A Constructional Reanalysis of Semantic Prosody

by

Philipp Stempel

A Thesis Submitted
in Partial Fulfillment of the Requirements for the Degree

Doctor of Philosophy

Approved, Thesis Committee:

Michel Achard, Chair
Professor of Linguistics

Robert Englebretson
Associate Professor of Linguistics

Simon Fischer-Baum
Assistant Professor of Psychological Sciences

Houston, Texas

December, 2019
ABSTRACT

A Constructional Reanalysis of Semantic Prosody

by

Philipp Stempel

This dissertation is a re-examination of semantic prosody within the framework of construction grammar. The basic idea of semantic prosody is that ostensibly neutral words can have a probabilistic tendency to co-occur with words that express either positive or negative evaluations. Sinclair (1987), for example, notes that the phrasal verb 'set in' tends to occur with nouns like 'decay' or 'despair'; Stubbs (1995a) finds that the verb 'cause' often collocates with words like 'problem' or 'damage'.

Even though semantic prosody is an important and much-discussed topic within corpus linguistics, it remains an elusive and contentious subject. There are numerous areas of disagreement in the research literature, e.g. what its function is, at what level it is located, or how it relates to other distributional phenomena such as collocation or semantic preference.

This dissertation proposes a novel account of semantic prosody to solve these issues by examining them through a different quantitative methodology and by contextualizing them within a different theoretical framework. A major limitation of previous studies has been their focus on word-level co-occurrences. Corpus evidence reveals that much of a word’s meaning is not invariant but instead differs across constructional contexts. Once this context-dependent nature of meaning is considered, a much clearer image of co-selection between linguistic items emerges.
Evidence from collostructional analysis (Stefanowitsch and Gries, 2003), a type of quantitative corpus methodology, demonstrates that constructional patterns are often associated with specific lexical fields, which allow for a detailed description of their semantics. Not only does this reveal that the evaluative meaning of semantic prosody reaches far beyond the simple positive/negative polarity posited in prior research; it also shows that semantic prosody surfaces in language use as an epiphenomenon of more complex evaluative semantics.

With the help of the methodology of quantitative corpus linguistics and the theoretical frameworks of construction grammar and usage-based linguistics, semantic prosody is readily explained as an ordinary part of meaning and does not pose a challenge to linguistic theory.
Contents

Abstract ii

List of Figures viii

List of Tables ix

1 Introduction 1

2 Prior Work on Semantic Prosody 6
   2.1 Foundational Research on Semantic Prosody 7
   2.2 Practical Applications of Semantic Prosody 16
   2.3 Issues and areas of disagreement 21
      2.3.1 Semantic transfer 21
      2.3.2 Semantic prosody and semantic preference 24
      2.3.3 Genre or register effects 26
      2.3.4 The hidden nature of semantic prosody 28
      2.3.5 The lexical priming perspective and the discourse perspective 30
      2.3.6 The relevant level of analysis 33
   2.4 Discussion 36

3 Theoretical Foundations 39
   3.1 Usage-based linguistics 40
      3.1.1 The primacy of context over isolated ‘lexicon meaning’ 43
      3.1.2 Distributional linguistics and the Firthian maxim 45
6.2 Lack of evidence ........................................ 166
6.3 Evaluation is often more complex than evaluative polarity 168
6.4 Some evaluations lack the positive/negative dimension 171
6.5 Some words are associated with opposite evaluations 174
6.6 The difficulty of assigning evaluative polarity ............ 177
6.7 Discussion ........................................... 181

7 Investigating the Claim of Hiddenness 184
7.1 Issues with claims of hiddenness ........................ 185
7.2 Psycholinguistic studies on semantic prosody ......... 199
7.3 Explicit knowledge of semantic prosody ............... 211
  7.3.1 Methodology ..................................... 212
  7.3.2 Results ........................................ 216
  7.3.3 Discussion ..................................... 224
7.4 Discussion ........................................... 228

8 Semantic Prosody from a Cross-Linguistic Perspective 231
8.1 Previous Studies ....................................... 232
  8.1.1 Research on Semantic Prosody in German ....... 232
  8.1.2 Other Contrastive Studies ......................... 235
8.2 Semantic prosody in German and English ............. 245
  8.2.1 Study 1: English ‘cause’ and its German counterparts 247
  8.2.2 Study 2: vollkommen .......................... 263
8.3 Discussion ........................................... 270

9 Conclusions ........................................... 277
9.1 A re-evaluation of semantic prosody within a constructional framework 278
  9.1.1 Central findings .................................. 278
  9.1.2 Common questions about semantic prosody....... 283
9.2 Semantic prosody in a broader theoretical context .......................... 290
  9.2.1 The primacy of context-dependent meaning .............................. 290
  9.2.2 Do speakers have mental representation of semantic prosody? ...... 294
  9.2.3 Is semantic prosody a useful concept? .................................. 296
9.3 Further avenues of research .................................................... 300

Bibliography ................................................................................. 305
viii

List of Figures

5.1

Collexeme clusters for ‘persistent’ in WordNet. . . . . . . . . . . . . . 102

5.2

Collexemes of persistent NP mostly belonging to the lexical field of
illness. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 103

5.3

Collexemes of persistent NP mostly belonging to the lexical fields
negative mental states and scarcity. . . . . . . . . . . . . . . . . . . . 104

5.4

Collexeme clusters for persistent NP found through hierarchical
cluster analysis of their word vectors. . . . . . . . . . . . . . . . . . . 105

5.5

Collexeme clusters for cause in the transitive construction, found
through a hierarchical cluster analysis of their word vectors. . . . . . 115

7.1

Correlation between corpus data and elicited judgments for ‘happen
words’ . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 219

7.2

Correlation between corpus data and elicited judgments for
‘amplifying intensifiers’ . . . . . . . . . . . . . . . . . . . . . . . . . . 221

7.3

Correlation between corpus data and elicited judgments for all items.

225


## List of Tables

5.1 Attitudinal stance found in 200 examples with the adjective ‘persistent’. 97
5.2 Collexemes with the strongest attraction to the noun slot in 
   \textit{persistent NP}. ................................. 101
5.3 The most common noun collocates for the verb ‘cause’ in COCA. 109
5.4 Constructions used with verbal ‘cause’. ................................. 111
5.5 Collexemes with the strongest attraction to the \textit{direct object slot} 
   in the transitive causative construction. ................................. 113
5.6 Collexemes with the strongest attraction to the \textit{theme slot (= caused 
   event)} in the prepositional dative causative construction. 118
5.7 Collexemes with the strongest attraction to the \textit{theme slot (= caused 
   event)} in the ditransitive causative construction. 119
5.8 Distinctive collexeme analysis of the \textit{caused event} slot in the 
   ditransitive compared to the prepositional dative. 123
5.9 Collexemes with the strongest attraction to the \textit{theme slot (= caused 
   event)} in the passive construction. 124
5.10 Collexemes with the strongest attraction to the agent slot (= causer) 
   in the passive construction. 126
5.11 Collexemes with the strongest attraction to the infinitival 
   complement construction. 128
5.12 Collexemes with the strongest attraction to the noun slot in \textit{cause for 
   NP}. ................................. 134
5.13 Collexemes with the strongest attraction to the noun slot in cause of NP. .......................................................... 137
5.14 Collexemes with the strongest attraction to the adjective slot in the there BE something ADJ about construction. ................. 141
5.15 Attitudinal stance found for different semantic fields in 200 examples of the there BE something ADJ about construction. ............. 143
5.16 Collexemes with the strongest attraction to the first verb slot in V NP into Ving. .......................................................... 150
5.17 Covarying collexeme analysis of the collexemes in the verb slots of the into Ving causative. .................................................. 152

6.1 The ten collexemes most attracted to the NP slot in the striking NP construction. ......................................................... 173
6.2 The ten collexemes most attracted to the adjective slot in the absolutely adj construction. .................................................... 176

7.1 Evaluative polarity for sentences produced by speakers in McGee (2012). ................................................................. 206
7.2 Speakers’ judgments and three different corpus measures: ‘Happen words’, Partington (2004) .................................................... 218
7.4 Speakers’ judgments compared to token frequencies. ......................... 224

8.1 The most common German verbs or structures corresponding to verbal ‘cause’ in the Europarl corpus. .......................... 248
8.2 Collexemes with the strongest attraction to the adjective slot in the vollkommen + adjective construction. ...................... 265
8.3 Collexemes with the strongest attraction to the verb slot in the
vollkommen + verb construction. . . . . . . . . . . . . . . . . . . . . . . . . . 269
Chapter 1

Introduction

The central aim of this dissertation is a fundamental reanalysis of a phenomenon referred to within corpus linguistics as *semantic prosody*. The basic idea behind semantic prosody is that some words, which do not appear to carry strong evaluation as part of their lexical meaning, have probabilistic tendencies to collocate with either negative or positive words. I will argue that this is a result of the dynamics between constructional and lexical meaning. The effects that emerge in corpus data at the lexical level are an epiphenomenon of this relationship. As such, a constructional approach is both able to provide a much more detailed account and avoid the analytical issues encountered within a lexical perspective that limits its investigation to the word-level.

