

**LEGE ARTIS**

Language yesterday, today, tomorrow

Vol. I. No 1 2016

**THE ELUSIVE ELLIPSIS – THE COMPLEX HISTORY  
OF A VAGUE GRAMMATICAL CONCEPT IN NEED  
OF EMPIRICAL GROUNDING***Katrin Menzel*<sup>1</sup>

Menzel, K. The elusive ellipsis – the complex history of a vague grammatical concept in need of empirical grounding // *Lege artis*. Language yesterday, today, tomorrow. The Journal of University of SS Cyril and Methodius in Trnava. Warsaw: De Gruyter Open, 2016, vol. 1 (1), June 2016. – p. 163-201. DOI: 10.1515/lart-2016-0004 ISSN 2453-8035

**Abstract:** This paper reports a part of the larger corpus-based study, which investigates English-German contrasts in text cohesion and discourse organisation. It has its focus on ellipsis-antecedent-relations that contribute to the cohesiveness of texts. The present work is intended to contribute to the development of a discourse-oriented contrastive grammar on the English-German language pair with relevance to theoretical and applied linguistics, translation studies and foreign language pedagogy.

**Keywords:** text linguistics, cohesion, ellipses, sentence fragments, grammatical categorisation, history of linguistics, English-German contrasts, corpus linguistics, frequency analyses.

**1. Introduction**

The aim of this paper is to elaborate on a specific aspect of the author's recently completed PhD dissertation (Menzel submitted) that was only briefly touched on in the dissertation and in previous publications on ellipses as cohesive devices in English and German corpus texts (Menzel 2014; Menzel forthcoming). The dissertation "Understanding English-German contrasts – a corpus-based comparative analysis of

---

<sup>1</sup> Parts of this study were funded by the German Research Foundation (DFG) in the framework of the project "German-English contrasts in cohesion – Towards an empirically-based comparison" (GECCo), GZ STE 840/6-1 and 6-2 und KU 3129/1-2, <http://www.gecco.uni-saarland.de/GECCo/index.html>

ellipses as cohesive devices" has its focus on endophoric ellipsis-antecedent-relations creating textual links and contributing to the cohesiveness and coherence of texts. Cohesive devices are indicators of lexico-grammatical relations within a text across clause or sentence boundaries. Textual cohesion in multilingual contexts is discussed in more detail in Menzel et al. (forthcoming). In the case of ellipsis, in contrast to other cohesive ties, a textual relation is not exactly set up between two actual text segments, but between a textual antecedent and a subsequent omission (cf. de Beaugrande 1991: 252). Some phrases and sentences are remnants of full constructions, of which certain parts have been annihilated by ellipses. The omitted elements are assumed to be present in the underlying syntactic structure. Elliptical structures can be used to avoid the reiteration of lexical material and the insertion of pro-forms functioning as substitutes.

We annotated ellipses as cohesive devices along with other types of non-cohesive ellipses and sentence fragments in an English-German electronic text corpus for a cross-linguistic study. In general, we expect to find more ellipses in English than in German. German, on the other hand, has probably higher frequencies of other types of irregular syntactic structures. Additionally, we assume that the German data will reveal more variability. In this study, we stress the need for a clear distinction between endophoric ellipses as linkages between different text segments on the one hand and other types of omissions and sentence fragments on the other hand (e.g. context-independent non-clausal units or ellipses with phrase-internal antecedents that do not primarily contribute to textuality and do not establish cross-clausal or cross-sentential endophoric relations). This paper is based on a relatively theory-neutral, broadly functional approach for describing ellipses underpinned with observation statements and authentic examples from a bilingual corpus of written and spoken language. In that way, we hope that it has the potential to arbitrate between competing academic theories that are interested in the nature of ellipses.

Due to the widely varying structures of ellipsis remnants, manual annotation, which we performed with the open-source tool MMAX2 (Müller & Strube 2006), turned out to be more efficient and reliable than semi-automatic procedures that we tested. The annotated patterns have been extracted from the corpus and evaluated with statistical methods to analyse the differences between languages and registers. We used the programming language and software environment MATLAB for various statistical analyses and for the visualisation of the data. So far, the most typical users of MATLAB are engineers, but the MATLAB working environment turned out to be well-suited for corpus-linguistic research.

The conceptualisation, annotation and analysis of ellipses as cohesive devices contribute to a larger research project "German-English contrasts in cohesion – Towards an empirically-based comparison" (GECCo). The GECCo project is a comprehensive corpus-based investigation in the field of contrastive linguistics with the aim to identify German-English contrasts and similarities in their systems of cohesion and their textual instantiations of these systems across languages (English vs. German), across registers (different text types and communication scenarios along the written-spoken continuum) and across production types (non-translated vs. translated texts).

One point that has not sufficiently been discussed yet in prior publications on the role of ellipsis in textual cohesion is the history of ellipsis as a grammatical concept and how this has influenced current ellipsis taxonomies and definitions. In this paper, we describe the origin of the concept of ellipsis and trace certain stages in its historical development. The historiography of linguistic concepts is not unimportant or negligible, as it helps to explain how certain definitions have evolved that are now applied in contemporary linguistic studies and put forward in pedagogical grammars. Ellipsis is a rather complex linguistic concept. It is related to concepts such as 'word',

'sentence' or 'cohesion' which would equally be difficult to sum up in one-sentence dictionary definitions.

The legacy of Greek and Latin rhetorical-grammatical categorisation and the thinking of seventeenth- and eighteenth-century prescriptivists from a particularly productive period for the development of grammatical norms and classifications are two important factors that have paved the way for our modern grammatical terms. Putting the concept of ellipsis into a wider historical context of Greek and Latin grammatical and rhetorical terminology and its reception and adaptation from the Renaissance to modern times is a useful step that will help to untangle some of the complex issues regarding this term. Ellipsis is still a concept that is difficult to pin down empirically in corpus-linguistic research. We will therefore also address some aspects of the operationalisability of the traditional ellipsis concept for empirical quantitative studies. The existing literature provides us with relatively broad or vague descriptions of ellipses which do not readily translate into an operational definition that we can apply to corpus data. This has hindered the development of clear hypotheses in quantitative terms in the past.

The discussion is organised as follows. After this introductory chapter, Chapter 2 starts by sketching the history of ellipsis as a grammatical concept from classical antiquity to modern times. Chapter 3 gives an overview on some prominent contemporary ellipsis taxonomies. In Chapter 4, we discuss challenges for empirical, corpus-based studies arising from existing ellipsis descriptions and present an annotation scheme for ellipses as cohesive devices and some findings from a cross-linguistic corpus study. Finally, in Chapter 5, we present our conclusions.

## **2. A brief historiography of ellipsis as a grammatical concept**

### *2.1 The grammatical heritage of classical antiquity*

To a certain extent, the concepts and categories we now use to analyse and describe modern languages have been shaped by the heritage of early philological, rhetorical

and even philosophical discourses. Therefore, this chapter provides a closer look at the emergence and historical development of ellipsis as a linguistic concept from classical antiquity to modern times.

The term ellipsis comes from the Greek word *ἐλλείπω* which means: *to leave out, fall short of, lack, be inferior to*. Grammarians have borrowed the term from early rhetorical theory where ellipsis was understood as a stylistic device used for linguistic brevity, conciseness and focus. The classical Greek term for ellipsis came into Latin usage about 2000 years ago. Later, English and German as well as other modern languages borrowed the grammatical sense of this word from the classical languages. Presumably this took place slightly later than the earlier recorded borrowing of the geometrical sense of the word, but it may have happened at the beginning of the 17th century, at a time when more and more educational institutions were established in Europe.

From the very beginning of the discussion on ellipses until now, there has been disagreement among scholars about an exact definition, possible subtypes and the clear distinction from other structures; it is impossible to miss these parallels between classical and contemporary discussions of ellipses. Smith (1986: 73) who examined ellipsis definitions of English grammarians from different centuries, each influenced to a different extent by the writings of early Greek and Latin scholars, compared the concept to a polymorphic chameleon not only changing across time, but also according to the language under discussion and according to the perception of the author. He called it a syntactic quicksand able to swallow all kinds of constructions (ibid. 360), which is a fitting and appropriate metaphor because there has always been a tendency to subsume a multitude of different patterns under the notion of ellipsis – from optional omissions of single words within complex structures to short one-word fragment, with numerous types of structures in between.

