On the foundations of the contrastive study of information structure
Carsten Breul

Abstract

The paper presents some ideas on the delineation of a more specific contrastive approach to information structure analysis within the larger and more loosely circumscribed comparative area. It will be argued that this delineation can be effected by emphasising the methodological role of the notion tertium comparationis. Ontological and methodological aspects of contrastive information structure analysis will be discussed, and benefits of taking a specifically contrastive approach to information structure will be pointed out. Finally, some lines of argumentation and observations from the preceding chapters of the present volume that can be construed as instances of contrastive information structure analysis will be briefly recapitulated.

1. Introduction

The articles in this volume are concerned with comparative approaches to issues of information structure. This chapter puts forward some ideas on the delineation of a more specific contrastive approach within the larger and more loosely circumscribed comparative area. I will be arguing that this delineation can be effected by emphasising the methodological role of the notion tertium comparationis, which has been prominent and the focus of much discussion in the history of contrastive linguistics and contrastive analysis (see e.g. Krzeszowski 1990: ch. 1, 2, Chesterman 1998: ch. 1 and the references given there). Note that I will be using the terms contrastive linguistics and contrastive analysis not as referring to a certain theoretical framework alongside those such as structuralist linguistics, (systemic-)functional linguistics, generative linguistics, and, respectively, to a style of analysis typical of these frameworks. Rather, I will be using them to refer to kinds of linguistic research where the concern with the question of the tertium comparationis in the comparative endeavour is obvious, either implicitly or explicitly. That is, the notion 'contrastive linguistics' is orthogonal to notions such as 'structuralist / functional / generative linguistics'.

In the remainder of this introductory section I will explain what the more general benefits are of taking a specifically contrastive approach to issues of information structure. The topic of the main section 2 is a discussion of ontological and methodological aspects of contrastive information structure analysis. Section 2 also contains some relativising remarks on the ideal methodological requirements on contrastive information structure analysis as discussed before. The point here is that keeping this ideal in mind and explaining in how far a prospective instance of contrastive information structure analysis matches it is more fruitful than not undertaking the analysis in the first place just because it may be impossible to heed the ideal completely. Section 3 presents
some lines of argumentation or observations from the preceding chapters of this book which I think are, or can be interpreted as, contrastive information structure analyses or observations.

'Contrastive information structure analysis' is not an established label in linguistics. There is research into information structure and there is contrastive analysis, and quite a few linguists have indeed investigated issues of information structure from a more or less explicitly contrastive perspective. Consequently, there has accumulated a considerable body of research whose results can be interpreted as results of contrastive information structure analysis. What seems to be missing so far for the identification of a more clearly demarcated research paradigm 'contrastive information structure analysis' is some ontological and methodological groundwork that provides us with the conceptual means, or criteria, for saying when or in which respects a certain work, or a part of it, represents an instance of contrastive information structure analysis. Some suggestions in this direction will be made below.

What do we need such criteria for, if extensive cross-linguistic, or comparative, research into information structural aspects has proceeded against the background of existing, well-established frameworks such as functional linguistics, generative grammar, or language typology? And why should we want to ponder in which respects and to what degree (part of) this research can also be said to instantiate work within a new research paradigm characterised by these criteria? The answer, I believe, lies in the specific benefits of contrastive linguistics as the overarching type of approach. Whatever the reputation of contrastive linguistics may have been after its popularity had drastically ebbed down in certain quarters not long after its inauguration in the late 1950s, there can be no doubt that investigating commonalities and especially differences between two languages on a sound methodological basis is interesting and important in several respects: for language teaching and learning (see e.g. James (ed.) 1996, Kortmann 1998, Sheen 1996), for translation practice and translation studies (see e.g. Albrecht 2005, esp. ch. 4, 5, Chesterman 1998, e.g. p. 198f.), for research into language typology (see e.g. Comrie 1986, Hawkins 1986, König 1990, 1992, 1996) and into language universals (including universal grammar).

That is, given that information structure can be conceived of as an object and a domain of linguistic study similar to objects / domains such as phonology, morphology and syntax, contrastive information structure analysis is worth having for the same reasons as it is worth having contrastive phonological, morphological and syntactic analysis. Contrastive analyses of information structure (a) focus on the linguistic phenomena and supply data that have to be integrated in theoretical frameworks that aim at cross-linguistic generalisations over the phenomena (notably generative grammar and language typology) and (b) lend themselves to application, especially in language teaching and learning and in translation, at least insofar as an awareness of these phenomena on the part of language teachers and learners or translators is considered to be beneficial.

It will have become apparent that, by suggesting some ontological and methodological ideas for a more explicitly framed new sub-domain of contrastive analysis, I do not intend to elevate the theoretical status of contrastive linguistics
within linguistics in general. Elevating the status of contrastive linguistics seems to be one of the objectives of Pan & Tham (2007). Two passages illustrating this may be given:

On the whole, if James has driven the discipline to develop both vertically (into text) and horizontally (into pragmatics and culture), taking the discipline in the height and breadth dimension, the inward movement of Chesterman (into language users' minds) has added the depth. With contrastive rhetoric pushing at the side and developments in Australia pushing from the back, contrastive linguistics is ready to grow in full dimensions as a magnificent discipline with high macro-research vision. This is historical momentum. (Pan & Tham 2007: 66f.)

Our account of the Western contrastive studies shows that many do not acknowledge their studies to be a branch of linguistics and choose to use the term 'contrastive analysis' to emphasize its applied significance in second language pedagogy. [...] the insistence of its belonging to applied linguistics will forever hinder the advancement of the discipline, particularly in the West and especially in the US where theoretical linguistics is so divided and challenging. It follows quite naturally then that contrastivists must feel inferior and fall outside of the horizon of theoretical linguists. To reverse the situation, a change in paradigm is required. (Ib.: 167)

Quite to the contrary, I am perfectly content with considering contrastive linguistics subservient to other paradigms with more fundamental or more general objectives (e.g. generative grammar, language typology) or with aims in the domain of application (e.g. language pedagogy, translation). The proposals to be made, however, may bring contrastive linguistics into a position to perform this task more effectively where aspects of information structure are concerned.

