SynSem research group
Oslo Center for Advanced Studies (CAS)

Mary Dalrymple
University of Oxford

6–8 February 2017

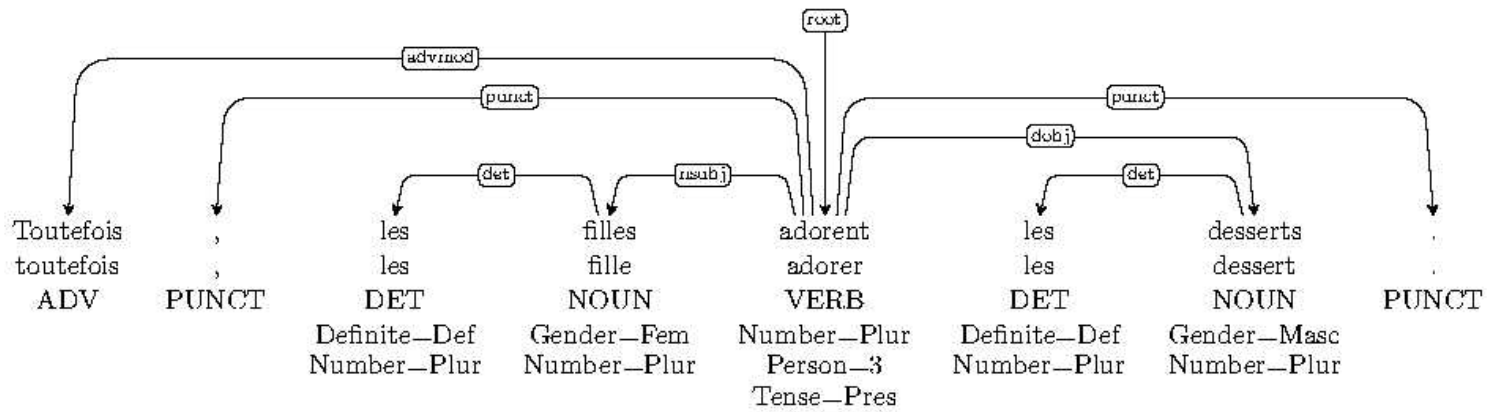
Simple syntactic feature values

Simple atomic values for common syntactic features:

PERS	1, 2, 3
GEND	M, F, N. . .
NUM	SG, DU, PL. . .
CASE	NOM, ACC, DAT, ERG, ABS, . . .

Simple syntactic feature values

Features in UD, Nivre (2015):



Issues in feature representation

Motivation for complex syntactic feature values?

- Feature resolution
- Natural classes of feature values
- Constructed feature values
- Feature indeterminacy

Complex feature value representations: Considerations

- Universality
- Relation between syntax and semantics
- Markedness
- Consistency of representation

Universality

- Is the representation and structure of the value of a feature determined on a language-by-language basis, or do all languages make use of the same representation?
- Can the same feature have a simple, atomic value in one language, and a complex value in another language? Can the same feature have different(compatible or incompatible) complex values in different languages?

Relation between syntax and meaning

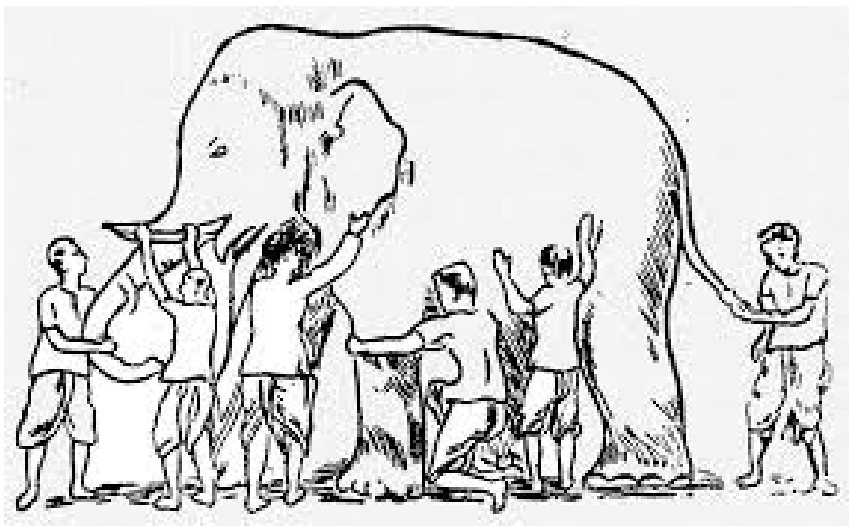
- For semantically motivated features such as PERS and NUM, do we expect the structure of a complex f-structure feature value to partially or completely reflect details of the meaning of the elements bearing the features?
- If there is a systematic relation between the representation of a syntactic feature value and the meaning of phrases bearing the feature, should this be attributed to historical factors, or do we require a synchronic theory of this relation?