Investigating the nature of the ellipsis concept and its complex evolution over the centuries may indeed sometimes seem like examining a chameleon concept. A term used both in rhetoric and syntax, ellipsis has always had a dual nature as linguistic deviation and rhetorical figure at the same time. Early sources on rhetoric and grammar already stressed the ambivalence of ellipses as examples of deviant or non-standard grammar that can also be used for reasons of emphasis, economy or style. On the one hand, ellipses were described as an elegant linguistic strategy to avoid unnecessary reiteration. As a figure of speech, ellipses were considered part of ornateness, one of the virtues of style. Although the exact or partial repetition of structures was also seen as a rhetorical device, in "Rhetorica ad Herennium" (4.18) – formerly attributed to Cicero, but of unknown authorship – it was pointed out that orators should not merely reiterate words to emphasise the meaning and content of a message. While such figures of speech may add linguistic material and contribute to the expressiveness and verbosity of a text, ellipses, in contrast, were seen as a means to avoid redundant elaborateness. At the same time, in the context of classical oratory, ellipses were frequently associated with casual speech and bad style. They were described as devices that violate stylistic virtues such as clarity, explicitness and propriety propounded by Aristotle and his disciple Theophrastus and reiterated by rhetoricians such as Cicero and Quintilian. This dual nature of ellipses and a certain acceptability cline for elliptical structures is something we still observe in current approaches that frequently distinguish between proper and improper ellipses.

Our corpus-based research is based on a definition of ellipses as omissions of constituents or parts thereof that are understood via complete syntactic deep structures. This approach is based on the general assumption that languages have 'deep' or underlying invisible structures, detectable only by inference. This means that certain types of phrases and sentences can be regarded as remnants of full phrases or sentences of which certain parts have been annihilated by ellipsis. The assumption that languages have both a surface level and an underlying structure is nothing new in linguistic

theories. The thought that a word can be regarded as being present in an underlying structure if it is understood, although the sign for it does not appear, probably goes back as far as ancient Greece, namely to Aristotle's "De Interpretatione" (Latin title by which the work is usually known), as has been claimed by Smith (1986: 23). Householder (1981: 17), on the other hand, suggested that it was Apollonius Dyscolus who invented the concept of an abstract underlying structure of language. Particularly the literature on ellipsis that appeared from the 1970s to the present took up the question whether ellipsis resolution is mainly syntactic or semantic in nature and whether ellipsis involves deletion from an underlying syntactic structure or not.

Several examples of ellipses provided by early grammarians already exhibit many aspects of modern discussions, for instance, whether grammatical categories such as ellipsis are broad or abstract enough to be used as cross-linguistic categories. In many cases, the internal structures of phrases or clauses containing an ellipsis site are language-specific, but have certain parallels to other languages. In Ancient Greek, for instance, we find several impersonal 3rd person verbs with no explicit syntactic subjects, e.g. *astrapei* [(it) is lightning] or *huei* [(it) is raining] where it has been suggested that an underlying definite subject (Zeus) can be understood. Additionally, the deletion or substitution of articles was seen as a typical elliptical phenomenon in Greek. This phenomenon has a certain parallel in modern languages that have articles, such as English and German, but since Latin had no articles – and obviously no equivalent to article ellipsis – grammarians of Latin such as Priscian started to subsume those structures under the term of ellipsis that were typical non-realised elements in Latin, e.g. pronouns and prepositions which were often redundant in Latin due to case-marking and verbal morphology, but could be inserted for emphasis.

The non-realisation of the copula verb in the present tense between a noun phrase and a predicative expression is an example for optional omissions in both Latin and Greek which were possible since the copula does not describe any specific verbal action. In

several modern languages, we can also attest the possibility of omitting the copula verb as well as other sentence elements with low lexical information value in specific contexts. In English and German, we find grammatically acceptable copula omissions preferably in examples of block language such as headlines (e.g., *Japans Wirtschaft auf Talfahrt / Japan's economy in a downward spiral*), but also in some aphorisms, slogans and fixed expressions (e.g. *Every man a king*). This omission possibility is different from zero copula in languages like Russian where no underlying present tense copula verb is understood in verbless sentences.

Nowadays, modern linguists continue to refer to the roots of grammatical terms in antiquity, although their meanings have constantly been extended and modified in order to be applied to the specific structures of various modern languages.

## *2.2 Effects of prescriptivism and the development of language norms*

The previous section has illustrated how our modern understanding of grammatical concepts is, at least indirectly, influenced by grammatical descriptions of classical languages. Another non-negligible factor that played a role in paving the way for our current understanding of ellipsis are the writings of early prescriptivist grammarians from the seventeenth- and eighteenth-century onwards who extensively discussed this concept in debates about language norms, correct language use and the notion of the complete and well-formed sentence.

This section will discuss some important stages in the development of language norms and modern grammatical categorisations. It deliberately has a strong focus on the evolution of grammatical description and grammatical theory-formation in the English-speaking world as the ellipsis annotation scheme we suggested in our study (Menzel submitted) covers omission possibilities in the English language to which we compare German textual omissions. Like in many recent cross-linguistic analyses, major references from English-speaking academic contexts – in our case, Halliday & Hasan's



ellipsis taxonomy from their standard reference book on textual cohesion (1976) – play an important role for the development of theoretical frameworks and linguistic descriptions as English has become the dominant lingua franca of science and research, which also has a certain influence on linguistic categorisation processes.

After Roman institutions, culture and literature had gradually lost their influence in Europe after the fourth century, Latin still remained the language of the church and the language of the learned élites of Europe well into the eighteenth century. Professors in most European universities were inclined to deliver their lectures in Latin until the late 1800s. Latin continued as part of the requisite curriculum of many universities and schools until the beginning of the 20<sup>th</sup> century. In fact, English 'grammar schools' originally were schools for mainly teaching classical grammar and literature; and the influence of Latin on the education system of German schools and on German grammar teaching some centuries ago was similarly strong as in English contexts. Early works on English and German grammar that appeared from the 16<sup>th</sup> century onwards were mainly written in Latin and often contrasted the structure of English and German with this language – then seen as the model language par excellence.

The earliest grammars of modern European language were by no means English ones. The first grammar of any vernacular modern European language was the Castilian Grammar by the Spanish humanist Antonio de Nebrija in 1492 that expressed pride and confidence in the Spanish language, worthy to be set alongside Latin and deserving an explanatory grammar. Linguistic hegemony and superiority of French compared to other languages was expressed by the Port-Royal grammarians from the middle of the 17<sup>th</sup> century on, when French – instead of Latin – became seen as the language of culture through Europe, spoken by the members of the aristocracy, high bourgeoisie and intelligentsia. English writers from this era such as Daniel Defoe were in favour of establishing an English language academy comparable to those that had been

established in Italy (1582), France (1635) and Spain (1713) (cf. van Gelderen 2006: 224), but no such academy was ever successfully established.

Compared with certain continental vernaculars which had official institutions for their regulation, English appeared relatively uncultivated, unstable and unregulated at that time. Therefore, the codification and standardisation of English usage that took place during the 18th century due to a substantial growth of the printing industry in England was not controlled by an official language academy, but rather by people representing the major English publishing houses concentrated in London and the East Midlands. In that period, judgments on the acceptability of grammatical constructions, such as ellipsis, were often made by individual grammarians who, in many cases, also worked as poets, literary critics, translators or editors. As a consequence of the growing print culture, the 18th century was characterised by a rapid standardisation process and the development of an influential prestigious form of English. Not only English vocabulary and spelling were transformed and codified, its syntax and grammar equally underwent certain standardisation processes. A growing national and international market for English-language dictionaries and grammars started to develop. The authoritarian and normative nature of 18th-century English grammars gave rise to a doctrine of correctness according to which expressions and syntactic structures such as elliptical sentences were strictly seen as either correct or incorrect. The newly fixed standard rules suitable for the speech of the literate classes, for literary and scientific publications as well as for legal and commercial discourse resulted in being distributed and used for educational purposes and were looked upon as a model. Knowledge of linguistic diversity and variation was rather incidental and unsystematic. Competing forms and alternative ways of expressing things led to concern about correct usage and were perceived as negative. The standard variety including strongly standardised syntactical rules began to be developed for a range of functions while non-standard use became more associated with simplicity and informal purposes. In sum, the newly developed normative rules changed how English was taught, judged, analysed and

printed with regard to domestic and international markets and educational systems. English prescriptive grammarians from this period were still heavily influenced by Latin grammatical concepts and categories and often attempted to make English grammar follow Latin rules: "Despite the triumph of English in all domains of use, Latin grammar continued to cast a long shadow over the grammatical analysis of the vernacular" (Finegan 1998: 547).