2. **Ontological and methodological aspects**

2.1 *The distinctive features of contrastive analysis in general*

The question of the nature of contrastive linguistics in general and of the way in which contrastive analysis proceeds or should proceed has been discussed extensively. The bottom line still is that contrastive linguistics is concerned with the comparison of very few, typically only two, languages with the aim of detecting commonalities and especially differences between them. As there are other research paradigms in which the comparison of languages plays a central role, notably language typology and generative grammar, it is necessary to comment on the distinctive features that set contrastive linguistics apart from these other comparative paradigms. These distinctive features are (a) the importance accorded to the question of the *tertium comparationis* of a given contrastive analysis; (b) the methodological constraint that a contrastive analysis has to involve what may be called the meaning/function side of language on the
one hand and the form side on the other, with the *tertium comparationis* being established on one of these sides and the comparison being carried out on the other (cf. Chesterman 1998: 52f.).

Of course, there is a *tertium comparationis* in cross-linguistic studies within the various frameworks of generative grammar too. In the principles and parameters framework, for instance, this is the set of syntactic features and the set of syntactic principles that are taken to be universal and taken for granted at the outset of the study, mostly implicitly so by stating, or allowing the reader to infer, in which more specific grammar-theoretical framework the study is situated. The point of the cross-linguistic study, then, is to provide evidence which corroborates the assumptions about the (universality of the) sets of features and/or principles, or which enforces a modification of these assumptions, or which demonstrates how surface differences between languages can be accounted for by different parameter settings. Kayne's (1996/2000) point already quoted in note 5 is reiterated by the following passages from the chapter "The new comparative syntax" in Haegeman & Guéron (1999):

Comparative studies of languages can help us to answer the question of what is language-specific and what is universal: in other words, what is variable from one language to the next and what is invariant across languages. (Ib.: 587)

Parametric variation is itself predetermined; the values which we can assign to a parameter are selections made within a restricted class of possibilities. The purpose of comparative research is to identify the parameters which have to be set by the learner and what kinds of settings there are. (Ib.: 596)

The blind spot in this type of approach is the question which lexico-grammatical forms from the different languages in focus one is to select for inclusion in the study of a given principle or parameter in the first place. (My two examples involving (1)-(4) below will illustrate this problem.) This question is usually not raised, and it is usually answered implicitly by taking structures filled with lexical items into account that are considered to be more or less cross-linguistically the same in terms of meaning. But this methodological step is not motivated by the theoretical framework. It may be thought that it is motivated by the additional assumption that the universal syntactic features and principles are associated with semantic content from the different languages in such a way that, say, the same features and the same syntactic operations conditioned by the same syntactic principles will result in the same meaning. But this, in fact, is rather an implication of the methodological decision to compare just those structures that seem to have (roughly) the same meaning; it is not an assumption which could, in principle, be falsified and which could be taken to motivate this methodological decision as long as it is not falsified. Now, what would it mean if it should turn out that those structures compared are not actually the same in meaning? Or what does it mean to say that those structures are only roughly the same in meaning? I will mention two cases that raise these questions in a more concrete and illustrative way.  


What, for instance, does the availability of KNOW and the unavailability of KENNEN in English and German passives mean for generative accounts of passives?

(1)  
   a. Everybody knew him.
   a'. He was known by everybody.
   b. Jeder kannte ihn.
   b'. *Er wurde von jedem gekannt. (Cf. the copulative-adjectival structure Er war jedem bekannt, which may serve as translation equivalent of (1a') in some contexts; but note also that Er war sehr bekannt does not have *He was very known as a potential translation equivalent.)

Does it mean that the syntactic principles involved in passive formation are essentially different in English and German? Does it mean that the syntactic principles are the same, but in different ways sensitive to syntactically relevant lexical characteristics of KNOW and KENNEN, due to different parameter settings? Or does it mean that the syntactic principles as well as potentially involved parameter settings are the same, while syntactically relevant lexical features of KNOW and KENNEN are different in just those respects that are relevant for passive formation? For a convincing analysis set in a universalist (notably generative) framework to be possible, questions such as these have to be raised and plausibly answered. In addition to triggering such questions, a contrastive syntactic approach provides the data on the basis of which answers to such questions are to be assessed.

To give another example: Imagine a generative syntactician investigating the ordering options among phrases that premodify the nominal head in German noun phrases. Confronted with data such as those in (2) below, where (2a) is preferred, but (2b) accepted as well, she may start by hypothesising that the order of adjectival phrase and numerical phrase in pre-head position is reversible in German.

(2)  
   a. (Wir verbrachten dort) drei schöne Wochen.
   b. (Wir verbrachten dort) schöne drei Wochen.

Bringing English into play, the syntactician may then observe that this language only allows the order numerical phrase > adjectival phrase.

(3)  
   a. (We spent) three pleasant weeks (there).
   b. (We spent) *pleasant three weeks (there).

At this point the syntactician may conclude her analysis by proposing a parametrical difference between German and English in this respect. However, she may also continue the investigation by taking a more contrastively oriented perspective, observing that nominal phrases with the German pre-head order adjectival phrase > numerical phrase actually seem to have English nominal phrases as translational and semanto-syntactic equivalents (see Krzeszowski 1990) that are slightly, but significantly, different in structure, as suggested by the examples in (4).
And while pollution incidents for 1991-92 reached an all-time high of 29,524 [...], there were a meagre 536 prosecutions. (British National Corpus (BNC), document CH6)

a'. … magere 536 Anklagen (not: 536 magere Anklagen '536 meagre prosecutions')

b. Terry, a biscuit kiln fireman, shaved a remarkable 12 minutes off his best ever time to finish in three hours five minutes. (BNC HBE)

b'. … bemerkenswerte 12 Minuten (not: 12 bemerkenswerte Minuten '12 remarkable minutes')

c. Situated in a former ex-Great Western Railway coach, No. 1160 liveried in chocolate and cream at platform one, the exhibition boasts a staggering 200 visitors per day during the operating season of the SVR. (BNC CKK)

c'. … erstaunliche 200 Besucher (not: 200 erstaunliche Besucher '200 staggering visitors')

