Canonical Morphology and Syntax

Dedicated to the memory of Anna Siewierska

Edited by
DUNSTAN BROWN,
MARINA CHUMAKINA,
and
GREVILLE G. CORBETT

OXFORD UNIVERSITY PRESS
Contents

Notes on contributors vi
Acknowledgements ix
Abbreviations x

1. What there might be and what there is: an introduction to Canonical Typology 1
   Dunston Brown and Marina Chumakina

2. A base for canonical negation 20
   Oliver Bond

3. Canonical morphosyntactic features 48
   Greville G. Corbett

4. Some problems in the typology of quotation: a canonical approach 66
   Nicholas Evans

5. Unpacking finiteness 99
   Irina Nikolaeva

6. The canonical clitic 123
   Andrew Spencer and Ana Luis

7. Passive agents: prototypical vs. canonical passives 151
   Anna Stiewierska and Dik Bakker

8. The criteria for reflexivization 190
   Martin Everaert

9. Possession and modification—a perspective from Canonical Typology 207
   Irina Nikolaeva and Andrew Spencer

10. An ontological approach to Canonical Typology: Laying the foundations for e-linguistics 239
    Scott Parrar

References 262
Language index 295
Index of authors 299
Subject index 307
Notes on contributors

Dirk Barker is Associate Professor of Linguistics at the University of Amsterdam and a senior researcher at Lancaster University. His research interests lie in the area of linguistic typology, functional theories of grammar, language contact and language change, and the application of the computer for linguistic purposes. His publications are on Functional Grammar, grammatical relations (with Anna Sievertsa), Otomi and other Latin American languages (with Ewald Hekking), and the reconstruction of language families (ASJP project). Recent major publications are Aspects of Language Contact, and Hispanisation (edited, with Thomas Stolz, Mouton de Gruyter 2006).

Olivier Bond is Lecturer in Linguistics in the Surrey Morphology Group, University of Surrey. His principal research interests lie in typology, morphosyntax, historical linguistics, and language documentation, with a current focus on the typology of negation and cognate object constructions. His publications, based on fieldwork on Eleme, an undescribed Benue-Congo language spoken in south-east Nigeria, cover topics such as agreement, logophoric reference, and valence.

Dunstan Brown is a member of the Surrey Morphology Group, University of Surrey, and Professor at the University of York. His research interests include autonomous morphology, morphology–syntax interaction, and typology. His recent work has focused on describing and understanding different aspects of morphological complexity, notably The Syntax–Morphology Interface: A Study of Syncretism (with Matthew Baerman and Greville G. Corbett, CUP 2005) and Network Morphology (with Andrew Hippiusley, CUP 2012).

Marina Chumakina is a Research Fellow in the Surrey Morphology Group, University of Surrey. Her work focuses on Nalch-Dagestanian languages and typology. She has done extensive fieldwork on the Archi language resulting in an electronic Archi Dictionary (together with Dunstan Brown, Greville G. Corbett, and Harley Quilliam).

Greville G. Corbett is Distinguished Professor of Linguistics, University of Surrey, and leads the Surrey Morphology Group. He works on the typology of features, as in Gender (1991), Number (2000), and Agreement (2006), all with CUP. Recently he has been developing the canonical approach to typology. He is one of the originators of Network Morphology, see The Syntax–Morphology Interface: A Study of Syncretism (with Matthew Baerman and Dunstan Brown, CUP 2005).

Nicholas Evans is Professor of Linguistics in the Australian National University College of Asia Pacific, Australian National University. He has carried out wide-ranging fieldwork on languages of northern Australia and Papua New Guinea, and has particular interests in the dialogue between language description, typology, and general linguistic theory. In addition to grammars of Kayardild (1995) and Bininj Gun-wok (2003), a dictionary of Dalabon (with Merlan and Tukumbu 2004), edited collections on polysynthesis, historical linguistics of north Australian languages, the semantic typology of reciprocals, he recently published the widely acclaimed book Dying Words: Endangered Languages and What They Have to Tell Us (2010) which sets out a broad programme for the field’s engagement with the world’s dwindling linguistic diversity.

Martin Everaert is Professor of Linguistics and director of the Utrecht Institute of Linguistics OTS. His research interests lie in the area of the syntax–semantics interface (reflexives, reciprocals), and the lexicon–syntax interface (idioms/colocations, and argument structure). His main publications include The Blackwell Companion to Syntax (co-edited with Henk van Riemsdijk 2006), The Unaccusativity Puzzle: Explorations of the Syntax-Lexicon Interface (co-edited with A. Alexiadou and E. Anagnostopoulou, CUP 2004), and Idioms: Structural and Psychological Perspectives (co-edited with E. van der Linden, A. Schenke, and R. Schreuder 1995).

Scott Farrow is the Natural Language Processing (NLP) architect at Stripe in Seattle, Washington and lecturer at the University of Washington. He is a computational linguist whose work concentrates on natural language processing, mobile applications development, language documentation, semantics, linked data, and ontology. He has published on the role of linguistic ontologies in discourse (with Beteman and Tenbrink 2007), ontologies and semantics (with Langendoen 2003), and use of linguistic ontologies for language description (2007).

Ana Luís is an Assistant Professor at the University of Coimbra. Her research interests lie in morphological and morphosyntactic typology, the morphology–syntax interface, and Creole morphology. Her typological work involves publications on Romanian, Udi (with Andrew Spencer 2004), Cappadocian Greek (2005), Pula and Swinhill (2004), among others. Working within the theory of Paradigm Function Morphology, she has published on cliticization and verbal inflection (2004).

Irina Nikolaeva is Professor of Linguistics at the School of Oriental and African Studies, University of London. Her primary research interests are morphology, syntax, and information structure, as well as the modelling of these areas within non-transformational grammatical theories. She has also documented and published grammatical descriptions of little studied endangered languages of Siberia, such as Yukaighir (1989, 1993, 1997, 2005), Udibe (with Tolkhaya, Mouton 2009, also in 2002, 2005).
2003), Ostyak (1995, 1999), and Nenets (with Ackerman and Salmisen, in preparation).

Anna Siewierska was Professor of Linguistics and Human Communication at Lancaster University. Her major research interests were language typology, the comparison of different theoretical frameworks, diachronic change, discourse pragmatics, and the morphosyntax of English dialects. Her typological work had centred on various aspects of clause structure, grammatical relations, case and agreement marking, word order, valency changing operations, and pragmatic functions. Major publications include Person (CUP 2004), Functional grammar (Routledge 1991), and Impersonal constructions: a cross-linguistic perspective (edited, with Andrej Malchukov, Benjamins 2011). Anna sadly died in 2011 and this volume is dedicated to her memory.

Andrew Spencer is Professor of Linguistics at the University of Essex. He is the author of Morphological Theory (1991), co-author of Clitics (to appear, CUP), and co-editor of The Handbook of Morphology (1998) and the Oxford Handbook of Case (2008). His research interests centre around the architecture of inferential-realizational theories of morphology, the morphology-syntax interface, and the notion of 'lexical relatedness'.

Acknowledgements

The canonical method in linguistic typology is a way of addressing the major problem of how we can determine the fit of language-specific constructions with cross-linguistically applicable categories. It deals with the problem by taking the criteria used to define particular categories or phenomena to create a multidimensional space in which language-specific instances can be placed. In this way, the issue of fit becomes a matter of greater or lesser proximity to a canonical ideal. This volume is a selection of papers from a conference on this topic held at the University of Surrey on 9-10 January 2009, sponsored by the University’s Institute of Advanced Studies, and the workshop on Canonical Typology held as part of the Eighth Biennial Conference of the Association for Linguistic Typology at Berkeley on 26 July 2009. We thank the participants at both events for making them such successes. Anna Siewierska took part on both occasions, and we dedicate the volume to her memory. We are grateful to the Economic and Social Research Council (grant number RES-062-23-0690) and the European Research Council (grant number ERC-2008-AdG-230268 MORPHOLOGY) for continuing support during work on this volume, and to John Davey and Julia Steer at Oxford University Press for their expert advice. Special thanks go to Penny Everson and Lisa Mack for their careful help in the preparation of the manuscript.
What there might be and what there is: an introduction to Canonical Typology

DUNSTAN BROWN AND MARINA CHUMAKINA

1.1 Introduction

Data comparability is a fundamental problem within typology and it has far-reaching implications for our ability to define natural languages. Haspelmath (2007) characterizes the problem in the following way:

Which are the right categories for a given language? ... For descriptive linguists, there would have to be a list that contains the pre-established categories that general linguists have figured out in some way. These would not necessarily have to be innate, but they would have to be universal in the sense that a descriptive linguist can be sure that the categories needed for describing his/her language are on the list. (Haspelmath 2007: 120–121)

The issue of correspondence between similarly named features in different languages is a substantial part of the data comparability issue. Corbett (2000a) calls this the Correspondence Problem:

The Correspondence problem: as typologists we need to be able to justify treating features and their values as comparable across languages. This is not straightforward, and yet a good deal of typology, including enterprises such as the World Atlas of Language Structures (WALS), depends upon it. (Corbett 2000a)

In relation to morphosyntactic features, the Correspondence Problem can be illustrated by the notion of past tense. Languages can divide the notional space of
Some problems in the typology of quotation: a canonical approach

NICHOLAS EVANS

'The idea that languages make a clear distinction between direct and indirect speech is for the most part a grammatical fiction'

(Ebert 1986: 158)

4.1 Canonicity and points of typological reference

Meaningful typological comparison of languages cannot occur without determinate points of reference—what typologists have traditionally called tertius comparationis. However, there are widely differing conceptions of what the status of these points is, and how they should be discovered and formulated.

