

CHRISTINA SANCHEZ-STOCKHAMMER

What Counts in Linguistics: Quantity vs. Quality

1. Introduction

Asked to define the meaning of *linguistics*, a large number of linguists will presumably qualify their discipline as "the scientific study of language," thereby following Crystal's (1985: 9) definition. This can be explained by the relatively strong focus on quantitative empirical research which has shaped most of the discipline in the past few decades (cf. e.g. Meyer 2009b, 212; Gries 2015, 725). However, this has not always been the case – as is reflected in Leisi's (1981) distinction between what he calls "traditional" and "modern" linguistics, with the former using qualitative philological methods and understanding itself as part of the humanities, whereas the latter is regarded as closer to the natural sciences, with a preference for quantification. Crystal (1985, 143) goes even so far as to postulate a great "psychological gap between linguistics and philology" and assumes that members of one of the two disciplines "would get very emotional" if they were mistakenly referred to as members of the other. This split characterisation of the discipline of linguistics can be explained by the fact that linguistics does not fit neatly into the distinction between the natural sciences and the humanities: while e.g. the physiological organisms which form the basis for language production can be studied with scientific methods, the communicative function of language concerns social aspects (Bartsch and Vennemann 1973, 15-16), which may profitably be analysed using other types of approach. What most of the currently used approaches in linguistics have in common is the observation of real-world data in order to describe actual language use: since "the predominance of generative linguistics, with its stance that experimental or observational evidence for one's judgments is not really required, has waned" (Gries 2015, 725), most current research in linguistics is based on observable empirical evidence (Meyer 2009b, 212), no matter whether it is of the quantitative or the qualitative type.¹ This suggests at least a certain degree of closeness in line of thought. What distinguishes the two approaches, therefore, are their aims on the next level of specificity as well as their methods.

The discussion of the dichotomy QUANTITATIVE vs. QUALITATIVE has a long tradition and can be traced back to Aristotle (Neuenschwander 2003, 3-4). Interestingly, there is an asymmetry in the associations that the two dichotomous concepts evoke: while the adjective *quantitative* neutrally expresses "relatedness to the measurement of amount, size etc.," which is countered by one of the meanings of the word *qualitative* (as a neutral way of designating the possession of a particular feature), the adjective *qualitative* also has another, markedly positive meaning, namely "excellence of standard or level" (www.oxforddictionaries.com). This explains why *quality* is usually associated with the ideas 'warm,' 'humane,' 'holistic' and 'concrete,' whereas *quantity* or *quantification* rather evoke 'cold,' 'technocratic,' 'isolating,' 'reductionist' and

1 While some linguists seem to consider mentalistic generative approaches as empirical, since they also attempt to understand reality, their sole reliance on intuition rather than language use in order to describe the language of ideal rather than real speakers makes them unempirical in the opposing view (cf. the discussions in Ballmer 1976, 43, 50; Crystal 1985, 103-107; Meyer 2009b, 210-212), which is also adopted here.

'abstracting' and therefore receive more negative evaluations (Neuenschwander 2003, 1). As a consequence, most people will presumably tend to agree with the out-of-context statement that quality is more important than quantity, even if they tend to work within quantitative research paradigms.

To provide a brief definition of the dichotomous concepts, quantitative research "involves data collection procedures that result primarily in numerical data which is then analysed primarily by statistical methods," whereas qualitative research "involves data collection procedures that result primarily in open-ended, non-numerical data which is then analysed primarily by non-statistical methods" (Dörnyei 2007, 24). The difference involves "the general ideological orientation underlying the study, the method of data collection applied, the nature of the collected data, and the method of data analysis used to process the data and to obtain results" (Dörnyei 2007, 24). In contrast to quantitative research, which proceeds in a linear fashion, qualitative analyses are iterative and "move back and forth between data collection, data analysis and data interpretation depending on the emergent results" (Dörnyei 2007, 243). Quantitative research mainly uses e.g. corpora, experiments, standardised tests or questionnaires in order to arrive at optimally objective quantifiable data that permit intersubjective verification (Gonzalez-Marquez, Becker and Cutting 2007, 59-60), whereas qualitative research prefers case studies, qualitative interviews, group discussions etc. in its attempt to understand the interlocutor holistically (Neuenschwander 2003, 11-12). Table 1 (which draws particularly on Dörnyei's 2007 detailed discussion of quantitative and qualitative research methods) sums up the typical differences between the two research paradigms:

	Quantitative research	Qualitative research
PERSPECTIVE	higher-level macro-perspective	context-sensitive micro-perspective
GOAL	test hypotheses in the search for generalisable trends	understand and describe individual reality in depth
RESEARCH DESIGN	linear; clear structure; ends after the analysis of a predetermined set of data	iterative; structure not entirely predetermined; ends when the point of saturation has been reached and no new ideas are found
CATEGORISATION AND CODING	predetermined, relatively rigid numerical categories	emergent, flexible, verbal coding and descriptions
DATA ANALYSIS	statistical tests of significance	reliance on the researcher's individual sensitivity
RESEARCHER INDEPENDENCE	aims at objectivity	subjective
APPROACHES TO DIVERSITY	large and representative samples intend to level out idiosyncrasies	focus on the unique individual

Table 1: Typical features of quantitative vs. qualitative research

As pointed out above, there is a general tendency to regard present-day linguistics as a purely quantitative discipline that strives to achieve objective results through the use of scientific methods and statistical analyses. However, one should not overlook the fact that language is a type of social practice, which differs inherently from the prototypical subjects of natural scientific research (cf. also Neuenschwander 2003). Since language use cannot be described deterministically but only probabilistically, it

makes sense to consider to what extent quantitative research is suited to the study of language at all and to discuss the relation between quantitative and qualitative approaches and their potential contribution to linguistic research. In spite of a number of inherent problems of quantitative linguistic research (discussed in Section 2), Section 3 makes a strong point that quantitative research is necessary for the study of language. Section 4 will argue for the place of qualitative research in linguistics and show that the two approaches are not mutually exclusive but complement each other well.

2. Inherent Problems of Quantitative Linguistic Research

Quantitative empirical methods are such a common and established part of present-day (non-generative) linguistic research that they are rarely ever called into question on a general level. There are, however, a number of slightly older texts (most of them in German) which point out the dangers of counting and which plead for more qualitative research in linguistics, namely Lüthy (1970), Leisi (1980), Leisi (1981) and Rissanen (1989). The following sections extend the arguments from these texts and adapt them to the second decade of the 21st century. While various types of example will be used, the focus will lie on the field of corpus linguistics as a currently particularly common form of quantitative empirical research.

2.1 Collecting Data and Crunching Numbers For Their Own Sake

One possible criticism on a very general level is that researchers carrying out quantitative descriptive research may collect large amounts of data, analyse the results descriptively and crunch numbers for their own sake. This also implies that some trivial facts may receive more focus than would seem appropriate (cf. Lüthy 1970). On the other hand, however, qualitative research is not immune against this problem either.

The remedy in this case is to follow the advice that is typically transmitted from more experienced to novice researchers, namely to devise research questions and hypotheses before engaging in the collection and analysis of linguistic data. It is only by having expectations about what will be found in the data that quantitative research can tell us whether our assumptions are supported by the facts or not. After all, as Crystal (1985, 110) puts it, "[c]ontrary to what many people believe, the basic aim of science is not to collect facts about the world," but to establish theories which aim at understanding and explaining reality or data (Crystal 1985, 110-111) – and in the case of linguistics, that corresponds to language systems and use. This typical scientific approach may be complemented by more exploratory quantitative bottom-up techniques, "in which the data [e.g. from corpora] are processed statistically in order to discover structures that had not necessarily been anticipated by the analyst" (Hilpert and Gries 2016, 44). However, the testing of large numbers of possible independent variables is not entirely unproblematic, since some may be significant purely by chance (Young and Karr 2011, 118), and the researchers' wish to find out something new may encourage "devising a plausible story to fit the statistical finding" (Young and Karr 2011, 119). Since data mining in the search for statistical correlations may result in nonsensical spurious correlations (such as the link between divorce rates and margarine consumption)² the interpretative effort of the analysts is of particular im-

2 Cf. <http://www.tylervigen.com/spurious-correlations>. 10 January 2017.

portance in this alternative type of empirical approach: the numbers take the lead initially, but they are then filtered by the researcher, who usually has some overarching research questions and expectations in mind.