Markedness

- Some feature values are traditionally classified as unmarked: for example, third person is usually classified as the unmarked value of the PERS feature.
- Should functionally unmarked properties be formally represented either by a negative value for a feature or by the absence of the feature, while marked values are represented by a positive value or by the presence of the feature?

Consistency of representation

- Do patterns of feature resolution, feature indeterminacy, syntactic compositionality, syncretism, and markedness converge on a single representation for the value of a feature?
- If not, which phenomena take precedence in determining how the value of a feature should be represented?



The PERSON Feature:
Resolution, markedness, natural classes of features

Person resolution

Person resolution (Corbett, 1983):

$$\begin{aligned} \text{first} \;\&\; \text{second} &= \text{first} \\ \text{first} \;\&\; \text{third} &= \text{first} \\ \text{second} \;\&\; \text{third} &= \text{second} \\ \text{third} \;\&\; \text{third} &= \text{third} \end{aligned}$$

- Should the representation of the person feature support a theory of feature resolution?
- Alternative: Stipulate the patterns in the table.

Person resolution

Sets as the value of the PERS feature: feature resolution is union or intersection.

Sag et al. (1985): Resolution by intersection.

First person:	{ }	
Second person:	{+XSP}	(+XSP: "excludes the speaker")
Third person:	{+XSP, +THP}	(+THP: "third person")

Criticised by Karttunen (1984) for not obeying markedness criterion: third person is the largest, most "marked" set.

Dalrymple & Kaplan (2000): Resolution by union.

First person:	{S,H}
Second person:	{H}
Third person:	{ }

Criticised by Vincent & Börjars (2007) for not obeying universality criterion: an additional value is needed for languages with an inclusive/exclusive distinction in the first person.

Natural classes of feature values

Present tense of the verb ‘say’, Icelandic, with 2/3 syncretism (Otoguro, 2015):

	SG	PL
1	segi	segjum
2	segir	segið
3		segja

- $1/3$ and $1/2$ syncretisms also attested. Should the representation of the person feature allow for the representation of natural classes of feature values?
- Alternative: Treat as homophony.

Sets and natural classes

A set-based representation also allows a treatment of syncretism and natural classes.

	First person:	$\{\ \}$
Sag et al. (1985):	Second person:	$\{+\text{XSP}\}$
	Third person:	$\{+\text{XSP}, +\text{THP}\}$

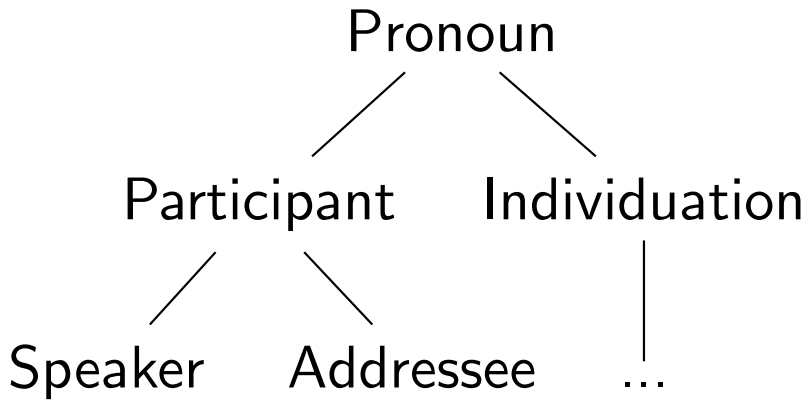
- 2/3 syncretism: $+\text{XSP}$ is a member of the set.