A crucial assumption of the canonical approach to typology is that it is sometimes useful to postulate canonical points which represent ideals never instantiated in any language. Obviously this should never be done lightly, since it risks proliferating unnecessary entities on the one hand and risks lessening the analyst's empirical responsibility to stay close to the data on the other. However, there are cases where taking this approach allows a clearer formulation of the design space within which systems attested in actual languages can be placed. (And in any case, since the dimensions are based on observable or recognizable phenomena, it's not the case that the researcher can think up just anything without constraint.) The payoff is that, even if a particular canonical phenomenon is nowhere exemplified perfectly, postulating them allows us to map actual cross-linguistic reality in a more satisfying and exhaustive way, and to put a finger on exactly what the ingredients are that are being mixed together in less canonical cases.

In this paper I use this approach to set up a revised typology of the grammar of quotation. The typology of quoted speech has long been a disorderly and unsatisfying area because of the huge number of ways that languages can deviate from the traditional ideals of 'direct' and 'indirect' speech. This has generated a profusion of terms for non-canonical types—semi-direct, semi-indirect, quasi-direct and so on—which do not correlate well with the many intermediate categories. Perhaps worse is the fact that typological research on quoted speech has generally been rather unclear about what 'true' indirect speech would look like. The result has been to obscure what canonical indirect speech would look like, by defining it through language-particular examples which frequently deviate from the essence of the definition.

The third and perhaps most serious problem with traditional classifications is that they overlook a third ideal type that I will call 'biperspectival' speech, where particular signs simultaneously reflect the viewpoint of both original and reporting speaker. The likely reason for this omission is that no language reported so far comes at all close to this type on all possible definitional dimensions—in other words it is a canonical ideal that is approached even less closely than indirect speech. But there are enough manifestations of it across a range of semantic dimensions—complex tenses in Western European languages being the most familiar—that the outlines of such an ideal are clear.

I construct the argument as follows. In section 4.2 I review key aspects of the traditional typological opposition between direct and indirect speech. In section 4.3 I examine a number of cases that have basically been categorized as direct speech, though with deviant elements, while in section 4.4 I move to deviations from the other pole, i.e. cases of indirect speech that fail to undergo the expected deictic re- evaluations on one or more dimensions. I will progress from cases where the behavioural split runs right through one semantic dimension (e.g. deictic shift in person, but not tense) to those where the split occurs on some subset within a single semantic dimension (e.g. deictic shift for the person of objects but not of subjects). In section 4.5 I pass to the new concept of 'biperspectival speech', exemplifying it with the dimensions of person (i.e. logophoric pronouns that simultaneously characterize their referent to speech-act roles in both primary and reported speech events) and tense (composite tenses that locate events in time with respect to both reported and primary speech events). Finally, in section 4.6 I review the implications of this new typology for a range of difficult cases and raise some outstanding problems.

1 I would like to thank Diensten Brown, Marina Chumakina, and Grev Corbett for organizing the Galápagos workshop on Canonical Typology (and subsequently a workshop on the same topic at the 2009 ALT meeting in Berkeley), thereby giving me the opportunity to present these ideas. I also thank the participants in these workshops for helpful discussion, and the following individuals for discussion or data pertaining to this paper: Amy Rose Daul, Robyn Loughman, Ger Jansen, Alan Rumsey, Stef Sperling, and Heito van der Vorst. I am also grateful to Marina Chumakina, Grev Corbett, Mishu Daniel, Alan Jones, Thomas Moktar, Stef Sperling, and two anonymous referees for comments on an earlier draft of this paper. I would also like to acknowledge the Australian Research Council for its support of research reported here, through grant DP0776326, 'Social Cognition and Language—the design resources of grammatical diversity'.

2 I use this term in a way synonymous with 'typological space'; on the more general use of 'design space' in evolutionary theory see Dennett (1995).
4.2 The traditional typology of quotation: direct, indirect, other

There is a long-standing tradition of distinguishing direct from indirect speech, which we can exemplify with the following pair of sentences:

(a) Grev said: 'I'm clearing up these typological conundrums.'
(b) Grev said that he was clearing up these typological conundrums.

This opposition is generally characterized along the following lines, here distilled into canonical statements. This represents my distillation of the existing literature, some representative quotes from which will be given later in this section.

(2a) Canonical direct speech:

C1. Canonical direct speech reproduces the original speaker's words, or at least words that are presented as if they were original speech.
C2. Canonical direct speech includes all linguistic particularities of the original (e.g. language or dialect choice).
C3. Canonical direct speech presents all deictically sensitive expressions from the perspective of the original speaker.

(2b) Canonical indirect speech:

C1. Canonical indirect speech assimilates the material presented by the original speaker to the reporting speaker's perspective.
C2. Canonical indirect speech abstracts away from all linguistic particularities of the original and replaces them with a stylistic representation consistent with the rest of the reporting speaker's stylistic choices.
C3. In canonical indirect speech, all deictically sensitive expressions are presented from the perspective of the reporting speaker.

Applying these definitions to our initial examples, (a) presents the quotation in a form in which Grev could have originally made it (whether he actually said this or not), including the use of the first person pronoun to refer to himself, and of present auxiliary am whose tense is calculated with respect to the moment of the original speech act. In (b), on the other hand, person is calculated from the perspective of the reporting speaker (thus placing Grev in the third person) and so is tense (placing the auxiliary in the past tense).

In addition, a full account of the direct vs. indirect speech contrast will note prosodic differences, represented graphically by quotation marks in English, and also syntactic differences such as the (optional) use of the complementizer 'that' in the indirect version. The degree of syntactic difference between direct and indirect speech, while rather minor in English, is far more significant in some languages, such as in Latin, (3) where in indirect discourse the verb of the quoted passage is placed in the infinitive and its subject in the accusative:

(3) Dic-o Cicero(aem) ed-ere ole-al

'say that Cicero is eating olives.' (Coulmas 1986: 19)

Returning to the existing literature on quoted speech, and pulling out just two representative general statements from this vast literature, that succinctly capture the difference between direct and indirect speech, consider the following:

The fundamental difference between the two lies in the speaker perspective or point of view of the reporter: in direct speech the reporter lends his voice to the original speaker and says (or writes) what he said, thus adopting his point of view, as it were... In indirect speech, on the other hand, the reporter comes to the fore. He relates a speech event as he would relate any other event, from his own point of view (Coulmas 1986: 2).

With direct speech reports (DSRs) the perspective of the original speaker is maintained; with indirect speech reports (ISRs) on the other hand, the perspective shifts towards that of the reporting speaker and his speech situation (De Roeck 1994: 332).

In what follows, it will be useful to adopt the following terminology and abbreviations:

(4) PSE: primary speech event (i.e. the speech act generating the overall utterance)
[Primary speaker / addressee etc.: the speaker/addressee etc. in the primary speech event]
RSE: reported speech event (i.e. the speech act being reported on in the utterance)
[Reported speaker/addressee: the speaker/addressee etc. in the reported speech event]

---

3 Though of course it is an oversimplification to assume there is total isomorphism between the graphical devices of quotation marks and the boundaries of actual prosodic contours: see Nunberg (1995) for a nuanced and interesting account of the complex relation that punctuation mediates between spoken and written form.
Note also that often it is thoughts, perceptions, intentions, desires and so on that are being represented as reported rather than speech, though for brevity's sake I will not generally make this explicit.5

With this terminology established, it should follow from the above characterizations that in direct speech all features of the quotation retain the deictic and other relevant values they would have in the RSE, while in indirect speech they are all recast with the values they would have if calculated from the PSE. We shall see in later sections, however, that this is far from being the case.

The above statement focuses on semantic features, particularly those of deictic inflectional categories, that distinguish direct from indirect speech. But there are also syntactic and prosodic features, summarised in Table 4.2 though they are not the

| Table 4.2: Common syntactic and prosodic differences between direct and indirect speech |
|----------------------------------------|----------------------------------------|
| Direct                                 | Indirect                               |
| Overt quotational clause optional      | Overt quotational clause obligatory     |
| Normal main clause morphology          | Various deviations from main clause syntax possible |
| Arguments remain in quoted clause      | Arguments may be raised into quotational clause |
| Quote can be interrupted by            | Quote not interrupted by quotational clause |
| quotational clause                     |                                        |
| 'John' she said 'has left.'            | 'John she said that had left.'          |
| Independent, 'Vivid' intonational      | Integrated intonation contour           |
| contour in quoted material             |                                        |

5 Commenting on an earlier version of this paper, Alan Jones (personal communication) points out that verbs that report thoughts, perceptions, intentions, desires, etc., are never to be followed by true 'direct speech' (at least unless the speaker of the primary speech event is also the cognizer of the reported speech event). In other cases the speaker can't in fact ever know anything for certain about the thing that they are reporting. However, what is primarily at issue here is how the language represents the quoted passage—just as in much quoted speech (made up to look like direct speech, and often taking liberties with content and form), many languages present thought, emotional reactions and so on as if they were direct speech.

6 And there are interesting intermediate cases here that I won't go into, where there is a clause fragment giving the speaker but no-quotative verb (of the type And he's 'Look at that!' and she's 'Oh my God!'), something like that has been grammaticalized as the primary quotative construction in Khmer (see section 4.2).

7 This issue has been particularly discussed within the literature on Spanish, perhaps because of the clear punctuational marking of intonational onsets in Spanish orthography, e.g. Verdín (1979: 65, cited in Kriwak 1986: 327): Tanto la frase introductora como la frase reproductora aparecen casi en el mismo plano respecto a la intonación, con un vehículo, por supuesto, que van grupos distintos y que están caracterizados respectivamente por sus terminaciones ascensante y descendente. 'Both the introductory phrase and reproduced phrase appear almost on the same plane with respect to intonation, without forgetting, of course, that they are distinct groups and are characterized respectively by their ascending and descending terminations.' [Translation mine] However, empirical testing of this claim, using acoustic data from actual speech, shows that the difference between these types is far more gradation than had been originally thought (Kriwak 1986).

primary focus of this article, since most characterizations of the direct vs. indirect speech dichotomy employ exclusively semantic criteria, such as perspective and point of view (examples will follow).

