A program for case features

Daniel Harbour—Queen Mary, University of London

$\mathfrak{WOTM3}$, $\mathfrak{Leipzig}$—June 2007

1. Aim

Aim 1. To argue that there are three features [±location], [±source], [±goal] and that these are the components of oblique case categories such as ablative, allative, locative, instrument, ... **Aim 2.** To compare this approach to case with that of Jakobson 1936 and contextualize in a broader research programs both for features and for case.

2. Rough syntactic/semantic outline

Syntactically, I suppose that there is a head that creates a semantic relation between two objects: $\mathrm{R}(x, y)$. In its bare state, R is semantically underspecified. However, it can be refined, by the features above, so that one of the objects is specified as being, say, a location that is a source and not a goal [+location +source −goal]. For current purposes, I'll be quite vague about the relation: the notion of applicative, as elaborated by Cuervo, or the notion of path, explored by Di Sciullo, will serve current purposes.

$[\![\mathrm{R}]\!] = \lambda x\ \lambda y \in D_e\ .\ x$ is in relation with y

$[\![+\mathrm{location}]\!] = \lambda x\ .\ x$ is a location

$[\![+\mathrm{source}]\!] = \lambda x\ .\ x$ is a source

$[\![+\mathrm{goal}]\!] = \lambda x\ .\ x$ is a goal

$[\![-\mathrm{F}]\!] = \neg[\![+\mathrm{F}]\!]$

$[\![\alpha\mathrm{F}\ \alpha'\mathrm{F}']\!] = [\![\alpha\mathrm{F}]\!] \wedge [\![\alpha'\mathrm{F}']\!]$

E.g., $[\![\mathrm{R}\ [+\mathrm{location}]]\!] = \lambda x\ \lambda y \in D_e\ .\ x$ is in relation with y and x is a location

3. Ambulatory introduction: Papua New Guinea

All data in this section from Foley 1986 (The Papuan Languages of New Guinea): 100−1

Some languages lead us to think of P elements as atomic, i.e., they lead us to posit four features [instrument], [source], [location], [goal].

(1)

Case/Preposition	**Alamblak**	**English**
LOCATIVE	-n	at
ABLATIVE	-pnë	from
ALLATIVE	-ko	to
INSTRUMENTAL	-e	with

However, other languages suggest overlap between these categories and so that they are not atomic but composite.

(2)

Case	Kâte	Selepet	Kunimaipa
LOC	**-o**	**-ɔn**	**-ha**
ABL	**-o**-nek	**-ɔn**-gebɔ	**-ha-nanga**
ALL	**-o**-pek	**-ɔn**-gen	-ti
INSTR	-zi	-ŋe	**-nanga**

3.1. Kâte (and Selepet)

In Kâte, **-o** is common to ABL, LOC and ALL. In fact, LOC **-o** is a substring of ABL **-o-nek** and ALL **-o-pek**. (This is also so for Selepet, *mutatis mutandis*.) So:[1]

(3)

Case	Features
LOC	[location]
ABL	[location source]
ALL	[location goal]
INSTR	[instrument]

Vocabulary Items: Kâte		
[location]	⇔	**-o**
[goal]	⇔	**-pek**
[source]	⇔	**-nek**
[instrument]	⇔	**-zi**

3.2. Kunimaipa

Kunimaipa shows that the instrumental is not atomic either: it, **-nanga**, is a substring of the ablative, **-ha-nanga**. So, we have to revise (3) to reflect this relationship:

(4)

Case	Features
LOC	[location]
ABL	[location source]
ALL	[location goal]
INSTR	[source]

Vocabulary Items: Kunimaipa		
[location]	⇔	**-ha**
[location goal]	⇔	**-ti**
[source]	⇔	**-nanga**

(By the Subset Principle, as **-ti** expresses [location goal], it bleeds occurrence of **-ha**.)

3.3. Kâte (and Selepet) again

The feature inventory revised for Kunimaipa now creates problems for Kâte (and Selepet). If INSTR is not [instrument] but [source], then, when we update (3), we have:

(5)

Case	Features
LOC	[location]
ABL	[location source]
ALL	[location goal]
INSTR	[source]

Vocabulary Items: Kâte		
[location]	⇔	**-o**
[goal]	⇔	**-pek**
[source]	⇔	**-nek**
[source]	⇔	**-zi**

[1] It is unclear, on synchronic grounds, whether to decompose Kâte **-pek**/**-nek** into **-p-ek**/**-n-ek**, Selepet **-gen**/**-gebɔ** into **-ge-n**/**-ge-bɔ** and to try to correlate the velars in each case with a single feature (with **p**/**bɔ** and **n** having come to play opposite roles in the two languages). Such a feature enjoys no support from the other languages considered here.