2.2 Uncritical Acceptance as True

Another essential danger of quantitative research is that its results are often uncritically accepted as true (Leisi 1980, 1981). Leisi (1981, 387) postulates that there is a general human tendency to accept numerical or formalised descriptions as objectively more correct than statements in natural language.³ A reason which he advances himself (Leisi 1980, 202) is that numbers give certainty due to their precision and reliability. According to Leisi (1980, 207), this uncritical belief in the truth of numbers has been reinforced by the introduction of computers. However, in view of the other problems discussed below, the numerical accuracy may only be superficial and thus "in-still a false sense of security in the accuracy of the results that are ultimately obtained" (Meyer 2012, 24; cf. also Leisi 1980, 202). This critical view is supported by Young and Karr (2011, 116-117), who find that many observational studies fail replication testing.

While this is certainly an important criticism of quantitative research, there is a remedy: by reminding themselves of this aspect, and by raising awareness in their students and the audience of their studies, researchers in linguistics can contribute to a more detached and critical perception of quantitative research.

2.3 Uncritical Acceptance of the Database

Yet another problem derived from the uncritical acceptance of quantification is the uncritical acceptance of the database as representative (cf. the discussion in Lange, this volume, regarding GloWbE). However, any corpus – even an authoritative one, whose design is supposed to make it balanced and thereby representative – can reflect a complete language only imperfectly (cf. Rissanen 1989, 17; Crystal 1985, 101).⁴ While certain insights gained from a corpus consisting exclusively of newspaper texts might be generalisable to language as a whole, this is not necessarily the case, and the only safe generalisations would seem to be those to the language of newspaper texts in general. As a consequence of the limitations imposed by imperfect corpora, the same must apply to studies conducted on the basis of such imperfect corpora. This problem is particularly pervasive in the case of corpora which are used in different studies by researchers other than their compilers – a tradition that started with the 1959 *Survey of English Usage* (Meyer 2009a, 10-12). In this case, it is particularly important for researchers to know the make-up of the corpus, in order to make sure that it represents the variety of language that they actually want to analyse (e.g. with

3 Cf. Leisi about Zipf's (1965) laws: "Eine solche formalisierte Darstellung, mit einem Bruch, einer Wurzel und einer Konstanten, besonders wenn sie noch in einer Kurve, zum Beispiel auf Logarithmenpapier aufgezeichnet ist, macht den Eindruck höchster Exaktheit" (1980, 207). This increase of acceptance in charts or diagrams is also observed by Fletcher according to whom "[n]umbers in datasets can be hard to grasp, but show someone two lines moving up or down in apparent unison and you're halfway to convincing them that one causes the other" (2014).

4 By contrast, corpora for specific registers of a language (such as newspaper texts, cooking recipes etc.) are more likely to provide a representative sample of the respective type of text.

regard to the register, period, geographical origin and modality of the texts).⁵ Otherwise, the database might be representative of a particular variety – but not the one under consideration in the research. Furthermore, quantitative studies which are based on corpora might be slightly outdated relatively quickly as soon as larger or better corpora have been compiled, with the early-generation corpora getting out of use. Alternatively, older and younger databases may coexist within the same corpus collection, e.g. in the ICE project (<http://ice-corpora.net/ice/>). Rissanen points out yet another risk of "corpus work and computer-supported quantitative research methods," namely that they "will discourage the student from getting acquainted with original texts, from being on really intimate terms with his material" (1989, 16). In view of the increasingly large size of the latest corpora,⁶ few researchers will have the time to read the corpus texts in their entirety before conducting research. According to the *British National Corpus* website, for instance, "[r]eading the whole [100-million-word] corpus aloud at a fairly rapid 150 words a minute, eight hours a day, 365 days a year, would take just over four years" (<http://www.natcorp.ox.ac.uk/corpus/index.xml?ID=numbers>). However, even those researchers working with smaller corpora such as the one-million-word ICE corpora (<http://ice-corpora.net/ice/>) are usually satisfied with an overview of the genres contained therein and do not necessarily wish to read even the list with the titles of the corpus texts (which is already a kind of summary). While acquaintance with the corpus is a necessary prerequisite, trusting the corpus compilers is part of standing on the shoulders of giants (or, rather, a human pyramid).

The remedy is awareness-raising (once again), as well as the clear documentation of material used in multiple studies. Knowing the limitations of the material and the methods used in linguistic research is essential for the achievement of relevant results (cf. also Meyer 2012, 40; Hilpert and Gries 2016, 53). However, even if corpora may represent a whole language only imperfectly, impressionistic qualitative samples are unlikely to reach a higher level of representativeness. Therefore, even if corpora can only provide an approximation to the valid, objective and representative analysis of language, they are better than nothing – and the resources available to quantitative English linguistic research are constantly being improved, with the addition of corpora covering new registers (such as blogs, text messages etc.), the refinement of search tools, and the extension of databases.

2.4 Incorrect Measuring Due to Categorisation Problems

The most important criticism advanced by Leisi (1980; 1981) is derived from the uncritical acceptance of yet another aspect of quantification, namely the categories that are counted as its basis (see also Mindt, this volume). Quantitative linguistic research requires the classification of linguistic entities as belonging into predefined categories, so that the number of members for each category can be counted. For quantitative research, the number of items to be classified needs to be larger than the

5 For instance, by carrying out a qualitative analysis of the context in which his MICASE hits for *I want to/wanna* occurred, Meyer discovered that 26 (and thus 50%) "were uttered by a single graduate student who was holding office hours with five students in one rather lengthy sample that was 29,635 words in length" (2012, 37-38), and that the data was therefore not as representative as required for his analysis.

6 Such as the open-ended NOW Corpus containing over 2.8 billion words from online newspapers; cf. <http://corpus.byu.edu/now/>. 10 January 2017.

number of categories – otherwise, one would end up with as many categories as there are items (which suggests a more qualitative type of approach). The main problem lies in the correct assignment of linguistic test items to categories: for instance, different items may be erroneously put into the same category, or similar items may be erroneously put into different categories (cf. Leisi 1981). To give an example, counting the number of nouns in a text may seem straightforward at first sight, but it requires the decision whether to include items such as *John* or *London* in the category of nouns, whether to assign them to a subcategory PROPER NOUNS or whether to create a distinct category PROPER NAMES (cf. Crystal 1985, 80). The problem is increased where highly technical linguistic definitions coexist with vague everyday non-technical uses of certain terms, such as *word* or *sentence* (cf. Crystal 1985, 78). Since categorisation is usually not self-evident but rather displays different degrees of arbitrariness, even exact counting may lead to incorrect results if the categories are ambiguous or underspecified (cf. Leisi 1981, 387), and the problem resides precisely in the unawareness of this ambiguity. Leisi (1980, 207) therefore criticises that the formulaic rendition of Zipf's law is seemingly exact but relies on a very questionable basis for the counting, since the number of meanings of individual words cannot be established objectively. He therefore notes the danger of a discrepancy between the accuracy that is applied to the counting procedure and to the processing of numbers, in contrast to the precision with which the semantic basis of the counting is determined in quantitative research (Leisi 1980, 207). The problem lies in the tacit assumptions that are made by the authors and readers of quantifications, which may vary with different perspectives in mind (cf. also the context sense 6 in Möhlig-Falke, this volume).