Many authors have pointed out the idealized nature of the above dichotomy, a topic I return to in the next sections. In one of the most systematic cross-linguistic studies of the phenomenon, De Roock (1994: 30) concludes that

the question, whether DSIs [=Direct Speech Reports] and ISIs [=Indirect Speech Reports] form a dichotomy definitely calls for a negative answer.

Similarly, Haberland (1986: 248), after a thorough study of quotation in Danish, writes:

This makes us ask if the clear-cut division between direct and indirect speech, mainly developed on the model of Latin and Classical Greek, and in different ways applied to a handful (but not even all) of the European literary languages (of which Danish is one), ever really was as clear as it appeared from the point of view of the study of the literature of those languages in certain historical periods.

Likewise, Coulmas (1986a: 6) states that:

a simple dichotomy of direct versus indirect cannot do justice to the complexities of reported speech.

De Roock and Haberland represent a large number of authors who point to the existence of less-than-canonical quotation types.

There is also a long tradition, in literary stylistics, of investigating the type of genre known variously as 'free indirect discourse' or 'represented speech' (Jespersen 1917, as style indirect libre in French (Bally 1912), and in German as verselbständigte Rede 'veiled speech' (Kalepky 1899), erlebte Rede 'lived/experienced speech' (Lotz 1932), or uneigentliche direkte Rede 'quasi-direct speech' (Lotz 1939). This can be exemplified by (5), and is characterized by the combination of direct-speech syntax (such as word order characteristic of direct-speech questions, i.e. when would he be... rather than when he would) with person and tense values characteristic of indirect speech.

(5) Dunstan lay awake thinking about the database. When would he ever be able to finish it?

Among the authors quoted above, some have seen free indirect discourse as a balanced mingling of two primary types (Tołbór 1894), some as a third primary type (Kalepky 1899), some as a deviant form of indirect speech (Bally, Jespersen), and

8 See Steever (2002) for a detailed treatment of the constructional differences between direct and indirect speech in Tamil, where the structural difference between direct and indirect speech is particularly striking.
others as a deviant form of quasi-direct speech (Lerch). The concerns of these authors differ from those of this chapter in being based on a primarily literary device, and moreover on one which generally omits a direct and integrated representation of the quoting clause. But it is clear that they form part of a long and distinguished tradition of assembling a corpus of intermediate cases lying between canonical direct and canonical indirect speech. The long and interesting tradition of narratological analysis of this phenomenon (see Banfield 1978; Tannen and Tannen 2003 for summaries) has been interested both in the transmission of others words in what Bakhtin (1981) characterized as double-voicedness, and in the way it can be used by a narrator to model the mind of another person in action (Mäldorf 2008).

On the other hand, a question that has not been asked, as far as I know, but which we are led to by a canonical approach, is: what should canonical indirect speech look like if the definition given in (5b) above is pushed to its logical extreme? Which language—if any—exemplifies it? I will delay answering this question until we have looked at less canonical quotation types in more detail, something we do in the next two sections.

4.3 Some deviations from canonical direct speech

I have already mentioned the fact that direct speech does not generally represent what was actually said. Tannen (1986) has made this point particularly clearly, arguing that “the term ‘reported speech’ is a misnomer” (p. 211), that most ‘of the lines of dialogue represented in storytelling or conversation...were probably not actually spoken’, (p. 314) and that therefore ‘[i]t cannot be the case that dialogue represented in oral storytelling is being reported exactly as it was spoken, unless the report is based on the deliberate memorization of a transcript which was based on a tape-recording of the talk’ (p. 313).

4.3.1 Construction through style and translation

In recognition of these problems, direct speech is normally characterized as representing what could have been said, rather than what was said. This move already introduces a role for the reporting speaker, since they are actually constructing rather than reporting speech: witness the telltale wording ‘lending his voice to the original speaker’ in Coulmas’ characterization of direct speech cited above. This makes it almost inevitable that aspects of the reporting speaker’s creative latitude will seep in to direct speech. In Dalabon, for example, the interjection ngale! ‘hey!’ is certainly possible in ordinary speech, but occurs at the beginning of quoted passages far more frequently than we would expect if it truly represented what was said (or thought). Without being standardized to the point where it is grammaticalized, it is nonetheless being deployed in a stereotyped, dramatizing way to make the narrative more vivid, but at the same time it serves the useful function of indicating the start of a quoted passage (6).

(6) ka-h-qing-bo-ng bika-h-na-ng yulu-kah yulu-djerring...
305.3-r-peg-go-pp 305/mL/306-110m-k-see-pp ground-loc earth-new
yulu-djerringa ka-h-na-ng
earth-new 305.3-r-see-pp
ngale! Nga-h-yo-ngyana djarra,
hey 155-5-r-sleep-fut here
kunbork ngi-h-m-tyan-kahn
corroboree 155-305-r-get-fut-purb 305.3-r-say-prs
‘He went along, and saw the ground, he saw newly-dug earth (of a fresh grave).
‘Hey! I’ll sleep here, so as to get the corroboree song’, he said.”

Another contribution of the speaker to the construction of ‘direct speech’ that we tend to take for granted in the English-speaking world, and in many speech communities more generally, is the translation of quotes into the language currently being used by the narrator. This, however, is far from a universal practice. In many Aboriginal speech communities, for example, a good narrator will reproduce the language choice of the characters as accurately as possible, even where the hearer may not understand the quoted language (see Evans 2011).

A nice example is the following text recorded by Allan Marett from Alan Maralung in 1956, about two ancestral beings who had composed the songs Maralung had just been singing. The primary narration is in Kriol, but the speech of the beings is primarily in Dalabon (aka Ngalkbon), the language they are said to have spoken (7). Narrative conventions that stipulate quotation in the original language thus bring the quoted speech closer to the ‘direct’ canonical ideal than is the norm with typical English narrative in which translation into the narrator’s language occurs even in ‘direct’ speech.

(7) wai $i$ $i$ bungkurdjibungkurdj $i$ imin $i$ lea $i$ kardoaljork well this olden-times 305.3-r say loc barn-owl
imin $i$ “what’s your name? marraah ngg-yng-ngu” 305.3-r say where name=305.3-r-poss
kuida “marrah mah ngg-yng-ngu” 305.3-r say like that
echna where whatabout name=305.3-r-poss 305.3-r-say
ask-im $i$ nim $i$ burman lenguwij, burman Ngalkbon
ask-tr name from language from Ngalkbon

8 Dalabon text, ‘How Djordi got Bongolinj-bongasilj’, recorded from Jimmy Wescum by Nicholas Evans, 2 June 2007, Beswick.
9 See Lichtenberg (2008: 129-34) for an explicit example of this point, where an original quotation in English is presented in the narrator’s language, Tagabapli of the Solomon Islands.
"marrah nyeg-ngug\" inin lagjad
where name-SG.POSS 3SG.ISR like.that

'Well, in the olden days, he said to the barn owl, he said "What's your name? 
Marrh nyegngug\", Echidna said "Marrh-nah nyegngug\" like that, he asked 
his name in language, in Ngalkhon, "Marrh-nah nyegngug\" he said.'

Deviations like translating into the language of narration, or using stereotyped expressions to make the narrative more vivid, are departures from canonical direct speech that are stylistic rather than grammatical. In canonical direct speech the original stylistic features (dialect, language) are maintained; in canonical indirect speech they are assimilated or translated to the style, dialect, and language of the framing utterance.

We now pass from these stylistic and lexical features to a more grammatical phenomenon in which constructions characterized as direct speech nonetheless make grammatical choices within the quotation that reflect some features of the primary speech event.

4.3.2 Second person magnetism

A number of languages exhibit a phenomenon whereby, in constructions where direct speech is structurally expected, second person values calculated from the primary speech event dominate over the values that would be expected from the reported speech event: I exemplify here with one example each from South and North America.

In KwaZu, an unclassified language of Rondônia state, Brazil, all quotation is basically direct. Person is calculated as in the Reported Speech Event: 'In KwaZu, no distinction is made between direct and indirect reports. Maybe it is even better to say that there is no indirect speech at all. Speech is quoted by literally repeating what has been said.' (van der Voort 2004: 402). Note that quotations are followed by a person/number marker identifying the subject of the notional quotational clause but with no overt quotative verb. This means that two independently valued subject suffixes are then stacked up (van der Voort 2002, 2004), like a grammaticalized version of the increasingly common English 'And then she's 'Oh my God!''. An example is (8); note that in this case the outer subject is third person singular and hence not overtly realized.

(8) 'marju kuku\'ny\'-da-ki-da\'ny\'-tse
Mario ill-3S-DECL-WANT-DECL
'Mario is going to say he is ill.' (lit. 'Mario is going to say "I am ill"')
(Van der Voort 2004: 403)

Van der Voort's analysis and examples make it clear that in just about every respect KwaZu quotations display all features of canonical direct speech. However, just where the subject of the quoted utterance is second person (with regard to the Primary Speech Event) 'the interpretation is one of indirect speech and the identity of the second person equals the hearer in the actual speech context instead of the (logically) expected speaker. So the quoted second person represents an exception' [to the previously discussed direct speech analysis of the quotation construction] (van der Voort 2004: 411). An example is (9). The original, uttered by Margarida, who is absent from the primary speech event, about the hearer in the primary speech event who was absent from the reported speech event, would have been the equivalent of 'he is ill'.