There are two entries for [source]. So, either this should create optionality or ineffability.

To fix this, we could posit another feature, but it is unclear what its semantics would be. A simpler solution (given previous work on features) is bivalence, which makes it possible to refer to 'not being a location', as well as 'to being a location'. That is: we upgrade from [F] to [±F].

(6)

Case	Features	Vocabulary Items: Kâte		
LOC	[+location −source −goal]	[+location]	⇔	**-o**
ABL	[+location +source −goal]	[+goal]	⇔	**-pek**
ALL	[+location −source +goal]	[+source]	⇔	**-nek**
INSTR	[−location +source −goal]	[−location +source]	⇔	**-zi**

The crucial point is that [±locative] enables us to distinguish INSTR and ABL. However, we can retain the overlap for Kunimaipa, whilst also retaining exactly the same feature specifications for the LOC, ABL, *etc.*

(7)

Case	Features	Vocabulary Items: Kunimaipa		
LOC	[+location −source −goal]	[+location]	⇔	**-ha**
ABL	[+location +source −goal]	[+location +goal]	⇔	**-ti**
ALL	[+location −source +goal]	[+source]	⇔	**-nanga**
INSTR	[−location +source −goal]			

4. Typology

Let's call the set of features just proposed X = {[±location], [±source], [±goal]} (X for Greek $\chi\omega\rho\alpha$ 'space (partially occupied)').

Principles of linguistic variation:

(8) For each feature category, such as X (or person, or number features), languages may differ in which subset of X they use. (There are no *ad hoc* combinatorial stipulations, or 'geometries' as they are euphemized.)

(9) A language exploits its X-features up to contradiction—that is, every non-contradictory feature bundle is a legitimate χ-value—provided that the language does not treat one such χ-category by different means.[2]

Hence, we predict 8 different X-systems:

(10)
a. { }
b. {[±location]}
c. {[±source]}
d. {[±goal]}
e. {[±location], [±source]}
f. {[±location], [±goal]}
g. {[±source], [±goal]}
h. {[±location], [±source], [±goal]}

[2]This proviso is intended to cover cases like, e.g., Dyirbal causatives, which are expressed clausally: the relevant features are not collocated with nominals, only with clauses. Presumably, the pronunciation of these features yields a complementizer, rather than a case suffix.

Within the most complex, we expect 8 different feature bundles, corresponding to 8 different oblique cases:

(11)

[+location +source +goal]	[−location +source +goal]
[+location +source −goal]	[−location +source −goal]
[+location −source +goal]	[−location −source +goal]
[+location −source −goal]	[−location −source −goal]

4.1. [+source +goal]

I want to suggest that [+source +goal] is contradictory. For something to be both a source and a goal suggests circular motion. However, I'm not aware of any language for which this is an oblique case category. Equally tellingly, there do not seem to be languages with directional systems (which, I assume, are also X-systems) in which circular motion is represented—even in the most complex cases.

This means that languages have maximally a six-way use of the features:

(12)

[+location +source −goal]	[−location +source −goal]
[+location −source +goal]	[−location −source +goal]
[+location −source −goal]	[−location −source −goal]

The new categories are [−location −source −goal] and [−location −source +goal]. If we confine attention to physical objects, it is hard to think what these could mean. The latter would seem to be the answer to *What is x for?*. But this seems to require a verb-based answer (e.g., *These boots are made for walking*). The former is just an object without any specific qualities, and so, would seem to be of little use. However, again, if we think of events, an interpretation again becomes available (see next section).

4.1.1. Kemtuk

In Kemtuk (data from Foley 1986: 149−52), 'the primary distinction is between direction towards [the] speaker and away from the speaker, but within this grand division, there is a further distinction between same level as or different level from the speaker and, if the latter, motion up or motion down.'

(13) **Kemtuk directionals/elevationals**

motion towards			motion away					
speaker's level	not speaker's level		near			far		
			speaker's level	not speaker's level		speaker's level	not speaker's level	
	up	down		up	down		up	down
-san	**-ban**	**-kan**	**-Ø**	**-bi**	**-si**	**-sa**	**-ba**	**-na**

So, for instance, **-si** indicates motion that is away from the speaker, to a location nearby, on a different level from the speaker, and downwards. Think of yourself up a ladder and passing a book down to someone.