The awareness of pre-existing categories is of particular importance whenever corpus research is based on a corpus that was not compiled by the analyst. Annotated corpora are always biased, since they impose an idea of structure on the texts (Meyer 2012, 29), and it is only through awareness of the annotation scheme that searches will retrieve the targeted items and not result in unwanted hits (cf. Meyer 2012, 27). This is of particular importance whenever results from two or more corpora are to be compared. Otherwise, the application of different coding schemes would lead to the comparison of results that are actually not comparable. For example, the MICASE Corpus transcribes *want to* and *wanna* as different linguistic entities (Meyer 2012, 36), whereas some other spoken corpora might group these under one transcription code and consider them mere pronunciation variants.

By extension, the issue of categorisation also applies to the format in which the texts are edited. For instance, orthographic variation (particularly in older texts) makes part of the targeted hits irretrievable, unless special programmes such as VARD 2 (<http://ucrel.lancs.ac.uk/ward/about/>) are used. For a study on the spelling of English compounds (Sanchez-Stockhammer, forthcoming), it was necessary to consider very precisely how end-of-line hyphenation from the original printed texts was coded in a set of corpora, since the hyphen present at the end of line within a compound such as *week-end* could either have been intended as a permanent hyphen by the author or happen to be a soft end-of-line hyphen used for better justification of the right margin. The corpora did vary in this respect (cf. the discussion in Sanchez-Stockhammer, forthcoming): the Brown Corpus compilers, for example, deleted hyphens at the ends of lines if they could only be line breaks (e.g. in *situa-tion*) but preserved them in words that were hyphenated elsewhere in the same sample text. For the remaining ambiguous 162 word types (including e.g. prefixations and neoclassical compounds),

decisions were based on the third edition of Webster's *New International Dictionary* (cf. Francis and Kucera 1979). The BE06 Corpus, by contrast, does not modify the formatting of the original texts in any way. Since it is based on the pdfs of texts that also appeared as print publications, this means that it may contain some unusual compound spellings which are the result of copying end-of-line hyphens from pdfs and which were presumably not intended as permanent hyphens by the authors of the texts (e.g. *funloving*).

Last but not least, categorisation problems may also occur on a more abstract level, namely that of statistical analyses: thus Baker and Egbert observe that the cut-offs for statistical significance, which have a very important influence on the findings of corpus studies, "are often imposed rather arbitrarily" (2016c, 206).

The remedy for all of the above is to aim at conscious categorisation (Leisi 1980, 203), to define categories explicitly and precisely, to apply the categories in a consistent manner (Crystal 1985, 78-79) and – where data is put to secondary use – to familiarise oneself with the underlying principles of categorisation.

2.5 Incorrect Measuring Due to Lack of Context

As pointed out above, researchers who carry out quantitative research, particularly with very large corpora, will usually not be able to read the entirety of the corpus texts, so that the automated corpus search "takes the analyst one step away from the texts being analysed" (Meyer 2012, 27). While qualitative research places particular focus on context and the individuality of data items, quantitative research by definition requires decontextualisation, at least to a certain extent: it is only by formulating abstract search patterns (e.g. "my_{ADJ} friend" in BNCweb) that more general information can be retrieved from a corpus (e.g. *dear, good, honourable, new* or *old* as possible adjectives in the expression *my_{friend}*). Counting is only possible by considering data items from a more abstract perspective that stresses similarities, e.g. by treating all instances of *my best friend* as the same type and thus regardless of whether they occur as subjects, objects or complements (if the search question does not require such a distinction).

However, one may argue that modern corpus analysis tools (such as AntConc) make up for the fact that researchers have not read the whole text material in that they usually allow their users to read some context to the left and right of their search word in the keyword-in-context (KWIC) lists. This user-friendly quick overview reduces the context to a few words (usually about five to ten) on each side of the target, so as to permit the display of each context on just one line of text, but it is often possible to expand the context by clicking on the file code (e.g. in Mark Davies' TIME Corpus; cf. <http://corpus.byu.edu/time/>). While the reduced context does not compare to a qualitative full-text study, the user-friendly format has the advantage of permitting to spot generalisations and inconsistencies at one glance and thus far more efficiently than would be possible by a manual analysis of the texts.

Even if purely quantitative studies carry the inherent danger of disconsidering context, one may assume that most quantitative researchers are aware of the importance of context for their results and either consider at least the reduced context of a representative number of corpus hits or – where that is not possible for some reason – point this out as a limitation of their study with direct implications for the generalisability of the results. This was, for instance, the case in a study that automatically computed the spelling of compounds in attributive position before nouns (in which

prescriptive accounts would expect hyphenation, e.g. *a well-known author*) and in predicative position after verbs (in which one would rather expect open spelling, e.g. *the author is well known*;⁷ cf. Sanchez-Stockhammer, forthcoming). Among the items incorrectly retrieved as attributive hits of the nouns *drive-in* and *by-product* in BNCwritten were *I am uncertain as to whether I am capable of laying a **drive in concrete** and after standardizing **by product** sector*.

The importance of actually reading the corpus hits with context becomes clear when comparing the use of punctuation marks in different registers to establish their value as potential register indicators (Sanchez-Stockhammer 2016). Contrary to expectations, a similar amount of commas occurred in comics (whose spatial restriction favours relatively short sentences) and academic texts (in which long and complex sentences were expected to contain far more commas), but a detailed text-based analysis revealed that commas in comics are mainly used after sentence-initial proper nouns with vocative function (*Toyman, you maniac!*) or after introductory interjections (*Man, would you look at THAT!*). Since these functions are unusual in academic writing, the refined data does lend support to the general hypothesis that punctuation is employed differently in distinct types of text – a result that a mere quantitative approach would have overlooked. The consideration of context thus permits to improve quantitative research by refining the search patterns.

2.6 Lack of Larger Perspective

According to Kanningeßer (1976, 107), academic disciplines have two aims, namely to describe and explain their respective subjects. While quantitative research may tend to focus on the aspect of description, the more summary explanations may rather be produced by qualitative research. Quantification necessarily requires operationalisation as a process whereby complex phenomena are broken down into smaller units and processes, e.g. when the analysis of personality is reduced to the combination of social roles and social positions (Lüthy 1970, 29-30). However, essential aspects may be lost if phenomena are split up into tiny quantifiable subphenomena (Neuenschwander 2003, 3). Sridhar therefore argues that in a unified, enlightening approach, languages would "not simply be used as sources of data (isolated sentences, decontextualized) for testing theoretical notions" but rather "be studied in all their complexity and uniqueness" (1993, 12). This is important, because the "compartmentalization" (Sridhar 1993, 12) of research has led to "real gains in knowledge but a real loss in understanding" (Sacks 1987, 40-41). Leisi (1981, 386) utters the related criticism that while there are many quantitative studies covering restricted aspects of linguistics in a detailed manner, there is a lack in generalising research that gives an evaluative overview of the results achieved so far. However, such meta-studies (which would not merely serve as guidance to novices in the field but contribute to the shaping of the whole discipline) are relatively subjective. Their contradiction of the objectivity principle advocated by quantitative research might therefore explain their relative scarcity in linguistics.

To conclude this point, the remedy is to keep the larger picture in mind even in the case of quantitative studies with very limited scope and to complement such research

7 Note that one may also argue that this latter example ends with a phrase rather than a compound (cf. Sanchez-Stockhammer, forthcoming (a)).

with summarising research of a more qualitative nature that provides a weighted overview.