These data suggest that the pre-head adjectival phrase > numerical phrase order in German nominal phrases corresponds rather to the somewhat peculiar English nominal phrases displayed in (4), also in cases such as those in (2b), corresponding to We spent a pleasant three weeks there. (These nominal phrases are peculiar, of course, in that they apparently feature an indefinite article in construction with a plural noun, which 'usually' leads to ungrammaticality; cf. *a 536 prosecutions; *a 12 minutes etc.) This contrastive observation, then, may lead the hypothetical generative syntactician to a significantly different analysis than the one mentioned above. In this alternative analysis based on a contrastive observation the German pre-head adjectival phrase > numerical phrase order is structurally brought in line with the syntactic structure of the respective nominal phrases in (4) – whatever their structure may be – rather than that in (3a). An analysis along these lines would certainly be more explanatorily adequate than simply postulating a parametrical difference as in the analytical approach concentrating only on data such as (2) and (3). The second line of analysis is contrastive in that it takes relevant translation equivalents as tertia comparationis into account (schöne drei Wochen / a pleasant three weeks) that may have significant implications for the syntactic analysis.9

Of course there are tertia comparationis in typological studies too. These are certain grammatical categories or concepts, such as 'word', 'word order', 'voice', 'tense', which are mostly taken from or rooted in traditional grammar of the western, Greek and Latin based, school. Typological studies aim at classifying languages according to the manifestation of such categories, potentially with the further aim of uncovering universals. (The notion 'classifying' in this context comprises the assignation of languages to certain positions on continuous classificatory scales, or to certain positions in spheres around a prototype, in those cases where the researcher rejects the existence of discrete classes.) Here as well, researchers often take a more or less implicit recourse to equivalence in meaning or function when they explore which classes to set up in the first place. This may
put the typological endeavour to a certain risk. On the one hand, a thorough analysis of the ways in which language B expresses what language A expresses by operating with or on a form that has been analysed as the manifestation of a certain grammatical category or concept makes this analysis look very much like a contrastive analysis. This may mean nothing more, but also nothing less, than that typology is dependent on contrastive analysis. On the other hand, the preliminary contrastive analysis may reveal that the putative grammatical category or concept that the typologist sets out to explore is so heterogeneous in nature cross-linguistically that it is questionable whether it can survive as a typological tertium comparationis. This is to say that contrastive analysis has a complementing function vis-à-vis typology as well, a point that has been emphasised in a series of articles by E. König (1990, 1992, 1993, 1996). Note also that contrastive analysis cannot be subsumed under typology, at least not as currently practised in the typological mainstream, as contrastive analysis does not necessarily share with typology the assumption that any of the grammatical categories or concepts used as typological tertia comparationis have a well-defined, or properly definable, identity in the first place.

2.2 The distinctive features applied to contrastive information structure analysis

What do the distinctive features (a) and (b) of contrastive linguistics mentioned at the beginning of section 2.1 mean in more concrete terms for contrastive information structure analysis?

The meaning/function side is manifested by the categories of information structure as identified by a given theory of information structure. By this I mean categories such as theme and rheme, given and new information, topic and comment, focus and background – as long as they have been characterised or defined independently of formal criteria (cf. further below) – or the three categories identifiability, activation and focus structure as thoroughly discussed especially by Lambrecht (1994). As my own conception of information structure is coined by Lambrecht's, I may quote his definition:11

INFORMATION STRUCTURE: That component of sentence grammar in which propositions as conceptual representations of states of affairs are paired with lexicogrammatical structures in accordance with the mental states of interlocutors who use and interpret these structures as units of information in given discourse contexts. (Lambrecht 1994: 5)

This definition allows one to see why, for Lambrecht (and the present author as well), identifiability and activation also figure under the comprehensive label 'information structure', alongside the more prominently discussed phenomena that he subsumes under the label 'focus structure'. To give a simple example: The lexicogrammatical alternatives displayed by (5a, b) below may, in an appropriate context, be due to nothing but a difference as to whether or not the producer of (5a) assumes that the policeman is identifiable by her addressee, and as to whether or not the producer of (5b) assumes that the referent of John / he is active in the mind of the addressee at the given point in the discourse.
(5) a. There is {a / the} policeman standing at the corner.
   (acceptable with the if there is only one policeman
   available in the discourse world)
b. {John / He} is ill.

However, as is well known, theories of information structure have not yet
developed to a stage where there is broad terminological and conceptual
consensus and where the open theoretical issues or questions could be formulated
in a manner that would make sense to every researcher. This is only to be
expected in a fairly young domain of investigation, and contrastive information
structure analysis may be one of the factors that help information structure theory
to reach the next stage in its development.

The form side is manifested by the forms, categories, operations, principles,
constructions identified on the levels of phonology, including intonational
phonology, morphology, syntax and the lexicon that have been associated with
categories of information structure by researchers with diverse theoretical
backgrounds. Here as well there is no consensus between representatives of
different theoretical schools about even the most elementary aspects of the nature
of such areas as phonology, morphology, syntax, or the lexicon. However, on a
more shallow, or 'surface', level, certain phonological, morphological, syntactic
and lexical phenomena seem to be unanimously associated with information
structural categories in at least some languages. I am thinking of certain
intonational phenomena (such as falling pitch accents or tunes), bound
morphemes (such as Japanese -wa and Korean -nun), constituent order
phenomena (such as phrase movement to clause initial position), constructions
(such as cleft-constructions) and lexical items (such as pronouns, definite articles,
focus and topic particles) about whose information structural relevance in some
way in certain languages there is no disagreement – even if disagreement may
begin as soon as one delves a little deeper into questions of the description and
explanation of these phenomena as such and of their relation to information
structural categories.

A variety of theoretical backgrounds is no problem in principle for
contrastive information structure analysis. What seems necessary, though, is that
the theoretical background chosen allow the researcher, first, to identify and
demarcate their tertium comparationis and, second, to motivate the assumption
that it applies to both languages under investigation. As to the second point, we
can recall the dilemma of those in the early days of contrastive linguistics who
adhered to the structuralist (of a certain brand) conviction that each language is
unique and has to be described in its own terms and who were interested in doing
contrastive analysis at the same time (see James 1980: 166f.). That is, in order to
avoid such a dilemma, it is necessary that one's theoretical background not
exclude the possibility that the tertium comparationis does in fact apply to both
languages involved. It would make the task of the contrastive analyst even easier
if their theoretical background not only allows for, but suggests that the respective
tertium comparationis is cross-linguistically, perhaps even universally, applicable.
It has been pointed out more than once that cross-linguistic or universal
applicability of a tertium comparationis is by no means guaranteed by the mere
fact that the same grammatical term has traditionally been used for the description of a linguistic phenomenon in language A and of a linguistic phenomenon in language B.¹²