There has always been a relatively subjective acceptance and rejection of certain grammatical structures, including elliptical constructions, particularly by grammar prescriptivists. In an attempt to standardise and stabilise the English language and to reduce it to a certain set of teachable rules, prescriptivist grammarians started to set up strict rules of grammatical usage condemning those forms and constructions that they considered to be improper. Elliptical constructions were extensively discussed by 18th-century English grammarians who made a distinction between omissions as a grammatical figure and a grammatical error – like Greek and Latin grammarians already did a long time ago. In his grammar "A Key to the Art of Letters", published at the threshold to the 18th century, Lane (1700, cited in Smith 1986: 388), for instance, described improper omissions of words in sentences. These 'suppressions', as he called them, are words that in his view would be necessary to full constructions. The 18th-century lexicographer and grammarian Ash equally distinguished between proper and improper omissions in his pedagogical grammar and explained some seemingly ungrammatical constructions by an assumed underlying structure. He saw the purpose of ellipsis in the avoidance of "disagreeable repetitions" and the expression of ideas "in as few words and as pleasing a manner as possible" (1763: 119).

As indicated above, despite increasing attempts at standardising the language, the concept of ellipsis varied extensively in early descriptions and often depended on the professional judgment by individual grammarians. The grammarian Greenwood for instance accepted the ellipsis of *that* and relative pronouns in subclauses in his early

18th-century works on grammar, while some later sources criticised these types of intrasentential omissions as being ungrammatical. From the 18th century on, copula omissions in spoken language are another example of omissions that became, in the opinion of some, indicators of an incorrect or defective use of English. Such omissions became particularly associated with the 'broken English' of speakers of English-based pidgins and creoles (Rickford 1998). In general, 18th-century authors saw omissions as a useful means when they helped to avoid repetitions, but criticised ellipses when they resulted in ambiguity or were used in a non-standard way. However, these early grammarians unduly inflated the ellipsis concept and suggested numerous constructions as being elliptical that would not or only marginally count as actual cases of ellipsis today even if it is possible to add redundant syntactic material in many cases (e.g. *I live [a life] in London.* – Lane 1700: 105; *turn to the right [hand]* – Greenwood 1711:220; *my house and [my] lands; the man was old and [the man was] crafty* – Ash 1763).

Despite the multitude of slightly or greatly varying ellipsis descriptions from that period, grammar writing in the 18th century involved a lot of plagiarism – sample sentences and passages were often copied from one grammar to another, including translated material from books originally written in Latin or French. The main goal at that time was not seen in describing language as used by the common people or developing innovative linguistic theories, but in suggesting (and selling) new and better methods for teaching the grammatical principles of the vernacular. This inevitably led to many terminological simplifications and generalisations (Smith 1998: 435).

In early descriptions of ellipsis in English, this concept is frequently associated with negative aspects such as harsh, violent or rough language as several quotations from the online Oxford English Dictionary (OED) quotation database in the dictionary entry for ellipsis indicate:

*1789 W. Belsham Ess. I. ii. 25: Violent ellipses and inversions of language.*

*1874 H. R. Reynolds John the Baptist ii. 112: Grammatical roughnesses or ellipses.*

The topics of classical grammar writing and early grammar writing on languages such as English and German are quite complex but help to explain the evolution of the current understanding (and misunderstanding) of the meaning of the term ellipsis. It is possible that particularly English reflects a strong influence of the grammatical heritage and analytical methods of early Latin and Greek grammarians that were enjoying a revival in a period of time when the standardisation of English received a lot of attention. Moreover, fewer deviations from the standard sentence pattern were tolerated from then on in English and this may explain a tendency to see ellipses and sentence fragments as inferior compared to other sentence types, which is probably more pronounced in English than with regard to German or other modern languages.

### *2.3 Ellipsis and the history of the sentence concept*

The historical development of 'ellipsis' as a linguistic concept is strongly related to debates about correct language use as shown in the previous section. It is particularly related to the notion of the correct, complete and well-formed sentence and the definition of its obligatory elements. However, the sentence is another complex concept whose meaning may seem evasive. Highly depending on the definition of such linguistic reference units, ellipsis is a relative concept as it refers to what is understood as a complete unit in which an omission can take place.

The current debate on the meaning of 'sentence' is thought to be dominated by approaches and definitions that occurred from the 1950s onwards, but Graffi (2001: 111-166) demonstrates that this topic was carefully investigated long before, which paved the way for modern definitions, just as the related concept of ellipsis has been evolving over a long period of time. There are definitions of the sentence which

classical tradition handed down to us (e.g. by Priscian who defined a sentence in rather philosophical terms as an arrangement of words that follows some criteria of combination having the capacity of expressing a complete thought) or the 18th-century Port-Royal grammarians' tripartite structural definition of sentential propositions with subject, copula and predicate. Ries (1931) listed more than 150 different sentence definitions, many of which date from the late 19th century alone. For some grammarians, the main characteristics of the sentence were its expression of thought and its propositional content. Others rather emphasised its form and its essential syntactic constituents. Jespersen (1933: 106) coined the label of 'nexus' for the subject-predicate grammatical structure and additionally defined the sentence on the basis of its communicative function as a complete and independent unit of communication. All syntactic theories had and still have to struggle with sentence types and utterances that do not seem to fit into the respective model. The issue of subsentential speech is not a new one; and in recent time it has been emphasised by several authors that genuinely subsentential phrases can be used to perform speech acts without being cases of ellipsis.

Nowadays, linguists typically refer to at least three senses of the term 'sentence' (Stainton 2006: 31):

- a. sentence<sub>syntactic</sub>: an expression with a certain kind of structure/form,
- b. sentence<sub>semantic</sub>: an expression with a certain kind of content/meaning,
- c. sentence<sub>pragmatic</sub>: an expression with a certain kind of use.

With reference to Stainton, Merchant (2010) put forward the question of whether there are also three kinds of ellipses as syntactic, semantic and pragmatic phenomena. The focus on different aspects in sentence descriptions allowed different understandings of sentence completeness to emerge in terms of grammatical, semantic or contextual completeness. The sentence and the ellipsis concepts today are discussed in syntactic, semantic and pragmatic terms, but also on the graphic and prosodic level. They include

factors such as speech prosody and the intonation which we naturally associate with the end of units in spoken language as well as punctuation used to mark the closure of an orthographic unit in written language. Ellipsis has raised a number of questions on how syntax and semantics, but also pragmatics and phonology, interact with each other.

It would be interesting to trace the main stages of the early history of grammar writing in Europe in more detail, but this chapter confined itself to the most relevant aspects of the grammatical heritage of classical antiquity and the past centuries demonstrating briefly the complexity of the evolution of grammatical norms and of ellipsis as a grammatical concept.

### **3. Contemporary ellipsis descriptions**

Both the ellipsis and the sentence concept have triggered large-scale discussions in modern linguistics and are important topics in grammar instruction with regard to the fields of syntax, stylistics or foreign language teaching. However, it can be observed that statements and exercises from pedagogical grammar workbooks and style guides for writers that are used in educational contexts tend to rely on broad dictionary-like ellipsis definitions; and the concept is not always treated adequately in classificatory terms. Due to such broad definitions, many current approaches are rather 'ellipsophile', in the sense that grammarians still tend to subsume a myriad of fragment types under the notion of ellipsis: non-sentential utterances, non-clausal units, different types of semantic implications, lexical and syntactic reductions, and omissions of single words, constituents and groups of phrasal constituents.

Some for instance, would also consider phenomena such as asyndeton, i.e. the absence of an explicit conjunction (e.g., *I came, I saw, I conquered*) or aposiopesis (an interruption or breaking off in mid-sentence) as types of ellipsis. Quite often such means of language economy and deviations from standard syntactic and grammatical

constructions cannot be used as linking elements in texts, and they are not always the result of an omission.

The broad and general meaning of the grammatical term 'ellipsis' according to the OED takes both syntactic and semantic aspects into account and explains ellipsis as "omission of one or more words in a sentence, which would be needed to complete the grammatical construction or fully to express the sense; *concr.* an instance of such a semantic or grammatical omission". No example for ellipsis is given in the quotations for this OED entry that would actually match this description, which may be due to the fact that many OED entries for words in common use lack recent illustrative quotations. If an entry goes back to the first edition of the OED (1884-1928), the quotation evidence will mostly reflect the material that was available to the editors at the time of writing. Nevertheless, this oft-cited ellipsis explanation that has become widely propagated through this well-known dictionary almost exactly follows the wording that we find in Fowler's "Dictionary of Modern English Usage" which was first published in 1926 and became a classic reference book with several editions published by Oxford University Press. This, in turn, closely resembles the characterisation of ellipsis from Ash's grammar 18<sup>th</sup> century grammar describing ellipsis as the "omission of some word or words which must be supplied, either to complete the sense, or to make out the grammatical construction of the sentence" (1769: 119).

