominative NPs can appear in positions other than the immediately preverbal one. Such NPs cannot be interpreted nonspecifically. Truly bare NPs do not lend themselves to a referential interpretation and thus are largely restricted to the immediately preverbal position (OBJ).

4.2 Semantic incorporation and N-V complex predicates

This way of interpreting nominative preverbl objects is reminiscent of the type of constructions that have been discussed by T. Mohanan (1995) for Hindi from the point of view of noun incorporation. T. Mohanan argues that the example in (35) must be analyzed as an instance of noun incorporation since it exhibits semantic, syntactic and phonological behavior which differs from that of canonical transitives.

(35) anil [gʰoṛe bec-taa hai]$_F$
Anil.M.Nom horse.M.Pl sell-Impf.M.Sg be.Pres.3.Sg
'Anil does horse-selling.'

However, rather than seeing these examples as syntactic noun incorporation, we propose to analyze them as semantic incorporation whose bare objects are base generated as unrestricted, nonspecific objects (OBJ) and hence cannot scramble. Note that the obligatory focussing of the semantically incorporated object results in predicate focus. The fact that the nominative object is in SpecVP and hence focused is seen by examples like (36). In (36), the locative *lahor=mẽ* cannot be the focus of the clause since the object already occupies SpecVP; instead, it can only be a constrastive focus, which in Urdu are licensed in situ via prosodic factors.

(36) anil [lahor=mẽ]$_{CF/*F}$ [gʰoṛe bec-taa hai]$_F$
Anil.M.Nom Lahore=in horse.M.Pl sell-Impf.M.Sg be.Pres.3.Sg
'Anil does horse-selling in Lahore.'

5 Conclusion

The difference between the distribution of nominative NP objects in Turkish and Urdu was thus seen to be minimal. The apparently conflicting functions of the immediately preverbal position in Urdu and Turkish argue for two independent constraints: a differentiation between semantically incorporating and non-incorporating objects and the licensing of focus in a way similar to that of case, thus confirming the close association between focus and Case licensing in two genetically unrelated but typologically similar languages, and perhaps in other head-final languages as well.

References

Bresnan, J. 1995. *Morphology Competes with Syntax: Explaining Typological Variation in Weak Crossover Effects.* In P. Barbosa, et al. (eds.) *Is the Best Good Enough? Proceedings from the Workshop on Optimality in Syntax.* MIT Press.

Butt, M. 1993. Object specificity and agreement in Hindi/Urdu. In *Papers from the 29th Regional Meeting of the Chicago Linguistic Society*, 89–103.

Choi, H. 1996. *Optimizing Structure in Context: Scrambling and Information Structure.* PhD thesis, Stanford University.

Dwivedi, V. 1994. *Syntactic Dependencies and Relative Phrases in Hindi.* PhD Thesis, University of Massachusetts at Amherst.

de Hoop, H. 1992. *Case Configuration and Noun Phrase Interpretation.* Phd Thesis, Rjiksuniversiteit Groningen.

Enç, M. 1991. The semantics of specificity. *Linguistic Inquiry* 22(1):1–25.

Eskenazi, J. 1996. Bare objects and the NP/DP distinction. Unpublished Ms., University of California, Santa Cruz.

Gambhir, V. 1981. *Syntactic Restrictions and Discourse Functions of Word Order in Standard Hindi.* PhD Thesis, University of Pennsylvania, Philadelphia.

Hoffman, B. 1995. *The Computational Analysis of the Syntax and Interpretation of "Free" Word Order in Turkish.* Phd Thesis, University of Pennsylvania, Philadelphia.

Horvath, J. 1995. Structural focus, structural case, and the notion of feature-assignment. In K.É. Kiss (ed.) *Discourse Configurational Languages*, 28–64. Oxford, Oxford University Press.

Kidwai, A. 1996. Word Order and Focus Positions in Universal Grammar. Paper presented at the *Table ronde internationale sur la Grammaire du focus.* U. Paris II and U. Paris X, February.

Kidwai, A. 1997. *Scrambling and Binding in Hindi-Urdu.* PhD thesis, Jawaharlal Nehru University, New Delhi.

King, T.H. 1995. *Configuring Topic and Focus in Russian.* Stanford, CA: CSLI Publications.

Kornfilt, J. 1995. *On Focusing and De-Focusing in Turkish.* Paper presented at the Generative Grammatik im Süden (GGS), Universität Jena.

Mohanan, T. 1992. Word Order in Hindi. Presented at Syntax Workshop, Stanford University.

Mohanan, T. 1995. Wordhood and lexicality: Noun incorporation in Hindi. *Natural Language and Linguistic Theory* 13(1):75–134.

Neeleman, A. and T. Reinhart. In press. Scrambling and the PF Interface. In M. Butt and W. Geuder (eds.) *The Projection of Arguments: Lexical and Compositional Factors.* Stanford, CA: CSLI Publications. To appear in 1998.

Reinhart, T. 1996. Interface Economy: Focus and Markedness. In Christopher Wilder, Hans-Martin Gärtner, and Manfred Bierwisch (eds.) *The Role of Economy Principles in Linguistic Theory.* Berlin: Akademie Verlag.

Vallduví, E. 1992. *The Informational Component.* New York: Garland Press.

Van Geenhoven, V. 1996. *Semantic Incorporation and Indefinite Descriptions.* Phd Thesis, University of Tübingen.

Verma, S.K. 1970. *Word Order in Hindi.* Achiv Orientalnini, Number 38:29–32.