

BERNARD COMRIE (Leipzig)

Reference-tracking: description and explanation*

Summary

Various typologically recurrent properties of reference-tracking systems can be given a coherent explanatory account in terms of functional, rather than formal criteria. Two principles in particular are proposed, namely that coreference is more likely to be marked than non-coreference in local domains (e.g. the arguments of a single predicate), whereas non-coreference is more likely to be marked than coreference in extended domains (e.g. across clause boundaries). For the former principle a cognitive explanation is proposed, while it is suggested that the latter principle has a discourse basis.

1. Introduction: Description and Explanation

The phenomenon of reference-tracking, the various linguistic devices by means of which a language can indicate whether reference is being made to the same or to a different participant, has been the subject of a rich literature over the last few years within both formal grammatical and functional approaches.¹ In this article, I want to suggest that much of the typological (cross-linguistic) variation that has been described in this literature can be brought together against the background of a conceptually relatively simple explanatory framework. This explanatory framework will involve consideration not only of the formal properties of reference-tracking systems, but more particularly of their conceptual and functional background. As a simple initial example, I will consider some data concerning the occurrence of reflexive pronouns in a number of languages, in particular the interaction of the occurrence of distinct reflexive pronouns and of grammatical person. Consider English examples (1)–(5):

- (1) *David_i hates himself_i.*
- (2) *You_i hate yourself_i.*
- (3) *I_i hate myself_i.*

* This paper is dedicated to EKKEHARD KÖNIG on the occasion of his sixtieth birthday (15 January 2001).

¹ A Spanish translation of this article appeared under the title “Seguimiento referencial: descripción y explicación” in LUQUE DURÁN, JUAN DE DIOS & PAMIES BERTRÁN, ANTONIO (eds.): *Estudios de Tipología Lingüística*, Granada, 1998: Granada Lingüística, 37–54. I am grateful to the editors for allowing publication of the English original. Versions of parts of the material contained in this article have been presented to various audiences; I am grateful to all those who have provided comments. I am also grateful to participants in the Workshop on Lexical Anaphors and Pronouns in South Asian Languages held at the University of Delhi in January of 1996 for insight into South Asian languages, and to UFFE BERGETON LARSEN for discussion of the Danish data.

(4) *David_i hates him_j.*

(5) *David_i hates me_j.*

These examples show that, irrespective of grammatical person, English requires a reflexive pronoun distinct from the ordinary personal pronoun whenever the object of a predicate is coreferential with the subject of that predicate. Thus we find a third-person reflexive pronoun in (1), a second-person reflexive in (2), and a first-person reflexive in (3). In (4)–(5), the object is not coreferential with the subject, and therefore we find ordinary (non-reflexive) pronouns; subscripts are used in the usual way, with identical subscripts indicating coreference, distinct subscripts non-coreference. We can thus establish the generalization that in both third person and non-third person, English has distinct reflexive pronouns.

Other languages show different systems. Interestingly, Old English had a different system, as can be seen in sentences (6)–(8):

(6) *Hē_i slōh hine_{ij}.*
'He_i hit himself_i/him_j.'

(7) *Ic_i slōh mē_i.*
'I_i hit myself_i.'

(8) *Hē_i slōh mē_j.*
'He_i hit me_j.'

From these examples, it emerges that Old English nowhere has a distinct reflexive pronoun. By comparing the non-third person object pronouns in (7)–(8), we see that the first-person object pronoun has the same form whether interpreted reflexively, as in (7), or non-reflexively, as in (8). The two possible interpretations of (6), with the third-person object pronoun either coreferential or non-coreferential with the subject, show that the same applies to the third person. Let us now compare similar examples from French, as in (9)–(14):

(9) *Satan_i se_i déteste.*
'Satan_i detests himself_i.'

(10) *Tu_i te_i détestes.*
'You_i detest yourself_i.'

(11) *Je_i me_i déteste.*
'I_i detest myself_i.'