The Duden, a German normative dictionary and standard reference for grammar and spelling rules gives a similarly broad explanation for 'Ellipse' as 'Ersparung von Redeteilen; Auslassungssatz' (omission of parts of an utterance, sentence involving an omission) where the nature of the unsaid elements is not specified. It gives a formulaic expression as an example of ellipsis that actually may have been perceived as an omission in the past, but has become a conventionalised phraseological unit: [*ich*] *danke schön* ([I] *thank you very much*). In another entry in the same source, the word *danke* in exactly this phrase is classified as a particle which shows that it must have



undergone a grammaticalisation process and is not a typical case of ellipsis in modern German.

In both the German and the English literature, widely used dictionaries, pedagogical grammar books and standard references on grammar and language use list all kinds of structures as ellipsis examples. Stang and Steinhauer (2014: 110-112), for instance, include non-verbal and non-clausal structures, semantic implications, formulaic expressions, phrases that have become disconnected from a preceding main clause (e.g., *Wir essen um acht. Auch ohne dich.* – [*We'll eat at eight. Even without you.*]), and other cases of non-standard syntax. However, often there is no need to regard all these cases as omissions.

Due to the conceptual confusion between anaphoric ellipses and textual omissions on the one hand and other means of brevity such as fragments and non-clausal units on the other hand, all these terms are frequently used interchangeably and synonymously, and many grammar books as well as stylistic guidebooks will advise the reader to avoid the usage of these structures altogether. Students can find numerous exercises to learn how to correct or 'repair' ellipses and fragments and to expand them into complete structures. If pedagogical books recommend using ellipses and fragments in English writing at all, they usually emphasize that these constructions should be used intentionally, but sparingly: "Ninety-nine percent of the time you should use complete sentences in your writing" (Provost 1988: 62). Those who consider ellipses to be synonymous with sentence fragments would probably point out that they are a typical phenomenon of spoken language and therefore acceptable in texts with feigned orality or emotional and persuasive language. Non-sentential block language can also be used as a technique in advertisement texts, especially when visual elements complete the thought, or in reduced registers with distinct syntactic properties, e.g. instructional writing, broadcast commentaries or weather forecasts.

Hartwell (1985: 120) showed that the rule for not writing sentence fragments is described in several widely-used college grammar texts where sentence fragments are traditionally seen as conceptual failures (i.e. not having conscious knowledge of the school grammar definition of a sentence) or performance errors. The school grammar approach focusing on complete sentences is also reflected in modern grammar checkers of electronic word processing programmes that tend to flag fragments with the suggestion "Consider revising". Users of word processing programmes find this feature misleading in many cases as the programmes tends to mark mainly non-errors as fragments that actually do not need any revision or correction. It is easy to find numerous similar pejorative judgements with regard to ellipses and fragments in German in prescriptive grammars and in books and articles containing writing advice and communication tips for native speakers and learners of German. Nevertheless, German style guides sometimes intuitively seem less strict with their recommendations.

In many existing typologies, ellipses categories tend to overlap or are treated as gradual notions. One of the most prominent contemporary classification systems for ellipsis types is the one suggested by Quirk et al. who assumed a continuum of ellipsis and proposed a definition of 'strict' ellipsis and less 'strict' forms of ellipsis (1985: 888ff.). A strict ellipsis is defined as an incomplete, grammatically defective structure that can be understood from the surrounding text. It involves verbatim recoverability of the missing elements, i.e. the omitted words are precisely recoverable. The elements left out have to be present in the text in exactly the same form while the insertion of the elided words results in a grammatical sentence with the same meaning as the original sentence. Less 'strict' forms of ellipsis that meet only some of these criteria fall under 'standard ellipsis', 'situational ellipsis', 'structural ellipsis', 'weak ellipses' or 'quasi-ellipsis'. Quirk et al. rank their definition criteria assuming that some are more important than others and describe ellipsis as a gradual phenomenon, which could lead us to think of ellipsis as a prototype continuum with core and peripheral members.

Greenbaum and Quirk (1990: 256) additionally distinguish three categories of ellipsis depending on the place where the ellipsis occurs within the construction. They divide omissions into ellipses of initial, medial or final elements. This has sometimes led to a rather simplified description of textual ellipses in pedagogical grammar books where *only* initial, medial and final ellipsis are given as categories regardless of other aspects (e.g., Downing 2014: 225).

Another oft-cited ellipsis typology is the one proposed by Klein (1981, building on Bühler 1934) who, with a particular focus on German, defined non-gradual subcategories of regular ellipsis where contextual information may be taken from preceding or following utterances, from the perceivable situation or from factual knowledge:

- text-type specific ellipses ("Textsortenellipsen"),
- ellipses as orders to perform actions ("Handlungsellipsen"),
- expressive exclamations ("expressive Ausrufe"),
- elliptical formulaic expressions ("elliptische Formeln"),
- lexicalised or conventionalised ellipses ("lexikalische Ellipse"),
- coordinate ellipses ("Koordinationsellipsen"),
- adjacency ellipses ("Adjazenzellipse"),
- ellipses due to processing difficulties ("verarbeitungsbedingte Ellipsen"),
- ellipses due to incomplete linguistic development ("entwicklungsbedingte Ellipsen"),
- other types of ellipses.

Nevertheless, not all of the categories, suggested by Klein, follow consistent rules for their resolution. Different dimensions of comparison with regard to form, context, recoverability, level of conventionalisation and reason or intention behind the use of elliptical structures are reflected in this classification scheme so that the categories do not have the same level of abstraction in the taxonomy.

A well-established branch that has been extremely productive on the topic of ellipsis during the last decades is transformational generative grammar (cf. Winkler 2005). In this framework, ellipsis subtypes tend to be illustrated from a theoretical point of view, and many examples that are given in the literature are relatively rare in actual language use, e.g., ellipses resulting in ambiguity and underspecified scope or cases where syntactic identity between the antecedent and the elided phrase do not seem to be respected. Such examples have frequently been constructed in the form of isolated, invented sample sentences for the purpose of a theoretical discussion or for testing their acceptability in psycholinguistic experiments. As it has always been the case, many people intuitively and subjectively distinguish between 'proper' and 'improper' uses of ellipses. This distinction depends on factors such as linguistic experience and stylistic understanding of the writer or speaker, but also on the text type or register within the spoken or written mode of discourse. However, in many cases, the acceptability of an ellipsis cannot be determined without reference to the surrounding particular linguistic co-text and extra-linguistic context.

Nowadays, the topic of ellipsis offers a seemingly limitless body of literature, but this has resulted in some confusion over what exactly ellipsis is. That is why several linguists have labelled ellipsis generally as a 'problem' (Ortner 1987), a 'category mistake' (Buss 2004), or even a 'plague' (Bühler 1934: 189).

After a thorough examination of the literature on ellipsis as a grammatical concept (of which we mainly addressed the earlier sources in this paper and not the multitude of publications on ellipsis that appeared from the 1970s onwards which were discussed in more detail in Menzel (submitted)), we decided to build on Halliday & Hasan's (1976) ellipsis subcategories for our corpus-linguistic study because Halliday & Hasan were among the few scholars who had focused on the possibility of using ellipses for the connection of sentences in their function as cohesive devices. They associated ellipsis with the textual metafunction in the theoretical framework of Systemic Functional

Linguistics (SFL). Cohesive devices are an important topic in the systemic functional model which emphasises that languages do not function in isolation, but in actual situations of use and in texts.

#### **4. Defining and measuring ellipses empirically**

The strong empirical focus in contemporary linguistics has drawn the attention of researchers to large textual corpora and quantitative methods as means for analysing discourse phenomena. The challenge with many linguistics terms is that they do not easily provide a basis for empirical research as they have evolved from philosophical and rhetorical concepts and still lack uniform and exact definitions. They have been and are still used as semi-popular terms and their usage is often inconsistent. The challenges of clearly distinguishing ellipses from different structures are, in a way, similar to any early or modern attempts of putting linguistic elements into clear-cut categories, which has led to the fact that many terms are now treated as gradual notions with prototypical and peripheral members. Existing typological schemes do not always place ellipsis subtypes on an equal footing and turn out to have grey areas or overlaps between their categories. This makes them highly non-operational for corpus linguistic analyses.