	First person:	$\{\text{S},\text{H}\}$
Dalrymple & Kaplan (2000):	Second person:	$\{\text{H}\}$
	Third person:	$\{\ \}$

- 2/3 syncretism: S is not a member of the set.

Another view of features: Feature geometry

Harley & Ritter (2002):



- Feature geometry encodes markedness relations and implicational relations.
- Feature geometry constrains crosslinguistic variation and paradigm-internal gaps and syncretisms.
- Not intended to account for feature resolution.

Person as features

Otoguro (2015):	First person:	$[+1, -2]$
	Second person:	$[-1, +2]$
	Third person:	$[-1, -2]$

- Allows underspecification: Can encode 2/3 syncretism as $[-1]$ (but not 1/2, or at least not easily; need an additional feature).
- Can account for feature resolution: Coordinate structure acquires positive features of conjuncts (similar to union analysis).
- Sadler (2011): set-based analyses can be directly translated to equivalent feature-based analyses, with a positive value representing the presence of an element, and a negative value representing its absence.

The NUM Feature: Constructed values

Constructed feature values

Constructed number in Hopi (Corbett, 2000):

- Pam wari
  that.SG run.NONPL
  'S/he ran.' (singular + nonplural = singular)
  1 1 or 2 1
- Puma yùutu
  that.NONSG run.PL
  'They ran.' (nonsingular + plural = plural)
  2 or more 3 or more 3 or more
- Puma wari
  that.NONSG run.NONPL
  'They (two) ran.' (nonsingular + nonplural = dual)
  2 or more 1 or 2 2

- Should the representation of the number feature allow partial specification from different sources?
- Alternative: Treat as ambiguity: e.g. nonplural form is ambiguous between singular and dual.

Number as features

- Sadler (2011) for Hopi: Nonsingular + nonsingular = dual.

Singular:	$[+\text{SG}, -\text{PL}]$
Nonsingular (unspecified for SG):	$[+\text{PL}]$
Nonplural (unspecified for PL):	$[+\text{SG}]$
Dual:	$[+\text{SG}, +\text{PL}]$
Plural:	$[-\text{SG}, +\text{PL}]$

- Arka (2011, 2012) for constructed number in Marori: More distinctions needed. The full set of feature values:

Singular:	$[+\text{SG}, -\text{PL}, -\text{AUG}]$
Dual:	$[-\text{SG}, -\text{PL}, -\text{AUG}]$
Limited plural:	$[-\text{SG}, -\text{PL}, +\text{AUG}]$
Generic plural:	$[-\text{SG}, +\text{PL}, -\text{AUG}]$
Large plural:	$[-\text{SG}, +\text{PL}, +\text{AUG}]$

A different view of features

Harbour (2014) for Mele-Fila:

	SG	DUAL	PAUCAL	PLURAL	GR.PLURAL
article	te	ruu		a	
pronoun	aia	raaua	raateu		reafa
ATOMIC	+	−	−	−	−
MINIMAL	+	+	−	−	−
ADDITIVE	−	−	−	−+	+

- ADDITIVE: additive closure.
- A feature can be specified as both − and +.
- Features encode operations on the denotation P of the noun: $-\text{ADDITIVE}(+\text{ADDITIVE}(P))$ gives a "lesser plural".

The CASE Feature: Indeterminacy

Feature indeterminacy

Case indeterminacy in German (Groos & van Reimsdijk, 1979):

- Er findet und hilft Papageien.
  he finds and helps parrots
  OBJ=ACC OBJ=DAT NOM,ACC,DAT,GEN
  'He finds and helps parrots.'

- The representation of the case feature must allow indeterminate specification.
- (No obvious alternative.)