2.8 Probability-Related Problems

The analysis of quantitative data requires the use of statistics, which makes probabilistic statements about the likelihood with which the results of a study are significant or merely the result of chance (cf. Gonzalez-Marquez, Becker and Cutting 2007, 60). As a consequence, statistical testing cannot really confirm a hypothesis (even if one frequently finds shorthand formulations of this kind) but only state that the data do not suggest that the hypothesis is incorrect. In the humanities, the results are hardly ever deterministic in the sense that one particular type of event necessarily leads to another type of event. The reason for this non-determinism is that language is a sub-type of human behaviour – and human behaviour escapes all attempts at deterministic prediction, as the objects of studies in the humanities are intricately linked to the subjective histories and the consciousness of human agents (Lüthy 1970, 37). To some researchers, this would seem to suggest that quantification by analogy to the natural sciences (which prototypically describe natural laws with permanent validity) does not make any sense for linguistics. However, even the natural sciences are sometimes non-deterministic (e.g. when it comes to the probabilistic location of atoms within atomic orbitals; cf. Vollhardt and Schore 2011, 25), and conversely, a number of linguistic studies provide relatively accurate probabilistic predictions of linguistic phenomena such as stress assignment in noun-noun compounds (Plag, Kunter and Lappe 2007), the use of *s*-genitive vs. *of*-genitive (Szmrecsanyi and Hinrichs 2008) or the most likely spelling of English compounds (Sanchez-Stockhammer, forthcoming).

Since the discussion of language on a more general level requires the combination of many aspects of linguistic and non-linguistic behaviour, about which we have only approximate knowledge, the combination of these slightly uncertain insights within larger models will blow up the margin of error, so that we may end up with a margin of error that is larger than the exact content (Lüthy 1970, 20). Furthermore, the fact that statistics needs to be used in quantitative research brings with itself yet another probability-related problem, namely that probabilistic results are actually unempirical, because probabilistic laws cannot be used to predict a particular state of affairs in one specific instance (Lass 1980, 19-21). Neither the fact that they do apply nor that they do not apply in one specific case can be used as evidence for or against a particular theory (cf. *ibid.*). Merely their repeated success or failure in predicting states of affairs can serve to validate or invalidate a hypothesis, but where to draw the numerical boundaries is also subjective. One-percent and five-percent thresholds are commonly used, but this is just a matter of convention – an arbitrary cut-off point which actually has important repercussions on what is considered statistically significant. One counterexample may disprove a scientific theory in the natural sciences, but in linguistics, a specific instance of language use may also be considered a performance error, and where to draw the boundaries between unusual or innovative uses and errors is unclear.

Nevertheless, in spite of all these issues, the use of statistical methods also has very important advantages, because it provides guidance as to which results are more important and generalisable than others.

3. Why Quantitative Research is Necessary

In spite of the challenges pointed out above, the following sections will argue that quantitative research is practically unavoidable in the present-day study of language (and English linguistics in particular). A very obvious type of advantage, particularly in corpus linguistics, is that the use of automated processes in the analyses should prevent any biases (e.g. social or political prejudices) on the part of the researcher that might otherwise skew the results (Baker and Egbert 2016b). Furthermore, computerised procedures consider all cases in a sample and are not affected by fatigue or occasional lack of attention (Baker and Egbert 2016b) – but of course quantitative research should not be equated with automated computerised processes in all instances. Automated computerised processes have resulted in ground-breaking linguistic research results, such as the Longman Grammar (Biber et al. 1999), or Biber's very accurate descriptions of registers (e.g. Biber 1988).

3.1 Illustrations are not Evidence

One of the main drawbacks of qualitative research is that one cannot be certain to what extent its results are based on a representative database. To put it in the words of Weinreich (1955, 539), "[i]llustrations are not evidence." The reason for this is that instances of unusual language use are more salient and therefore more likely to be noticed by human analysts, so that their hand-picked examples may overemphasise specific types of construction, particularly those supporting the analyst's view. Leisi (1955), for instance, gives a relatively large number of examples for the dissociation of the English vocabulary compared to the consociation of the German vocabulary, such as the morpho-semantically motivatable German *Fluss|pferd* compared to its opaque English translation equivalent *hippopotamus*. However, a contrastive empirical analysis of the 2,500 most frequent words of both languages (Sanchez 2008) reveals that English is actually no dissociated language and that it may therefore make sense to test even assumptions that seem obvious.⁸ Another inherent danger of qualitative research is that rare features in small samples – particularly if these are not truly representative – risk being overemphasised (Baker and Egbert 2016c).

In view of the technological progress that has been made in computing in the last years and decades, e.g. with regard to processing capacity and functionality of the hardware and software, the size of corpora has been ever growing – so that the textual basis "cannot be studied by mere eyeballing" (Gries 2015, 275) anymore in the majority of cases. One consequence of the availability of large quantities of data is a growing expectation that researchers who are carrying out innovative studies will also make use of the available data, which simply do not lend themselves equally well to qualitative research, and the sole reliance on illustrative examples will seem less convincing to a readership with that expectation.

Viewed from the opposite perspective, one could paraphrase the idea that illustrations are not evidence more positively by arguing that applicability requires generalisability: impressionistic evidence only permits qualitative statements of the type 'the following type of construction'⁹ can be observed in the English language.' While such

8 In some cases, however, the tested obvious assumptions simply turn out to be correct. For instance, an analysis of the linguistic complexity of the register of comics/cartoons found the expected simplicity in this type of text (Sanchez-Stockhammer 2012a).

9 In the sense of a Goldbergian form-meaning pairing (cf. Goldberg 1996, 68).

information is certainly of interest, e.g. to language learners whose native language does not permit such a construction, that specific target group will not be able to use the construction in an appropriate manner unless it is given some information on the construction's frequency in specific contexts. In qualitative studies, this function may be fulfilled by hedging, with more hedges (such as the use of modal verbs) implying lower frequency of use, but the frequency-based distributional probabilities in quantitative studies provide more accurate usage information. For instance, Sanchez-Stockhammer and Arendholz (2012) correlate the expressions that are used to refer to the times of day in English and German (e.g. *morning, afternoon, evening, night*) with the 24 hours of the day. The analysis of the DWDS Core Corpus yields hits for both *Morgen* and *Vormittag* in combination with practically all numbers up to twelve, which would suggest a synonymous use in a qualitative approach. However, the quantitative frequency distribution rather suggests a complementary use: while *Morgen* co-occurs with similar likelihood with the times from 4 to 9 a.m. (10% - 14%), *Vormittag* also occurs in combination with 9 a.m. but shows a clear preference for 10-11 a.m., which corresponds to its literal semantic reading 'before noon' (Sanchez-Stockhammer and Arendholz 2012, 145-146). The more accurate results of the quantitative analysis thus overthrow the incorrect impression which would have been gained from a mere qualitative treatment of the data.

3.2 Small But Important Correlations are Easily Overlooked

According to Newmeyer, "[s]cience, by definition, is the search for order in nature" (1983, 41). The same can presumably be said of much empirical linguistic research, which is preoccupied with the search for patterns in language use. While it is certainly possible to identify and discuss such patterns on a qualitative basis alone, quantitative and particularly computerised quantitative research have several advantages in this respect. For instance, there may be relatively unobvious but still significant (and therefore important) correlations between variables, such as the use of a particular construction and the gender of the language user that only become obvious when carrying out a statistical analysis. This is particularly important where "two contextual features interact in such a way that, for instance, they only have an effect if they co-occur, but not if they occur in isolation" (Hilpert and Gries 2016, 39). Hilpert and Gries conclude that "[o]bservations of these kinds are difficult, if not impossible, to make on the basis of individual examples" and that "quantitative corpus analysis thus works like a magnifying glass, allowing the researcher to detect phenomena that would not otherwise be open to inspection" (ibid.).

3.3 Quantification Forms the Basis for Expectations

Moreover, the different frequencies of contrasting phenomena, variables etc. in quantitative research are an aid in the interpretation of the relative meaningfulness of data, since weighted importance can be assigned on the basis of the numerical values. The observation of frequency distribution for representative sets of data is then used as the basis for the extrapolation to other, similar contexts and samples. Quantification is thus the necessary basis for linguistic expectations and predictions on a general level. While some of these are more speculative than others (cf. the discussion of the question whether it is possible to predict linguistic change in Sanchez-Stockhammer 2015), the (sub)conscious formulation of expectations about what will happen more

or less soon (e.g. when trying to catch a ball) is actually part of regular everyday human behaviour (cf. e.g. Gigerenzer 2007). In the domain of language, all users sooner or later form expectations in the form of pre-statistical but still somehow quantitative intuitions, e.g. about what kind of language is appropriate in a specific situation.