(9) maga\'rDa kuku\'ny\'-xa-\'ki-tse
Margarida ill-2S-DECL-DECL
'Margarida says you're ill.'

As a result, quotations with second person subjects are ambiguous: thus they allow a direct-speech interpretation (10a), where the subject was the addressee in the reported speech event, and an indirect-speech interpretation (10b), where the subject is the addressee in the primary speech event (and would have been third person in the reported speech event):

(10) tokorit-xa-\'ki-tse
tired-2S-DECL-DECL

(a) 'She said that I am tired.'

'my original speech was 'you're tired', addressed to a person X who was addressee in the reported speech event and is now the speaker'

(b) 'She said that you are tired.' (original speech was 'he/she is tired', about a person X who is now the addressee in the primary speech event)

A very similar phenomenon is found in Slave (Rice 1989: 1274), an Athabaskan language of Canada. Slave has the typologically unusual feature that verbs taking quotations or other complements divide into two types:

- Direct discourse-determining verbs, whose complements are direct quoted discourse; person in their complements is calculated with respect to the reported speech event

---

11 I am grateful to Henk van der Voort for supplying me with these additional examples, which are not in his 2004 grammar, and which are based on additional information he obtained from Mário KwaZu in January 2009.

12 A more explicit way to express this would be: 

-mar\'t\'-xa-\'ki
tired-2S-DECL-DECL

'She said to me 'You're tired.'
Some problems in the typology of quotation: a canonical approach

Slave (Nez Perce) or the speaker (Nez Perce) to trump the person value that they would have as calculated from the perspective of the reported speech event. This leads the construction to depart from canonical direct-speech status those values of its person features that are susceptible to second person attraction, even though in all other respects the construction is direct speech.

The phenomenon of speech-act participant attraction in these languages raises interesting questions for future typological research. Based just on this data one might propose a hypothesis that if it only targets one participant, it will be the addressee in the primary speech act; if it extends to a second participant, it will be the speaker in the primary speech act; and also that attraction also reaches from the primary speech act outwards to the reported speech act rather than the other way around. Though I am not aware of languages that falsify this implicational formulation, it would be interesting to test this against a wide range of languages.

4.4 Some deviations from canonical indirect speech

We now turn to the other pole of the traditional opposition, namely indirect speech. First we look at cases where deviations from canonical indirect speech are merely stylistic (section 4.4.1), though cross-linguistic comparison shows us that what is generally regarded as style in some languages (e.g. English) has important grammatical consequences in others (e.g. Tamil). We then pass to a range of grammatical deviations, that show how varied are the dimensions in which languages can resist total assimilation to the reporting speaker’s point of view despite using constructions which satisfy the definitional requirements for indirect speech in some regards: person shift but no tense shift in Russian (section 4.4.2), person shift but no shift away from modal or vocative values in Jamsay (section 4.4.3), honorific shift but no tense shift in Japanese (section 4.4.4). In each of these cases the behaviour splits right along an indexical dimension. But there are also cases where a single dimension is itself split between the two poles: spatial deictics may split by coding site (directional prefix vs. adverb), as we will see in section 4.4.5, and person can behave differently according to grammatical relation in a number of languages (section 4.4.6).

4.4.1 Emotional colouration—style or grammar?

One language’s style is another’s grammar, and it is revealing to see how choices which appear merely stylistic in one language reappear as grammatical constraints in another.

To begin with the stylistic, consider the following example from the novel The Believers. This is set in New York but with one character (Audrey) who is signaly British and whose speech is consistently depicted using rather stereotypical British lexical choices not employed by any of the American characters, such as ‘I rather
think that…' In the following example, which in fact involves reported thought rather than reported speech, this particular stylistic marker is used by the narrator, who otherwise lacks such British markers, even though all other markers (person, tense) have been assimilated to the narrator's standpoint as in canonical indirect speech. (Naturally, in order to process the sentence in this way, we need to know that this use of rather is not normally characteristic of the authorial voice.)

(11) Audrey, who rather thought her mother-in-law had a point, remained silent. (Heller 2008: 87)

A common and less literary manifestation of a similar stylistic leakage concerns the reporting of swearing: in (12) the emotive word bloody is used in direct speech, but in (13) it survives into otherwise canonical indirect speech (again with person and tense shift).

(12) John said: 'Of course I'll bloody come.'

(13) John said that of course he'd bloody come.

The retention of bloody here might be regarded, on one reading of (13), simply as a stylistic carryover from the original speech.14 But there are languages like Tamil (Steever 2002), in which 'attitude' transcends stylistics to become a grammatical category, marked by a subset of the auxiliary verbs. Interestingly, in Tamil there is a general grammatical constraint against employing attitudinal auxiliaries in indirect speech.15 Attitudinal auxiliaries that are not available in indirect discourse constructions include: telviyang 'lose [main verb use]; damn it! (auxiliary use)', ozhittan 'purge [main verb use]; go and… (auxiliary use)' and po- 'go [main verb use]; get (uncontrolled bad outcome) (auxiliary use)'.

(14) po- y telai
  go-CF lose(IMP)
  'Go away, damn it!'

(15) po- y telai
  go-CF lose(IMP) say-CF tell-PST-3SGM
  'He said, “Go away, damn it!”'

14 For me and several other speakers of Australian and British English who I have checked it with, (13) allows two interpretations—the bloody could express the primary speaker's exasperation with John's decision (i.e. the emotional evaluation is reclaimed from the reporting speaker's standpoint, as expected in indirect speech), or the reported speaker's words—in which case it carries over the original speaker's emotional colouration into indirect reported speech.

15 And note that this constraint does not block the use of other, non-attitudinal auxiliaries such as valthu 'place; do for future use'.

Some problems in the typology of quotation: a canonical approach

(16) “[po-y telai-yar-k]”
    go-CF lose-INF-LIKR
tell-PST-3SGM

'The said to go damn it.'

This Tamil data ups the ante in our definition of canonical indirect speech. Cross-linguistically, the definition of indirect speech now needs to be elaborated so as to exclude attitudes of the reported speaker, since it presents the event from the viewpoint of someone else. If we add this element to our cross-linguistic definition of canonical indirect speech, English sentences like (13), though canonical if only person and tense are considered, become less so once the word bloody is taken into account.

A further example of a language where RSE-attitudinal elements—this time in the form of diminutives in -in-, shown here in bold—intrude into what would otherwise be regarded as indirect speech comes from Italian. Consider the following example from a novel by Carlo Fruttero (my translation):

(17) La baddante è gentile, è lei che mi porta su e giù di peso per la scala, una rusa robustissima, anche se mi ruba i pistacchi, la vacca, o la vedo, ma la tengo un po' alla larga, parla solo per diminutivi, le mie manine, i miei piedini, il mio bastoncino, che invece, toc toc toc, è un bastone da pastore, papà l'aveva preso a un mercato a Grasse o a Monasque (Fruttero 2006: 172–3).

'The caret is kind, she's the one who carries me up and down the stairs, a big solid Russian lady, even if she steals my pistachios sometimes, the cow, I see her, but I keep her at a bit of a distance, she only talks in diminutives, my hands, my feet, my walking stick, whereas really, thump thump thump, it's a great big stick fit for a shepherd, that dad got for me at a market in Grasse or Monasque.'

In neither Italian nor English would these cited examples be regarded as pure indirect speech. On the other hand, the fact that the intrusions from the RSE aren't aspects of the core grammatical system means that traditional characterizations of indirect speech in Italian or English make no stipulation about how these values should be specified. But if our goal is to provide a cross-linguistic canon for indirect speech, we need to add the specification that canonical indirect speech lacks any lexical emotional colouration or diminutives and other evaluatives attributable to the original speaker.

4.4.2 Person shift but no tense shift: Russian

Russian is a well-known example of a language where person is assimilated to the Primary Speech Event in indirect speech, but where the behaviour of tense is more problematic.

The c here is a form of phonological sandhi produced by tight linkage of some constituents to the following verb.
In Russian, tense generally remains that of the Reported Speech Event (Borras and Christian 1971: 144–5): 17

(18) On skazao-l
he say.pfv-pst.m.sg that go.pfv-pst.3sg

'He said that he would go.' (Literally: 'He said that he will go.')

(19) fa sprosi-l ego,
I ask.pfv-pst.m.sg him

lit. 'I asked him, whether he is reading War and Peace.'

(Lit. 'I asked him, whether he is reading War and Peace.')

In Russian, then, indirect speech is canonical as long as one considers only person, but not once one considers tense. The term 'semi-indirect' has been used by authors describing other languages with comparable phenomena, e.g. Avar, where the tense of the embedded speech remains unchanged but there is adjustment of person values (Hewitt and Crisp 1986), though this term has also been used for different phenomena (see section 4.4.6).

4.4.3 Person and tense shift, but retention of imperative mood and vocatives

Jamsay, a Dogon language (Niger-Congo) spoken in Mali (Heath 2008), illustrates two other dimensions on which languages can deviate from canonical indirect speech: mood and vocatives. Quoted speech in Jamsay has a number of traits typical of indirect discourse, most importantly person shifts in free pronouns and subject-agreeing verb forms. It uses logophoric—see below—for arguments which are first person in the reported speech event but third in the primary speech event. However, it retains two important features of the reported speech event: modal categories (especially imperatives) and hortatives) and also vocatives—these latter have the unusual characteristic that

In indirect discourse, the vocative takes the form of a third person human independent pronoun, usually 3SG wô. Therefore even 'indirect' quotations may begin with what is literally 'well, hello, hello...' intended to suggest an original utterance like 'well, Seydou...,' or 'well, Mother...' (Heath 2008: 569).