A much-discussed example in the research literature is the verb ‘cause’. It has been argued that ‘cause’ possesses a negative semantic prosody because it tends to collocate with nouns such as ‘problem’, ‘trouble’, ‘accident’ or ‘disaster’ (e.g. Stubbs, 1995a). A deeper exploration of the data that goes beyond simple word-to-word collocations, however, reveals a more complex picture. ‘Cause’ occurs with a number of different constructions such as the transitive, the ditransitive or the prepositional dative construction, all of which have different distributions in text: Some of them
occur with lexical fields describing medical problems, while others describe a person’s internal or external states. The semantics of ‘cause’ are not context-independent but instead vary based on the constructional context. Another example is the adjective ‘persistent’, which tends to convey negative evaluations in attributive contexts, but not when used predicatively.

My reanalysis of semantic prosody comprises the investigation of three central claims that are commonly made in the research literature. First, semantic prosody is described as a matter of co-occurrences between words. I will present data to show that the constructional context needs to be taken into account. Second, semantic prosody is regarded as a two-term system of positive and negative evaluations. I argue that evaluative semantics are far more complex and nuanced than that. A constructional account reveals patterns of co-selection that allow for an analysis that is both richer in detail and less subjective in its interpretation. I also argue that there is no evidence to postulate a two-term system of evaluation, as corpus data do not support a separation of semantic prosody from semantic preference. Third, semantic prosody is described as hidden from speakers’ introspection or intuition. I will argue this claim poses the wrong question and has never been empirically verified.

Semantic prosody is often seen as an elusive concept. It is a contentious topic with many open questions and areas of disagreement among corpus linguists. I will argue that semantic prosody is not an enigmatic phenomenon or particularly challenging to linguistic theory. When considered within the framework of usage-based linguistics,
it becomes clear that it is a natural part of the evaluative semantics of partially conventionalized patterns that emerge through frequent co-occurrence in language use. As such it is not a separate linguistic category, but can be accounted for as an ordinary part of language. Within the right theoretical framework and with the help of the right research methodology, formerly problematic issues naturally fall into place and are readily solved.

The structure of this dissertation is as follows: Chapter 2 introduces the phenomenon of semantic prosody through a review of the prior research literature. I outline its foundational research, its applications in different linguistic fields, and discuss problems with current approaches that warrant a fundamental reanalysis of the data.

In Chapter 3, I summarize the theoretical framework used in this dissertation. I introduce usage-based linguistics and constructional approaches to grammar, both of which inform the reanalysis of semantic prosody in chapters 5 – 8 and contrast them to the linguistic theories underlying most of the prior research on semantic prosody. I also outline basic concepts of evaluation and stance in language.

Chapter 4 deals with linguistic evidence and research methodology. I describe the tenets of quantitative corpus linguistics, which is the broad methodological stance adopted here. I describe the practicalities of data retrieval and preprocessing. I also detail collexeme analysis, the specific method of quantitative analysis used throughout much of chapters 5 – 8.
In chapter 5, I make the case for a constructional approach to the study of semantic prosody. I show where prior studies that focused on word-level lexical semantics fell short and demonstrate how the different constructional contexts words occur in show different distributions in text, and, as such, different semantics. I provide several studies that illustrate and corroborate my claims, including both reexaminations of examples prominently discussed in the prior research literature and novel case studies.

Chapter 6 is a critical exploration of the idea that the function of semantic prosody is based upon a two-term system of positive and negative evaluations. I contend that this commonly repeated claim that is usually taken for granted is not supported by corpus data. Instead, the evaluative meaning of semantic prosody is far more complex and often defies a binary system. While a positive/negative polarity can be an important dimension of evaluation, it does not necessarily possess a privileged role. Instead, corpus evidence shows that it is often the result of other evaluative components of meaning, not their antecedent.

In Chapter 7, I examine the commonly made claim that semantic prosodies are hidden from speakers’ intuitions. I point out several issues with this claim, most importantly its imprecise description and the lack of empirical evidence. I present a study that tests awareness of semantic prosody and comes to the conclusion that speakers do in fact possess some explicit knowledge of it, which generally seems to be in line with what we know about speakers’ awareness of other frequency-based linguistic phenomena.
Chapter 8 is a contrastive investigation into the differences in semantic prosody between German and English. I assess to what extent the findings established for the English language also hold up in another language, and to what extent semantic prosody is language-specific. The data show that while there are cross-linguistic differences in specific patterns, the general principles described for English are attested in German as well. Additionally, a comparison of corresponding structures between the two languages in a parallel corpus further supports the context-dependent nature of semantic prosody.

Finally, chapter 9 reviews the research findings and recasts the questions discussed in the literature within a constructional framework. Theoretical, methodological and practical implications are discussed and avenues for further research are suggested.
Chapter 2

Prior Work on Semantic Prosody

Semantic prosody is a complex and somewhat elusive concept. This section approaches its definition through an overview of its research and idea history. In section 2.1, I outline the development of the concept through a chronological review of its foundational research. I describe the most important case studies and trace the development of theories about semantic prosody. At this point, I provide a largely impartial summary of the prior literature. In later chapters, I will critically discuss the issues that the data show to be in need of reexamination and propose alternative explanations that aim to better account for these issues.

In section 2.2, I describe the wide range of linguistic fields that have found practical applications for semantic prosody, including literary studies, second language learning, and natural language understanding. In 2.3, the most essential characteristics of semantic prosody are outlined and contentious issues and definitional differences are contextualized and problematized, with reference to later chapters that deal with these issues in my own research. Section 2.4 summarizes the key points of this chapter and sets the scene for the subsequent chapters.
2.1 Foundational Research on Semantic Prosody

The first remarks that sparked the idea of semantic prosody can be traced back to Sinclair’s discussion of the phrasal verb ‘set in’. He profiles some of its distributional characteristics, e.g. its tendency to occur in clause-final or sentence-final position, and the relatively short average length of the clauses it occurs in (Sinclair, 1987, p.74). More importantly, however, he points out that it tends to occur in contexts that seem to carry negative evaluations:

“The most striking feature of this phrasal verb is the nature of the subjects. In general, they refer to unpleasant states of affairs. […] The main vocabulary is rot, decay, malaise, despair, ill-will, decadence, impoverishment, infection, prejudice, vicious (circle), rigor mortis, numbness, bitterness, mannerism, anticlimax, anarchy disillusion, disillusionment, slump. Not one of these is desirable or attractive.” (Sinclair, 1987, pp. 74-75)

The relevance of these findings needs to be contextualized within Sinclair’s fundamental understanding of language production. He postulates two competing grammatical principles, the principle of idiom and the open-choice model, also called the open-slot principle (cf. Sinclair, 1991, inter alia). The maximally reductionist open-choice model essentially corresponds to a traditional lexicon-and-grammar model. It describes language as consisting of two separate modules – a set of lexical words and a set of grammatical rules. The idea that words do not carry semantic restrictions is what makes this model ‘open-choice’: Grammatically correct sentences are generated
by combining words according to the rules of grammar. These syntactic rules are
independent of semantic constraints and do not account for idiomatic phrases: As
Sinclair puts it, “at each point where a unit is completed (a word or a phrase or a
clause), a large number of choices opens up and the only restraint is grammaticalness”

The idiom principle on the other hand regards language as consisting of semi-
preconstructed phrases: “a language user has available to him or her a large number
of semi-preconstructed phrases that constitute single choices, even though they might
appear to be analyzable into segments” (Sinclair, 1991, p. 110). Sinclair acknowledges
that the idiom principle “has been relegated to an inferior position in most current
linguistics, because it does not fit the open-choice model” (ibid). He believes that
both principles are valid but that the idiom principle is the more important one,
accounting for a much greater portion of language in use. Sinclair (1991) does not
provide any empirical evidence for this claim which seems to be solely based on his
intuitions as a corpus linguist. That being said, his intuitions appear to be correct.
Altenberg (1998) studies English phraseology in the London-Lund Corpus of Spoken
English (cf. Svartvik, 1990) and estimates that more than 80% of all words in the
corpus are part of recurrent word combinations (ibid, p. 102). The author also notes
that the great majority of these phrases are not completely fixed, but instead “occupy
a position along the cline between fully lexicalized units and free constructions [...],
and they illustrate very clearly the difficulty – or impossibility – of making a sharp
distinction between lexicon and grammar” (ibid, p. 121) – a description that is very much in line with the idiom principle.