(12) *Satan_i le_j déteste.*
'Satan_i detests him_j.'

(13) *Satan_i te_j déteste.*
'Satan_i detests you_j.'

(14) *Satan_i me_j déteste.*
'Satan_i detests me_j.'

In the first and second persons, French behaves like Old English, with no distinction between object pronouns that are interpreted as coreferential or non-coreferential with the subject (as shown by examples (10)–(11) and (13)–(14)). In the third person, however, French

behaves like (Modern) English, and has a distinction between reflexive *se* (as in (9)) and non-reflexive *le* (to cite the masculine singular form, as in (12)).

An obvious question that arises at this point is the following: Are there languages that show the inverse pattern from French, i.e. that would have distinct reflexive pronouns in the non-third person, but lack this distinction in the third person? Although syntactic paradigms such as those illustrated above have been investigated for a vast number of languages from different language families and from different regions, and although numerous instances of languages patterning like English, Old English, or French are known, no language has been found that would have the inverse pattern from that of French. The typological variation is set out in (15), where the asterisked row is the logical possibility not found, and claimed to be in fact impossible for human language:

(15) Occurrence of distinct reflexives

	<u>Third person</u>	<u>First/Second person</u>
English	yes	yes
Old English	no	no
French	yes	no
*	no	yes

The generalization incorporated in (15) can be reformulated as an implicational universal, as in (16):

- (16) If a language has distinct reflexive pronouns in the non-third person, then it will also have distinct reflexive pronouns in the third person.

So far, the generalization of (16) has been presented descriptively, but no suggestion has been offered as to why it should hold. Why should there be languages like French, but no languages like the inverse of French (the asterisked row of (15))? From a descriptive point of view, it is no easier to formulate the generalization (16) than to formulate its inverse, which would allow the asterisked row but disallow French. If, however, we broaden our horizons by considering not only the form of reflexive pronouns, but also their function, then the reason for preferring (16) to its inverse soon becomes apparent. Reflexive pronouns serve the function of selecting one among a range of possible interpretations for a pronominal element, namely, in the case of the reflexive, coreference with the subject. In the third person this is clearly functionally useful, since third person pronouns can in principle refer to any entity in the universe of discourse other than the speaker and the hearer. If we compare English sentences (1) and (4), or French sentences (9) and (12), it is clear that the mere choice of a reflexive versus a non-reflexive pronoun carries a clear difference of meaning. By contrast, in the first and second persons, given that the speaker and the hearer do not normally change in mid-sentence, it makes little difference from a functional viewpoint whether a language requires one to say 'I hit myself' with a reflexive pronoun, as in (Modern) English, or to say 'I hit me', as in Old English and French – the first-person pronouns must either way be coreferential. Thus, the distribution captured in generalization (16) is a special case of a more general principle of economy in language: more explicit expressions are more likely to be used where they are more useful. Note that this does not require a language to use reflexive pronouns anywhere – Old English lacks distinct reflexive pronouns – nor does it prevent a language from using reflexive pronouns where they are semantically redundant – (Modern) English requires them in all grammatical persons – but it does define where reflexive pronouns are more likely to occur, and this is borne out by the cross-linguistic data.

In the body of this article I will be examining a somewhat more complex instance of the interaction of the form and function of pronominal elements, namely the correlation between particular pronominal forms and the DOMAIN that includes the pronominal element and its antecedent. In speaking of domains, the main feature under discussion will be the extent to which the domain is LOCAL versus EXTENDED. The range from local to extended is a scale, with the most local domain being, for instance, the arguments of a single predicate, and progressively more extended domains bringing in the adjuncts of the predicate, elements in other clauses, and finally (as discourse is considered) elements in other sentences. The particular examples in sections 2 and 3 will serve to clarify this, but for the moment it will be useful to state in (17) and (18) the two complementary generalizations that will be illustrated in detail. The explanatory bases for these generalizations will be considered in sections 2 and 3.