2.2.1 Tertium comparationis on the meaning/function side
As already pointed out above, the tertium comparationis may, in principle, be stated on the meaning/function side or on the form side. Stating a tertium comparationis on the meaning/function side presupposes that the respective meaning/function is in a certain sense 'equivalently present' in both languages to be compared. Strictly speaking, in order to prevent a vicious methodological circularity, a theory of meaning/functions is required that is independent of categories that serve the linguistic description of an individual language. This requirement is not fulfilled by the Hallidayan variety of the notions theme and rheme, for instance, as it conflates both formal and meaning/function aspects in the definitional characterisation of these notions. Halliday (1985/1994: 37) writes:

In other languages, of which English is one, the theme is indicated by position in the clause. In speaking or writing English we signal that an item has thematic status by putting it first. […] The Theme is the element that serves as the point of departure of the message; it is that with which the clause is concerned. The remainder of the message, the part in which the message is developed, is called in Prague school terminology the Rheme.

This conflation – form: first vs. second position; meaning/function: expressing "that with which the clause is concerned" vs. expressing that by which "the message is developed" – renders these notions unsuitable for serving as tertium comparationis in contrastive information structure analysis. They already presuppose what may actually be a result of a contrastive information structure analysis, namely that there may be a systematic relation between the linear position of a formal unit in a linguistic expression on the one hand and its information structural value (meaning/function) on the other hand. Actually, the fact that this presupposition does not hold in a language such as English, as in cases like (6B) (from Lambrecht 1994: 223), where small capitals signal the word that carries the primary sentence accent, is then accounted for in terms of markedness in the Hallidayan framework.

(6) A: I heard your motorcycle broke down.
B: My car broke down.

Here my car cannot be said to be "that with which the clause is concerned" (see quotation from Halliday above) and where broke down cannot be said to be "the part in which the message is developed" (see ib.) – rather the reversed characterisation is correct. Such cases are conceived of as being 'marked' in the Hallidayan framework (see e.g. Halliday 1985/1994: 59). Markedness and unmarkedness in fact reintroduce the separation between the formal and the meaning/function sides, as they provide for the situation that the initial element in a clause is not "that with which the clause is concerned" (the marked case for
themes) and the rest apart from the initial element is not "the part in which the message is developed" (correspondingly the marked case for rhemes). This allows one to state, for instance, that while (6B) is marked, its Italian counterpart in (7) is unmarked in the Hallidayan sense.

(7) Si è rottà la mia MACCHINA.

However, this simply translates back as saying that in the English case the initial clausal element (in the Hallidayan sense), my car, is not "that with which the clause is concerned" whereas in the Italian case it (si è rottà) is. The argumentation has come back to square one in a full circle. We could have spared ourselves this circle and started right away with a syntactic analysis of the interesting observation that, whatever information structural category is equivalently expressed by my car on the one hand and la mia macchina on the other hand – the category focus for Lambrecht (1994) – it seems to correlate with a syntactic difference manifested by the different positions of these phrases in surface structure. And, as Lambrecht (1994) has shown, this leads to the further interesting observation that the same lexical material arranged in the same predicate-argument configuration as in sentences (6B) and (7) requires a different syntactic configuration in Italian while disallowing a different (surface) syntactic configuration in English if the information structural value of my car / la mia macchina is changed:

(8) What happened to your car?
   a. My car broke DOWN.
   b. La mia macchina si è ROTTA.

Lambrecht's (1994) characterisation of the information structural category identifiability, for instance, which is based on Chafe (1976), fulfils the requirements of equivalence and language independence in that it is embedded in a theory of cognition. Lambrecht (1994: 77f.) writes:

I will postulate the cognitive category of identifiability, using a term once suggested by Chafe (1976). [...] an identifiable referent is one for which a shared representation already exists in the speaker's and the hearer's mind at the time of the utterance, while an unidentifiable referent is one for which a representation exists only in the speaker's mind.

Identifiability being characterised as a cognitive category implies its universality and hence its applicability to the languages involved in a contrastive information structure analysis. The notion 'representation' entails no commitment to the effect that the relevant cognitive processes operate with structures that have a linguistic format; and that these processes operate with structures that have a format corresponding to the linguistic structures that are specific to a certain language can be safely ruled out. Two comments with respect to the identifiability example are in order:

First, we are confronted with only a rudimentary cognitive theory of identifiability here, of course. But the question of how elaborated the theory has to be within which the tertium comparationis is to be stated is of minor importance for the methodological point aimed at here. A theory as rudimentary as the one
about identifiability expressed by the quotation from Lambrecht (1994) just given
is methodologically sufficient for contrastive analyses since it can be used to ask
the question which formal means there are in languages A and B that have the
function of expressing the respective values (± 'referent is identifiable') of the
identifiability parameter so characterised. An answer to the question would
constitute the preparatory stage of a specific instance of contrastive information
structure analysis. The contrastive analysis proper would consist of the cross-
linguistic comparison of the formal means identified within a descriptive
grammatical framework that allows for the description of the grammar of both
languages involved. Of course, what kinds of descriptive grammatical frameworks
that may be is a difficult question in its own right. But whatever choice one makes
at this point will probably have interesting and insightful implications not only
with respect to the languages involved but also with respect to the potentials and
limitations of the grammatical framework itself.

Second, with respect to tertia comparationis for contrastive information
structure analyses that are stated on the meaning/function side, such as
identifiability, there will often arise a problem of operability. That is, it will often
be difficult to find the criteria and means for identifying expressions or utterances
from the languages to be contrasted that express identifiability (and other such
categories for that matter). Even if there are such criteria and means developed
within the domain of science in terms of which identifiability is defined, cognitive
psychology, their application by a contrastive linguist will, for various reasons,
seldom be possible. Sometimes a – at least preliminary – solution to this problem
may present itself in the form of theoretical considerations within a given theory
of information structure in connection with linguistic observations. For example,
it is argued in Breul (2008b) that the DP in as for DP expressions of the kind in
(9), exemplified for English in (10), are identifiable by virtue of the assumption
that they are necessarily active in the sense of Lambrecht (1994) in connection
with the fact that the activeness of a referent entails its being identifiable in
Lambrecht's theory.