An example is (20) (Heath 2008: 569). Person shift is shown in both quoted stretches by the use of the logophoric pronoun (substituted for an original first person) and a third person pronoun (substituted for an original second person), as well as in the third person vocative wô in the second line. On the other hand, the second verb ('kill') remains in the imperative, as in the original quoted speech, and the subjunctive vocative form mentioned above is used.

(20) [ënë anà jë wô-rô jëna-lô] wô[21]
LOG God for 3SG-DAT request.pfv-l-3SG-S saying
[ë wô enë të[22] wô-yô] wô
3SG-S LOG shoot kill.imp.m.sg saying

'He said, 'I hereby beg you in the name of God, don't shoot and kill me!' (lit. 'I begged him for (=in the name of) God (L) said, don't shoot L! (L) said,' where L stands for the subject of the logophoric pronoun. Note that the logophoric pronoun is used here because this passage, in its turn, is embedded in the middle of a lengthy dialogue treated as quoted speech.)

For this reason, Heath characterizes such quotative complements as having a mix of "direct" and "indirect" discourse features (Heath 2008: 567), though he notes that it is still possible to distinguish between direct discourse and the (non-canonical) indirect discourse exemplified above on the basis of whether pronouns undergo person shift.

4.4.4 Honorary shift but no tense shift: Japanese

In Japanese, indirect speech is so non-canonical that some authors have claimed the language lacks indirect speech altogether. 23 Pronouns are used so sparingly in

[21] There are just a few constructions in Russian (e.g. predicative constructions) where tense is calculated from the vantage point of the primary speech event:

One čuvatnox-ta, kak on by-d' vvedenov
she fed.pfv-pst.m.sg, now he be-pst.m.sg agitated.m.sg

'She felt how agitated he was.' (Borras and Christian 1971: 145)

[22] Further problems with the choice between second person pronouns and other ways of referring to persons will be discussed in Footnote 21.

[23] The appearance of imperatives in indirect discourse has been reported for other languages as well: Old Icelandic and Swedish (Platnack 2007), Slovenian (Bus 2005), and Manambu (Akhrenvold 2006: 395, 2006b: 49).

[24] In Jamsay, verbs take third person forms in agreement with logophoric subjects, suggesting that for the purposes of verb agreement person has shifted from first to third in indirect speech.


[26] See e.g. Levin (1975: 231): 'Das japanische kennt keine Sätze mit indirekter Rede und auch keine abhängigen Fragetsätze. Jeder Ausspruch wird in der Form der direkten Rede als Zitate gegeben und unterscheidet sich syntaktisch nicht von den üblichen Satsaufbauten.' (Indirectness is translatable as 'Japanese has no clauses with indirect speech and also no dependent questions-clauses. Every utterance is given in the form of direct speech, as a quoted clause, and is not syntactically distinct from a regular clause form.'). As Couperin (1966b: 176) points out, 'Levin's statement is correct inasmuch as indirect speech is not marked by a subjunctive mood, a special word order, or any other overt syntactic feature. However, consideration of other factors, such as the honoraic behaviour discussed below, and the differential availability of the complementizers to and koto (Maynard 1946) can be invoked to show that there is nonetheless a contrast between direct and indirect speech, albeit of a very non-canonical type.
Japanese that they rarely provide evidence relevant to person shift. Tense values are those of the RSE, and quoted material can take an independent topic marker wa in the quoted clause, something not normally found in subordinate clauses (Nariyama 2003). But there is nonetheless a shift to the standpoint of the PSE, on a semantic dimension that is not familiar from European languages.\textsuperscript{23} Honoficocity. Essentially, all sites at which honorificity can be coded are reassessed from the viewpoint of the reporting speaker rather than the original speaker, and this can produce far-reaching changes to the form of the quoted material because of the large number of honorificity-sensitive coding sites. Maximally, the rejigging of honorificity can involve:

(a) removal of addressee honorifics from quoted speech
(b) addition of subject honorifics on the verb, reflecting viewpoint of reporting speaker
(c) addition of honorific prefix on subject, again reflecting viewpoint of reporting speaker

This can be illustrated by the following pair of sentences:\textsuperscript{24} (21a) is direct speech, and (21b) indirect, with affected sites marked bold. In (21a) ryooshin 'parents’ doesn’t take an honorific prefix, nor is the verb in subject honorific form, because the subject is in the speaker’s in-group/family; on the other hand the verb takes the addressee honorific suffix -mas, which the original speaker (Yasukawa) uses to express polite-ness to his interlocutor. In (21b), on the other hand, the verb in the reported clause is in the plain form (with no -mas suffix). In contrast, the word ryooshin ‘parent’ takes the honorific prefix go- to express the primary speaker’s deference to the referent, and this respect is also shown by the choice of the subject-honorific verb irasshaku ‘be, come (of honoured subject)’.

\textsuperscript{23} Greer Corbett and Marina Chumakina, in commenting on a draft of this paper, suggested that this phenomenon is far from being limited to Japanese, and indeed would be found in Russian. Consider a situation where an adult, Kelya, visits the house of his friend Sasha in the house are Sasha’s parents, to whom Kelya uses the polite vy, and Sasha’s two children. Sasha leaves, saying to his father something like ‘Dad, you [by [intimate] read to the older one] and Kelya can play with the younger one;’ Once Sasha has gone, his father asks Kelya ‘What did he say?’. The appropriate response from Kelya would be Os alosh, Kelya vy ioni so storion. ‘He said that you [you should read with the older one], adjusting the second person pronoun to the vy appropriate to the current speech act away from the vy used in the original. Another example suggested by Marina Chumakina involves the choice between titles. Again say I am on close terms with Aleksandr Ergenevic Kibrik and refer to him with the diminutive Sasha as an index of my close acquaintance. I then say to a third person something like Sasa no poshe ‘Sasha isn’t going to the conference.’ If that third person is on more formal terms with A.E.K., they would normally render this into indirect speech using a title appropriate to their relationship. Nik shalal, On Aleksandr Ergenevic no poshe using a first name + -t patronymic formula instead of the diminutive.

These facts are interestingly parallel to the Japanese situation. Russian honorificity, that is—whether manifested in second person pronoun choice or in titles—is adjusted to the standpoint of the PSE in indirect discourse, though the difference is that in Japanese honorific adjustment is the primary dimension on which adjustment is made, whereas in Russian it is normally taken to be personal (and to my knowledge the adjustment of honorificity in Russian indirect speech has not received significant discussion in the literature).

\textsuperscript{24} Examples from Horikawa 2005: 53

4.4.5 Mixed spatial deixis in Nev Perce

Nev Perce, as already mentioned, is a language for which some analysts (e.g. Aoki 1979) have claimed that all speech is direct. However, there is some assimilation of person deixis, in the case where referents are the speaker or the addressee of the primary speech event (section 4.3.2 above). Primarily on this basis, Deal (2009) argues that it does indeed have some form of indirect discourse. In addition to the calculation of some person values in reported speech from the primary speech event, there is an interesting phenomenon involving spatial deixis. Locational adverbs like koni ‘there’ (22) and kine ‘here’ (23) are anchored with respect to the PSE, whereas the collocative verbal suffix -m ‘towards’ is anchored with respect to the RSE (examples from Deal 2009):

(22) Context: we are not at the café, I am reporting on the activities of my friend who is there
hi-hi-ne 6ykalos-na ke koni hi-w-s-iis-5
35-say-ppp everyone-oni rel there 35-be-ipv-s-pl-ppp
\textsuperscript{25} e-nego-sepekkis-s5-5
wife 55>3081-obj-pl-introduce-ipv-ppp
\‘He said he’s introducing his wife to everyone who’s here.’

[Original utterance: ‘I’m introducing my wife to everyone who’s here’, so person is not adjusted to the PSE, but the locational adverb is.]

(23) [We are in Lapwai, a policeman is stationed 3 miles to the north in Spalding, a speeding car from the direction of Lapwai comes towards him.]

\textsuperscript{25} Example from Horikawa 2005: 155
Thus in Nuxalk, two semantic dimensions exhibit only partial assimilation to the RSE—within person, referents who are speaker or addressee in the RSE are assimilated, but not others, and within the dimension of space, locational adverbs are assimilated but not directional suffixes.

### 4.4.6 Person deixis mixed across arguments

Another intriguing manifestation of partial assimilation to the RSE’s standpoint comes from languages in which some argument positions undergo person assimilation and others don’t. (This phenomenon has been called ‘semi-indirect speech’ by Loughnane (2003, 2005) in her work on Golin, and ‘semi-direct speech’ by Alkhennvald (2008a) in her cross-linguistic survey on similar phenomena in Manambu and other languages.) This phenomenon has been reported in Classical Egyptian (Kamernell and Peust 2002), Tamil (Lehmann 1989), and a number of Papuan languages, including Usan (Reesink 1993), Golin (Loughnane 2003, 2005), Dom (Tida 2006), and Manambu (Alkhennvald 2008a, b). Alkhennvald (2008a: 393) characterizes this phenomenon in the following way:

> in semi-direct speech, the reference for some participants is shifted, while for others it is not (Alkhennvald 2008a: 393).

However, this is a problematic formulation, since it is not the reference that shifts, but the person values, so that a better formulation would be: the person deixis of reported-clause participants is shifted in some syntactic positions, but not in others.

We can exemplify this phenomenon with the following example with Golin (24).