In linguistics, only quantitative research permits the testing of hypotheses in the strictest sense. As a consequence, this type of research is necessary if researchers want to check or refine their models by testing whether the distributional patterns observed in a particular sample also hold elsewhere. For instance, Sanchez-Stockhammer (forthcoming) derives an algorithm for the spelling of English compounds from all variables that are statistically significant for a selected sample of compounds and then successfully tests the predictive accuracy of that spelling algorithm for other lists of compounds which serve as test samples.

3.4 Frequency is Involved in Mental Processing of Language

The language users' conscious or subconscious formation of linguistic expectations builds a bridge to yet another very important argument in favour of quantitative research, and that is the widespread view that frequency (and thus quantification) plays an important role in language processing. According to Hasher and Zacks (1984), humans automatically store information on event frequency in their minds. By analogy to the digital collections of text analysed in linguistic research, present-day linguistics commonly compares the human mind to a corpus, which supposedly contains the language that each person has encountered so far (cf. Taylor 2005). The stronger representation in memory explains why frequent linguistic patterns are evaluated as more acceptable than less frequent patterns (Bybee 2003, 13) and why frequent irregular forms (such as *keep/kept*) are likelier to persist than less frequent irregular forms (e.g. *weep/wept*), which more readily undergo regularisation (in the example, to *weeped*; cf. Bybee 2006, 715). Frequency data can also explain why certain constructions are avoided due to their low likelihood of occurrence (Stefanowitsch 2006). Support for the importance of frequency of co-occurrence also comes from evidence on the physiological level of the brain, as neurons that are repeatedly activated at the same time develop stronger synaptic links between each other (Pulvermüller 2002, 19).

While linguistic intuition is also based on the frequency-based mental corpus, it may be skewed by the researchers' own views and the exposition (that may be more or less frequent) to their own constructed examples, thereby raising their level of acceptance. As a consequence, we may conclude that if we want to find out how language works in the mind, we need to take frequency information into account – and that means carrying out quantitative linguistic research.

3.5 Modern Linguistics is Evaluated by the Criteria of Natural Sciences

All of the arguments discussed above have in common that they emerge from the requirements of linguistic research as such. This line of argument is complemented by an external reason for quantitative research in linguistics, namely the present-day tendency to evaluate linguistic research by the criteria of the natural sciences: Leisi (1980, 201) already discusses the ambiguity of the expression *the academic subjects that count*, which may either be paraphrased as 'the important academic subjects' or 'the academic subjects which deal with numbers.' The importance conferred on numbers in present-day academia and society may also partly be due to the now omni-

present use of computers, since these are particularly good at processing numbers and can easily calculate relations between entities (Lüthy 1970, 17; Leisi 1980, 202). Moreover, the high esteem demonstrated for academic subjects which quantify their objects of study may be due to the higher degree of precision that is reached when results are expressed in numbers (Leisi 1981, 387), since more precise statements lend themselves better to testing and thus validation or definitive refutation (Popper 1979, 356), thereby making academic progress easier and faster. Lüthy (1970, 18) notes that the requirement of quantitative research to formulate hypotheses and results accurately can have positive effects on the humanities with their tradition of discursive, less formalised accounts. While traditional linguistics did simply not make certain claims regarding precision (Leisi 1981, 386), many researchers in present-day academia seem to be of the opinion that quantification is a prerequisite for academic scholarly analyses (Neuenschwander 2003, 1-2), so that traditional types of research may appear outdated and unscientific to them (Leisi 1981, 379) – cf. also Rutherford's frequently quoted scientific credo that "[q]ualitative is nothing but poor quantitative" (Neuenschwander 2003, 19).

4. Why Qualitative Research has a Place as well

In spite of all of the above, it will be argued in the following that qualitative research has its place in present-day linguistic research, and even an important one.

4.1 Alternative Quantitative Analysis may not be Possible or less Suitable

For instance, there are some types of research question that lend themselves to qualitative rather than quantitative types of research. Since quantitative research attempts to control and standardise the conditions under which language is used in the research settings, there are certain questions in linguistics that cannot be answered that easily from a quantitative perspective (cf. also Kohnen, this volume). For example, context-specific and social phenomena are more easily observed when language is studied in natural settings (Phakiti and Paltridge 2015, 13). Some research questions simply do not require counting – e.g. Rundell's (1998) account of trends in learner lexicography or the exploration of what may be considered instances of hybridity on various levels of language (Sanchez-Stockhammer 2012b). This is not to say that quantitative studies are not possible in these fields, but these would constitute the next step (cf. below).

There are even empirical research questions that cannot be approached from a quantitative perspective but merely with qualitative descriptive approaches: thus the authentic citations that are traditionally used in lexicography do not permit generalising statistical analyses (Herbst 2010, 33). As a consequence, these short examples (many of which are drawn from literary sources) form a corpus that can be used to answer qualitative questions, e.g. about the earliest mention or the orthographic changes of individual words. In grammar writing, such pre-electronic corpora were also traditionally used to provide examples (cf. Meyer 2012, 25).

In other cases, quantitative analyses are made impossible by the lack of sufficient data: thus the cross-medial comparison of the dialogues in *Tintin* comicbooks and their filmic adaptation by Steven Spielberg (Sanchez-Stockhammer, forthcoming (b)) had to be limited to a qualitative discussion, since the film script underwent such important modifications that only a very small number of dialogues was still similar

enough to the comicbook to permit a comparison. If quantitative research is to apply statistical testing, then sample size and the proportion of large-enough expected counts also need to surpass critical thresholds (cf. Moore and Notz 2006, 496).

4.2 Quantitative and Qualitative Linguistic Research are not Mutually Exclusive

As pointed out above, the scope of linguistics is extremely wide, ranging from system-related aspects and issues in the teaching of foreign languages to historical developments and the processing of language in the human mind. Since the variety of research questions that may be asked in those different fields necessarily calls for a multitude of methods, qualitative research is obviously legitimised as one of the many conceivable ways of finding out about language, in spite of a predominance of quantitative approaches in present-day linguistics. While some of the points discussed above may convey the impression that linguistic researchers are forced to choose between purely quantitative and purely qualitative approaches, the two are actually not mutually exclusive, and they can be combined in a number of ways with regard to data collection, analysis and interpretation (cf. Ivankova and Greer 2015).

The various methods may either be used simultaneously or successively, e.g. when quantitative test data is subsequently complemented by qualitative interviews in order to find support for explanations, or when a questionnaire for large quantitative analyses is based on a small number of qualitative interviews (Phakiti and Paltridge 2015, 14). The successive combination of quantitative and qualitative research also involves the possibility of multiple alternation in several steps (e.g. in Baker 2016). In multiple level analysis, the qualitative characteristics of a subsample are discussed in detail and compared statistically to a larger group (Dörnyei 2007, 273).

A recent and very comprehensive triangulating study is that by Baker and Egbert (2016a), in which ten researchers were asked to apply a different method each to the same corpus of Q+A forum responses in order to find out how language use differs between four varieties of world English (India, Philippines, United Kingdom, United States) and between three topic areas (society & culture, family & relationships, politics & government). Interestingly, they find that the results of qualitative analyses correlate very highly with quantitative analyses and yield complementary results. As a consequence, Baker and Egbert conclude that "[o]bvious limitations aside, a strictly qualitative approach to corpus linguistics seems to have a great deal of insight that naturally complements other approaches" (2016c, 200) and therefore permits to represent the complex system of human language(s) more completely and holistically. In the triangulation of the data, Levon's (2016) qualitative study turns out to be involved in more shared results than all of the other, mainly quantitative, studies (but also in one partial disagreement). While one might have expected that generalisable findings should be retrieved by different methods and that triangulation should therefore mainly result in converging evidence, Baker and Egbert (2016c) make the interesting observation that about three-quarters of the discoveries in their triangulating study were only made by a single author each, and even shared results occurred in about two to four of the ten studies only. As a consequence, they conclude that the value of data triangulation lies in providing more comprehensive, complementary evidence (which would be missed if only a single method, e.g. the researcher's favourite one, were used) and that the different methods used all lead to Rome – but to different parts thereof.