(9)  a. as for DP
    b. quant à DP
        (French equivalent of a.)
    c. was DP {angeht / (an)betrifft} (German equivalents of b.)

(10) My encounters with girls were destined always to end in rejection
      until I'd left my teens behind me. I caught up a bit during the '60s
      when I became the oldest teenager in town – in fact I was in my
      early thirties. As for love, I fell easily and often. (CH8 1762)

The comparison of attested English DPs in as for DP expressions with their
translationally equivalent German counterparts in corresponding was DP angeht
expressions yields interesting observations concerning contrasts in definite article
usage in English as opposed to German and raises interesting syntactic, semantic,
and diachronic questions (see Breul 2008b).

Lambrecht's (1994) information structure category 'activation' (see ib.: 93ff.),
derived from Chafe (1974, 1976, 1987), is equally characterised in terms of
cognition and thus independently of language. Lambrecht's (1994) information
structure category 'topic relation' is defined as "the relation of aboutness between a proposition and a discourse entity"; and his 'focus relation' is characterised as one where "the relation between a focus denotatum and a proposition is taken to be non-recoverable and unpredictable at the time of an utterance" (ib.: 335f.). Although the notions 'proposition' and 'denotation' may be conceived as being dependent on the human language faculty in general, they are certainly not tied to specific languages. This makes Lambrecht's information structure categories suitable as tertia comparationis on the meaning/function side. Let us now turn to the question of tertia comparationis on the form side.

2.2.2 Tertium comparationis on the form side

The situation presents itself similarly here: Strictly speaking, stating a tertium comparationis on the form side presupposes that the respective form or structure is in a certain sense 'equivalently present' in both the languages to be compared. It seems that those linguistic theories which are essentially concerned with showing how forms and structures of individual languages can be derived from universal linguistic features and principles can be made use of in this connection. If, for instance, we assume that a certain syntactic operation or construction is syntactically conditioned in the same way in languages A and B in terms of universal syntactic features and principles, we may say that this operation or construction is syntactically equivalent in both languages. We may then set out to investigate what role this syntactic operation or construction plays in the expression of information structural categories in language A and in language B and we may note commonalities and differences.

This approach can be easily exemplified by making use of one point made by Lambrecht in the present volume. We may assume that the English and French cleft constructions in (11) below (cf. Lambrecht this volume) are syntactically equivalent (at least in the relevant aspects14) in that they can be derived from universal features and principles – an assumption that forms our tertium comparationis.

(11)  

   a. It's Isabelle that gave it to me five years ago.  
   b. C'est Isabelle qui me l'a donnée il y a cinq ans.

Lambrecht's observation is that the French cleft sentence provided with a suitable intonation constitutes a perfectly acceptable reply to a context utterance like (the French counterpart of) I like your shirt, do you remember where you bought it? Whereas the English cleft sentence would be inappropriate in this context with any intonation. This, then, is a contrastive information structural observation that involves the information structural category focus structure on the meaning/function side. The context utterance puts certain constraints on the focus structure of the reply; the English cleft sentence does not comply with these constraints, while the French one does.

There is the problem that, most generative approaches apart, some grammatical theories have built into their very structure the assumption that the formal side, syntax, for instance, cannot be conceived of independently from the meaning/function side, information structural categories, for instance. In a strictly
functionalist grammatical framework, an item or pattern on the form side is only identified if it is associated with a certain meaning or function, and the formal characteristics of this item or pattern are assumed to be determined by this meaning or function. According to Nichols (1984: 97), "[f]unctionalists maintain that the communicative situation motivates, constrains, explains, or otherwise determines grammatical structure, and that a structural or formal approach is not merely limited to an artificially restricted data base, but is inadequate even as a structural account." For a contrastive information structure analysis that is based on such a grammatical theory it does not make sense to state its tertium comparationis on the form side. This is because a meaning or function must have been identified before an item or pattern on the form side can be said to have been identified (for otherwise there would be non-functional criteria for identifying formal entities; the existence, though, of such non-functional criteria is denied). Consequently, in such a framework it does not make sense to speak of a formal entity that is equivalent in two languages and whose potentially different characteristics on the meaning/function side are to be investigated. Thus, a contrastive information structure analysis which is based on a linguistic theory which denies the independence of the form from the meaning/function side is only possible if the tertium comparationis is stated on the meaning/function side. That is, in such a framework contrastive information structure analysis is restricted to the basic question what the formal means are of expressing a given information structural category in language A as opposed to language B.

2.3 Relativising the methodological ideal

In the preceding discussion of the role of tertium comparationis and equivalence in contrastive information structure analysis the hedge expression 'strictly speaking' has been used twice. I said: (a) "Strictly speaking, in order to prevent a vicious methodological circularity, a theory of meaning/functions is required that is independent of categories that serve the linguistic description of an individual language." (b) "Strictly speaking, stating a tertium comparationis on the form side presupposes that the respective form or structure is in a certain sense 'equivalently present' in both the languages to be compared." What is intended by this hedge is this: Given that the methodological requirements just sketched are considered reasonable in principle – which cannot be expected to be the case for everybody – they represent an ideal achieving which may be very hard or even impossible at the present stage for the majority of potential research questions. I do think, however, that it is an ideal which is worth keeping in mind if progress is to be made on the way towards the goal of contrastive linguistics mentioned earlier, namely to assist other research paradigms in their more general or fundamental objectives or in their applied aims. That is, reflecting on the tertium comparationis and equivalence problems in each contrastive information structure analysis adds to its quality. As far as other aspects of the contrastive methodology as well as the stages of a contrastive analysis are concerned, I agree with Chesterman's (1998: ch. 1) proposals. His "contrastive functional analysis" (CFA) methodology
derives directly from Popper's (e.g. 1972) philosophy of science. Popper argues that all growth of objective knowledge proceeds as a form of problem-solving, in which hypotheses (tentative theories) are suggested, tested and refuted, giving rise to revised hypotheses which are in turn tested and revised, and so on in an endless process of conjecture and refutation. On this view, the goal of cross-linguistic comparisons is to propose and test falsifiable hypotheses, and their theoretical value lies precisely in this function (see also Janicki 1990). (Chesterman 1998: 53)

It is true, my insistence on the heuristic value of reflection about the tertium comparationis and the question of equivalence in contrastive information structure analysis (and other branches of contrastive analysis for that matter), does not appear to accord well with other passages in Chesterman's (1998) discussion of his CFA methodology. For example, he writes

The starting point for a given CFA-type analysis is a perception, made by a linguist, a translator, a language learner. This is a perception of a similarity of some kind […]. It is this perception, not some assumed equivalence, that provides the initial comparability criterion. For the language learner, this initial perception is the potential trigger for interference. For the contrastivist, it is the reason why X and Y are worth comparing. It is significant that this initial perception is often vague, unspecified: one task of contrastive research is to clarify and specify such perceptions. (Ib.: 55f.)