In our empirical study on ellipses as cohesive devices, we were not able to use a clear existing definition from the literature and had to specify and define the concept in operationalisable terms. In quantitative research, if we want to call a concept a variable, we have to determine clearly observable indicators. Thus, we partly revised and elaborated on Halliday & Hasan's definitions that were still unsatisfactory in their applicability to real discourse data from different text types in English and German and not only to idealised or prototypical isolated utterance pairs. Halliday & Hasan's point of view was that being able to give a theoretical definition of ellipsis does not mean that for every instance we can always recognise whether it is elliptical or not (1976: 168). They described ellipses broadly as "something left unsaid" or "something

understood" (ibid.: 142). Here again, we see that a familiar concept people also use in everyday speech, which is often vague, has been extended to new domains and to linguistic analysis.

Present-day pedagogical grammar books and textbooks on cohesion and discourse analysis give rather vague or sketchy definitions of cohesive ellipsis or illustrate this phenomenon with examples which in our opinion do not always reflect this concept correctly. In Martin and Rose (2013: 167), ellipsis in its function as a cohesive device is broadly defined as "referring to participants by leaving them out", for which the following example of a non-repetition of pronouns is given: *Suddenly, at strange times, they would become restless. Abruptly mutter the feared word 'trip' and drive off.* The second part of this example is a sentence fragment where some elements have become isolated from their preceding clause and punctuated as if they are an independent clause. The non-repetition of subjects and auxiliary verbs, which has sometimes been called 'subject ellipsis' in the literature, is a purely syntactic and virtually obligatory phenomenon in clause complexes. It does not contribute to non-structural text-forming relations. This broad view on cohesive ellipses as implicit or unsaid textual elements has to be narrowed down to a precise definition and confined to reasonable limits, particularly if we want to apply the definition in an empirical study. Otherwise it would be of little descriptive worth and we would run into several practical difficulties.

Nevertheless, the three main categories of ellipsis in Halliday & Hasan's description – nominal, verbal and clausal ellipses – are a good starting point for the description of cohesive ellipses in English and German. In general, they provide a suitable framework for a cross-linguistic analysis of ellipses with possible textual antecedents as both English and German have noun phrases, verb phrases and clauses where certain elements can be omitted that are deducible from the co-text. Where necessary, we made adaptations or clarifications in the description of these subcategories. It is necessary to subsume the multitude of categories suggested in the literature under such general

categories to cover the variety of different omission possibilities in English and German corpus texts from a wide spectrum of communication scenarios on written-spoken and formal-informal continua.

The bilingual GECCo corpus that we annotated for ellipses consists of texts and transcriptions from a broad range of written and spoken registers and text types. The written part of the corpus is a comparable corpus and, at the same time, a bi-directional parallel corpus. It contains comparable texts of English and German as well as their sentence-aligned German and English translations. Parallel corpora consist of texts in a language and their translations in another language. Comparable corpora consist of original texts in two languages that are similar with regard to their sampling frame, balance and representativeness. Our corpus is comprised of academic lectures, political essays, fictional texts, texts from internet forums, instruction manuals, political essays, popular science texts, and letters to shareholders, prepared speeches, tourism leaflets and corporate websites.

We developed a fine-grained annotation scheme of nominal, verbal and clausal ellipses that aims to ensure that all cases found in the corpus can be placed clearly in only one category in order to provide the basis for a meaningful quantitative analysis. Due to space constraints, we cannot describe the annotation scheme in detail in this section, but refer the reader to Menzel (submitted) for a full description. In our study on ellipsis, we were mainly concerned with regularly incomplete structures derived from regular phrasal and clausal structures – a 'proper' use of ellipsis. In the annotation scheme, we defined cohesive ellipsis as a phenomenon where the remnant of a syntactic omission is left grammatically incomplete to create an incomplete nominal or verbal phrase or an incomplete clause. We included the omissions of head nouns within noun phrases, of modal, auxiliary or lexical verbs within verb phrases and the omissions of constituents from entire clauses in our definition and focused on those cases of nominal and verbal/clausal ellipses that endophorically establish textual links. That means that

the content of the ellipsis site can be recovered from its textual antecedent that occurs in a different clause or sentence.

The following example from our corpus illustrates the use of cohesive ellipsis-antecedent relations in English and German:

(1) *Many distilleries welcome visitors to see the whisky-making process, and sample the finished product. Eight of the most famous [ ] are to be found on the world's only Malt Whisky Trail. (corpus ID: EO\_TOU\_004) / Viele Brennereien veranstalten Werksführungen, bei denen Sie auch das Endprodukt probieren können. Acht der berühmtesten [ ] liegen auf dem einzigen "Malt Whisky Trail" der Welt. (GTRANS\_TOU\_004)*

(2) *Sandy went off, and this enabled Sarah to ask if there had been a young man with the girls. Yes, there had [ ]. (EO\_FICTION\_009)*

We excluded exophoric ellipses without textual antecedents that refer exophorically to extra-linguistic elements and can be recovered from to the situational context. Moreover, we excluded cases such as clause-internal ellipsis-antecedent relations (3) as well as those ellipses that are merely the result of coordination or subordination and cannot refer back to textual passages longer than the coordinated or subordinated structure, for instance gapping structures (4).

(3) *Our economies are the most interconnected and interdependent [ ] in the world. (EO\_ESSAY\_010) / Unsere Volkswirtschaften sind die am stärksten miteinander verflochtenen und voneinander abhängigen [ ] der Welt. (GTRANS\_ESSAY\_010)*



(4) *Thus, additional financial burdens are certain, but additional jobs are not [ ].* (ETRANS\_ESSAY\_019) / *Also: die zusätzlichen Belastungen sind sicher, aber die zusätzlichen Jobs [ ] nicht [ ].* (GO\_ESSAY\_019)

Examples (3) and (4) can be explained by syntactic rules alone and do not actually establish textual links as cohesive devices. They do not occur within independent grammatical structures, but in a predicative expression within the same sentence or in the second conjunct of a coordinated structure. Non-cohesive ellipses have been annotated separately in order to be analysed for comparative purposes and to clearly distinguish them from cohesive ellipses.

Additionally, the annotation scheme covers other types of fragments such as the above-mentioned non-clausal structures or sentence fragments where elements have become isolated from the rest of a statement. We clearly distinguish between ellipses as omissions within and across sentence boundaries and other types of fragments or independent non-clausal units and non-sentential utterances that equally serve as means of language economy, but do not necessarily involve an omission, do not have the potential to be used as cohesive devices.

Some authors would consider such structures as 'improper' cases of ellipses. They have sometimes been described as incorrect usage of punctuation or as grammatical errors in standard written English and German, based on the assumption that such constructions cannot stand alone. Nevertheless, they are sometimes used intentionally as staccato sentences to capture the immediacy of oral narration or to create unexpected linguistic contrasts in advertisement texts or fictional texts. They can also signal a speaking pause, a pause for effect, emphasis, or reflection, which is a typical device in political speeches and political essays in our corpus data. We view these fragment structures as being conceptually different from elliptical structures. Due to stricter word-order conditions and the fact that sentence fragment seem to be frowned upon

more commonly in English, we expect fragments to occur more frequently in the German data. They have been annotated in the data to compare their frequency with that of elliptical structures involving omissions.

In the following, we will briefly present some results and implications of our corpus study. Additional information on the data and the analytic procedures have been described in other project publications (Menzel submitted, Kunz et al. forthcoming) where more statistical analyses were conducted to compare English and German, written and spoken language and individual text types as well as non-translated and translated texts.

As expected the overall frequencies of cohesive ellipses were relatively low in our data in both English and German compared to other types of cohesive devices. Surprisingly, the absolute numbers of cohesive ellipses in English and German non-translated texts turned out to be exactly identical with 397 occurrences in each corpus section of non-translated texts. If we compare the normalised frequencies, as the corpus sections have slightly different sizes (English: ca. 408 thousand tokens, German: ca. 414 thousand tokens), we see that English texts have only marginally more cohesive ellipses than German texts. There are about 9.7 cohesive ellipses per 10,000 tokens in English compared to 9.6 in German. Furthermore, our data suggest that the German data have a higher variability than the English data with more German texts having either no cohesive ellipses at all or rather high values, while the English texts are more similar to each other across registers.

The bar chart in Figure 1 visualises the results for the frequencies of cohesive nominal and verbal/clausal ellipses, non-cohesive nominal and verbal/clausal ellipses and fragments in English and German non-translated texts and shows clearly that the frequencies of the ellipses types in the English and German data are very similar.