Sinclair sees semantic prosody as an example for lexical or semantic co-selection and thus for the insufficiency of the open-choice principle. He seems to believe that this is a common phenomenon in language, providing the verb ‘happen’ as an example:

“[...] many uses of words and phrases show a tendency to occur in a certain semantic environment. For example, the word happen is associated with unpleasant things – accidents and the like.” (Sinclair, 1991, p. 112)

It should be noted that Sinclair was not necessarily the first linguist to make these observations. The idea that speakers’ knowledge of lexical items also includes contextual or collocational knowledge undoubtedly existed prior to Sinclair’s statements. For example, in a seminal ESL paper on lexical competence, Richards (1976, p.83) declares: “For many words we [...] know the sort of words most likely to be found associated with the word.” The idea that collocates affect the meaning of the node word is even older, going back to the work of 19th century French philologist Michel Bréal. Ullmann (1962) Bréal’s account in his 1897 Essai de Semantique as follows:

“Habitual collocations may permanently affect the meaning of the terms involved; by a process known since Bréal as ‘contagion’, the sense of one word may be transferred to another simply because they occur together in many contexts.” (Ullmann, 1962, p. 185)
Ullmann gives the grammaticalization of French negative particles ‘personne’, ‘rien’, ‘pas’, ‘point’ as examples of Bréal’s concept of contagion:

“[...] Latin ‘persona’ has become a negative particle in French. This change is due to purely linguistic reasons: contiguity with the negative particle ne has ‘infected’ this word, in the same way as rem, passum, punctum, by a process which Bréal has termed ‘contagion.’” (Ullmann, 1962, p. 185)

Additionally, the attitudinal or evaluative dimension of collocational meaning has also been discussed. As Bublitz (1996, p. 22) points out, researchers in phraseology have long known of the existence of words with collocates that can be classified as generally negative. He mentions Gläser (1986, p. 40), who found that ‘commit’ tends to occur with actions that typically have negative connotations. Nonetheless, Sinclair’s remarks are historically important because they directly influence Louw’s foundational work which establishes semantic prosody as a subject of scientific inquiry. Additionally, much of the early work on semantic prosody is done by researchers in the Sinclairian tradition of British text linguistics.

Louw first introduces the term ‘semantic prosody’ into the academic discourse in his seminal paper from 1993 that sparks almost all subsequent research into the topic. He credits Sinclair with naming it and mentioning it to him first in personal communications as early as 1988 (Louw, 1993, p. 158). Louw goes on to explain that Sinclair named the concept in analogy to speech prosody in the Firthian sense, as a
phonological coloring that transcends segmental boundaries (Firth, 1957), e.g. nasal consonants passing on their nasal quality to their surrounding vowels. Semantic prosody is taken to be a semantic coloring that spreads beyond word boundaries, conceptually similar to phonological co-articulation, which spreads to its surrounding segments:

“The nasal prosody in the word *Amen* would be an example: we find that the vowels are imbued with a nasal quality because of their proximity to the nasals *m* and *n*. In the same way, the habitual collocates of the form *set in* are capable of colouring it, so it can no longer be seen in isolation from its semantic prosody, which is established through the semantic consistency of its subjects.” Louw (1993, p. 158-159)

In a much-cited description, Louw (1993, p. 157) characterizes semantic prosody as a “consistent aura of meaning with which a form is imbued by its collocates.” One of Louw’s important contributions to the literature on semantic prosody is the concept of transfer of meaning between co-occurring words. Sinclair’s account, taken by itself simply describes a phenomenon of lexical co-selection: Some words have the probabilistic tendency to co-occur with words that carry either a positive or negative evaluation. Louw goes a step further and assumes a *semantic transfer* from the collocates to the node word: The word takes on some meaning components of its common neighbors and retains this meaning in other contexts as well. The theoretical basis for this is the Firthian maxim, “You shall know a word by the company it
keeps” (Firth, 1957, p. 11), which is essential to all distributional approaches in linguistics (cf. Chapter 3). Concerning the phrasal verb set in, Louw states “the habitual collocates of the form set in are capable of colouring it, so it can no longer be seen in isolation from its semantic prosody, which is established through the semantic consistency of its subjects” (Louw, 1993, p. 159).

Louw is interested in stylistics, and as such describes the stylistic potential of semantic prosody in creating an ironic effect through a “collocative clash” (Louw, 1993, p. 157). Skilled writers may purposely violate the semantic prosody of a lexical unit to create an ironic effect. He provides a sentence from the novel ‘Small World’ by David Lodge: “The modern conference resembles the pilgrimage of medieval Christendom in that it allows the participants to indulge themselves in all the pleasures and diversions of travel while appearing to be austerely bent on self-improvement” (p. 164). According to Louw, the irony is created by the collocative clash between ‘bent on’, which has a negative prosody, and ‘self-improvement’, which is a transparently positive word (it seems to me that there is also a clash between ‘austerely’ and ‘bent on’). At the same time, a collocative clash could also indicate insincerity (in the form of a sort of between-the-lines criticism) or may even accidentally reveal a “speaker’s real attitude even where s/he is at pains to conceal it” (p. 157). Louw provides the example of someone using ‘symptomatic of’ (p. 170) in a positive context, even though corpus data show that it has an overwhelmingly negative semantic prosody.
Finally, Louw also makes basic observations about the context-dependence of semantic prosody. He notices that the phrasal verb ‘build up’ occurs in two different contexts with two different semantic prosodies (1993, p. 171). In intransitive constructions, it has a negative semantic prosody (e.g. ‘dirt builds up’) while in transitive constructions, it has a positive prosody (‘he built up his confidence’). Louw explains this difference with the fact that the intransitive construction implies control on the part of a human actor, while the transitive construction does not. However, he still attributes the semantic prosody to the verb itself and describes it as having a mixed prosody.

Stubbs (1995a) discusses semantic prosody in the context of collocational studies. He points out that words have distinctive semantic profiles and prosodies and that their strength of association can be quantified. He stresses the importance of using statistical measure over raw frequencies, which is an important methodological contribution. Stubbs (1995a) introduces the verb ‘cause’ into the literature, which has perhaps become the single most studied and discussed word in all research on semantic prosody. His investigation of ‘cause’ in the LOB corpus (the Lancaster-Oslo-Bergen corpus, a one million word corpus of British English that replicates the design of the Brown Corpus) finds that ‘cause’ occurs in 80% negative, 18% neutral and 2% positive evaluations. He concludes: “CAUSE acquires guilt by association. At some point, the word itself acquires unpleasant connotations and affects” (Stubbs, 1995a, p. 51).
Another historically relevant article is Bublitz (1996). His theoretical account of semantic prosody is clearly directly influenced both by Louw (1993) and by Sinclair’s account of the idiom principle. He describes the semantic transfer from the collocates to the node word as a ‘halo’ and emphasizes that it creates a semantic expectation in the listener, clearly in line with Sinclair’s idiom principle:

“Apparently, words can have a specific halo or profile, which may be positive, pleasant and good, or else negative, unpleasant and bad. Whenever some such word is uttered, it prompts hearers to expect a following word with a clear (un)pleasant sense. […] It opens not just any slot but one with a definite semantic set-up.” (Bublitz, 1996, p. 11)

Bublitz’s main contribution is the reanalysis of some of Sinclair’s earlier examples. Similar to Louw’s example of ‘build up’, Bublitz (1996, p. 17) notes that ‘happen’ also has two different senses: The “occurrence-meaning” (p. 19), e.g. “this accident happened” (p. 20) and the “by-chance-meaning” (p. 18), e.g. “he happens to be a good friend of mine” (ibid). He finds a negative semantic prosody only for the ‘occurrence-meaning’ but not for the ‘by-chance-meaning’.