This CFA methodology therefore differs from the traditional one in its interpretation of the tertium comparationis. Traditionally, this has been taken as the starting-point of a comparison; however, as suggested above, this view risks circularity, in that some kind of equivalence is both assumed at the start and arrived at in the conclusion. In the methodology proposed here, the starting-point for a comparison is not an equivalence but a perceived similarity: the starting-point is this perception. The perception is then refined and operationally defined as a similarity constraint, specifying the acceptable range of similarity. The relation of identity (equivalence) occurs first in the initial hypothesis to be tested, and perhaps also as part of the result of such testing. (Ib.: 59)

However, there does not seem to be too deep a rift between Chesterman's view and the one sketched in section 2 above. Of course, the trigger for the detection of an interesting topic for a contrastive information structure analysis is a perception of similarity or dissimilarity as described by Chesterman. I consider the reflection on the tertium comparationis and equivalence issues to be the crucial part of the stage where this perception is "refined and operationally defined as a similarity constraint" (Chesterman quotation above).15

There is another respect in which the ideal that requires a clear demarcation of the tertium comparationis on the meaning/function side may have to be relativised. The point is that a category of information structure is never manifested independently of other aspects of meaning. For example,
identifiability and activation have to do with denotations or referents of expressions; focus structure, or the dimensions of topic-comment, of focus-background, and of theme-rheme have to do with the way in which propositional information is expressed. 'Denotation', 'referent' and 'proposition' essentially being semantic notions, there is no way of cutting off these meaning aspects from the respective category of information structure. Consequently, any consideration of the ways in which such a category is formally expressed in two languages can only proceed against the background assumption that these necessarily involved meaning aspects do not interfere with the contrastive analysis. And this presupposes a reliance on being able to keep these meaning aspects equivalent, with nothing more at our disposal than the usual criteria of translational and semanto-syntactic equivalence (or similarity)\(^\text{16}\) as they have been discussed and shown to be problematic in the contrastive linguistics literature (see e.g. Chesterman 1998: ch. 1, Krzeszowski 1990). My point, however, is not to make a contribution to the discussion of whether or not there is equivalence on the meaning/function side between languages. The point is rather that being forced to consider this question in connection with setting up the tertium comparationis of a contrastive analysis is precisely the heuristic specificity of contrastive linguistics by which it differs from other comparative approaches.

### 3. Examples of contrastive information structure analyses in the present volume

This section illustrates how the ontological and methodological points about contrastive information structure analysis presented above are at work in some of the contributions to the present volume. In addition to the reference to Lambrecht's work (this volume) in section 2.2.2 above, I will briefly refer to four more examples.

The main point of Cohen's contribution with respect to contrastive information structure analysis – 'main' in my construal for the purposes in the present paper, not necessarily in Cohen's – is based on a tertium comparationis on the meaning/function side. It is actually the building up of the tertium comparationis which provides the bulk of Cohen's article. Against the gist of several previous studies, the author presents a unified, monosemous, account of the function of intensive reflexives (IRs) in English. She explains how her account relates to information structure, an important aspect of the explanation being as follows:

[T]he scope of the PNself [i.e. post-nominal IR] marks the referent as an anchor to which the newer information should be linked, the entry under which new information is inserted. The VPself [i.e. post-verbal IR] marks the predicate similarly, thereby marking the set based on it as the anchor entry. The PAUXself [i.e. post-auxiliary IR] takes scope over the informationally poor auxiliary. In this case, the IR signals that both the predicate and the referent are discourse-old and already activated, thereby marking them as anchor entries, while highlighting the connection
between them as the new information in the discourse. (Cohen this volume: sect. 4.2)

Cohen argues that her theory of the function of IRs in English applies equally well to Hebrew IRs. This step in the line of argumentation may be considered a contrastive observation in its own right, based on a tertium comparationis on the form side, i.e. on the identification of the lexico-grammatical category IR in both languages, and revealing a cross-linguistic commonality rather than a difference. However, my construal of Cohen's main contrastive observation affects precisely formal aspects of IRs:

While IR scope effects are evident in both languages, they differ somewhat in the specific linguistic marking of this scope. As in English, some Hebrew IRs mark their scope by linear position. Thus, the bare PNself and the b-marked VPself take scope backwards over the preceding segment […]. However, two important differences must be considered: the wider range of positions open to the Hebrew IR and its occurrence with a preposition. Unlike English, Hebrew requires prepositional marking with some IRs. As noted in section 2, bare IRs can occur with any nominal antecedent and must immediately follow it, and so are identified as PNself. In contrast, b-IRs require subject antecedents. (Cohen this volume: sect. 4.1)

Thus, strictly speaking, we cannot take the step of argumentation just mentioned to be a legitimate contrastive observation in its own right, embedded in the main one. The form side of the IRs in Hebrew in comparison to English is at issue in the main contrastive observation, revealing a difference. Consequently, it cannot be made the tertium comparationis of another contrastive analysis aiming at the meaning/function side. The challenge that Cohen's main contrastive observation sets, a challenge inviting attempts at falsification and thus being of significant scientific value, is this: Can the claim that the differences between English and Hebrew IRs on the form side do not correspond to differences on the information structural meaning/function side be maintained in the light of further evidence?

The contribution by López contains several lines of argumentation that constitute instances of contrastive information structure analysis. I will pick one of them for illustration: On the meaning/function side, López makes a distinction between 'givenness' and 'discourse anaphoricity'.

(12)  

a. Context: I'm wearing a red coat. What are you wearing?  
   (i) I'm wearing a BLUE coat.  
   (ii) I'm wearing a blue SHIRT.

b. Context: What kind of coat are you wearing?  
   (i) I'm wearing a BLUE coat.  
   (ii) # I'm wearing a blue SHIRT.