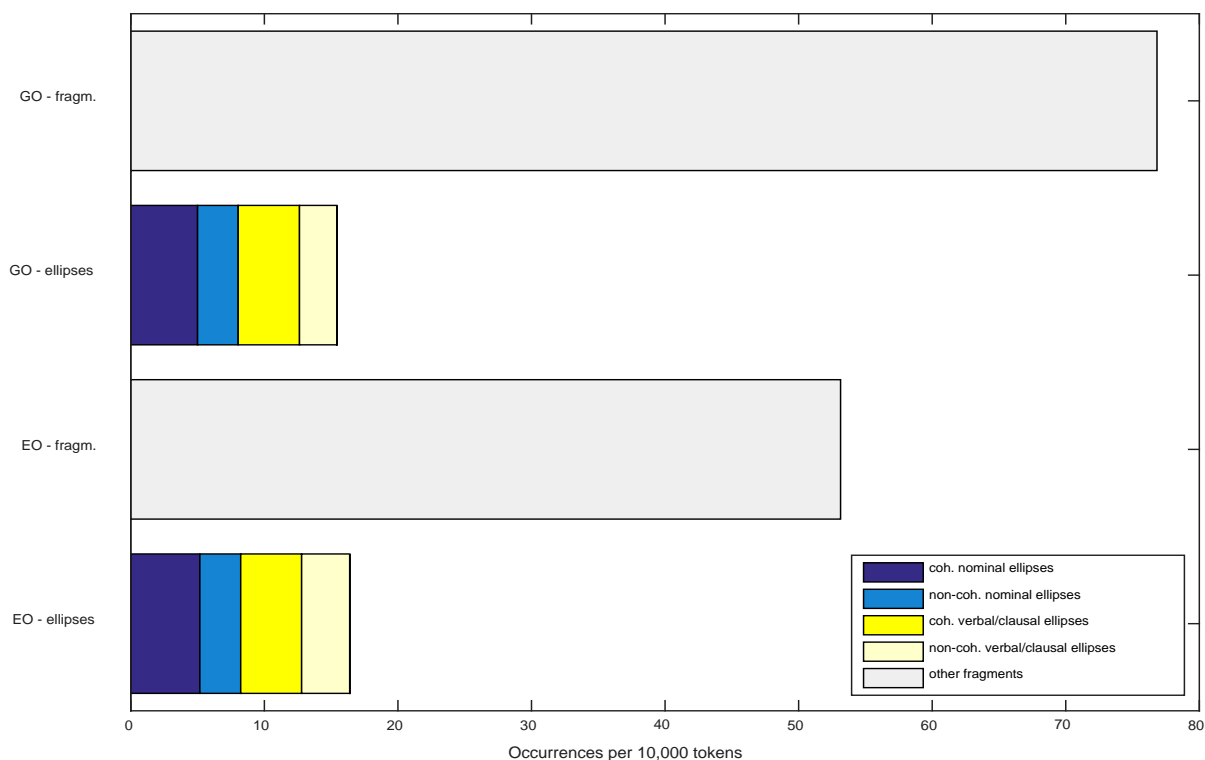


Figure 1. Normalised frequencies (per 10,000 tokens) of ellipses (cohesive nominal and verbal/clausal ellipses, non-cohesive nominal and verbal/clausal ellipses) and fragments in the corpus sections of non-translated texts (EO: English originals, GO: German originals)

There are more striking differences between the two languages if we look at the frequencies of other fragments in the texts. Fragments are much more frequent than ellipses, but there are marked differences between English and German. They occur considerably less frequently in English than in German. Here again, the data suggest a generally higher variability and more extreme outliers for the German data (Figure 2).

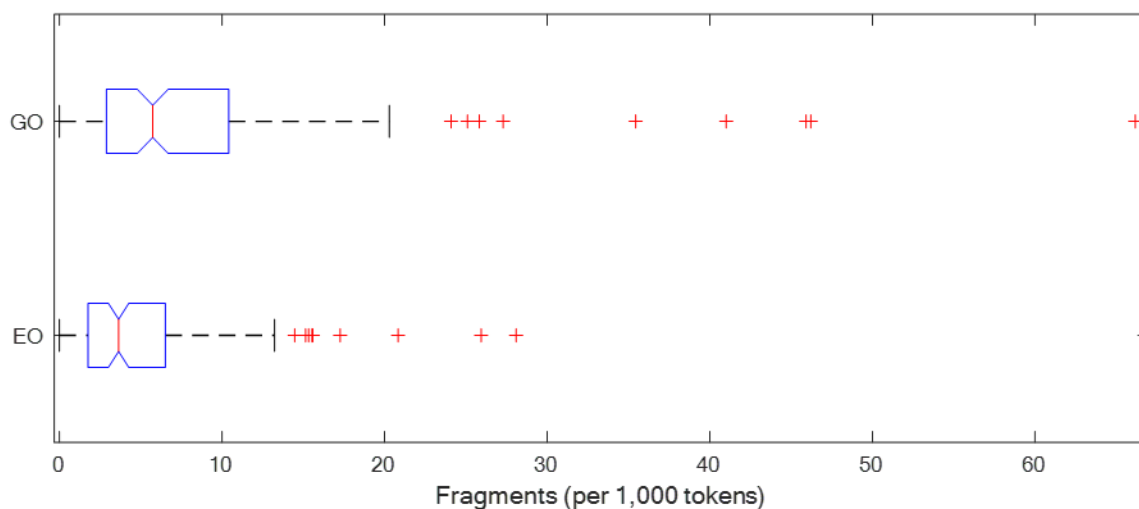


Figure 2. Frequencies of fragments in individual texts in EO and GO (per 1,000 tokens)

Figure 2 is a box plot in which the central mark in each box indicates the median. The bottom and top edges of the boxes indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points that are not considered outliers. Outliers are plotted using the '+' symbol. The notches in each box represent the 95% confidence interval around the median. MATLAB plots the notches at the median plus or minus 1.57 times the interquartile range divided by the square root of the number of observations.

Figure 2 shows that the median values for fragments are 3.64 (EO) and 5.74 (GO). We therefore have a difference of about 2 between the medians of the two groups. As there is a non-overlap of the notch intervals, Figure 2 indicates that the medians for fragments in EO and GO are significantly different at the 5% significance level. While not being a formal test, the comparison of the notches provides a rough measure of the significance of the differences between the values. Thus we can conclude with 95% confidence that the true medians of the 'populations' of English and German texts differ.

The results indicate that German – probably due to its freer word order and more flexibility for structuring textual information – makes use of sentence fragments more

often than English and, for instance, places many dislocated sentence elements outside the sentence boundary to signal a pause, to emphasise certain constituents or parts of complex sentences or to make a prepared text sound more natural where English, on the other hand, often puts the same information into a new syntactically complete sentence. Ellipses as textual omissions and cohesive ellipsis-antecedent relations as particular discourse phenomena that contribute to the cohesiveness of texts are relatively rare in both languages, but are used with similar frequencies in both corpus sections.

These results have several practical implications for language learners and professional writers such as journalists or translators. It is necessary to be aware of the functions and differences between ellipses as omissions within and across sentence boundaries and other types of fragments or independent non-clausal units to be able to use these structures appropriately and in language- and register-typical frequencies.

## **5. Conclusion**

The aim of this paper was to shed light on some of the reasons behind the heterogeneity of attitudes with respect to ellipsis and the relatively loose or even contradictory definitions that can be found in the literature and to illustrate how quantitative corpus-linguistic methods can assist qualitative text analysis. As is the case with many grammatical terms, we can observe certain static and dynamic aspects in the evolution of ellipsis as a linguistic category. Certain types of ellipses establish textual links; other types of omissions are subject to strict locality restrictions and depend solely on syntactic relations. Many other types of sentence fragments and means of language economy have been inadequately subsumed under the ellipsis concept since it was first introduced.

Maintaining vague or excessively broad definitions for ellipsis as a linguistic concept leads to difficulties for empirical, corpus-based studies. On a general level, the aim of

this study is to suggest a model that makes it possible to specify and define the traditional ellipsis concept in operationalisable terms for empirical quantitative studies. Additionally, we would like to lay the foundations for a discourse-oriented contrastive grammar on the English-German language pair with relevance to theoretical and applied linguistics, translation studies and foreign language pedagogy.