(24b) 

- **Primary event:**
  - **Speaker**: [x <i>]
  - **Hearer**: [hit y <i>]
  - **Others**: said x <2nd>

- **Reported event:**
  - **Speaker**: na siwa
  - **Hearer**: dian-ge

Golin is reminiscent of the problems we encountered with second person attraction in Kwa & Slave (second and first person attraction in Nuxalk). But in those examples the determinant was one role in the primary speech act, whereas here the mixing of strategies is determined by syntactic position (e.g. subject vs. object in Golin). Moreover, there are other grammatical features which systematically co-occur with the mixing of person values, such as the distal and reportive verbal suffixes in Golin:

- the main features of semi-indirect reported speech and thought [in Golin—N.E.] are the distal marker -n, the reported marker -w (for semi-indirect reported speech only), absolute spatial deixis, and mixed relative–absolute person encoding via pronouns and nominal suffixes (Loughnane 2005a:25).

---

23 In fact, the Golin situation becomes more complicated once we look at other person values. For example, where the subordinate clause subject is first person in the PSE and speaker in the RSE, there is a mismatch between actual and free pronominal representation: ‘He said [of me] “he will go’ is rendered (very roughly) as ‘he [I do go-he] said-he’. See Loughnane (2003, 2005) for a thorough exploration of the full set of interaction possibilities.

24 Golin is typologically unusual in that it is the first person rather than the third person which has a zero realization in the subject agreement paradigm—see Lee (2005).

---

---
A related but interestingly different phenomenon is found in Egyptian, for various phases of its history:

In Late Egyptian reported speech, one of the grammatical person roles within the utterance may be chosen with respect to the deictic setting of the embedding context, but never more than one (Kammerzell and Peust 2002: 308).

In Egyptian, that is, the constraint is not based on argument position, but on a ration of one argument per quoted clause, though in some cases there is vacillation between PSE- and BSE-based person assignments even within a single passage. I lack the space to work through all the attested combinations here—see Kammerzell and Peust (2002) for the details—but will illustrate their proposition with two examples from their article, representing different historical phases of Egyptian. In (25), written around 1200 BCE, the quotation actually vacillates between a direct-speech first-person-plural verb first time around (preceded by an imperative), then followed by a verb in the third-person-plural, calculated from the PSE.

(25) jw-s-gd
    n-f
    mj
    jr-mn
    comp-3p-sg-say-inf
    for-3m-sg
    come-inf
    make-s-1pl
    n-n
    wr-t
    sgr-w
    for-1pl
    hour-f
    sleep-stat-3pl

    ... and she said to him: “Come and let us spend some time making love.”
    (lit. ‘and she said to him, “Come and let us spend some time they making love”’)

In (26), from the later Coptic phase of Egyptian, the Emperor is addressing a monk who had been given the assignment of curing his sick daughter. The monk had resorted to rather unusual methods, as revealed by the account given to the Emperor by his healed daughter: the monk is referred to using the second person (calculated from the PSE), while the daughter is referred to using the first person (calculated from the RSE).

(26) a-ė ’jê:6
    gar na-j nêl i ‑ têrê
    ptl-3f-sg-say-inf,inf
    for-sg
    pos-1sg-daughter
    jo
    hâh n- ‘têp
    sa-k-os’padzê
    comp many of-occasion
    hab-2msg-kiss-inf
    n-ta-tap’tê
    awo jo
    sa-k-on’to:k
    prep-poss-1sg-mouth
    comp
    hab-2m-sg-sleep-inf
    hîn-u ‘poj
    n-’wot
    nêm-u
    n-ta-w jî
    ’th-s
    on-indef-bed
    attr-single
    with-3sg
    in-art-p-night
    all-3p-s

    ‘My daughter has told me that you frequently kissed her on the [my] mouth and that you used to sleep with her [me] in a single bed all night.’ (Drescher 1947:1, lines 16-18, cited in Kammerzell and Peust 2002).

Languages like Golin and Egyptian appear to have grammaticalized, in somewhat different ways, the tendencies reported on in section 4.3.2 above for first and second person participants in the PSE to exert their influence on the person values assigned to the quoted clause. This turns person-based attractions into systems based on which argument is involved, as in Golin, or on the selection of a single argument to shift, as in Egyptian. Languages can yield to these forces in many ways—they can grammaticalize person-based attractions (Kweza, Slave), or argument-based assignments (Golin), or stipulate that one referent can shift (Egyptian). In each case they end up with a grammatically systematic set of rules that is neither canonically direct nor canonically indirect.27

But even in languages that have not grammaticalized a resolution of these competing tendencies, ‘performance errors’ that exhibit similar tendencies can be found in actual discourse, as in the following English example:

(25) X had become confused about which house his daughter-in-law Y lives in, knocked on the neighbour’s door, and had been directed to the daughter-in-law Y’s house.

Later he tells his daughter-in-law what had happened.

‘They told me: “Oh, you’ve got the wrong house, you live next door.”’

X Y

[Example overheard by N.L. in Canberra, June 2009]

Cases like this, of course, raise the issue of how easy it is to make an analytic distinction between person attraction as a grammatical phenomenon and person attraction as a performance error, particularly in corpus-based studies.

4.4.4 Summary

The cases discussed in this section all involve structures that depart from canonical indirect speech. Some select certain semantic dimensions to shift to the primary speaker’s viewpoint (person of pronouns in Russian and of vocatives in Jamars, honorifics in Japanese), while leaving other dimensions with their regular values—tense in Russian and Japanese and imperatives in Jamars. Others make the split within a single semantic dimension, conditioned in various ways—by grammatical

27 It appears that, in general, such mixed patterns always result from the pull of the speech act participants in the primary speech event. Nonetheless, one occasionally finds examples like the following—which would normally be considered incorrect—in which the mismatched person shifts must be accounted for either by attraction to the first person in the reported speech act (if regarded as basically indirect speech), or by re-evaluation of a RSE first person to third person in the light of its role in the primary speech act (if regarded as basically direct speech).

In two games on Wednesday, Japanese pitcher Yukiko Ueno threw a total of 13 innings, but she showed no ill effects as she played seven innings and got the gold medal for Japan. ‘My strong belief to win, was what carried her through the fatigue,’ Ueno said. Canberra Times, 24 August 2008, C6.
relation in Golin, by the choice of a single argument to shift in Egyptian, or by
whether the spatial information is in a spatial adverb or a directional prefix in Nez
Perce. Whatever the system, the language employs a principled mix of shifted and
unshifted values when the ensemble of deictically sensitive dimensions is considered.
It should be emphasized that this by no means exhausts the ways dimensions can be
mixed and matched—for example it is also possible to assess space with respect to
the reported speech event, but tense with respect to the primary speech event, as
Heath (1980) reports for Nunggubuyu. A full survey of all the ways of splitting
dimensions, as well as a search for any implicational statements that hold between
them, is a challenging task for future research.

Systems of the type described above can, evidently, be treated neither as straight-
forward direct nor as straightforward indirect speech.98 Of course it would be
possible to adopt a compromise term like ‘semi-direct speech’ or ‘semi-indirect
speech’, but the problem with this is that it can denote so many types of mixture:
‘ssemi-indirect speech’ has been used, for example, for systems that shift person but
not tense, as in Avar, but also for systems that shift person in some grammatical
persons but not others, as in Golin or Manambu. What is clearer is that these
systems all depart from both canonical direct and canonical indirect speech, but that
there are many ways of doing this.

Complex as these systems are, they do not exhaust the ways in which languages
can deal with the interwoven perspectives of reporting and original speaker. After all,
in all of the cases discussed above, the chosen value is either that of the primary
speech event or of the reported speech event. In the next section we turn to another
option: to choose a value that is a function of both speech events.

4.5 Another reason the typology of quoted speech needs revision:
biperspectival speech

Traditional dichotomies, as we have seen, contrast direct and indirect speech as
representing the viewpoint of the original speaker (in direct speech) or of the
reporter (in indirect speech). However, there is a third possibility which needs to
be made explicit: that of ‘biperspectival speech’ in which constructions simulta-
aneously represent two distinct viewpoints (cf. Evans 2007).99 Before seeing how this

98 This applies to the use of the terms in cross-linguistic comparison, where fixed reference points
for the terminology are important. In enclitic descriptions of particular languages, on the other hand, it may be
useful to use the terms for whatever construction type is used to encode the opposition in the deictic way:
person (and honorific) shift in Russian, honorific shift in Japanese, and so forth.
99 As several colleagues who read this chapter pointed out, this term is not perfect: the contrast between
direct and indirect speech is not only about perspectives or viewpoints, but also about how far the original
form is maintained (cf section 4.3.1), which is not a matter of perspective but is better characterized along
the traditional lines of form vs. content, de direct vs. de re and so on. Nonetheless, once it comes to the
might apply to quoted speech, let us consider an example from quite another part of
the grammar: that of ‘irrational kin terms’, which simultaneously calculate kin
relationships from two points at once. These are widespread in Australian languages
(see e.g. Merian 1989) and also attested in the Amazon (Lea 2004). An example from
Binin Gun-wok (Evans 2003) is the following pair of terms, each of which relates the
referent by two different kin relationships—one with respect to the speaker, and one
with respect to the hearer.

(28a) al-gerring ‘the one who is your mother and my daughter, given that I am your
mother’s mother’

(28b) al-doingu ‘the one who is your daughter and my mother, given that I am your
dughter’s child’

Alternatively, kin terms in some Australian languages build in a calculation of two
different kin relationships to the same person (the speaker) at two different points
in time, as with the following Iwaidja term:

(29) rangalalik (converse bukunwar): someone who I (a male) used to call ‘sister’s
child’ and who, following the death of his or her father, I now call ‘child’
(niowmay)10

Such constructions establish the need—indeed the necessity of quoted speech—to allow
the encoding of two perspectives at once. Could it not be the case that there are types
of quoted speech that also do this? I will call this ‘biperspectival speech’, as a third
term in the typology ‘direct’ vs. ‘indirect’. As with direct and indirect speech, it is
more helpful to examine each coding dimension one at a time, rather than assuming
there is a single choice which then plays out in all dimensions.