- (17) Languages are likely to have special marked forms that indicate coreference within the most local domain (the predicate and its arguments), possibly extending to more expanded domains.
- (18) Languages are likely to have special marked forms that indicate non-coreference, or less expected coreference (absence of ‘topic continuity’) within more extended domains (across clause boundaries), possibly extending to more restricted domains.

2. The local domain

For most purposes, we can think of the most local domain as being that domain which includes only the arguments (subject and objects) of a single predicate. If we look at the cross-linguistic distribution of reflexive pronouns with respect to domain, then it turns out that all languages of requiring reflexive pronouns at least somewhere have obligatory reflexive pronouns in this most local domain. Different languages of this type will then differ from one another in terms of how far the obligatory or optional use of the reflexive pronoun extends beyond this most local domain. In English, for instance, reflexive pronouns are required in the most local domain, as shown in (19):

- (19) *John_i saw himself_i in the mirror.*

A soon as we get beyond this most local domain, however, English typically requires a non-reflexive pronoun, as in (20)–(23):

- (20) *John_i heard steps behind him_i.*
- (21) *John_i saw his_i book.*
- (22) *John_i told Mary to make him_i some tea.*
- (23) *John_i says that Mary loves him_i.*

In each of (20)–(23), the antecedent is subject, but the coreferential noun phrase is an adjunct in (20), a possessor in (21), separated from the antecedent by a non-finite clause boundary in (22), or by a finite boundary in (23). And in each of these configurations, English disallows a reflexive pronoun. The only kind of example where English sometimes allows reflexives other than in the most local domain is with some adjuncts, e.g. with benefactives, as in (24), although most adjuncts behave as in (20):

(24) *John_i bought a book for himself_i.*

Barring these few border-line cases, the generalization for English is that reflexives are obligatory in the most local domain (arguments of a single predicate), and are disallowed beyond this domain (with the result that sentences like (21)–(23), neglecting the subscripts, are ambiguous – the ordinary personal pronouns could in principle refer to someone else other than John).

Russian is like English in presenting a two-way distinction between ordinary personal pronouns and forms of the reflexive pronoun *sebjā*. However, the cut-off point between the two is very different in Russian as compared to English, as can be seen in (25)–(29):

- (25) *Volodja uvidel sebjā v zerkale.*
 Volodya saw REFL in mirror
 ‘Volodya_i saw himself_i in the mirror.’
- (26) *Volodja uslyšal za soboj šagi.*
 Volodya heard behind REFL steps
 ‘Volodya_i heard steps behind him_i.’
- (27) *Volodja videl svoju knigu.*
 Volodya saw REFL book
 ‘Volodya_i saw his_i book.’
- (28) *Volodja velel Vere vskipjati’ sebe/emu čaj.*
 Volodya told Vera to:boil REFL/him tea
 ‘Volodya_i told Vera to make him_i some tea.’
- (29) *Volodja skazal, čto Vera ljubit ego/*sebjā.*
 Volodya said that Vera loves him/REFL
 ‘Volodya_i said that Vera loves him_i.’

These Russian examples closely parallel English examples (19)–(23). In Russian, however, reflexives are obligatory not only in the most local domain (arguments of the same predicate, as in (25)), but also for adjuncts (as in (26)) and for possessors (as in (27)). They are moreover optional across certain non-finite clause boundaries, as in (28), which allows both reflexive and ordinary pronouns with the same reference-tracking interpretation. The domain of reflexivation in Russian does not, however, extend across a finite clause boundary, so that only the ordinary pronoun is possible in (29).

Perhaps even more interesting data are provided by some languages that have not a two-way, but rather a three-way distinction among relevant pronominal forms, namely languages that distinguish ordinary personal pronouns, ordinary reflexive pronouns, and so-called emphatic personal pronouns. Compare the Danish data in (30)–(34):²

- (30) *Peter kritiserer sig selv.*
 ‘Peter_i criticizes himself_i.’
- (31) *Sofie lagde bøgerne bag sig.*
 ‘Sofie_i put the books behind her_i.’