Indeterminate features

- Special formal treatment of indeterminate features, Ingria (1990): *Papageien* has case $\textsc{nom} \vee \textsc{acc} \vee \textsc{dat} \vee \textsc{gen}$. Case is not checked via equality, but with a special nondistinctness check $\approx$.

Indeterminate features

- Indeterminate features are overspecified, Bayer (1996): Arguments can be overspecified for case, and predicates can be strengthened to require overspecified arguments.

$$\dfrac{\dfrac{\dfrac{\text{findet}}{\text{VP/NP[acc]}}}{\dfrac{\text{findet}}{\text{VP/(NP[acc]}\wedge\text{NP[dat])}}}\ \text{und}\ \dfrac{\dfrac{\text{hilft}}{\text{VP/NP[dat]}}}{\dfrac{\text{hilft}}{\text{(VP/NP[acc]}\wedge\text{NP[dat])}}}}{\text{VP/(NP[acc]}\wedge\text{NP[dat])}}\quad\dfrac{\text{Papageien}}{\text{NP[acc]}\wedge\text{NP[dat]}}}{\text{VP}}$$

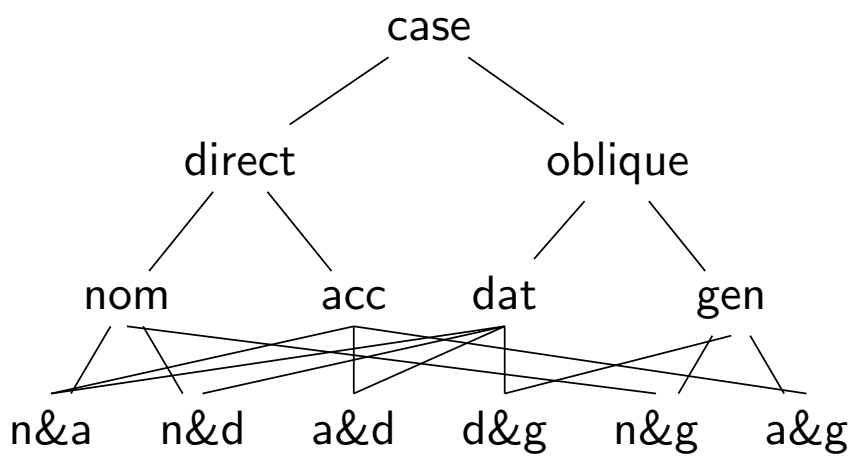
Indeterminate features

- Indeterminate features as sets, Dalrymple & Kaplan (2000): *Papageien* has case $\{\text{NOM},\text{ACC},\text{DAT},\text{GEN}\}$. Case is checked via a set-membership test imposed by the predicate.

findet	und	hilft	Papageien
$\text{ACC}\in\text{OBJ CASE}$		$\text{DAT}\in\text{OBJ CASE}$	$\text{CASE}=\{\text{NOM},\text{ACC},\text{DAT},\text{GEN}\}$

Indeterminate features

- Indeterminate features as underspecification in a type hierarchy, Sag (2002):



findet: $\textsc{comps}\ \textsc{case}= \boxed{1}$, acc$\leq \boxed{1}$

hilft: $\textsc{comps}\ \textsc{case}= \boxed{1}$, dat$\leq \boxed{1}$

Papageien: (no case specification)

findet und hilft Papageien: $[\textsc{case} = \textsc{a\&d}]$

Indeterminate features

- Indeterminate features as feature structures, Dalrymple et al. (2009): *Papageien* is unspecified for case; case is checked through assignment of a + value for the case that the predicate requires.

findet	und	hilft	Papageien
OBJ CASE ACC=+		OBJ CASE DAT=+	CASE=[ACC +,DAT +]

Summary

- The analysis of resolution, constructed feature values, and syncretism involving certain features (person, number, gender) can probably be finessed by assuming homophony or ambiguity, but complex values might provide a more satisfying analysis of these phenomena.
- Indeterminacy of the case feature cannot be treated by simple feature values; complex values are needed.
- Complex values for features should be justified crosslinguistically as well as by examining multiple phonomena in the same language.