## References

- Ash, J. (1763). *Grammatical institutes, or, an easy introduction to Dr. Lowth's English grammar*. London: printed for E. and C. Dilly.
- Beaugrande, R. de (1991). *Linguistic theory: the discourse of fundamental works*. London: Longman.
- Bühler, K. (1934). *Sprachtheorie. The representational function of language*, 1990, Amsterdam: John Benjamins.
- Buss, M. (2004). Die Ellipse – ein linguistischer Kategorienfehler? In *TRANS – Internet-Zeitschrift für Kulturwissenschaften* 15. Available at: [http://www.inst.at/trans/15Nr/06\\_2/buss15.htm](http://www.inst.at/trans/15Nr/06_2/buss15.htm)
- Downing, A. (2014). *English grammar: a university course*. 3rd ed. London-New York: Routledge.
- Finegan, E. (1998). English grammar and usage. In *The Cambridge history of the English language. Vol. 4, 1776-1997*. Romaine, S. (ed.). Cambridge: Cambridge University Press, p. 536-88.
- Gelderen, E. van (2006). *A history of the English language*. Amsterdam: John Benjamins.
- Graffi, G. (2001). *200 years of syntax. A critical survey. Studies in the history of the language sciences* 98. Amsterdam: John Benjamins.
- Greenbaum, S. & Quirk, R. (1990). *A student's grammar of the English language*. Harlow: Longman.
- Greenwood, J. (1711). *An essay towards a practical English grammar*. Menston: Scolar Press [reprint 1968].

- Halliday, M.A.K & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Hartwell, P. (1985). Grammar, grammars, and the teaching of grammar. In *College English* 47(2), p. 105-127.
- Householder, F.W. (1981). Apollonius Dyscolus and the origins of deep structure. In *The syntax of Apollonius Dyscolus, translated and with commentary by F.W. Householder*. Amsterdam: Benjamins, p. 4-17.
- Jespersen, O. (1933). *Essentials of English grammar*. London: Allen & Unwin.
- Klein, W. (1985). Ellipse, Fokusgliederung und thematischer Stand. In *Ellipsen und fragmentarische Ausdrücke* (1)31. Meyer-Herrmann, R. & Rieser, H. (eds.). Tübingen: Niemeyer, p. 1-24.
- Kunz, K., Degaetano-Ortlieb, S., Menzel, K. et al. (forthcoming). GECCo – an empirically-based comparison of English-German cohesion. In *New ways of analysing translational behaviour in corpus-based translation studies. TILSM series*. De Sutter, G., Delaere, I. & Lefer, M.-A. (eds.). Mouton de Gruyter.
- Duden-Online: 'Ellipse'* Available at:  
<http://www.duden.de/node/664771/revisions/1616025/view>
- Duden-Online: 'danke'* Available at:  
<http://www.duden.de/node/852563/revisions/1339761/view>
- Lane, A. (1700). *A key to the art of letters*. Menston: Scolar Press [1969 reprint].
- Martin, J.R. & Rose, D. (2003). *Working with discourse: meaning beyond the clause*. London/New York: Continuum.
- Matlab documentation: 'boxplot'* Available at:  
<http://de.mathworks.com/help/stats/boxplot.html>
- Menzel, K. (submitted). *Understanding English-German contrasts – a corpus-based comparative analysis of ellipses as cohesive devices*. PhD dissertation, Saarland University, Saarbrücken.
- Menzel, K, Lapshinova-Koltunski, E. & Kunz, K (forthcoming). Cohesion and coherence in a multilingual context. In Menzel, K., Lapshinova-Koltunski E., & Kunz,

K. (eds.) *New perspectives on cohesion and coherence*, Berlin: Language Science Press, p. 1-11.

Menzel, K. (forthcoming). *Annotating cohesive ellipses in an English-German corpus*. In *Systemic functional linguistics in the digital age*. Gardner, S. & Alsop, S. (eds.). London: Equinox.

Menzel, K. (2014). Ellipsen als Stil- und Kohäsionsmittel in deutschen und englischen politischen Reden. In *Germanistische Mitteilungen. Zeitschrift für Deutsche Sprache, Literatur und Kultur, 40.1: Deutsch kontrastiv*, p. 31-50.

Menzel, K. (2014). Ellipsen als Stil- und Kohäsionsmittel in deutschen und englischen politischen Reden. In *Germanistische Mitteilungen. Zeitschrift für Deutsche Sprache, Literatur und Kultur, 40.1: Deutsch kontrastiv*, p. 31-50.

Merchant, J. (2010). Three kinds of ellipsis. In *Context-dependence, perspective, and relativity*. Recanati, F., Stojanovic, I. & Villanueva, N. (eds.). Berlin: de Gruyter, p. 141-192.

Müller, C. & Strube, M. (2006). Multi-level annotation of linguistic data with MMAX2. In *Corpus technology and language pedagogy. New resources, new tools, new methods*. Braun, S., Kohn, K. & Mukherjee, J. (eds.). Frankfurt: Peter Lang, p. 197-214.

*OED Online: 'ellipsis'*, Available at: <http://www.oed.com/view/Entry/60527>

*OED Online, Frequently asked questions, Why are there no recent illustrative quotations for many words in common use?* Available at:

<http://public.oed.com/about/frequently-asked-questions/#illustrate>

Ortner, H. (1987). *Die Ellipse. Ein Problem der Sprachtheorie und der Grammatikschreibung*. Tübingen: Niemeyer.

Provost, G. (1988). *Beyond style: Mastering the finer points of writing*. Cincinnati, Ohio: Writer's Digest Books.

Quirk, R., Greenbaum, S., Svartvik, J. et al. (1985). *A comprehensive grammar of the English language*. London: Longman.



- Rickford, J.R. (1998). *The creole origins of African-American vernacular English: Evidence from copula absence*. In *African-American English: Structure, history, use*. Mufwene, S., Rickford, J.R., Bailey, G. et al. (eds.). London: Routledge, p. 154-200.
- Ries, J. (1931). *Beiträge zur Grundlegung der Syntax. Teil 3: Was ist ein Satz?* Prag: Taussig & Taussig.
- Smith, R.D. (1986). *A syntactic quicksand: ellipsis in seventeenth and eighteenth century English grammars*, Ph.D. Dissertation, Leiden, Delft: Eburon.
- Smith, R.D. (1998). Eighteenth-century linguistics and authorship: The cases of Dyche, Priestley, and Buchanan. In *Advances in English historical linguistics*. Fisiak, J. & Krygier, M. (eds.). Berlin/New York: de Gruyter, p. 435-441.
- Stainton, R. (2006). *Words and thoughts: Subsentences, ellipsis, and the philosophy of language*. Oxford: Oxford University Press.
- Stang, C. & Steinhauer, A. (2014). *Duden Ratgeber – Handbuch Zeichensetzung: Der praktische Ratgeber zu Komma, Punkt und allen anderen Satzzeichen*, Berlin: Dudenverlag,
- Sundby, B., Bjørge A.K., & Haugland, K.E. (1991). *A dictionary of English normative grammar 1700-1800*. Amsterdam/Philadelphia: John Benjamins.
- Winkler, S. (2005). *Ellipsis and focus in generative grammar*. Berlin/New York: de Gruyter.

<p style="text-align: center;"><b>Contact data</b></p> <p>Katrin Menzel, lecturer and researcher at the Department of Applied Linguistics, Translation and Interpreting at Saarland University, Campus, 66123 Saarbrücken, Germany, k.menzel@mx.uni-saarland.de</p>		<p style="text-align: center;"><b>Fields of interest</b></p> <p>Corpus linguistics, contrastive syntax and morphology, discourse analysis, translation studies, language variation, historical linguistics</p>
---	---	--

## **Résumé in English**

This paper reports a part of the larger corpus-based study, which investigates English-German contrasts in text cohesion and discourse organisation. It has its focus on ellipsis-antecedent-relations that contribute to the cohesiveness of texts. In the case of ellipsis, in contrast to other cohesive devices, a textual relation is set up between a textual antecedent and a subsequent omission and not between two elements of the text's surface structure. In this study, we show that existing vague or excessively broad definitions for ellipsis as a linguistic concept lead to difficulties for empirical studies. We developed a corpus annotation scheme that makes it possible to specify and define the traditional ellipsis concept in operationalisable terms for a cross-linguistic quantitative study. One point that has not sufficiently been discussed yet in prior publications on the role of ellipsis in textual cohesion is the history of ellipsis as a grammatical concept and how this has influenced current ellipsis taxonomies and definitions. This paper traces certain stages in the complex evolution of ellipsis as a grammatical concept. It addresses the emergence and historical development of ellipsis as a linguistic concept from classical antiquity to modern times and discusses challenges for empirical, corpus-based studies arising from existing ellipsis descriptions. The challenge with many linguistics terms is that they do not easily provide a basis for empirical research as they have evolved from philosophical and rhetorical concepts and still lack uniform and exact definitions. This paper will also briefly cover some details of the development of the annotation scheme we used and discuss some findings from a contrastive analysis of the English-German 'GECCo' corpus. Our initial assumption was that English uses more omissions and hence more cohesive ellipses than German. Surprisingly, the frequencies of cohesive ellipses in English and German turned out to be almost identical. Furthermore, our data suggest that the German data have a higher variability than the English data. The present work is intended to contribute to the development of a discourse-oriented contrastive grammar on the English-German language pair with relevance to theoretical and applied linguistics, translation studies and foreign language pedagogy.