Hoey’s theory of lexical priming is another important concept in the idea history of semantic prosody. It is employed by some researchers, e.g. Stewart (2009b) and Partington (2015) to explain the idea of semantic transfer. Lexical priming, a formulation of the distributional hypothesis (see chapter 3) that is centered around the acquisition of speakers’ mental concepts through collocational structures:
“[A]s a word is acquired through encounters with it in speech and writing, it becomes cumulatively loaded with the contexts and co-texts in which it is encountered, and our knowledge of it includes the fact that it occurs with certain other words in certain kinds of context. The same applies to word sequences built out of these words; these too become loaded with the contexts and co-texts in which they occur.” (Hoey, 2005, p. 8)

Semantic prosody then is the evaluative component of this ‘loading of context’. If a word frequently occurs within a negative or positive context, it is claimed that this polarity will eventually “rub off” on the node word (Leech, 1981, p. 16). As such, semantic prosody becomes a mechanism of stance-taking. As Louw (1993, p. 170) states, a sequence like symptomatic of prepares the hearer for what is coming next; it guides the hearer to interpret the attitudinal evaluation of the next noun phrase through comparisons with its common collocates such as ‘disease’ or ‘problem’.

The vast majority of research into semantic prosody has been done through corpus studies, but there are a small number of studies that aim to study mental representations of semantic prosody through psycholinguistic experiments (e.g. Guo et al., 2011; Ellis et al., 2007; Ellis and Frey, 2009; Hauser and Schwarz, 2016, 2018). I discuss these studies at some length in chapter 7.
2.2 Practical Applications of Semantic Prosody

The combination of its evaluative meaning and its supposed hiddenness (or non-obvious nature) has led to an interest in semantic prosody in various fields. Literary scholars have investigated the way authors exploit semantic prosodies for artistic effects (e.g. Adolphs and Carter, 2002; Spencer, 2011). Most literary studies focus on irony being conveyed through purposeful violation of prosodic primings (Louw, 1993; 2000; Burgers et al., 2011; Benna et al., 2014). O’Halloran (2007) discusses semantic prosody in the context of using corpus evidence to support and substantiate literary evaluation.

Drawing on Louw’s early work, researchers in the field of critical discourse analysis believe that the study of semantic prosody can be used to reveal speakers’ hidden (or even subconsciously held) attitudes and beliefs (cf. Cotterill, 2001; Mautner, 2009). As such, Koller and Mautner (2004, 223) believe that semantic prosody “is at least as exciting a concept for the critical discourse analyst as it is for the lexicographer and the grammarian.” Már mol and Almela (2016) and Soto-Almela and Alcaraz-Már mol (2017) study the usage of the words inmigración and inmigrante in the Spanish media. They find that newspapers both right and left of the political center use the two terms in overwhelmingly negatively evaluated contexts, typically conveying a sense of illegality, secrecy, and irregularity or describe it as something that needs to be opposed or curbed.

There are also suggestions to use semantic prosody in computational approaches.
An example is sentiment analysis (sometimes referred to as opinion mining, subjectivity analysis or appraisal extraction), an emerging subfield within the discipline of natural language understanding that aims to automatically categorize the overall evaluation expressed in a text as either positive or negative. A typical real-world application of sentiment analysis is gauging public opinion by mining the internet (e.g. news websites or social media) for texts about a given topic. Feng et al. (2011) and Feng et al. (2013) are examples for research that proposes to include semantic prosody in sentiment analysis.

Most work on semantic prosody has been conducted in English but there are a number of studies on other languages as well, including Portuguese (Berber Sardinha, 2000), Italian (Tognini-Bonelli, 2001), Chinese (Xiao and McEnery, 2006; Wei and Li, 2014), German (Dodd, 2006) and Norwegian (Ebeling, 2014). So far, all studies of languages other than English have also found semantic prosody to exist in those languages. When comparing approximately synonymous or equivalent word pairs between two languages (i.e. words in one language and their translations in another), the results are mixed: Some word pairs have similar distributional characteristics, but many do not. Research in translation studies comes to similar results although it typically emphasizes the differences rather than the similarities and argues that translators need to be aware of cross-linguistic variation in semantic prosodies (Dam-Jensen and Zethsen, 2008; Oster and van Lawick, 2008; Volanschi, 2012, inter alia).

McGee (2012) studies the use of semantic prosody in written texts produced by
translation students and English teachers, all of whom are native speakers of Arabic. He arrives at two important conclusions. First, their use of semantic prosody is often different from native speakers of English, leading to infelicitous prosodic clashes. Second, the awareness of semantic prosody is not predicted by the level of English training. This means that speakers do not learn semantic prosody through formal instruction; instead just seem to pick it up incidentally, through exposure to English outside of the classroom. Two studies test to what extent native speakers of Farsi are aware of semantic prosody in English. Ahmadian et al. (2011) test students of English at various language institutes and Elahi and Rahbar (2018) works with translators in training. Both find the awareness of semantic prosody to be lacking. Ahmadian et al. (2011, p. 294) conclude:

“[L]earning individual words and their meanings does not suffice to achieve great fluency in a second language. Knowing the way words combine into chunks (collocations) […], as well as being aware of the conditions of semantic prosody is necessary. Moreover, […] from the very beginning, learners’ attention should be turned to these kinds of combinations (words) and conditions (semantic prosody), and students should be constantly acquainted with an increasing number of collocations.”

In a similar line of research, semantic prosody has been studied in the context of second language learning. Gries (2009, p.17) finds that semantic prosody “has implications for […] foreign language teaching since, for example, if a foreign language
learner uses a word \( w \) without being aware of \( w \)’s semantic prosody, this may result in comical situations or, more seriously, communicating unwanted implications.”

Hunston (2002) makes the argument for teaching semantic prosody (not specifically limited to L2 teaching) based on its potentially hidden nature:

“Because semantic prosody is not always part of a speaker’s conscious knowledge of a language, it may be something that learners are not taught. Because of the kind of meanings that it can convey, however, it is an important aspect of language. Vocabulary teaching needs to take account of semantic prosody, and can do so only if the approach is phraseological rather than word-based.” (Hunston, 2002, p. 142)

L2 learners have the added difficulty that they are often exposed to little natural language in use, making it difficult for distributional categories to emerge in their minds. Most research here has been done by ESL researchers in the context of vocabulary pedagogy (Zhang, 2009; Guo et al., 2011; Lee, 2011; McGee, 2012; Zhao, 2017, inter alia). Key findings include the facts that L2 learners often lack knowledge of semantic prosody and that semantic prosodies are not adequately represented in bilingual dictionaries.

Ji and Wu (2000) studies the translation of three lexical items, ‘set in’, ‘rife’, and ‘propaganda’ (all of which have negative semantic prosodies) in three different English-Chinese dictionaries. They find that the Chinese translations either lacked the negative prosody or even possessed a positive prosody. Lee (2011) conducts a
similar study for English-Korean bilingual dictionaries and finds that translations did not correspond well in terms of their semantic prosodies. Siepmann (2005) compares bilingual dictionaries for English, French and German and also finds mismatches in evaluative polarity between translations.