For López (this volume: sect. 1), the expression coat in (12ai) is 'given' – by virtue of being only 'accidentally' occurring again, after having been mentioned in the preceding question, as part of the larger focus expression; cf. the appropriateness of (12a) as a reply to (12a). In (12bi) coat is 'discourse anaphoric' – by virtue of
being mandatorily coreferential with an antecedent in the previous discourse and not being a part of the focus expression; cf. the inappropriateness of an alternative like (12bii) as a reply to (12a). These two information structural categories, 'givenness' and 'discourse anaphoricity', can be applied to both English and Catalan, that is, each of them may serve as a tertium comparationis on the meaning/function side. English and Catalan show interesting non-equivalences on the form side in the manifestation of these information structure categories. For example, a Catalan 'discourse anaphoric' constituent, in contrast to an English one, is obligatorily clitic right-dislocated, as shown by (13) (see ib., and p.c.).

(13) A: What kind of coat do you have?
B: (i) En tinc un de BLAU, d'abric.
    'I have a BLUE coat.'
(ii) # Tinc un abric BLAU.
    have.1st a a coat blue

Also, whereas an English 'accidentally given' constituent is obligatorily 'deaccented', as is the case with coat in (12ai), a phonological rule to the same effect does not exist in Catalan, as shown by (14), where blau receives the main sentence accent despite its being 'given' (see ib.: sect. 4).

(14) A: Mary drove her blue convertible. What did John drive?
B: Va conduir el seu sedan BLAU.
    Past drive.inf the her/his sedan blue
    'He drove her/his blue SEDAN.'

The method employed by Skopeteas & Fanselow is based on a tertium comparationis on the meaning/function side. For each of the languages investigated (American English, Québec French, Hungarian and Georgian), data are gathered by an elicitation procedure that targets utterances where, in a canonical sentence structure, (a) the subject would be an identificational focus expression, (b) the object would be an identificational focus expression, (c) the subject would be a non-identificational focus expression, (d) the object would be a non-identificational focus expression. The visual stimulus is held constant across languages, and the verbal context is created by questions that are semantolexically and translationally equivalent. I would like to highlight from among the various very interesting results gained by this procedure the following observation: As can be expected on the basis of the previous literature (see the references given by Skopeteas & Fanselow), there is a much stronger tendency in (Québec) French than in (American) English of using a cleft construction in those cases where the elicitation procedure targets utterances where, in a canonical sentence structure, a subject would be an identificational focus expression. What I found surprising and in need of further investigation beyond the suggestion put forward by Skopeteas & Fanselow (this volume: sect. 5.2) is the fact that, even in French, in 26 % of these cases the sentence structure used is the canonical one rather than the cleft construction. One would have expected a much lower percentage.
Gast's approach in his contribution to the present volume is the one that is most explicitly and specifically contrastive in that he discusses the *tertium comparationis* issue and devotes much care to establishing 'sub-informativity' as his *tertium comparationis* on the meaning/function side. For Gast, *tertia comparationis* on the meaning/function side constitute the ideal for contrastive information structure analysis: "Ideally, such a 'third of comparison' should be defined on a purely notional basis. It constitutes the invariant in the process of language comparison, while variation is expected in the formal means to encode the relevant categories." (Gast this volume: sect. 1) This view follows naturally either from scepticism as to the existence of universals or cross-linguistic constants on the form side, as suggested by the discussion in section 2.2.2 above, or from scepticism as to whether universals or cross-linguistic constants on the form side can be made to work as *tertia comparationis* in contrastive information structure analysis.

From among the many interesting observations by Gast about the commonalities and differences between English and German as far as lexical, syntactic and intonational realisations of sub-informativity are concerned, I would like to draw particular attention to the following one:

both English and German have contours that are used in contexts of 'sub-informativity', but they are used at different levels of generality: the English fall-rise is a general marker of 'incompleteness', and therefore covers 'sub-informativity' [...] as one of its functions, whereas the German root contour is a rather specific marker of 'context-changing sub-informativity'. (Ib.: sect. 7.3)

This remark reminds me of Hawkins's (1986, 1988) well-known claim about the tighter fit in German than in English between syntactic surface structures on the one hand and semantic relations between predicates and their arguments on the other. It seems worthwhile investigating whether there is also a tighter fit in German than in English between intonational patterns on the one hand and the information structural categories they are associated with on the other. Gast's observation concerning the English fall-rise and the German root contour points in this direction.

### 4. Conclusion

The specific *tertia comparationis* for contrastive information structure analysis are either information structural categories from the meaning/function side, or forms, structures, operations, principles from the form side assumed to be equivalently present in the languages involved and to be relevant for the expression of information structural categories in at least one of them. The background *tertium comparationis* of contrastive information structure analysis is translational and/or semanto-syntactic equivalence, as it is for contrastive linguistics in general as soon as meaning bearing units are involved (i.e. except for contrastive phonetics and phonology). By 'background' I mean that the range
of forms, structures etc. considered on the form sides has to be constrained by considerations of translational and/or semanto-syntactic equivalence (see e.g. Krzeszowski 1990). At the same time, the phenomena which cause translational and/or semanto-syntactic equivalence to manifest itself as a fuzzy and relative concept may point to linguistically interesting facts, also in the domain of contrastive information structure analysis.
Notes

1 I am grateful to Edward Göbbel, Alex Thiel and an anonymous reviewer for corrections and valuable suggestions. All remaining errors are mine.


3 The various papers in the present volume provide examples and many references to comparative research on information structure within functional, generative, and typological frameworks. For a quite recent overview of research on information structure, making reference to various languages, see Erteschik-Shir (2007). The publications of the Collaborative Research Centre (Sonderforschungsbereich, SFB) "Information Structure: The Linguistic Means for Structuring Utterances, Sentences and Texts" (see http://www.sfb632.uni-potsdam.de/main.html) and the references contained in these publications provide a wealth of relevant bibliographical material. So do the articles collected in Schwabe & Winkler (eds) (2007).

4 According to Kortmann (1998: 138f.), the decline of contrastive linguistics was much more pronounced in the USA than in Europe. Disappointment generated by exaggerated expectations concerning the benefits of contrastive linguistics (CL) for the purposes of language teaching and learning affected the more pedagogically oriented CL researchers in the USA more strongly than the more theoretically and descriptively oriented CL researchers in Europe.