² The Danish data are primarily taken or adapted from JAKUBOWICZ (1994) and VIKNER (1985), the latter not directly available to me.

- (32) *Peter tog sin frakke.*
 'Peter_i took his_i coat.'
- (33) *Marie hørte Eva kritisere sig/hende.*
 'Marie_i heard Eva criticize her_i.'
- (34) *Marie siger at Eva kritiserer hende/*sig.*
 'Marie_i says that Eva is criticizing her_i.'

The relevant pronominal forms in these examples are emphatic reflexive *sig selv*, ordinary reflexive *sig* (possessive: *sin*), ordinary pronouns *ham* 'him', *hans* 'his', *hende* 'her' (object), *hendes* 'her' (possessive). In Danish, the emphatic reflexive is required in the most local domain (arguments of the same predicate, as in (30)), and occurs only here. (Elsewhere, use of this form would only be possible if clear emphasis is intended, e.g. for contrast.) In the next set of domains, namely extending to adjuncts (as in (31)), to possessors (as in (32)), and optionally across certain non-finite clause boundaries (as in (33)), the ordinary reflexive is found, although across these non-finite clause boundaries the ordinary pronoun is also possible. Across a finite clause boundary, as in (34), only the ordinary pronoun is possible. The Danish pattern is thus that the most marked form, the emphatic reflexive, is used in the most local domain; the next most marked form, the ordinary reflexive, is used in intermediate domains; while the least marked form, the ordinary pronoun, is used for even more extended domains.

The Danish pattern can also be seen in Japanese, though with different cut-off points. In Japanese, the relevant forms are emphatic reflexive *zibun zisin*, ordinary reflexive *zibun*, and ordinary pronoun *kare* 'he', although zero anaphora is much more common than the ordinary pronoun and native-speaker judgments on the ordinary pronoun are often subject to considerable variation. The relevant data are set out in (35)–(38):³

- (35) *Taroo wa zibun (zisin) o butta.*
 Taro TOP REFL REFL ACC hit
 'Taro_i hit himself_i.'
- (36) *Taroo wa zibun ni issatu hon o katta.*
 Taro TOP REFL for one book ACC bought
 'Taro_i bought a book for himself_i.'
- (37) *Taroo wa zibun no heya de benkyoo-suru.*
 Taro TOP REFL GEN room in study
 'Taro_i studies in his_i room.'
- (38) *Hiroshi wa Mitiko ni zibun no heya de benkyoo-suru yoo-ni itta.*
 Hiroshi TOP Michiko to REFL GEN room in study to said
 'Hiroshi_i told Michiko_j to study in his_i/her_j room.'

In (35), illustrating the most local domain (arguments of the same predicate), either the emphatic reflexive or the ordinary reflexive is possible, though some speakers show a preference for the emphatic reflexive. Sentences (36)–(37) show that the ordinary reflexive is used for adjuncts and possessors. Japanese does not have a clear distinction between finite

³ The judgements given here are those of Akiko Kumahira Comrie; there is a fair amount of idiolectal variation on some points, all, as far as I am aware, consistent with the generalization in (17).

and non-finite complement clauses, and in any event the dependent clause in (38) is finite (with the finite verb form *-suru*), and even across this finite clause boundary Japanese at least allows the ordinary reflexive; note that (38) is thereby ambiguous, since the antecedent of the ordinary reflexive pronoun can be either the understood subject of its own clause ('for Michiko to study in self's room') or the subject of the main clause.