The proponents of mixed methods research design or data triangulation¹⁰ argue that it contributes to higher quality, since the different methods "can support, complement or expand" each other (Phakiti and Paltridge 2015, 14) and permit a cross-validation of each other's results (e.g. in Baker and Egbert 2016c, where more agreement than disagreement was found to obtain between the results of the individual studies). However, this view requires a research paradigm that is open to compromise: by contrast to positivism, which assumes that there is a "quantifiable and measurable" "set of immutable laws or theories that govern reality" (Phakiti and Paltridge 2015, 16) and – at the other extreme – constructivism with its emphasis on multiplicity and interdependence of realities, according to which "social reality (e.g. cultures, cultural objects, institutions, values) cannot be reduced in the same manner as physical reality" (Phakiti and Paltridge 2015, 17),¹¹ the postpositivist research paradigm favoured here lends itself to the quantitative-qualitative compromise, since it "believes reality can only be approximated and cannot be perceived with total accuracy" but still views objectivity "as an ideology to guide researchers" (Phakiti and Paltridge 2015, 16) in their aim of getting closer to the truth (if that exists). Note, however, that since qualitative and quantitative methods involve very different ways of thinking, most researchers prefer either the one or the other (Dörnyei 2007, 47). One challenge of mixed methods research therefore is that "some of the material will by definition be read and evaluated by a potentially unsympathetic (or even hostile)" audience, who "will find different arguments and evidence convincing and apply different quality criteria in judging the value of a study" (Dörnyei 2007, 300).

Strictly speaking, however, the quantitative and qualitative approaches intermingle anyway and cannot be separated as neatly as one might initially be led to believe: thus Dörnyei (2007, 25) points out that most qualitative research collects some numerical information (such as the age of subjects), and quantitative research also tends to collect some qualitative information (such as the gender or nationality of language users). Quantification cannot really be separated from qualitative research anyway, since quantitative research hypotheses usually have their origin in an intuitive impression gained through the qualitative consideration of linguistic phenomena in context,¹² while bottom-up, explorative research methods reverse this order and "start with the statistical processing of raw data, which then yields results that function as a stepping stone for a qualitative analysis" (Hilpert and Gries 2016, 44). Furthermore, there is presumably very little purely frequency-based research in linguistics that does not consider context at all – and be it in the interpretation of the results, when explanations for the behaviour of the data are sought (cf. Hilpert and Gries 2016, 47).¹³ Baker and Egbert make the point that

10 The latter is a more general cover term for the use of various methods to answer one research question and thus not necessarily requires the combination of qualitative with quantitative methods, since it may merely involve the combination of methods within one of the two research paradigms (Phakiti and Paltridge 2015, 15).

11 Cf. also Lyons, who is very critical about the widespread "fiction of homogeneity," which pretends that there is such a thing as "the" English language, and that this supposedly "homogeneous, determinate and well-defined" system can be measured (1990, 26).

12 Cf. also Crystal, according to whom "[w]e always begin an investigation with *some* vague, half-formulated ideas and assumptions" (1985, 124).

13 While outliers are problematic for statistical analyses, their qualitative analysis can offer interesting insights into "the reason for the deviance" or "shed light on broader issues" (Dörnyei 2007, 272).

it would be unfair to cast corpus linguistics as a merely quantitative form of analysis," because "the patterns need to be interpreted and explained by human researchers, and this involves close reading of the texts in a corpus, often abetted again by corpus tools which can present the texts or sections of them in ways that make it easier for human eyes to process. (2016b, 2)

An ideal quantitative approach would therefore seem to combine efficiency with careful analysis (cf. also Meyer 2012, 25).

Conversely, qualitative research is not possible without a specific type of quantification: one of the characteristic features of qualitative research is the fact that it is particularly dependent on the competence of the analyst (Dörnyei 2007, 41). Since competence can be considered as closely related to the researcher's experience with this type of approach, the type of text etc., what this actually means is that the particularly good qualitative researchers are those who have acquired a large amount of knowledge on their topic. Viewed from this perspective, quantification enters qualitative research through the back door, via the researcher's intuition that is based on extensive reading etc. Furthermore, even qualitative studies occasionally use quantification in the form of numbers or other expressions denoting size – only not as systematically as one would expect in quantitative studies. For example, Levon (2016) uses quantifiers like *some* or *many* and states that all four questions of a particular category display a particular quality (even if he merely lists "representative examples" elsewhere). In addition, if there are several qualitative studies on the same subject, which converge in their results, this may also be considered a kind of quantitative support, in the sense that repeated quality yields quantity.¹⁴ To conclude, quantitative and qualitative research are not as mutually exclusive as one might initially believe.

Linguistic changes result in qualitative differences such as the loss of certain lexical items or patterns, and "such differences can be quantified as observed frequencies becoming zero" (Hilpert and Gries 2016, 38). In this respect, all qualitative findings can be expressed as digital statements of existence from a quantitative perspective (though with an uneconomical number of categories), and conversely, quantitative research can be considered to make qualitative statements about occurrence or non-occurrence. Qualitative and quantitative data can thus be transformed into the respective other type, so as to permit a joint analysis with statistical or qualitative methods (Dörnyei 2007, 269).

Yet another way out of the philologist's dilemma, which combines quantitative and qualitative methods, is suggested by Meyer (2012, 39-40), who advocates focusing on specific parts of a corpus (if that is permitted by its search interface) and analysing those in a philological way: for example, his search for the construction *you should* in the Corpus of Contemporary English returned more hits than were manageable within a qualitative approach (23,865), and he therefore successively reduced the sample to the spoken section of the corpus (with 6,526 hits) and then to the period from 2010 to 2011 – which finally yielded a manageable sample size of 413. A similar item-reduction strategy was employed by Levon (2016), who used the statistics-based corpus tool ProtAnt (Anthony and Baker 2015) in order to select the twelve

14 For instance, Vennemann (2012) argues that the Germanic runes have their origin in the Phoenician writing system (rather than, as is more commonly believed, directly in the Latin alphabet) by using a relatively large number of qualitative findings (such as similarities in form and the conventional order of the runes in the runic alphabet) to support his hypothesis.

most prototypical texts from a larger corpus and then analysed the selected twelve texts using qualitative methods.

5. Conclusion: How to Solve the Philologist's Dilemma

Linguistics studies language, but the scope of the discipline is extremely wide, and it would be difficult to argue that there is one "linguistics proper:" apart from the study of the parts of the system of language (phonology, morphology, word formation, syntax) both from a synchronic and a diachronic perspective, linguistics comprises a multitude of branches, such as clinical linguistics (which overlaps with medicine), sociolinguistics (which cannot be separated completely from anthropology), computational linguistics (which overlaps to a large extent with computer science) etc. This explains why the methods that are used in linguistics come from many different fields, so as to best serve the needs of specific research questions. Consequently, we find philological studies of written texts, experiments, brain scans, corpus studies, questionnaire studies, oral interviews with informants and many more, but for most subdisciplines, quantitative empirical research is currently the norm. As early as 1980, Leisi argues that traditional linguistics (which does not quantify) has been almost exclusively replaced with exact, quantitative "linguistic science" – but lately, there has been a certain tendency towards "an increasing visibility and acceptance of qualitative research" (1980, 201), not only in applied linguistics (cf. Dörnyei 2007, 36), but also in the fields of physics or in mathematics, in which problems are solved by using geometrical approximations which are then modelled quantitatively in the next step (Neuenschwander 2003, 9-11; 14-16).