References I

Arka, I Wayan. 2011. Constructive number systems in Marori and beyond. In Miriam Butt & Tracy Holloway King (eds.), *On-Line Proceedings of the LFG2011 Conference*. Stanford: CSLI Publications. URL `http://web.stanford.edu/group/cslipublications/cslipublications/LFG/16/lfg11.html`.

Arka, I Wayan. 2012. Verbal number, argument number and plural events in Marori. In Butt & King (2012). URL `http://web.stanford.edu/group/cslipublications/cslipublications/LFG/17/lfg12.html`.

Bayer, Samuel. 1996. The coordination of unlike categories. *Language* 72(3). 579–616.

Butt, Miriam & Tracy Holloway King (eds.). 2012. *On-line proceedings of the LFG2012 conference*. Stanford: CSLI Publications. URL `http://web.stanford.edu/group/cslipublications/cslipublications/LFG/17/lfg12.html`.

Corbett, Greville G. 1983. Resolution rules: Agreement in person, number, and gender. In Gerald Gazdar, Ewan Klein & Geoffrey K. Pullum (eds.), *Order, Concord and Constituency*, 175–206. Dordrecht: Foris.

Corbett, Greville G. 2000. *Number*. Cambridge, UK: Cambridge University Press.

References II

Dalrymple, Mary & Ronald M. Kaplan. 2000. Feature indeterminacy and feature resolution. *Language* 76(4). 759–798.

Dalrymple, Mary, Tracy Holloway King & Louisa Sadler. 2009. Indeterminacy by underspecification. *Journal of Linguistics* 45. 31–68.

Groos, Anneke & Henk van Reimsdijk. 1979. Matching effects in free relatives: A parameter of core grammar. In Adriana Belletti, Luciana Brandi & Luigi Rizzi (eds.), *Theory of Markedness in Generative Grammar: Proceedings of the 1979 GLOW Conference*, 171–216. GLOW, Pisa: Scuola Normale Superiore di Pisa.

Harbour, Daniel. 2014. Paucity, abundance, and the theory of number. *Language* 90(1). 185–229.

Harley, Heidi & Elizabeth Ritter. 2002. Person and number in pronouns: A feature-geometric analysis. *Language* 78(3). 482–526.

Ingria, Robert J. P. 1990. The limits of unification. In *Proceedings of the 28th Annual Meeting of the ACL*, 194–204. Pittsburgh: Association for Computational Linguistics.

Karttunen, Lauri. 1984. Features and values. In *Proceedings of the International Conference on Computational Linguistics (COLING84)*, 28–33. International Committee on Computational Linguistics.

References III

Nivre, Joakim. 2015. Towards a universal grammar for natural language processing. In *International Conference on Intelligent Text Processing and Computational Linguistics*, 3–16. Springer.

Otoguro, Ryo. 2015. Generalising functional categories in LFG. In Miriam Butt & Tracy Holloway King (eds.), *On-Line Proceedings of the LFG2015 Conference*. Stanford: CSLI Publications. URL `http://csli-publications.stanford.edu/LFG/20/lfg15.html`.

Sadler, Louisa. 2011. Indeterminacy, complex features and underspecification. *Morphology* 21. 379–417.

Sag, Ivan A. 2002. Coordination and underspecification. In Jong-Bok Kim & Stephen Wechsler (eds.), *Proceedings of the 9th International Conference on Head-Driven Phrase Structure Grammar*, 267–291. Stanford: CSLI Publications. URL `http://cslipublications.stanford.edu/HPSG/3/hpsg02.htm`.

Sag, Ivan A., Gerald Gazdar, Thomas Wasow & Steven Weisler. 1985. Coordination and how to distinguish categories. *Natural Language and Linguistic Theory* 3(2). 117–171.

References IV

Vincent, Nigel & Kersti Börjars. 2007. Feature resolution and the content of features. In Annie Zaenen, Jane Simpson, Tracy Holloway King, Jane Grimshaw, Joan Maling & Christopher Manning (eds.), *Architectures, Rules, and Preferences: Variations on Themes by Joan W. Bresnan*, 293–315. Stanford: CSLI Publications.