Incidentally, in languages that distinguish between an ordinary reflexive and an emphatic reflexive, it is not always the case that the distinction is marked in the pronominal forms themselves (as in Danish and Japanese). In Tamil, for instance, the emphatic reflexive involves adding an auxiliary verb *koL-* (past tense *koN-T-*), whose literal meaning is 'hold' and which requires the main verb to be in a converbal (gerundive) form, as can be seen by comparing (39) and (40):⁴

(39) *kumaar raajaa tann-aip parrip peec-in-aan en-ru ninai-tt-aan.*
 Kumar Raja REFL-ACC about talk-PST-3SM that think-PST-3SM
 'Kumar_i thought that Raja_j was talking about him_i/himself_j.'

(40) *kumaar raajaa tann-aip parrip peec-ik koN-T-aan en-ru ninai-tt-aan.*
 Kumar Raja REFL-ACC about talk-CVB hold-PST-3SM that think-PST-3SM
 'Kumar_i thought that Raja_j was talking about himself_j.'

In (39), the ordinary reflexive is used, and as in Japanese sentence (38) there is ambiguity as to whether the antecedent of the ordinary reflexive is the subject of its own clause or the subject of the main clause. In (40), the emphatic reflexive forces an interpretation with the antecedent in a more local domain, namely in the same clause. However, in all examples known to me the emphatic reflexive is formed by adding an extra morpheme (or morphemes) – be it pronominal or verbal – relative to the ordinary reflexive.

The data presented so far in this section can be summarized as follows: If a language distinguishes different degrees of markedness in the general range of reflexive and ordinary pronouns, such that the more marked forms indicate coreference, then the most marked form will be used in the most local domain(s), the least marked form in the most extended domain(s), and the intermediate form (if there is one) in intermediate domains. Or more generally: the greater markedness of a pronominal form correlates with greater locality of the domain. The precise cut-off point varies from language to language, but the distribution must be consistent with this correlation. Now the question arises: Why do we find this correlation?

A number of linguists working within a number of different frameworks, and going back well into traditional grammar, have formulated an explanatory hypothesis that covers these data. The essence of the hypothesis is that the normal situation, in terms of our conceptualization of the world, is for an action that includes more than one participant to have an agent acting on a patient that is a distinct entity from the agent. In other words, the most natural situation is for the arguments of a predicate to be non-coreferential. Where this expectation is not met, i.e. where the arguments are coreferential, then this is where a language is most likely to mark coreference. As we move to more and more extended domains, the expectation of non-coreference is relaxed, so that, as we have seen, at some particular point an individual language will decide to shift from reflexive to ordinary pronoun even in cases of coref-

⁴ A detailed empirical and theoretical discussion of reflexivization in South Asian languages is included in Lusr et al. (1999).

erence. The emphatic reflexive can be viewed as an even more distinctive means of indicating coreference in very local domains. This overall hypothesis is formulated in slightly more concise form in (41):

(41) Hypothesis on local domain

In the most local domain – the arguments of the same predicate – the expectation is that noun phrases will be non-coreferential. Special marking is most likely to be given when this expectation is violated, i.e. when noun phrases are coreferential in this domain they are most likely to be marked reflexive. The extent to which the use of reflexives may extend beyond the most local domain may vary cross-linguistically. Implicational universal: If a language uses reflexives anywhere, it will use them in the most local domain.

Before we leave markedness of pronominal forms and the local domain, it will be useful to consider some data that at first sight seem problematic, in that they seem to contradict a formal account of the distribution of more and less marked forms, but fit naturally into place once a more conceptual-functional approach is adopted. In Danish, for instance, some verbs require the ordinary reflexive even in the most local domain, as in (42):

(42) *Han vasker sig.*
'He washes (himself).'

In (42), the emphatic reflexive *sig selv* would receive an emphatic, contrastive interpretation. Why does (42) behave differently from (30)? The distinction seems to be that washing is something that one normally does to oneself, whereas criticizing is not something that one normally does to oneself. Note that the most natural English translation of (42) is not in fact the overtly reflexive 'he washes himself', but rather the intransitive 'he washes' (or 'he gets washed', although this last version does not exclude the possibility of someone else washing him); many other languages would use not a reflexive pronoun but an explicit shift in diathesis, such as a derived intransitive or middle voice. The overall hypothesis is that in general coreference of arguments of the same predicate is unexpected, and is therefore marked. But if this approach is correct, then it is natural that in conceptually exceptional cases, where the coreference is not unexpected, the form may shift to match this shift in expectations. And this is just what we find in the Danish example.