We have seen above that the quantitative research paradigm is not without its problems, since the human tendency towards the uncritical acceptance of numbers as true brings with it that neither the database nor the categorisation are called into question. At the same time, however, it would be hard to argue entirely against quantitative research, because only this paradigm permits generalisable, relatively objective data that can also be used to explain the workings of the mind: since a growing body of cognitive linguistic research suggests that linguistic knowledge emerges from the representation of stored forms, with frequency playing a crucial role, one may argue that quantitative linguistic studies are essential in understanding how language works in the mind – and that has been one of the central aims of the study of language for a long time. In view of the arguments that speak in favour of qualitative research methodology, a combination of both quantitative and qualitative research would therefore seem particularly reasonable.

While the rigorous application of statistical methods may at first sight seem hard to align with a rigorous location in context, at least from the perspective of quantitative research, one solution to the philologist's dilemma consists in doing what most corpus studies presumably do anyway, namely to consider at least the surrounding context of every hit. This manual post-editing of search results, which is aided by modern corpus tools that permit a quick overview, can thus be considered a compromise between quantitative and qualitative research. While some might argue that it is necessary to become acquainted with the full text of a corpus before carrying out research on it, this is not feasible for the clear majority of researchers working with multi-million-word corpora – and, I would argue, not necessary, as long as one has gained a representative impression of the corpus based on manuals with the description of the make-up and the coding conventions and a number of short text passages

from the corpus. It makes sense to trust the compilers, but to gather information on the limitations of the corpora that are used. Furthermore, the amount of philological effort required for the interpretation of corpus-linguistic findings will vary depending on the phenomenon under consideration: thus the investigation of the syntactic behaviour of certain types of adjective within noun phrases will require less context for a meaningful interpretation than the pragmatic analysis of the situations in which English native speakers tend to apologise compared to German native speakers.

Yet another way out of the philologist's dilemma, which was discussed above, is the systematic and statistically aided reduction of representative samples to such a size that the resulting text corpus can be analysed using qualitative methods (Meyer 2012, 39-40). Alternatively, various types of method may be used within the same study in order to look for converging evidence, or be used in successive steps in order to develop hypotheses (qualitative), test them (quantitative) and search for explanations (qualitative). But, as we have seen, in many respects, quantitative and qualitative considerations cannot be separated neatly anyway, since e.g. the competence of qualitative researchers is based on their frequency-based experience.

Statistics is no end in itself, but merely an instrument which is used in order to determine how reliable a study's results are, and whether observations are meaningful or the result of chance. The numbers do not express causality but need to be interpreted by a human researcher in order to acquire meaning. As Wolpoff famously put it: "I have been in rooms with data and listened very carefully. The data never said a word" (1975, 15). This contribution therefore argues in favour of quantitative studies that test hypotheses with a strong focus on the consideration of the qualitative linguistic context and with the larger picture in mind.

To conclude, purely quantitative research may result in isolated figures whose scope is unclear and which are therefore meaningless. Conversely, purely qualitative research may produce a vague narrative about the existence of phenomena or constructions, whose relative importance is unclear. As a consequence, the qualitisiation of quantitative research and the quantisation¹⁵ of qualitative research appear to be a good way out of the philologist's dilemma. If we look closely, we see that the gap between quantitative and qualitative research is not too wide – and thus bridgeable – whenever real-world empirical data are analysed in order to describe actual language use.

Works Cited

- Anthony, Laurence, and Paul Baker. "ProtAnt: A Tool for Analysing the Prototypicality of Texts." *International Journal of Corpus Linguistics* 20 (2015): 273–293.
- Baker, Paul. "Gendered Discourses." *Triangulating Methodological Approaches in Corpus Linguistic Research*. Eds. Paul Baker and Jesse Egbert. New York: Routledge, 2016. 138-151.
- , and Jesse Egbert, eds. *Triangulating Methodological Approaches in Corpus Linguistic Research*. New York: Routledge, 2016a.

15 When qualitative data is quantised, certain themes are assigned numerical codes, e.g. frequency counts of their occurrence in the sample, or scales that evaluate "the intensity or degree or weight of a narrative account" (Dörnyei 2007, 269-270), as when open-ended survey questions are summarised in larger categories (cf. 273).

- , and Jesse Egbert. "Introduction." *Triangulating Methodological Approaches in Corpus Linguistic Research*. Eds. Paul Baker and Jesse Egbert. New York: Routledge, 2016b. 1-19.
- , and Jesse Egbert. "Research Synthesis." *Triangulating Methodological Approaches in Corpus Linguistic Research*. Eds. Paul Baker and Jesse Egbert. New York: Routledge, 2016c. 183-208.
- Ballmer, Thomas T. "Inwiefern ist Linguistik empirisch? Eine allgemeine Explikation des Empiriebegriffs mit spezieller Berücksichtigung der sich für die Linguistik ergebenden Konsequenzen." *Wissenschaftstheorie der Linguistik*. Ed. Dieter Wunderlich. Kronberg: Athenäum, 1976. 6-53.
- Bartsch, Renate, and Theo Vennemann. "Linguistik." *Linguistik und Nachbarwissenschaften*. Eds. Renate Bartsch and Theo Vennemann. Kronberg: Scriptor, 1973. 9-20.
- Biber, Douglas. *Variation across Speech and Writing*. Cambridge: Cambridge University Press, 1988.
- , Stig Johansson, Geoffrey Leech, Susan Conrad, and Edward Finegan. *Longman Grammar of Spoken and Written English*. Harlow: Pearson, 1999.
- Bybee, Joan. *Phonology and Language Use*. 2nd ed. Cambridge: Cambridge University Press, 2003.
- . "From Usage to Grammar: The Mind's Response to Repetition." *Language* 82.4 (2006): 711-733.
- Crystal, David. *Linguistics*. Harmondsworth: Penguin, 1985.
- Dörnyei, Zoltán. *Research Methods in Applied Linguistics: Quantitative, Qualitative, and Mixed Methodologies*. Oxford: Oxford University Press, 2007.
- Fletcher, James. "Spurious Correlations: Margarine Linked to Divorce?" BBC, 26 May 2014. Web. 11 June 2016.
- Francis, W. Nelson, and Henri Kucera. *Manual of Information to Accompany A Standard Corpus of Present-Day Edited American English, for Use with Digital Computers*. Providence, RI: Department of Linguistics, Brown University, 1979.
- Gigerenzer, Gerd. *Gut Feelings: The Intelligence of the Unconscious*. London: Penguin, 2007.
- Goldberg, Adele. "Construction grammar." *Concise Encyclopedia of Syntactic Theories*. Eds. Keith Brown and Jim Miller. Oxford: Pergamon, 1996. 68-71.
- Gonzalez-Marquez, Monica, Raymond B. Becker, and James E. Cutting. "An Introduction to Experimental Methods for Language Researchers." *Methods in Cognitive Linguistics*. Eds. Monica Gonzalez-Marquez, Irene Mittelberg, Seana Coulson and Michael J. Spivey. Amsterdam: John Benjamins, 2007. 53-86.
- Gries, Stefan Th. "Quantitative Methods in Linguistics." *International Encyclopedia of the Social and Behavioral Sciences*. Ed. James D. Wright. 2nd ed. Vol. 19. Amsterdam: Elsevier, 2015. 725-732.
- Hasher, Lynn, and Rose T. Zacks. "Automatic Processing of Fundamental Information: The Case of Frequency of Occurrence." *American Psychologist* 39.12 (1984): 1372-1388.
- Herbst, Thomas. *English Linguistics: A Coursebook for Students of English*. Berlin: de Gruyter Mouton, 2010.
- Hilpert, Martin, and Stefan Th. Gries. "Quantitative Approaches to Diachronic Corpus Linguistics." *The Cambridge Handbook of English Historical Linguistics*. Eds.