(14) genam mea iti- **si**-l- u
I FUT give-SI-FUT-I
'I will give it to you down there near here'

4.1.2. Georgian

Strikingly, in Georgian (*Georgian: A Structural Reference Grammar*, Hewitt 1995: 148−9), which has hither and thither directionals, **mo-**, **mi-**, these are permitted to cooccur. However, the meaning is still not one of circularity, but of toing and froing.

(15) Mi-mo-di- s tav-is otax- ši
fro-to- go-3SG.PRES self-GEN room-in
'He is pacing back and forth in his room'

Notice that this requires a multiplicity of events. Generally, when such elements combine, they qualify a single event: **ča-mo-di-s** signifies that his going (**dis**) is both hither (**mo**) and downwards (**ča**): 'He's coming down'. Circular motion would be a single event and, so, more in line with the general system; yet, notwithstanding, this singular-event interpretation is not available.

4.2. Zero-feature systems: Yimas

If a language eschews all X features, then R alone is active and, so, realizable. Such a language is Yimas, where all ablative, allative, etc, are all pronounced as **-ɨn**/**-nan**. $X_{Yi.} = \{\ \}$.

4.3. One-feature systems: Kewa, Dani, Iatmul

Kewa conflates ABL, ALL, LOC, contrasting this supercategory with INSTR. These correspond to the natural classes defined by [±location]: [+location] = {ABL, ALL, LOC}, [−location] = {INSTR}. So, we can capture the distinctions made by Kewa by supposing that it $X_{Ke.} = \{[\pm\text{location}]\}$.

(16)

Case	**Features**
LOC/ABL/ALL	[+location]
INSTR	[−location]

Vocabulary Items: Kewa		
[+location]	⇔	**-mé**
[−location]	⇔	**-para**

Dani conflates ABL with INSTR, and ALL with LOC. This leads us to $X_{Da.} = \{[\pm\text{source}]\}$.

(17)

Case	**Features**
ABL/INSTR	[+source]
LOC/ALL	[−source]

Vocabulary Items: Kewa		
[+source]	⇔	**-nen**
[−source]	⇔	**-ma**

Iatmul conflates ABL, INSTR and LOC, and contrasts these only with ALL. This leads us to $X_{Ia.} = \{[\pm\text{goal}]\}$.

(18)

Case	**Features**
ABL/INSTR/LOC	[−goal]
ALL	[+goal]

Vocabulary Items: Kewa		
[+goal]	⇔	**-(ŋk)ət**
[−goal]	⇔	**-mpa**

4.4. Two-feature systems

Damn...

4.5. Plenary systems: Australian P-like cases

Dixon (*The Languages of Australia*, 1980: 298−302) describes a variety of Australian languages with the cases LOCATIVE, ALLATIVE, ABLATIVE. These languages also have DATIVE, PURPOSIVE, CAUSAL. Dixon suggests that the two triplets are in 'semantic correspondence' (my paraphrases below):

(19)

	Location	**Event**	
LOC	location of/at entity	event for/at entity	DAT
ALL	motion leading to entity	action leading to event	PURP
ABL	motion coming from entity	action coming from event	CAUS

(20)

Case	**Features**	**Features**	**Case**
LOC	[+location −source −goal]	[−location −source −goal]	DAT
ALL	[+location −source +goal]	[−location −source +goal]	PURP
ABL	[+location +source −goal]	[−location +source −goal]	CAUS

(Note that CAUSE supplants INSTRUMENTAL as the latter is not an event. Such languages must therefore handle INSTR in a different fashion. They do so by conflating it with either LOC, as in Yidiny and Western Desert, or ERG, as in Dyirbal or Warlpiri. In the former case, LOC must therefore be viewed as a default—for which Dixon argues on general semantic grounds. The treatment of CAUSE in the Papuan languages above is neatly captured. If causes are non-locational sources, then we predict, correctly, that Dani conflates ABL/CAUS/INSTR, Kewa CAUS/INSTR, Iatmul ABL/CAUS/INSTR/LOC. Yimas folds it into its default category.)

(21)

Warlpiri			
LOC	-ngka∼-rla	-ku	DAT
ALL	-kurra	-ku	PURP
ABL	-ngurlu	-jangka	CAUS

Vocabulary Items: Warlpiri		
[+location +source]	⇔	**-ngurlu**
[+location +goal]	⇔	**-kurra**
[+location]	⇔	**-ngka∼-rla**
[−location +source]	⇔	**-jangka**
[−location]	⇔	**-ku**

(22)

Yidiny			
LOC	-da∼-la	-nda	DAT
ALL	-da∼-la	-gu	PURP
ABL	-mu∼-m	-mu∼-m	CAUS

Dyirbal			
LOC	-da∼-ŋga	-gu	DAT
ALL	-gu	-gu	PURP
ABL	-ŋunu	—[3]	CAUS

Therefore $X_{Wa.} = \{[\pm\text{locative}], [\pm\text{source}], [\pm\text{goal}]\}$ $(= X_{Yi.} = X_{Dy.})$

[3]Expressed by full clause.