3. The extended domain

In section 2, I presented the hypothesis on the local domain by starting with data and working towards the hypothesis. In section 3 on the extended domain, I will adopt the inverse order of presentation, in part because the general line of argumentation is already clear, in part because this leads to an overall clearer presentation of the hypothesis and data.

The basic explanatory principle underlying my account of the extended domain is that in the extended domain, the expectation is for referential continuity, or, as it is often called in the literature, topic continuity. That is, it is expected that reference will be made back to a participant who has already been introduced into the discourse, moreover to one that is salient in the discourse, for instance by being topic (or its grammaticalized counterpart: subject). What is unexpected is the introduction of new participants, or of referring back to less

salient participants, and it is in these cases that more marked forms are used. We can therefore formulate a hypothesis on the extended domain as in (43), paralleling the hypothesis on the local domain in (41):

(43) Hypothesis on extended domain

The unmarked situation is to have 'referential continuity' ('topic continuity'). Special marking is used to indicate deviation from this pattern, i.e. either non-coreference or unexpected coreference (coreference with an unexpected antecedent).

Simple examples illustrating this generalization can be presented first of all from a number of European languages in which, in cases of referential continuity, a more abbreviated expression is possible by omitting all overt reference to the continuing participant. Russian examples (44)–(45) show that where there is subject continuity in a dependent clause after the verb 'to want', it is obligatory to omit the subject of the dependent clause, which must take on infinitival form, as in (44), where Yura is the understood subject of 'give':

- (44) *Jura xočet peredat' časy djade.*
 Yura wants to:give watch to:uncle
 'Jura wants to give the watch to Uncle.'

Where there is no such continuity, the subject of the dependent clause must be expressed, and that clause requires subjunctive form, as in (45):

- (45) *Jura xočet, čtoby ty peredal časy djade.*
 Yura wants that you give watch to:uncle
 'Yura wants you to give the watch to Uncle.'

Something similar happens with certain adverbial clauses, except that here it is only optional in cases of subject continuity to omit the subject of the dependent clause and use a non-finite, converbal (gerundial) verb form, as can be seen by comparing (46) with the synonymous (47):

- (46) *Kogda ja kupil bilet, ja sel v avtobus.*
 when I bought ticket I sat in bus
 'When I had bought a ticket, I got on the bus.'
- (47) *Kupiv bilet, on sel v avtobus.*
 having:bought ticket he sat in bus
 'Having bought a ticket, he got on the bus.'

When the two subjects are non-coreferential, as in (48), only a finite dependent clause is possible:

- (48) *Kogda ja kupil bilet, mama sela v avtobus.*
 when I bought ticket mother sat in bus
 'When I had bought a ticket, mother got on the bus.'

In yet further cases, omission of a referentially continuous subject serves as the most natural expression of this referential continuity. In (49), for instance, on the interpretation of the English translation where the subjects are coreferential, it is natural in Russian to drop the subject of the dependent clause; this is indicated in (49), and below in (50), by translating omitted Russian subjects by means of English subject pronouns in parentheses:

- (49) *Ona skazala, što odenetsja sama.*
 she said that will:dress self
 'She_i said that (she_i) would dress herself.'