Lee (2011) studies the writing of Korean English students and finds that they are often not aware of prosodies (e.g. ‘utterly’ and ‘be bent on’ have positive prosodies in their writings). Several studies have been conducted on awareness of semantic prosody and preference among Chinese learners of English. Wang and Wang (2005) and Naixing (2006) find that Chinese speakers do not use ‘cause’ in a way that corresponds to its semantic prosody. Lu (2005) finds similar results for ‘gain’ and ‘obtain’ as does Li and Liu (2017) for the intensifiers ‘completely’, ‘totally’ and ‘absolutely’. A common theme in these articles is the insufficiency of language learning materials, in particular dictionaries. Zhang (2009) summarizes the conclusions from this line of research:

“Awareness of semantic prosody not only can be greatly beneficial in interpreting a text producer’s hidden attitudes but can also help language learners understand how to use lexical items appropriately. For vocabulary learning, therefore, ESL/EFL learners need to master not only a lexical item’s spelling, meanings, and grammatical features, but also its semantic prosody.” (Zhang, 2009, p. 9)
2.3 Issues and areas of disagreement

Even though semantic prosody has seen a great amount of interest within corpus linguistics and related fields over the last 25 years, it remains an elusive and contentious concept. There are still numerous issues that either lack a clear consensus or are points of open disagreement or confusion among researchers. In the past decade, several articles with the express purpose of critiquing the concept of semantic prosody or clearing up some of the confusion around it have been published (e.g. Whitsitt, 2005; Hunston, 2007a; Bednarek, 2008; Morley and Partington, 2009; Stewart, 2009b; Partington, 2015), some of which also include suggestions to rename semantic prosody to *discourse prosody* (Stubbs, 2001), *emotive prosody* (Bublitz, 2003) or *evaluative prosody* (Partington, 2015, 2014a). Still, many issues remain unresolved. Researchers argue about what kind of evaluative meaning semantic prosody is, how it should best be assigned, whether it is connotational in nature or not, what its relationship with collocation and semantic preference is, what its function or purpose is and at what level it resides. All main issues are introduced in this section. Since I am not presenting any evidence in this chapter, I will refer to later chapters in this dissertation where I will address them.

2.3.1 Semantic transfer

One of Louw’s most essential concepts is the notion of semantic transfer – a word is imbued with evaluative meaning through its habitual co-occurrence with its collo-
cates. Many, though certainly not all researchers draw upon this concept. Some of the sharpest criticism of semantic transfer, and to some extend semantic prosody in general, is found in Whitsitt (2005). One of the points Whitsitt (2005) criticizes is a lack of empirical evidence. Louw emphasizes that semantic transfer is a diachronic process – words become gradually associated with prosodic primings through co-occurrences with collocates over time: “Prosodies are undoubtedly the product of a long period of refinement through historical change” (Louw, 1993, p. 164). Bublitz (1996, p. 13) agrees: “[W]e know from lexical semantics that constantly using a word in the same kind of context can eventually lead to a shift in its meaning: the word adopts semantic features from an adjacent item. The meaning of the latter ‘rubs off’ onto the meaning of the former.” The problem is that neither of them actually look at diachronic evidence, instead, they make claims about historical processes based on synchronic data. Whitsitt (2005) summarizes this issue as follows:

[N]o matter how long one looks at what is the synchronic use of a word like set in, there is no evidence for assuming that we can see the results of a diachronic process of *imbuing.*” (Whitsitt, 2005, p. 296)

Additionally, Louw and many other researchers in semantic prosody always talk about the collocates coloring the node word. There are of course no node words in language. A node word only exists as an analytical category. If language is an interconnected system, “un système ou tout se tient” as Meillet (1908, p. 407) puts it, should semantic transfer between co-occurring items not be bidirectional?
Another issue that Whitsitt discusses at length and criticizes heavily is Louw’s use of metaphor and analogy. Instead of formulating a clear-cut definition, Louw describes semantic transfer as a process of *imbuing* and semantic prosody as an *aura of meaning* (Louw, 1993, p. 157). The use of figurative language may be due to the somewhat elusive and intractable nature of semantic prosody, but on the other hand, it seems to perpetuate these issues. One example of analogy is particularly egregious: As Whitsitt (2005, p. 291) points out, Louw’s analogy of co-articulation (prosody in the Firthian sense) is not valid. Speech sounds only influence each other immediately and locally; co-articulatory effects do not carry over to other contexts. Vowels are nasalized in English when their articulatory gestures overlap with those of nasal consonants but they are not *instilled* with nasalization that then exists in other contexts. The vowels in Louw’s example of the word ‘amen’ are nasalized in this word but not in other words.

This almost makes one wonder if Louw misinterpreted or at least reinterpreted Sinclair’s use of the word ‘prosody’. Unlike Louw, Sinclair does not talk about semantic transfer. If we understand prosody not in the Firthian sense as co-articulation but instead in its far more common use as any suprasegmental feature, it becomes much more in line with Sinclair’s work. Sinclair emphasizes on multiple occasions that semantic prosody does not just belong to individual words; instead it serves a discursive function and underlies large units of linguistic expression, up to the level of the utterance: “the initial choice of semantic prosody is the functional choice which
links meaning to purpose; all subsequent choices within the lexical item relate back to the prosody” (Sinclair, 2004, p. 34). A similar concept is sometimes referred to as *evaluative harmony* or *evaluative cohesion* (cf. Partington, 2017). The analogy of prosody as a suprasegmental feature that stretches across an entire utterance may be a more apt description of at least one perspective on semantic prosody (referred to as the discourse perspective below). Ultimately though, Sinclair first coined and described semantic prosody in personal communication with Louw, so we have no clear way of knowing, how he intended the analogy to be interpreted. The question of semantic transfer is discussed in more detail in the context of a constructional account in chapter 5 and during the conclusion in chapter 9.

### 2.3.2 Semantic prosody and semantic preference

Semantic preference is a concept that features prominently throughout Sinclair’s work. It describes the tendency of lexical items to collocate with words from common semantic fields (also called lexical fields); e.g. the verb ‘cure’ shows a semantic preference towards the lexical field of illness. As Sinclair (1998, p. 16) puts it, “Semantic preference is the restriction of regular co-occurrence to items which share a semantic feature, for example, that they are all about, say, sport or suffering.” Some earlier research does not clearly distinguish between semantic prosody and semantic preference, e.g. Hunston (1995, p. 137) who provides the following description: “a word may be said to have a particular semantic prosody if it can be shown to co-occur typically
with other words that belong to a particular semantic set.” Similar statements are found in earlier works by Hoey and Sinclair.

Most research now makes an explicit distinction and considers semantic preference and prosody two different though closely related categories. An idea that is often repeated is that semantic prosody is like semantic preference but on the level of positive or negative evaluations. Bednarek (2008, p. 121) describes the two as “arguably very similar phenomena: the only difference is that [semantic prosody] is a more general semantic group than the groups usually listed in [semantic preference] studies.”

Partington (2004, p. 149) mentions different strategies to distinguish the two; in the one view, “semantic prosody is a sub-category, or a special case, of semantic preference, to be reserved for instances where an item shows a preference to co-occur with items that can be described as bad, unfavourable or unpleasant, or as good, favourable or pleasant.” He believes while this description “probably works as a rule of thumb in most cases” (ibid), it misses some fundamental differences between the two. An important point is that “prosody is at a further stage of abstraction than preference” (p. 150) and that the semantic prosody is to communicate evaluative and attitudinal meaning. My issue with these descriptions is that they are purely theoretical. In chapter 6, I argue that semantic prosody cannot be separated from semantic preference through corpus data, making its status as an independent category in its own right questionable. This point is revisited in my final evaluation of the concept in chapter 9.
2.3.3 Genre or register effects

It is often claimed that semantic prosodies are subject to a genre effect, i.e. they show different tendencies for evaluative polarity in different linguistic registers (Tribble, 2000; Fuentes, 2001; Stubbs, 2001; Partington, 2004; Nelson, 2006; Louw and Chateau, 2010; Wachter, 2012; ¨Unaldı, 2013, inter alia). Tribble (2000, p. 86) refers to this as ‘local prosodies’: “words in certain genres may establish local semantic prosodies which only occur in these genres, or analogues of these genres.” Partington (2004, p. 153) agrees that it is “highly likely that the quality and strength of the prosody of a good many items will differ from genre to genre or from domain to domain.”