5 As to the relevance of a comparative approach that focuses on differences for the study of universals as conceived of in generative grammar, I may quote Kayne (1996/2000: 3):

Comparative syntax can be thought of as that facet of syntactic theory directly concerned with the question of how best to characterize the properties of human languages that are not universal. Put another way, comparative syntax directly addresses the question of how best to understand the notion of parameter taken to underlie syntactic variation.

The study of differences among languages must obviously proceed in tandem with the study of what they have in common, that is, with the study of the principles of Universal Grammar (UG) that interact with language specific parameters to yield observed variation. Similarly there is every reason to believe that the search for universal syntactic principles cannot proceed without close attention being paid to syntactic variation.
See the overview in Pan & Tham (2007: ch. 4) and the works referred to there. Pan & Tham's (2007: 208) own proposal for a "definition" of contrastive linguistics is as follows:

Contrastive linguistics is a branch of linguistics maintained by foundations in philosophical linguistics, comprising aspects in theoretical and applied studies with an object [sic] to contrast two or more languages or dialects to describe the similarities and, particularly, the differences for an explanation in view of the relations between human language and its spirit, so as to promote advancement in general linguistics and facilitate the exchange and understandings of cultures and civilizations for human harmony.

The examples involve (lexico-)syntactic topics – rather than information structural ones – in order for me to be able to connect them to the mainstream generative view of the relation between language specificity, universality and parametricity as laid down in the quotations from Kayne (1996/2000) and Haegeman & Guéron (1999) given above. There is no mainstream generative view of the elements and principles of information structure, nor of its place and role in – or in relation to – the architecture of grammar.

I have found no reference to (analyses of) the peculiar type of construction exemplified by a pleasant three weeks in the recent comprehensive survey volume Noun Phrase in the Generative Perspective (Alexiadou & Haegeman & Stavrou 2007). In the Cambridge Grammar of the English Language (Huddleston & Pullum 2002), this construction type is described as involving a recategorisation or respecification of a plural measure phrase as singular (see ib.: 346, 353f.).

By "relevant translation equivalents" I mean translation equivalents that are semanto-syntactic equivalents in Krzeszowski's (1990) sense at the same time.

For instance, it may be debated whether the impressive amount of knowledge that has been accumulated by typologists concerned with 'grammatical voice' have left anything coherent of this notion that may qualify it for the status of grammatical category and thus as typological tertium comparationis. See e.g. Klaiman (1991), who insightfully surveys and classifies different voice system, concluding:

Plausibly, then, what is common to different types of voice systems may be that wherever voice alternations occur, they encode alternative assignments of arguments to positions which have superior ranking at some grammatically significant level of organization, be that of relational structure, information structure or some other level. (One alternative level which will be taken up momentarily is ontological structure.) (Ib.: 262f.)

The author hopes that, for all its tentativeness, the present study might provide a fruitful basis for an enhanced understanding on an intriguing grammatical category whose nature has long seemed obscure. (Ib.: 271)
It is not obvious that the generalisation over quite heterogeneous "grammatically significant level[s] of organization" suggested in the first quotation may serve as a definition or characterisation of a grammatical category.

11 My reference to Lambrecht (1994) here to the exclusion of other theories of information structure reflects a personal view to the effect that this work is (still) outstanding by its precision, clarity, homogeneity and comprehensiveness in explaining what information structure is about in general. There are other works that are highly insightful and analytically deep on specific topics of information structure, many of which mentioned in the present volume. But they do not fit my "metalevel" (as an anonymous reviewer puts it) purposes in this paper.

12 Among others, this point was made by Lattey (1982: 133) and König (1993: 290f.). According to the latter:

A substantial part of the established terminology used in language-specific descriptions is totally hostile to and unsuitable for language comparison, since it treats certain categories as something sui generis and thus inaccessible to comparative statements. This is most obviously true of all the terms that identify a category purely on the basis of its form. The labels "-ing form" and "there sentences" are clear cases in point, but I would also regard the term "expanded form" instead of "progressive form" as unsuitable for any comparative enterprise. Terms like "gerund", "gérondif", "gerundio" and "gerundium" are notorious examples of the use of the same term or at least very similar terms for very different phenomena [...].

13 In previous work (Breul 2004, 2007), I argued that there is an underlying syntactic difference between such cases as (6B) and (8a):

(6B) [FocP [My CAR]1[+foc] [Foc C[-foc] [IP t1 broke down]]]
(8a) [FocP [My car]1[-foc] [Foc C[-foc] [IP t1 broke DOWN]]]

In both cases the phrase my car moves to the specifier position of a functional phrase above IP, called FocP. The movement is triggered by the checking requirement of a syntactic feature, [+foc] in the first case, and [-foc] in the second case. [+foc] is paired with a semantic focus feature (my car is an identificational focus expression); [-foc] is paired with a semantic topic feature (my car is a topic expression).

14 Of course, there are contrastive aspects involved in the syntactic structures of the sentences in (11) (such as the differences in the structures of the verb phrases, gave it to me versus me l'a donnée) whose investigation belongs to the domain of contrastive syntactic analysis. The implication that these differences are irrelevant may be challenged. Challenges of presupposed or implicated assumptions of this kind will potentially advance our understanding of the phenomena involved.

15 As far as English / German contrastive information structure analyses are concerned, I may point to work by Doherty (see note 2) where this methodological step – refining perceptions of (dis)similarity and defining them as
a similarity constraint – is carried out with utmost expertise, leading to highly insightful and interesting results.

16 Chesterman's (1998) Contrastive Functional Analysis is "based on a notion of similarity rather than one of identity; it explicitly relies on translation competence" (ib.: 40). On the relativity inherent in similarity, which is substituted by Chesterman for equivalence in other researchers' conceptions, see ib.: 5-16.

17 "A declarative sentence \( S \) is sub-informative relative to a strategy \( Q \) (containing \( S \)) iff \( S \) does not answer all questions in \( Q \)." (Gast this volume: sect. 3.2)
References

Collaborative Research Centre (Sonderforschungsbereich, SFB) "Information Structure: The Linguistic Means for Structuring Utterances, Sentences and Texts". <http://www.sfb632.uni-potsdam.de/main.html>.