As predicted by the hypothesis formulated in (43), this possibility can be pursued into even more extended domains, e.g. across sentence boundaries, as in (50), where the grammatically possible subject pronouns referring back to Ol'ga Ivanovna are systematically omitted:⁵

- (50) *Ol'ga Ivanovna ne ljubila dumat' o neprijatnom i počti nikogda ne dumala. Izbegala razgovorov o boleznjax, a kogda mužu ili dočeri slučalos' xvorat', govorila s nimi tak, točno oni vse vydumyvajut. Razumeetsja, pri ètom okružala ix samym zabotlivym uxodom.*
 'Olga Ivanovna did not like to think about unpleasantness and almost never did. (She) avoided conversations about illnesses, and when her husband or daughter happened to be poorly, (she) talked with them as if they were making everything up. Of course, at the same time (she) surrounded them with the most attentive care.'

Parallel phenomena, though often differing considerably in grammatical detail, can also be adduced from languages from other parts of the world. A widespread phenomenon in some parts of the world is switch-reference, whereby a dependent clause is marked, usually on its verb, to indicate whether its subject is the same as or different from the subject of the main clause, thus giving rise to a morphological distinction between same-subject and different-subject verb forms. In many languages with switch-reference systems, there is no markedness difference between same-subject and different-subject morphology, i.e. both have overt markers, but there is no way in which we can say that the one is more marked than the other. Some languages, however, do have such a markedness distinction, and in all such cases it is the same-subject morphology, indicating subject continuity, that is unmarked, as for instance in Siroi, a Papuan language of New Guinea. In Siroi (WELLS 1979), same-subject is indicated simply by using the dependent form of the verb of the dependent clause, with no overt indication of coreference and no other indication of person-number, although the only possible interpretation is with coreferential subjects, as in (51):

- (51) *Mbanduwaŋ ngur-mba buk-ng-ina.*
 bow break-DEP throw-TH-3SG:PST
 'He_i broke and (he_i) threw away the bow.'

Where the two subjects are non-coreferential, as in (52), we find two additional pieces of morphology relative to (51): First, the dependent verb, though still bearing the morpheme indicating that it is dependent, also bears indication of the person-number of its object. Secondly, there is an overt different-subject marker *le*.

- (52) *Agan ndende kusna-niŋ-mba min-na le teg puro-na.*
 thing various ask-3PLU-DEP be-3SG:PST DS fowl arrive-3SG:PST
 'She was asking them various things and the fowl arrived.'

A somewhat less obvious instance of the relevance of referential continuity to the choice between more and less marked pronominal forms can be seen in a number of European lan-

⁵ This example is taken from TIMBERLAKE (1993: 872); some of the earlier examples are taken or adapted from this source.

guages – and surely in other languages too, although the relevant European languages have been more thoroughly investigated from this viewpoint. We may start with the contrast in English between (53) and (54), where the pronouns in lower capitals indicate emphatic pronouns, here pronounced with greater emphasis:

(53) *John_i punched Bill_p, and then he_i kicked him_p.*

(54) *John_i punched Bill_p, and then HE_j kicked HIM_i.*

The difference in interpretation between (53) and (54) is quite striking. Without emphatic stress, in (53), the interpretation is that the grammatical relations and semantic roles of the first clause carry over into the second, i.e. that John also kicked Bill – a kind of continuity. The more marked emphatic forms in (54) indicate a break in this continuity, more specifically reversal of the roles, so that now it is Bill who is kicking John.

In some European languages – at least Dutch, German, and Russian – such a break in continuity can be indicated by using demonstrative pronouns in environments where the ordinary personal pronouns are also grammatically possible. I have elsewhere examined the Dutch data in more detail (COMRIE, in press), and will here consider only one brief example. In Dutch, the contrast is between ordinary personal pronouns such as *hij* ‘he’ and the more marked demonstratives such as *deze* ‘this’, *die* ‘that’; in the English translation, I will use small capitals to translate the Dutch demonstratives. The ordinary pronouns are used in cases of topic continuity, and in Comrie (forthcoming) it is argued that in Dutch it is specifically topicality, rather than say subjecthood, that is decisive, although the details are not relevant to our present concern. In cases of breach of topic continuity, the ordinary pronouns are not excluded, but there is a strong tendency in Dutch to prefer the demonstratives here; at least in written Dutch, the demonstratives cannot be used where there is no breach of topic continuity. Consider example (55):⁶

(55) *Eschenfelder had Erasmus gevraagd, hem een psalmverklaring te willen wijden (de vorm, die Erasmus in de laatste jaren meermalen koos). In het eind van 1535 herinnerde deze zich dat verzoek.*

‘Eschenfelder_i had asked Erasmus_j to dedicate a psalm to him (the form that Erasmus chose several times in his last years). At the end of 1535 HE_j remembered that request.’