**Key words:** text linguistics, cohesion, ellipses, sentence fragments, grammatical categorisation, history of linguistics, English-German contrasts, corpus linguistics, frequency analyses.

### **Résumé in German**

Dieser Artikel berichtet über ein korpuslinguistisches Forschungsprojekt, in welchem der Vergleich zwischen dem Englischen und Deutschen hinsichtlich Textkohäsion und Diskursorganisation im Mittelpunkt steht. Ziel dieser besonderen Untersuchung ist eine kontrastive Analyse von elliptischen Strukturen, die als potentielle Textverknüpfungsmittel mit bereits vorerwähnten Textelementen in Verbindung stehen. Es wird der Frage nachgegangen, wie sich das Englische und das Deutsche im Hinblick auf den Gebrauch von Ellipsen unterscheiden oder ähneln. Ellipsen unterscheiden sich von anderen Textverknüpfungsstrategien dadurch, dass hierbei nicht zwei Elemente an der Textoberfläche in Verbindung stehen, sondern ein Antezedens mit einer darauffolgenden Auslassung in einer syntaktisch unvollständigen Struktur verknüpft ist. Ein Aspekt, der in der Vergangenheit im Hinblick auf Ellipsen als Kohäsionsmittel noch nicht umfassend diskutiert wurde, ist die komplexe Geschichte des Ellipsenbegriffs an sich, die einen großen Einfluss auf aktuelle Ellipsenbeschreibungen und -taxonomien hatte. In diesem Artikel werden schwerpunktmäßig einige wichtige Aspekte der Grammatikgeschichte des Begriffs diskutiert, die sich bis in die Antike zurückverfolgen lassen. Frühe Grammatikographen haben Begrifflichkeiten geprägt, die teilweise aus philosophischen und rhetorischen Diskursen hervorgegangen sind, und Diskussionen angestoßen, die unsere heutigen Grammatikmodelle und unser Grammatikverständnis nicht unerheblich beeinflusst haben. Begriffe wie "Ellipse" sowie verwandte Konzepte wie "Wort" oder "Satz" gelten immer noch als relativ vage oder können sehr breit gefasst werden. Das macht sie zu schwierig zu beschreibenden Strukturen, die mit empirischen Methoden schwer erfassbar sind. Ein wesentliches Ergebnis dieser Arbeit

besteht darin, dass sie den Begriff der kohäsiven Ellipse in Abgrenzung von anderen Phänomenen präzisiert und operationalisierbar macht. In diesem Artikel gehen wir auch kurz auf das entwickelte Annotationsschema ein und erläutern einige Ergebnisse des Vergleichs der englischen und deutschen Daten aus dem GECCo-Korpus. Ursprünglich erwarteten wir, im Englischen mehr kohäsive Ellipsen zu finden, was sich jedoch nicht bestätigte. In den deutschen Daten ließ sich allerdings eine größere Varianz beobachten. Mit dieser Untersuchung soll eine diskursorientierte vergleichende Grammatik der englischen und deutschen Sprache angeregt werden, welche sowohl für die Sprachforschung von Relevanz ist, aber auch für den Fremdsprachenunterricht und die Ausbildung von Übersetzern.

**Stichwörter:** Textlinguistik, Kohäsion, Ellipsen, Satzfragmente, grammatische Kategorisierung, Geschichte der Sprachwissenschaft, Sprachvergleich Englisch-Deutsch, Korpuslinguistik, Häufigkeitsanalysen.

### **Résumé in French (translation by Olivier Landeville)**

Le présent article fait état d'un projet de recherche basé sur la linguistique de corpus et axé sur la comparaison entre l'anglais et l'allemand concernant la cohésion textuelle et l'organisation du discours. L'objectif de cette étude spécifique est de procéder à une analyse contrastive de structures elliptiques qui, en tant que liens textuels potentiels, sont en relation avec des éléments textuels déjà précités. La question est de savoir de quelle manière l'anglais et l'allemand se différencient ou se ressemblent au regard de l'emploi d'ellipses. Les ellipses se différencient d'autres liens textuels par le fait que la relation textuelle s'établit entre une antécédence et une omission subséquente au sein d'une structure syntactique incomplète et non pas entre deux éléments qui sont en relation à la surface du texte. L'évolution complexe du terme "Ellipse", qui a une grande influence sur les descriptions et taxonomies elliptiques actuelles, est un aspect qui, jusqu'à présent, n'a pas été abordé de manière approfondie au regard des ellipses en tant que moyens de cohésion. Le présent article traite principalement de certains

aspects importants de l'évolution grammaticale du terme qui remonte à l'Antiquité. D'anciens grammaticographes ont marqué de leur empreinte des termes en partie issus de discours philosophiques et rhétoriques et ont suscité des débats qui ont eu une influence non négligeable sur nos modèles grammaticaux actuels et notre compréhension grammaticale. Des termes tels que "Ellipse" et des concepts apparentés tels que "Mot" ou "Phrase" sont encore considérés comme étant relativement vagues ou peuvent avoir un sens très large. Cela fait d'eux des structures difficiles à décrire et à comprendre avec des méthodes empiriques. L'un des principaux résultats de cette étude réside dans le fait qu'il apporte des précisions sur le terme d'ellipse cohésive et le rend opérationnalisable. Par ailleurs, le présent article aborde brièvement le schéma d'annotation qui a été développé et fournit des explications sur certains résultats obtenus par la comparaison des données anglaises et allemandes extraites du corpus GECCo. Au départ, nous nous attendions à trouver en anglais davantage d'ellipses cohésives mais ceci n'a pas été confirmé. Les données allemandes ont toutefois présenté une plus grande variance. La présente étude a pour objectif de contribuer au développement d'une grammaire axée sur le discours et comparée de l'anglais et de l'allemand, celle-ci étant pertinente aussi bien pour la recherche linguistique que pour l'enseignement des langues étrangères et la formation des traducteurs.

**Mots-clés:** linguistique textuelle, cohésion, ellipses, fragments de phrases, catégorisation grammaticale, histoire de la linguistique, linguistique contrastive anglais-allemand, linguistique de corpus, analyse de fréquences.

### **Résumé in Russian (translation by Ekaterina Lapshinova-Koltunski)**

В статье представлена часть исследования, проводимого в рамках корпусного контрастивного анализа текстовой когезии и организации дискурса в немецком и английском языках. Главной целью работы является изучение отношений между эллиптическими конструкциями и их антецедентами, влияющими на когезивность текста. В отличие от других инструментов когезии, явление

эллипсиса не включает отношений между двумя структурами на поверхности текста. Мы показываем, как строится отношение между antecedентом и последующей эллиптической конструкцией. Следует отметить, что существующие определения эллипсиса являются слишком нечёткими, что затрудняет эмпирический лингвистический анализ. Мы разработали схему разметки эллипсиса, что даёт возможность уточнить традиционные определения эллипсиса в терминах операционализаций для сопоставительного количественного корпусного анализа двух языков. В предыдущих публикациях еще недостаточно обсуждена роль эллипсиса в текстовой когезии в контексте истории эллипсиса как грамматического концепта и его влияние на современные таксономии и определения. Мы анализируем возникновение эллипсиса и его историческое развитие как грамматического концепта от классической античности до наших дней; это дает возможность обсудить сложные, но интересные эмпирические исследования, основанные на корпусной лингвистике, связанные с существующими определения эллипсиса. Одной из проблем развития эллипсиса как лингвистического понятия является отсутствие единого и чёткого определения, т.к. используемая терминология была заимствована из риторики и философии. В работе также вкратце описываются разработка схемы разметки эллипсиса и результаты контрастивного анализа на основе англо-немецкого корпуса GECCo. Нашей изначальной гипотезой являлось преобладание когезивных эллиптических конструкций в английских текстах в отличие от немецких. Однако оказалось, что их количество в английских и немецких текстах практически одинаково. Кроме того, наши данные показывают, что в немецких текстах используются более разнообразные конструкции, чем в английских текстах. Представленная работа является вкладом в развитие дискурсивной контрастивной грамматики английского и немецкого языков, предназначенной для теоретической и прикладной лингвистики, теории перевода и преподавания иностранных языков.

**Ключевые слова:** лингвистика текста, когезия, эллипсис, фрагменты предложений, грамматическая категоризация, история лингвистики, сопоставление английского и немецкого языков, корпусная лингвистика, частотный анализ.

Article was received by the editorial board 28.05.16;

reviewed 11.06.16 and 29.06.16.

Similarity Index 5%.