Stubbs (2001, p. 106) for example notes that the adjective ‘lavish’ has a “neutral-to-good” semantic prosody in general, but a negative prosody in journalism. ¨Unaldı (2013) claims that ‘pose’ is more negative in academic discourse than in general English. Hunston (2007a) finds that the verb ‘cause’ is less negative in scientific registers. She argues however that while a register effect does indeed exist, the phenomenon is better explained in terms of involved participants and their semantic roles. She uses a similar line of argumentation as Louw (1993) in his example of ‘build up’. When humans occupy the patient position in causations, there is usually an unfavorable evaluation: Such situations often imply a lack of control and/or volition on the part of the human patient, which is typically not a desirable situation. This is however not the case when the patient is an inanimate or even abstract entity, as is typically the case in academic discourse, especially in the STEM fields. The ‘genre effect’ is
thus a secondary phenomenon and a function of the constellations of semantic roles that are usually employed in a particular register:

[T]he attitudinal meaning associated with CAUSE applies only when the ‘caused entity’ concerns animate beings, their activities and goals. Where the ‘caused entity’ is an inanimate object unrelated to human goals no attitudinal meaning is implied. If a register makes more use of the second phenomenon than the first, it will appear that in that register CAUSE has no attitudinal meaning. In other words, rather than suggesting that register can make attitudinal meaning appear or disappear we might argue that particular registers select one lexical phenomenon more frequently than another.” Hunston (2007a, p. 263)

Hunston’s explanation seems reasonable but things may be more complicated: Wei (2002) finds that ‘cause’ actually has a stronger negative prosody in academic English than in normal English (while ‘career’ has a weaker positive prosody in academia). Louw and Chateau (2010, p. 762) find that even in academic discourse ‘cause’ has an overwhelmingly negative prosody (although they do not compare it to normal English). The authors note that they used *The Corpus of Contemporary American English* or *COCA* (Davies, 2008), which contains various areas of study, not only hard science in its academic section. This seems to indicate that we need to carefully consider what constitutes a register. Yang and Chen (2016) investigate the verb ‘cause’ in three different academic fields. They find its semantic prosody to be overwhelm-
ingly negative in social science (96%) and applied science (86%) but almost neutral in natural and pure science (54%). Different areas of science have widely different styles of writing so it is not surprising that different that they would show register effects between them. This does not contradict Hunston’s account of register differences being explainable in terms of the involved entities. It makes sense that social sciences deal more with humans while natural sciences deal more with inanimate objects.

2.3.4 The hidden nature of semantic prosody

A major theme in the research literature on semantic prosody is the claim that it is hidden from speakers’ intuitions. The idea is that speakers are not explicitly aware of the fact that words like ‘cause’ or ‘happen’ tend to occur with negative collocates while words like ‘provide’ or ‘attain’ are more likely to occur with positive collocates. Semantic prosody is believed to be inaccessible to introspection; it is considered to be a phenomenon that only surfaces in corpora. Louw (1993, p. 173) for example believes that “for thousands of years [semantic prosodies] remained hidden from our perception and inaccessible to our intuition.”

There are several problems with the concept of hiddenness. There is very little research into this question. Most researchers seem to take its validity for granted based on the idea that they lack reliable intuitions themselves. Additionally, it is not clear if semantic prosody is always hidden. One important question is whether it is a property of all words or just some. In other words, does it make sense to talk
about the semantic prosody of a word like ‘horrible’ or is *semantic neutrality* (Stewart, 2009b) necessarily required? There is nothing in its definition that excludes obviously evaluative words from possessing semantic prosody.

This issue rarely discussed, and few studies deal with clearly evaluative lexis. Some exceptions include Channell (2000) who discusses the word ‘self-important’ and Stubbs (2001) who mentions ‘reckless.’ Lewandowska-Tomaszczyk (1996, p. 157) is a rare example of a researcher who explicitly addresses this issue: “I would see no reason […] not to use the term semantic prosody for the cases when the semantic load of an item is quite explicit.” Beyond this, a review of the case studies literature certainly conveys the impression that semantic prosody is a feature of ‘seemingly neutral’ words. The main reason seems to be that hiddenness is a major *selling point* for semantic prosody; it is part of what makes it intriguing and what draws researchers from fields as diverse as critical discourse analysis and second language teaching to it. As Stubbs (2001, p. 198) puts it, “evaluative meanings are often inexplicit, less clear-cut and, at least sometimes, deniable […] [which] makes them important to study for practical purposes.” Depending on what words we allow in the analysis of semantic prosody, it moves from being hidden to being on a cline of obviousness. I dedicate chapter 7 in its entirety to the question of hiddenness.
2.3.5 The lexical priming perspective and the discourse perspective

The prior research presented here so far can be broadly categorized based on the two different facets of semantic prosody it is focused on. On the one hand, it can be described as a distributional phenomenon. On the other hand, it has a communicative function: It is used to evaluate, or express attitudes or stances and it plays an important role in establishing *evaluative harmony, consistency or cohesion* in discourse. As Hunston remarks, “*A spell of fine weather set in* sounds very odd” (Hunston, 2002, p. 142). This is what Louw (1993) calls a prosodic clash. It is due to the evaluative inconsistency in this sentence: ‘Fine weather’ clashes both with ‘a spell of’ and ‘set in’.

The distinction between these two aspects of semantic prosody has led to some disagreement and confusion in the literature. It has been couched in the question “where precisely semantic prosodies reside” Morley and Partington (2009, p. 144). Hunston (2007a), (p. 251) claims that there are two opposing camps: One camp that sees it as a discourse function and another camp that interprets it within the lexical priming perspective and considers it to be a matter of co-occurrences (and thus a property of the node word). She mentions Sinclair (2004) and Stubbs (2001) as examples for the former position and cites Partington (2004) as exemplary for the latter position – a characterization with which Morley and Partington (2009, p. 144) explicitly disagree. To them, this is a false dichotomy and both views are just different perspectives on the same phenomenon; both views emphasize different
aspects of semantic prosody but are by no means mutually exclusive. The lexical priming perspective describes how words locally acquire semantic prosody and the discourse perspective describes how semantic prosody is used. They interact, both in that discourse sets the wider context within which words gain prosodic meaning, and in that lexical primings shape discourse through their lexical restrictions.

There is clearly some confusion in the literature, even between eminent researchers such as Hunston and Partington. Semantic prosody is a complex and nuanced issue and unfortunately, its descriptions are often vague and, as Whitsitt (2005) critically points out, make use of metaphorical language that seems to obscure rather than elucidate the issue. Some authors, including Partington (2004) adopt both the lexical priming and the discourse perspective but tend to blur the lines instead of explicitly distinguishing between them. Hunston can hardly be faulted for misrepresenting some viewpoints within a literature that requires the detailed exegesis of research articles, and she is certainly correct that some researchers focus more on one aspect than on the other.

Typically, researchers do not define what exactly they mean when they talk about what semantic prosody is a feature of or where it resides. One particularly problematic description is that some researchers assign semantic prosody to the larger text unit. It is often not clear what exactly they mean and there are at least three ways in which such statements can be interpreted. First, there is Sinclair’s notion of the lexical item, which he considers to stretch beyond the level of a single word. For example,
Sinclair (1996) discusses the semantic prosodies and preferences of idiomatic multi-word phrases such as ‘the naked eye’ or ‘true feelings’. This perspective still seems very similar to the lexical priming perspective, it is just slightly expanded to include idioms and not just words.

Second, some researchers also refer to the attitude or evaluation conveyed in an individual sentence as the semantic prosody of that sentence. This use of the term can be confusing and will be avoided here. I only refer to semantic prosody as a tendency that has been established over several related instances. Individual sentences can convey evaluative meaning but do not possess a semantic prosody. In some cases, what the authors, who use such descriptions, may intend to say is that semantic prosody is found in a sentence or that it operates upon a sentence in its function of creating evaluative harmony. This simply puts this type of description into the camp that emphasizes the discourse function of semantic prosody. Again, accounts of semantic prosody can be vague and open to interpretation.

Third, there is the issue of how far the relevant context of the lexical item goes and what type of context should be considered. Ellis et al. (2007) and Hauser and Schwarz (2016) only study collocates within a narrow search window, while Sinclair (1998) tends to analyze items in their entire context, e.g. pointing out that ‘budge’ tends to occur in negated clauses, carrying a sense of difficulty. This is a question with practical implications, that is, at what level semantic prosody should be interpreted, which is especially relevant if the local evaluation differs from the attitude conveyed.
in the context of the larger utterance. This question is further discussed in the next
section.

2.3.6 The relevant level of analysis

Broadly speaking, there are two different ways to analyze semantic prosody, first by
reading concordances, and second, by looking at collocational displays. Reading con-
cordances means analyzing semantic prosody within entire sentences or utterances
and describing it based on the overall attitude expressed in these contexts. Col-
locational displays on the other hand only involve looking at the most frequently
co-occurring words. The contrast between these two approaches is particularly pro-
nounced when there is a difference between the evaluation expressed in the sentence
and the evaluative valence of the collocates. A simple example is negation, e.g. ‘he
did not cause any problems’. The overall attitude of this statement does not convey
that anything bad happened, but a collocational analysis would only pick up on the
word ‘problem’. I could be argued that this is not necessarily a problem. Studies
that compare the data of psycholinguistic experiments with corpus data, such as Ellis
et al. (2007) or Hauser and Schwarz (2018) often only define semantic prosody in
terms of collocations. They see it as a local phenomenon and assume that a speaker’s
mental representation of a word’s semantic prosody is an abstraction that averages
the evaluative polarities of its collocates. We do not know whether or not that is
the case. A different possibility may be that speakers also remember in what kind of
overall evaluations words (or larger linguistic items) occur.