The topic of this passage, as is clear from the broader context, is Eschenfelder. The main information in the first sentence is that Eschenfelder had asked Erasmus to dedicate a psalm to him (i.e. Eschenfelder). The second sentence then says that he remembered the request in late 1535. But wait! Who remembered the request? The use of *he* in English is ambiguous: It could be that Erasmus remembered Eschenfelder’s request, but it could also be that Eschenfelder remembered his own earlier request to Erasmus (and decided to remind Erasmus of that request). The use of the marked demonstrative (*deze* ‘this’ in the Dutch version is a clear signal to discount the interpretation with topic continuity, i.e. with Eschenfelder as antecedent, and rather to pick what is in discourse terms the less likely antecedent, namely the less topical noun phrase Erasmus. Dutch thus has a clear and unequivocal means of expressing the interpretation, and this means relies crucially on the correlation between less marked forms and referential continuity, more marked forms and lack of referential continuity.

⁶ This example is from J. HUIZINGA’s *Erasmus* (3rd edition, 1936).

4. Conclusion

The material and generalizations presented in this article can be interpreted on two levels. On the more specific level, the article provides an overarching account of a number of specific observations concerning the occurrence of more and less marked pronominal forms, the overarching hypothesis being that more marked forms tend to occur where the antecedent-pronoun relation is less expected, the less marked form where the antecedent-pronoun relation is more expected. However, on the more general level this is simply a particular instantiation of a more general methodological position, namely that an explanatory account of linguistic phenomena will often have to go beyond the purely formal statement of the generalizations and look at the cognitive and functional principles that underlie those generalizations.

Abbreviations

ACC	accusative
CVB	converb
DEP	dependent
DS	different subject
GEN	genitive
PLU	plural undergoer
PST	past
REFL	reflexive
SG	singular
SM	singular masculine
TH	thematic suffix
TOP	topic

References

- COMRIE, BERNARD (Forthcoming): Pragmatic binding: demonstratives as anaphors in Dutch, in: *Berkeley Linguistic Society Papers 23*.
- JAKUBOWICZ, CELIA (1994): *Sig en danois: syntaxe et acquisition*, in: OBENAUER, HANS-GEORG & ZRIBI-HERTZ, ANNE (eds.), *Structure de la phrase et théorie du liage*. Saint-Denis: Presses Universitaires de Vincennes.
- LUST, BARBARA, & GAIR, JAMES & SUBBARAO, K. V. & WALI, KASHI (eds.) (1999): *Lexical anaphors and pronouns in selected South Asian languages: a principled typology*. (Empirical Approaches to Language Typology 22). Berlin: Mouton de Gruyter.
- TIMBLAKE, ALAN (1993): Russian, in: COMRIE, BERNARD & CORBETT, GREVILLE G. (eds.), *The Slavic languages*. London: Routledge, 827–886.
- VIKNER, STEN (1985): Parameters of binder and binding category in Danish, in: *Working Papers in Scandinavian Syntax 23*: 1–58.
- WELLS, MARGARET A. (1979): *Siroi grammar*. (Pacific Linguistics B-51). Canberra: Department of Linguistics, Research School of Pacific Studies, Australian National University.

BERNARD COMRIE
 Max Planck Institute for Evolutionary Anthropology
 Inselstraße 22
 D-04103 Leipzig, Germany
 comrie@eva.mpg.de