5. Broader conceptualization

5.1. Jakobsonian case decomposition

(23) **Jakobson's 1936 *Kasuslehre***

Case	[±subject]	[±governed]	[±oblique]
NOM	+	−	−
ACC	−	+	−
GEN	+	+	+
LOC	−	−	+
DAT	−	+	+
INSTR	+	−	+

Objection 1. Jakobsonian features are unlike other features. They are internal to the syntax. Indeed, they redundantly recapitulate what the syntax states. **Objection 2.** Cases have semantic import. Informally, we want statements such as $[\![\text{INSTR}]\!] = \lambda x.\ \text{source}(x) \wedge \neg\text{location}(x)$. How do we get from features descriptive of syntactic structure to truth conditions in the model? Intuitively:

$$[\![\ \begin{array}{ccc} & \text{X} & \\ \text{Y} & & \text{X} \\ & \text{X} \quad \text{Z} & \end{array}\ ]\!] \neq \text{source}([\![\text{x}]\!]) \wedge \neg\text{location}([\![\text{x}]\!]).$$

5.2. Grammatical Case alternations

General point: oblique cases are grammatically related (syntactically or morphologically) with core cases (accusative, ergative, nominative, ...)

5.2.1. Case alternations

Partee: accusative∼partitive in Finnish; Svenonius: dative∼accusative in Icelandic; ...

(24) Russian genitive of negation

a. Ja nashël knig**u**
I found book.**ACC**
'I found a book'

b. Ja ne nashël knig**i**
I not found book.**GEN**
'I didn't find a book'

(25) Slavic instrumental

a. Marcellus byl soldat
Marcellus.nom was soldier
'Marcellus was a soldier'

b. Moĭ otec byl soldat**om**
my father was soldier.**INSTR**
'My father was a soldier'

Coulson 1992 *Sanskrit: A complete course for beginners*, ch5:

(26) a. sa lekhaṃ likhita-vān
3**MASC.NOM** letter.MASC.SG.**ACC** written.PPART-have.**MASC.SG**
'He has written a letter'
b. tena likhitaḥ lekhaḥ
3MASC.**INSTR** written.PPART.**MASC.SG** letter.**MASC.SG.NOM**
'He has written a letter' (tena likhito lekhaḥ)

"There are, however, a number of past participles that may have both an active and a passive sense. Thus *pīta*, like 'drunk' in English, can be applied both to the drink and to the drinker (though in Sanskrit there is no necessary implication of intoxication). Similarly, *praviṣṭa* 'entered' or 'having entered', *vismṛta* 'forgotten' or 'having forgotten'. Thus with an active construction [(27a)] and with a passive construction [(27b)], the meaning of both versions being 'And Rāma entered the city'."

(27) a. Rāmaḥ api nagaram praviṣṭaḥ
Rāma.**MASC.SG.NOM** and city.**NEUT.ACC** enter.PPART.**MASC.SG**
'Rama entered the city' (Rāmopi nagarampraviṣṭaḥ)
b. Rāmena api nagaram praviṣṭaṃ
Rāma.MASC.SG.**INSTR** and city.**NEUT.NOM** enter.PPART.**NEUT.SG**
'Rama entered the city' (Rāmenāpi nagarampraviṣṭam)

5.3. Case syncretism

(28) Sanskrit case collapse:
a. {NOM}, {ACC}, {INSTR}, {DAT}, {ABL}, {GEN}, {LOC} singular
b. {NOM ACC}, {INSTR DAT ABL}, {GEN LOC} dual
c. {NOM}, {ACC}, {INSTR}, {DAT ABL}, {GEN}, {LOC} plural

Semantic overlap:
a. {ACC}, {INSTR}, {ABL (GEN)}, {LOC} time
b. {GEN (DAT)} giving

(29) Ergative is a structural case (Bruening 2007) but, in Dyirbal and Kalkatungu, ERG = INSTR

6. Conclusion

Conclusion 1. The featural components of semantically contentful cases are semantically contentful. **Conclusion 2.** Semantically contentful cases are syntactically and morphologically related to structural cases and, so, form a unified system.

Final note. Why this is a program: exhaustive typology, exhaustive case inventory.