The great majority of all work on semantic prosody focuses on interpreting it in its entire evaluative context, not just as a local phenomenon (Hunston, 2007a; Bednarek, 2008; Morley and Partington, 2009, inter alia). A common argument concerns words that are in a sort of contrastive relationship with their collocates, e.g. ‘cure’ or ‘alleviate’. They tend to collocate with illnesses, but they would not be described as negative because they describe easing or overcoming these illnesses. This leads to a related question: Why do these words not take on the negative prosodies of their collocates? This question has been posed by several researchers, including Stewart (2009b), who points out:

“[W]e need to establish why happen, break out, etc. have proved vulnerable to the evil influence of their bad co-occurrences and have thus acquired a bad ‘aura’, while alleviate and relieve are to all intents and purposes incorruptible.” [Stewart (2009b, p. 74)

Whitsitt (2005) takes up this issue to argue against Louw’s notion of a semantic transfer, noting that at the very least, it cannot be a general principle:

“One need but consider verbs like alleviate, heal, relieve, soothe, etc., all perfect candidates for semantic prosody since they all habitually appear in the company of clearly unpleasant words, yet it seems clear that a word like alleviate, to take one example, certainly does not come to have an unpleasant meaning because of that company.” (Whitsitt, 2005, pp. 296f)
Bednarek (2008, p. 130) makes a similar point about the verb ‘tackle’, noting that it collocates with items describing difficult problems or tasks. It does not have a negative connotation; if anything, ‘tackling problems’ is a good thing. The key to this issue appears to be the type of semantic relationship between the node word and its collocates.

Hunston (2010, pp. 57ff) explains the above contradictions by identifying three kinds of co-occurrence relationships, exemplified by ‘terrible’, ‘alleviate’ and ‘cause/bring about’. ‘Terrible’ collocates with words like ‘mistake’ or ‘pain’. It confirms the polarity of its collocates. ‘Alleviate’ on the other hand reverses the polarity. ‘Cause’ and ‘bring about’ are “associated with polarity but might not be identified as evaluative out of context” (p. 58). They do not necessarily impose their own evaluation, but they do not contradict or reverse the evaluation of their collocates. Partington (2004) makes a similar case.

“Simply being primed to appear in the environment of collocates of a certain evaluative sense, good or bad, is not a sufficient condition for an item to acquire the same sense. If the relationship between the item and its collocates is one of opposition or detraction, then the combination does not acquire the evaluative sense of the collocates. For example, relief and ease, which are inherently favourably evaluative, collocate very frequently with unfavourable words like debt, pain, poverty, suffering and so on. The overall meaning of “the relief of suffering”, “easing the pain” remains, of
course, highly favourable.” (Partington, 2004, pp. 154-155)

There is a third way of dealing with the differences between utterance-level evaluations and collocational polarity: defining them as two different things. Semantic prosody is then defined as the positive/negative evaluation found in the semantic fields a word collocates with. It is part of the compositional evaluative meaning of the larger utterance, but it is not the same (and it can be reversed). This position is sometimes hinted at but is relatively rare. In some appear to be a contradiction of some of her earlier descriptions, Hunston (2010, p. 58) discusses the evaluative reversal through words like ‘alleviate’ to argue that “semantic prosody should not be interpreted as an automatic ascription of evaluative polarity.”

I will revisit the issues raised in this section at two different points: First, in chapter 6 when I argue against an understanding of semantic prosody in terms of evaluative polarity and, second, in chapter 7 when I discuss research into the proposed psychological reality of semantic prosody.

2.4 Discussion

The research on semantic prosody can roughly be divided into five categories or areas of study. During the initial phase which includes Sinclair (1987), Louw (1993), Stubbs (1995a), Bublitz (1996) and Partington (1998), the phenomenon is first described through a few relatively small case studies, and fundamental theoretical claims are put forward. In the second phase articles like Whitsitt (2005), Hunston (2007a), Morley
and Partington (2009) and Louw (2000) problematize some of these theoretical issues. It becomes clear that there are some areas of disagreement and confusion. At the same time, there is a third category of research that does not deal with theoretical issues but instead applies semantic prosody to other linguistic fields. Much research is focused on translation studies and second language learning as in Dam-Jensen and Zethsen (2008), Zhang (2009) and Guo et al. (2011). Closely related to this is a fourth group that studies semantic prosody in a cross-linguistic context, typically contrasting two languages, e.g. Tognini-Bonelli (2001) and Xiao and McEnery (2006). Finally, a fifth and very small group of studies aims to investigate the psycholinguistic basis for semantic prosody, e.g. Ellis et al. (2007), Ellis and Frey (2009), Hauser and Schwarz (2016) and Hauser and Schwarz (2018).

As I have described in this chapter, there are still several open questions, areas of confusion and disagreement concerning the treatment of semantic prosody. Despite the extensive amount of research literature that exists, this suggests that semantic prosody has not been adequately explained so far. As I will argue in this dissertation, the most glaring issue is that researchers have considered the data solely from a word-centered perspective, i.e. they only study the collocational associations between words and do not take any differences between grammatical patterns into account. There are in fact examples in the literature where researchers have come across obvious constructional differences but have essentially chosen to ignore them and find a way to attribute them to the lexical level instead, as in Louw’s (1993) example of ‘build
up’ or Bublitz’s (1996) account of ‘happen’. As I argue from chapter 5 onward, as the data show, these examples are not the exception but the rule.

Throughout this dissertation, many of the intractable issues surrounding semantic prosody, such as the semantic interplay between collocates, their relationship to the wider context or the question at what level evaluative meaning should be analyzed and to what types of linguistic units it should be ascribed are recontextualized within a constructional framework. I argue that these problems are much more readily solved within this framework. Moreover, semantic prosody is accounted for as an ordinary part of language and is not the odd and elusive issue that it is sometimes portrayed as.

Before I can make this argument, I first need to outline the theoretical and methodological foundations. In the next two chapters, I introduce both the relevant theoretical background and the necessary methodological tools needed to tackle these questions. In the following chapters, I deal with the main issues one after the other. In chapter 5, I propose a constructional account. In chapter 6, I argue that the evaluative semantics conveyed by patterns of co-selection cannot be broken down into a dichotomy of positive and negative evaluations. In chapter 7, I investigate the concept of hiddenness based on corpus evidence and data from experiments. In chapter 8, I aim to contribute to the question of cross-linguistic equivalence of semantic prosody with a contrastive study of English and German. In chapter 9 I synthesize all these findings and arrive at a final description of semantic prosody.
Chapter 3

Theoretical Foundations

In this chapter, I provide a summary of the theoretical foundations that my work in this dissertation is based on. The chapter is divided into three sections, comprising usage-based linguistics, construction grammar and research into evaluation and stance-taking. For each topic, I first outline the basic tenets required for a basic theoretical understanding; then I focus on specific ideas that are of particular importance for my work. I point out relevant similarities and differences between these approaches and Sinclairian text linguistics. Where needed I cite empirical research as evidence for the statements made in the section.

In section 3.1, I provide a description of usage-based linguistics, an approach to linguistic inquiry that posits that linguistic structure emerges through language use. Important ideas that are emphasized in this section include context-dependent meaning, frequency effects, and the Firthian maxim.

In section 3.2, I present the fundamentals of construction grammar, a family of theories that regards language as a system of form-meaning pairings. I describe the overlap between usage-based and constructional approaches and focus on the idea that grammatical constructions carry meaning independent of their lexical instantiations.

In section 3.3, I briefly review work on evaluative language. I clarify the termi-
nology that will be used in the remainder of this dissertation and describe how it relates to semantic prosody and its communicative function. I also explain why I avoid describing evaluative meaning as peripheral or connotative.

3.1 Usage-based linguistics

The usage-based model of language sees linguistic structure as the conceptual organization of language which emerges from and is continuously redefined by speakers’ experience with language use (Hopper, 1987; Givón, 1995; Barlow and Kemmer, 2000; Tomasello, 2003; Langacker, 2008; Bybee, 2010). It postulates a dynamic, two-way relationship between language structure and language use: Every time a speaker engages in any linguistic interaction, their knowledge of language structure (or grammar) informs their language use; in turn, every exposure to language in use shapes their linguistic mental representations.