4.5.1 Logophoricity

One phenomenon involving a double deictic perspective on reported speech and
thought in some languages, especially in West Africa, involves 'logophoric' pro-
nouns, as exemplified by the following examples from Ewe (Clements 1979):11

semantic, deictically sensitive dimensions we have been focusing on in this paper—tense, person,
location, mood, hortenclitic, emotional evaluation—perspective becomes the dominant factor, and
I have been unable to come up with a better term. For the reasons outlined in the body of the paper,
I do not think semi-direct or semi-indirect is an effective alternative, since these cover so many different
kinds of phenomena and, more importantly, are simply a way of alternating between the two perspectives
rather than triangulating from both at once.

38 There is a second meaning for this term, also reflecting a calculation of a kin relation between a
single pair of people from two points in time: it can be used for someone who used to call me my
wife (mother’s brother) and who, following the death of his/her father, now calls me bujag (father).
39 See Amele (2004) for an interesting account of cultural practices in West Africa involving logophoric
pronominals, which includes a number of examples of unambiguously logophoric uses, rather similar to
free indirect speech, where mediating speakers repeat messages emanating from a third party.
(30) Kofidi be e dixo
Kofidi say 350 leave
Kofidi said (s)he left. [Orig. utterance: e dixo 's)he left']

(31) Kofidi be yä-dixo
Kofidi say 1D0-leave
Kofidi said he, left. [Orig. utterance: yë dixo 'I left']

The logophoric pronoun yë- in this example is best analysed as representing reference to a person who was the speaker in the reported speech event, but is third person in the primary speech event. To use it correctly thus involves the simultaneous calculation of person values for both the reported and primary speech events—a clear example of biperspectival speech on the person dimension.

4.5.2 Complex tenses in Western Indo-European languages

The concept of biperspectival quoted speech can also be applied to tense. It is well known that Western Indo-European languages employ complex tenses in indirect speech, and it has long been established that the representation of the semantics of these tenses requires reference to two temporal reference points, one derived from the RSE and one from the PSE. Consider the following pair:

(32a) John said: 'I will be at work.'
(32b) John said he would be at work.

The semantics of would be in this sentence is generally represented along the lines shown in (33): note that the narrated event is after the Report frame, which is before the Speech event.

(33)

Secondary (report) frame  Narrated event  Speech event
RSE                  PSE

(Reference point)  Situation  Present

In (34a) and (34b), on the other hand, the narrated event is before the Report frame, which is before the Speech event (35):

(34a) John said: 'I have been at work.'

(34b) John said he had been at work.

(35)  

Narrated event  Secondary (report) frame  Speech event

Though there are a range of approaches to this phenomenon, all agree that complex tenses simultaneously locate the action with respect to both the PSE and the RSE. Thus for Comrie (1985), the reported event is located with respect to the RSE, which is in turn located with regard to the PSE (36):

(36)

On the other hand, for Reichenbach (1947) and Hornstein (1977), the reported event is located with respect to both the RSE and the PSE:

32 See Bull (1976) for the original discussion; also Comrie (1985).
33 The second line gives the terminology employed by Comrie (1985: 54 ff.).
On either analysis, selection of an appropriate verbal syntagm (i.e. verb plus auxiliaries) depends on integrating information from both the primary and the reported speech events. In that sense it is inaccurate to characterize the construction in terms of either direct or indirect speech, each of which implies a single perspective (original or reporting speaker’s). Postulating a third, biperspectival type of speech allows a much more accurate representation of the semantics.

4.5.3 Other possible complex dimensions

I lack the space to exhaustively survey all the semantic dimensions in which biperspectival speech is found in quoted speech and thought. Many well-known cases invite analysis in this way. Should German subjunctives in quoted speech, for example, be treated as a biperspectival value, e.g. ‘primary speaker passes on modal evaluation to someone else; reported speaker asserts the truth of the proposition’? Or can it be assimilated to indirect speech, assessing the modality simply from the point of view of the primary speaker (‘primary speaker passes on modal evaluation to someone else’). Rather similar is the situation with the future and perfective infinitives in classical Greek, which occur almost exclusively in reported speech (Goddwin 1892, Noonan 2007). These are traditionally regarded as occurring in ‘indirect’ speech. But would it be better to regard them as a biperspectival modality, with different values for the primary and reported speaker? Or can they be assimilated entirely to a value given by the primary speaker (e.g. ‘event commanded by someone to be the case’)? Both the German and Greek examples seem to be prima facie cases of biperspectival modality, but more detailed analyses than I can give here would be needed to settle the argument.

One can likewise ask whether biperspectival values exist on other dimensions. Do any languages have biperspectival spatial systems comparable to complex tenses or logophoric pronouns—with meanings like ‘far from where the primary speaker is now, but close to where the reported speaker was at the moment of speech’? Complex honorific systems are also imaginable. A far-reaching cross-linguistic investigation may turn up other examples on other dimensions (e.g. evaluation—believed to be true by the reported speaker, but known to be false by the primary speaker).

In canonical biperspectival speech, then, morphemes on every deictically sensitive dimension would exhibit values that require simultaneous reference to both the RSE and the RSP—person would encode values like ‘neither speaker nor hearer in this speech act, but speaker in the reported speech act’ (like logophoric pronouns), tense would encode values like ‘before the current speech event but after the reported speech event’, location would encode values like ‘far from where I am now but close to where the speaker in the RSE was then’, modality would encode values like ‘I believe this proposition is false though the speaker in the RSE claimed it to be true’, honorificity would encode values like ‘person who I exhibit respect to but with whom the speaker in the RSE was on intimate and informal terms’, and emotional evaluation would encode values like ‘what the speaker in the RSE considered a good thing but which I consider a disaster’.

That is what the biperspectival canon would look like. An immediate objection would be: no language has a system anything like it, so why postulate it? There are two good reasons: first, as argued above, there are examples of phenomena in at least some dimensions—logophorics in the dimension of person, and complex tense in the case of tense—so that it is not simply a hypothetical way of doing things. Postulating this third canonical type, even if it is currently unattested, is a useful heuristic in helping us look for other instantiations as we extend our interest to the many poorly described languages of the world. The second reason is that it crystallizes the difference between this logically clear but empirically rare strategy and what is an empirically commoner but logically much more variegated strategy, that of alternating between perspectives in ways that generate the immense variety of so-called ‘semi-direct’ or ‘semi-indirect’ types which are constructed by mixing together some morphemes encoding the perspective of the RSE and some encoding the perspective of the RSP, a phenomenon we revisit in the next section.

4.5.4 Biperspectival values vs. mixed values on a single dimension vs. distribution across dimensions

Throughout this chapter we have been exploring various ways in which quoted speech mixes semantic values from the original speaker’s perspective, the reported speaker’s perspective, and a biperspectival perspective. It is worth recapitulating these here in a more principled way:

(a) dimensional split. Values (direct, indirect, biperspectival) may be mixed across semantic dimensions. Thus Russian assesses tense from the reported speaker’s perspective, but person from the primary speaker’s perspective; Japanese assesses tense from the reported speaker’s perspective, but honorificity from the primary speaker’s perspective. English ‘indirect speech’ arguably assesses person from the primary speaker’s perspective, but tense from the perspective of both primary and reported speakers.

(b) site split. Values on a single semantic dimension may be mixed by being realized in different ways at different coding sites. Colin does this for the...
dimension of person: it selects subject values from the reported speaker's perspective, but object values from the primary speaker's perspective.

(c) value split. Even while holding dimension and site constant, languages may vary their choices of perspective according to particular values on a dimension. A clear example of this is 'second person attraction' in languages like Slave and KwaZulu. Here quoted speech is normally assessed from the reported speaker's perspective on all dimensions, but within the dimension of person it adopts the primary speaker's perspective for referents who are addressees in the primary speech act.

Though useful, this conception of the typology faces a boundary-definition problem, hinging on what level of sign we are examining. If we are just looking at the morpheme, the sites are easy to identify but we face the standard problems of analysis that confront complex signs, as gestalt-like constructions. Consider an English string like *would have been*—if we just look at individual words we might say that the *have* (or *been*), by coding perfectivity, is taking the primary speaker's perspective, while the *would* is taking the reported speaker's perspective—this would then be a case of *site split*. On the other hand, we could take *would have been* as a single complex sign—this is implicit in the way that the semantics of English complex tense is usually analysed. Then we have a dimension split, in which tense is straightforwardly biperspectival (as argued above), but that this biperspectival value is realized on a complex sign.

Similar arguments apply, *mutatis mutandis*, to the analysis of person in a language like Goli. We could identify a complex argument-coding construction mapped out on various sites, in which case Goli would exhibit a dimensional split (with person coded in a biperspectival way across the complex syntagm comprising the various argument slots). Or we could go for simpler signs (object slot, subject slot, etc.), in which case we would have a case of *site split*, which is the preferred analysis I have adopted for Goli in this paper. The reasons for preferring one analysis over another are not straightforward, but a key factor will obviously be the degree to which, on independent grounds, the individual signs concerned can be shown to form a composite sign.

These two cases show that the concept of dual perspective itself has a clear canon: that the morpheme realizing it should be a simple sign. Logophoric pronouns like Ewe ye, being monomorphic, clearly conform to this canon. Complex tenses in English and other Western European languages depart from it by involving multi-word syntagms, and as far as I know a monomorphic complex tense marker has yet to be reported in the typological literature.