- Merja Kytö and Päivi Pahta. Cambridge: Cambridge University Press, 2016. 36-53.
- Ivankova, Nataliya V., and Jennifer L. Greer. "Mixed Methods Research and Analysis." *Research Methods in Applied Linguistics*. Eds. Brian Paltridge and Aek Phakiti. London: Bloomsbury, 2015. 63-81.
- Kanngießer, Siegfried. "Spracherklärungen und Sprachbeschreibungen." *Wissenschaftstheorie der Linguistik*. Ed. Dieter Wunderlich. Kronberg: Athenäum, 1976. 106-160.
- Lass, Roger. *On Explaining Language Change*. Cambridge: Cambridge University Press, 1980.
- Leisi, Ernst. *Das heutige Englisch: Wesenszüge und Probleme*. Heidelberg: Winter, 1955.
- . "Die Semantik und die zählenden Wissenschaften." *Sprachwissenschaft* 5 (1980): 201-214.
- . "'Traditionelle' Linguistik." *Die Neueren Sprachen* 80.5 (1981): 378-388.
- Levon, Erez. "Preference Structure and Stance-Taking in Online Question-Response Sequences." *Triangulating Methodological Approaches in Corpus Linguistic Research*. Eds. Paul Baker and Jesse Egbert. New York: Routledge, 2016. 152-166.
- Lüthy, Herbert. *Die Mathematisierung der Sozialwissenschaften*. Zürich: Arche, 1970.
- Lyons, John. "Linguistics: Theory, Practice, and Research." *Georgetown University Round Table on Languages and Linguistics 1990*. Ed. James E. Alatis. Washington, D.C.: Georgetown University Press, 1990. 11-30.
- Meyer, Charles F. "Pre-Electronic Corpora." *Corpus Linguistics: An International Handbook*. Eds. Anke Lüdeling and Merja Kytö. Berlin: Mouton de Gruyter, 2009a. 1-14.
- . "The Empirical Tradition in Linguistics." *Journal of English Linguistics* 37 (2009b): 208-213.
- . "Textual Analysis: From Philology to Corpus Linguistics." *English Corpus Linguistics: Crossing Paths*. Ed. Merja Kytö. Amsterdam and New York: Rodopi, 2012. 23-42.
- Moore, David S., and William I. Notz. *Statistics: Concepts and Controversies*. New York: W.H. Freeman, 2006.
- Neuenschwander, Erwin. "Einführung." *Wissenschaft zwischen Qualitas und Quantitas*. Ed. Erwin Neuenschwander. Basel: Birkhäuser, 2003. 1-32.
- Newmeyer, Frederick. *Grammatical Theory: Its Limits and its Possibilities*. Chicago, IL: University of Chicago Press, 1983.
- Phakiti, Aek, and Brian Paltridge. "Approaches and Methods in Applied Linguistics Research." *Research Methods in Applied Linguistics*. Eds. Brian Paltridge and Aek Phakiti. London: Bloomsbury, 2015. 5-25.
- Plag, Ingo, Gero Kunter, and Sabine Lappe. "Testing Hypotheses about Compound Stress Assignment in English: A Corpus-Based Investigation." *Corpus Linguistics and Linguistic Theory* 3.2 (2007): 199-233.
- Popper, Karl R. *Objective Knowledge: An Evolutionary Approach*. Rev. ed. Oxford: Clarendon Press, 1979.
- Pulvermüller, Friedemann. *The Neuroscience of Language: On Brain Circuits of Words and Serial Order*. Cambridge: Cambridge University Press, 2002.

- Rissanen, Matti. "Three Problems Connected with the Use of Diachronic Corpora." *ICAME Journal* 13 (1989): 16-19.
- Rundell, Michael. "Recent Trends in English Pedagogical Lexicography." *International Journal of Lexicography* 11.4 (1998): 315-342.
- Sacks, Oliver. *The Man Who Mistook his Wife for a Hat and other Clinical Tales*. New York: Summit, 1987.
- Sanchez, Christina. *Consociation and Dissociation: An Empirical Study of Word-Family Integration in English and German*. Tübingen: Narr, 2008.
- Sanchez-Stockhammer, Christina. "Comicsprache – leichte Sprache?" *Die Sprache(n) der Comics*. Ed. Daniela Pietrini. München: Meidenbauer, 2012a. 55-74.
- . "Hybridization in Language." *Conceptualizing Cultural Hybridization: A Transdisciplinary Approach*. Ed. Philipp Stockhammer. Berlin: Springer, 2012b. 133-157.
- . "Can We Predict Linguistic Change? An Introduction." *Can We Predict Linguistic Change? Studies in Variation, Contacts and Change in English* 16. Ed. Christina Sanchez-Stockhammer. Helsinki: VARIENG, 2015. <<http://www.helsinki.fi/varieng/series/volumes/16/introduction.html>> 5 June 2016.
- . "Punctuation as an Indication of Register." *Variational Text Linguistics: Revisiting Register in English*. Eds. Christoph Schubert and Christina Sanchez-Stockhammer. Berlin: de Gruyter, 2016. 139-167.
- . *English Compounds and their Spelling. (Studies in English Language)*. Cambridge: Cambridge University Press, forthcoming(a).
- . "How Comics Communicate on the Screen: Telecinematic Discourse in Comic-to-Film Adaptations." *Telecinematic Stylistics*. Eds. Christian Hoffmann and Monika Kirner-Ludwig. London: Bloomsbury, forthcoming(b).
- , and Jenny Arendholz. "'Abend ward, bald kommt die Nacht...!' Die Korrelation von Tages- und Uhrzeiten im deutsch-englischen Vergleich." *Lebende Sprachen* 57 (2012): 139-161.
- Sridhar, Shikaripur N. "What Are Applied Linguistics?" *International Journal of Applied Linguistics* 3.1 (1993): 3-16.
- Stefanowitsch, Anatol. "Negative Evidence and the Raw Frequency Fallacy." *Corpus Linguistics and Linguistic Theory* 2.1 (2006): 6-77.
- Szmrecsanyi, Benedikt, and Lars Hinrichs. "Probabilistic Determinants of Genitive Variation in Spoken and Written English: A Multivariate Comparison across Time, Space, and Genres." *The Dynamics of Linguistic Variation: Corpus Evidence on English Past and Present*. Ed. Terttu Nevalainen. Amsterdam: John Benjamins, 2008. 291-309.
- Taylor, John R. *The Mental Corpus: How Language is Represented in the Mind*. Oxford: Oxford University Press, 2005.
- Vennemann, Theo. "Germanische Runen und phönizisches Alphabet." *Germania Semitica*. Eds. Theo Vennemann gen. Nierfeld and Patrizia Noel Aziz Hanna Berlin: de Gruyter, 2012. 529-590.
- Vollhardt, Kurt Peter C., and Neil E. Schore. *Organische Chemie*. Transl. by Kathrin-M. Roy. Weinheim: Wiley-VCH, 2011.
- Weinreich, Uriel. "Review of Stephen Ullmann's *Précis de sémantique française*." *Language: Journal of the Linguistic Society of America* 31 (1955): 537-543.

- Wolpoff, Milford. "Discussion [of David Pilbeam and J. Rimas Vaišnys, "Hypothesis Testing in Paleoanthropology"]." *Paleoanthropology, Morphology and Paleoecology*. Ed. Russell H. Tuttle. The Hague: Mouton, 1975. 15-18.
- Young, S. Stanley, and Alan Karr. "Deming, Data and Observational Studies: A Process out of Control and Needing Fixing." *Significance* 8 (2011): 116-120.
- Zipf, George Kingsley. *The Psycho-Biology of Language: An Introduction to Dynamic Philology*. Cambridge, MA: MIT Press, 1965.