These considerations lead us back very naturally to the broader typology of quoted speech. The traditional dichotomy of direct vs. indirect speech rested on an assumption that a large number of semantic dimensions would pattern together—that recalculation of person and of tense values, for example, would go together. More precisely, it assumed that this happened at each of the three levels outlined above—of semantic dimension, of coding site, and of individual feature values. The classical conception of direct speech, in other words, is that all values, at all coding sites, in all dimensions, would be calculated from the reported speaker's perspective, while indirect speech would take the primary speaker's perspective for all values, coding sites and dimensions. Each of these defines a canonical type of reported speech at the level of the whole quoted-speech construction. In our third canonical type, biperspectival speech, all morphemes would take both perspectives, as indicated in the last section. And, between these three points, lie an immense number of mixed types where individual dimensions, sites, or values split in which of the three canons they draw their strategy from.

The cases discussed in this paper have shown that splits of one type or another are widespread, perhaps even the norm cross-linguistically. In the next section I conclude the chapter by sketching how a canonical approach can best develop a typology of this exuberant field of cross-linguistic variation.

4.6 Conclusion: towards a canonical approach to the typology of reported speech

'It should be noted that the heart and soul of indirect discourse' (Yoleishnik 1975: 150)

It should now be clear how useful a canonical approach is for dealing with a phenomenon that exhibits so many ways of deviating from the idealized norms still encountered in the literature. In the course of the chapter I have proposed definitions for three canonical types (direct, indirect, and biperspectival) that draw attention to a number of levels of specificity from the construction down through the dimension and the coding site to the value.

These canonical types were reached by deductive extrapolation from attested examples. Though drawn from fourteen languages across all continents, they only exemplify a tiny fraction of the possibility space. A substantial task for the typology of reported speech is to chart the actual behaviour of a much wider sample of languages, for the whole set of relevant semantic values, against this possibility space. In the meantime, it seems clear that the three canonical types sketched above have widely differing patterns of empirical instantiation against the data generally reported in the quoted-speech literature.

(a) canonical direct speech is widely attested, and is hence easy to exemplify from many languages where all dimensions behave in a canonical way
(b) canonical indirect speech turns out to be quite rare. In most languages said to possess indirect speech, only some dimensions are involved (say, person in Russian but not tense). In other widely quoted cases—so-called indirect speech in English, French, or Italian are good examples—we are dealing
The status of the perspective and indirect speech in this typology (also illustrated in the Appendix) is determined by the fact that they do not have a direct equivalent. This is the case with direct speech, in some cases unattested, or barely attested, and with orientation (e.g. Russian tense direct, others indirect), honorificity indirect).

### Table 3.2 Matrix illustrating some of the ways the three canonical types can be combined

<table>
<thead>
<tr>
<th>Dimension</th>
<th>1</th>
<th>2.1</th>
<th>2.2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>FSE value</td>
<td>[a, b.]</td>
<td>[c, d.]</td>
<td>[c, d.]</td>
<td>[e, f.]</td>
<td>[g, h.]</td>
<td>[j, k.]</td>
</tr>
<tr>
<td>Kosovo</td>
<td>[a, b.]</td>
<td>[c, d.]</td>
<td>[c, d.]</td>
<td>[e, f.]</td>
<td>[g, h.]</td>
<td>[j, k.]</td>
</tr>
<tr>
<td>Canonical direct speech, e.g. (as)</td>
<td>[a, b.]</td>
<td>[c, d.]</td>
<td>[c, d.]</td>
<td>[e, f.]</td>
<td>[g, h.]</td>
<td>[j, k.]</td>
</tr>
<tr>
<td>Canonical indirect speech (no attested case)</td>
<td>[a, b.]</td>
<td>[c, d.]</td>
<td>[c, d.]</td>
<td>[e, f.]</td>
<td>[g, h.]</td>
<td>[j, k.]</td>
</tr>
<tr>
<td>Canonical bireferential speech (no attested case)</td>
<td>F([0, \beta, 0], [a, b, i])</td>
<td>F([\tau, \delta, 0], [c.d, l])</td>
<td>F([\tau, \delta, 0], [c.d, l])</td>
<td>F([e, \varphi, 0], [e.f, l])</td>
<td>F([\tau, \delta, 0], [g.h, l])</td>
<td>F([l, m, n], [j, k, l])</td>
</tr>
<tr>
<td>Speech with one direct dimension but others direct, e.g. Russian (tense direct, others indirect), honorificity indirect</td>
<td>[a, b.]</td>
<td>[c, d.]</td>
<td>[c, d.]</td>
<td>[e, f.]</td>
<td>[\tau, \delta, 0]</td>
<td>[j, k.]</td>
</tr>
<tr>
<td>Speech with one bireferential dimension, one direct dimension, and the others indirect (no exact example)</td>
<td>[a, b.]</td>
<td>[c, d.]</td>
<td>[\tau, \delta, 0]</td>
<td>[e, f.]</td>
<td>[\tau, \delta, 0]</td>
<td>[j, k.]</td>
</tr>
<tr>
<td>Speech with one dimension split by coding site (Galin, with person split between subject (direct) and object (indirect))</td>
<td>[a, b.]</td>
<td>[c, d.]</td>
<td>[c, d.]</td>
<td>[e, f.]</td>
<td>[\tau, \delta, 0]</td>
<td>[j, k.]</td>
</tr>
<tr>
<td>Speech with one dimension split by value (e.g. Kwa or Slave)</td>
<td>[a, b.]</td>
<td>[c, d.]</td>
<td>[c, d.]</td>
<td>[e, f.]</td>
<td>[\tau, \delta, 0]</td>
<td>[j, k.]</td>
</tr>
</tbody>
</table>
Table 4.2 summarizes the range of phenomena considered above, indicating schematically how the behaviour of various canonical and mixed types can now be placed in the typology we have been elaborating. Relevant semantic dimensions are simply numbered, followed by a further number when they have more than one coding site (e.g. subject and object for person, as 4.2.1 and 4.2.2). Roman letters are used for value sets calculated from the perspective of the primary. Greek letters for value sets calculated from the perspective of the Reported speech act, while a function from both value sets is used where the calculation is biperspectival.

It will readily be appreciated what a huge number of ways languages can have of realizing quoted speech. The three canonical types are the situations when all dimensions, all coding sites, and all values behave the same way — i.e. all being calculated from either the PSE, the RSE, or from both perspectives at once. Once things get more mixed the number of possibilities is astronomical: the full possibility space (impossible to show here) would be the product of all dimensions (person, tense, mood, honorificity, space, evaluation, etc.), times all coding sites (e.g. subject, object, possessor, etc.), times all values (e.g. first vs. second vs. third person). Each of these is able to choose between direct, indirect, or biperspectival behaviour. This is the reason for the immense number of possible mixed types. With a design-space matrix of this type, we can go on to see which particular mixtures are actually attested, which are not, and what the implicational statements may be (e.g. that if there is a person magnetism of the type outlined in section 4.3.2 it will target speech act participants over non-participants).

The complexity of representing the words and thoughts of others and relating them to the perspective of ourselves and our interlocutors lies at the heart of our ability to coordinate, distinguish, and calibrate the jostling versions of a partly shared social world. In the evolution of language, as in biological evolution more generally, the more factors are at work in shaping structure the more complex and varied are the grammatical solutions that emerge. This makes the need for a supple yet comprehensive typology especially acute in the realm of quoted speech. A canonical approach, I have argued, is signally helpful in moving us towards that goal.²⁸

²⁷ See e.g. Nidel (1994, 2004) for simulations of the evolution of plant morphology under different numbers of selective factors.
²⁸ An anonymous referee, while accepting the value of multifactorial typologies in characterizing the full design space, suggests we can do this without needing a canonical approach. I am sure that, for some imaginatively gifted individuals, this is true. But by the same token my own experience both as a typologist and as a descriptive linguist is that it can be difficult to imagine just what hypothetical linguistic types look like as long as they are simply viewed as abstract dimensions for putting the typological cards on the table, so to speak. I believe that canonical typology can help us conceptualize what is going on much more clearly, all the more so when we are dealing with phenomena for which we don’t have ready, existing reference points.

Unpacking finiteness

IRINA NIKOLAEVA

5.1 Introduction

Typology is usually thought of as the study of grammatical patterns that systematically occur across languages and the search for possible explanations for them. The standard procedure in typological research can be described as follows (simplified here after Croft 2000: 14): (i) determine the particular semantic-pragmatic structure or situation type that you want to explore, (ii) examine the morphosyntactic constructions or strategies used to encode that situation type, and (iii) search for dependencies between them and other linguistic factors, such as other structural features, external functions expressed by the construction in question or both. This procedure implies a strong relationship between form and function. The assumption is that there is a set of cross-linguistic semantic-pragmatic entities instantiated in languages by various grammatical means. Grammatical patterns are identified and classified based on these universal semantic concepts (Stassen 1985: 14; Croft 1990: 11–22, 1995: 88; Song 2001: 10–13, and others).

Finiteness has no obvious semantic-pragmatic corollary and therefore cannot be easily equated across languages using this traditional procedure. In typology it is usually understood either as a fuzzy cross-linguistic meta-category with no universal content or a kind of functional tendency, but in this chapter I suggest that canonical typology (Corbett 2005, 2006, 2007a) offers more rigorous comparative tools. The idea is to identify canonical ideals, i.e. the best uncontroversial instances of a linguistic category (rather than semantic-pragmatic situation), to construct a set of relevant parameters and to investigate languages and constructions against them. This provides a clear analytical tool for measuring the actual phenomena of individual languages against well-established criteria.

In some sense canons are not unlike ‘comparative concepts’ (Hapeman 2007; see also Dryer 1997; Croft 2001; and Gil 2003), which are claimed to be convenient abstractions ensuring comparability across grammars. They are not part of the actual language structure and do not have psychological reality for the speakers. Like