As such, linguists working within this paradigm reject the strict separation of grammar and usage emphasized in other linguistic traditions, as seen in the Saussurian dichotomy of langue and parole or the Chomskyan competence/performance distinction. Consequently, the concept of what constitutes linguistic evidence is also a different one. Other approaches eschew studying language usage as too noisy, too unsystematic and ultimately irrelevant, and instead consider the introspection of a trained researcher to be the gold standard of linguistic evidence. Usage-based approaches instead believe that the study of authentic language data can yield valuable
insights into the structure of language as it reveals how the latter emerges from the former.

A fundamental construct of the usage-based model is Langacker’s notion of the usage event which Langacker (2000, p. 9) defines as “an actual instance of language use” and “an utterance characterized in all the phonetic and conceptual detail a language user is capable of apprehending” (ibid). Usage-based approaches postulate that language structure arises in a bottom-up manner, meaning that “usage events are the source of all linguistic units” (Langacker, 2008, p. 220). Tomasello (2000, pp. 61f) describes this as follows: “In usage-based models of language [. . .] all things flow from the actual usage events in which people communicate linguistically with one another. The linguistic skills that a person possesses at any given moment in time – in the form of a ‘structured inventory of symbolic units’ – result from her accumulated experience with language across the totality of usage events in her life.”

Many usage-based theories emphasize the importance of frequency and repetition on the emergence of linguistic categories in the speaker’s mind. Bybee (2006, p. 1) summarizes this as follows:

“A usage-based view takes grammar to be the cognitive organization of one’s experience with language. Aspects of that experience, for instance, the frequency of use of certain constructions or particular instances of constructions, have an impact on representation that is evidenced in speaker knowledge of conventionalized phrases, and in language variation and
Langacker (2000) defines the role of frequency in terms of *entrenchment*, which he describes as a process of ‘routinization’ or ‘automatization’ giving rise to linguistic units:

“Through repetition, even a highly complex event can coalesce into a well-rehearsed routine that is easily elicited and reliably executed. When a complex structure comes to be manipulated as a ‘prepackaged assembly’, no longer requiring conscious attention to its parts or their arrangement, I say that it has the status of a unit.” (Langacker, 2000, pp. 3f)

An important prerequisite to Langacker’s notion of entrenchment is his concept of the *rule/list fallacy*. Langacker describes it as an important version of the *exclusion-ary fallacy* which more generally states that “one analysis, motivation, categorization, cause, function, or explanation for a linguistic phenomenon does not necessarily preclude another” (Langacker, 1987, pp. 28f). The rule/list fallacy specifically (p. 29) states that any linguistic item that can be fully accounted for through regular rules of grammar cannot be part of the mental lexicon. According to the fallacy, plural nouns such as ‘shoes’ or ‘walls’ that can be derived by the general noun-pluralizing rule of English would not be listed in an ‘optimal grammar’. But as Langacker points out, this is a ‘gratuitous’ assumption that does not necessarily correspond to the psychological reality of language. There is no reason, why the two could not coexist.
It is possible for fully regular forms to reach unit status through entrenchment and thus be represented in the lexicon. This means that the mental lexicon does not only consists of single words or lexical primitives but also redundantly stored prefabricated items.

As mentioned before, Sinclair, while not associated with usage-based linguistics, shares this view: “a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analyzable into segments” (Sinclair, 1991, 110, emphasis mine). I will add that there is no reason to believe that our organization of language is driven by a principle of parsimony of grammatical formalisms; in fact, the realities of language use and language acquisition (briefly discussed below) make such an assumption exceedingly unlikely.

3.1.1 The primacy of context over isolated ‘lexicon meaning’

A corollary of the idea that linguistic form follows usage is the primacy of contextual meaning. Words are not encountered in isolation but in the context of usage events. They are learned in their contexts and with their contexts. As such, it is not the case that they secondarily acquire contextual meaning; instead, the contextual variants should be considered as primary. The lexical semantics of a word thus emerge when a speaker abstracts away from the differences of its context-dependent variants. Langacker summarizes this view as follows:
“A lexeme is not precisely the same in all its environments. Since elements are always shaped by the contexts in which they occur, it is only by abstraction away from contextual variation that a constant representation emerges. [. . . A] variant enters into a kind of ‘ecological system’ with its structural context and does not necessarily exist outside that habitat. I am suggesting that these context-dependent variants may be more fundamental than the context-neutral schematizations we regard as primary. In a ‘bottom-up’ account of this sort, the polysemy of lexical items should be expected as the normal state of affairs. Whether the contexts are structural, collocational, or pragmatic, they inevitably shape the construal of symbolic structures and thus give rise to semantic variants.” (Langacker, 2000, p. 35)

This is a crucial point that centrally informs my treatment of semantic prosody. My dissertation aims to move away from a lexical, or word-centered, perspective that considers the context-neutral word as the primary unit of meaning, and move towards an account that studies the semantics of context-dependent units, in this case lexical items within their constructional context as primary.

Since it is impossible to study any speaker’s experience with language in its totality, usage-based linguists instead turn to the study of authentic linguistic data as an approximation of the former. Text corpora containing collected instances of language use serve as a proxy of a generalized speakers’ language experience. This encom-
passes any level of analysis from very fine-grained and in-depth studies of language in interaction on the micro-level (e.g. Ford, 2005; Geluykens, 1992) to large-scale, computational studies of huge data sets on the macro-level (e.g. Baayen, 2002; Manning and Schütze, 1999). Large quantitative corpus studies often follow a frequentist approach were counts of linguistic items in text stand in for the frequencies speakers are exposed to in ambient language. Halliday famously summarizes this idea as such: “frequency in text instantiate[s] probability in the system” (Halliday, 1993, p. 3).

This leads to another important point: In usage-based linguistics, much of grammar is not strictly categorical but instead probabilistic. It should be noted however that while the existence of a relationship between usage frequencies and linguistic structure is well-supported by evidence, the exact nature of this relationship is not entirely clear; e.g. probabilities and continuums in the real world can lead to categoricity in the conceptual system (cf. Kronrod et al., 2016, for a review of categoricity effects).

### 3.1.2 Distributional linguistics and the Firthian maxim

The notion of the *distribution* of a linguistic item is an important concept in quantitative approaches to the study of language. Harris, an early proponent of distributional linguistics describes the distribution of a linguistic item as “the sum of all its environments” (Harris, 1954, p.146). An important idea in this context is the *distributional hypothesis*, which states that the meaning of a linguistic item can be understood through its distribution in language use. The most famous formulation of this idea
is probably the Firthian maxim, “You shall know a word by the company it keeps” (Firth, 1957, p. 11). In a similar vein, Landauer and Dumais (1997, p. 218) state: “a representation that captures much of how words are used in natural context will capture much of what we mean by meaning.”

An important corollary of this is that distributional properties can be used as a measure of synonymy. The basic idea is that there is a principled relationship between semantic similarity and distributional similarity. What makes this relationship so attractive is that it provides corpus linguists with a quantitative way of studying semantics through distributional properties. This point has been made by countless researchers; typical statements include “words which are similar in meaning occur in similar contexts” (Rubenstein and Goodenough, 1965, p. 627) or its reverse form, “words that occur in the same contexts tend to have similar meanings” (Pantel, 2005, p. 126). There is one important caveat: While distributional linguistics can yield powerful insights into semantics, things are perhaps not quite as straightforward as some researchers make them out to be. The two descriptions by Rubenstein and Goodenough (1965) and Pantel (2005) are not synonymous: The statement \( \text{similar meaning} \rightarrow \text{similar distribution} \) is not logically equivalent to the statement \( \text{similar distribution} \rightarrow \text{similar meaning} \). It is conceivable that while all semantically similar words have similar distributions in text, not all words that have similar distributions in text are also semantically similar. This is not simply theoretical conjecture either – such examples are attested: The two verbs ‘alleviate’ and ‘exacerbate’ have very
similar collocational profiles and while their meanings are clearly related in some way, they are not at all similar (at least not if we understand similarity as near-synonymy), but instead constitute near-antonyms. There does not appear to be any in-depth research into this issue, so it is unclear how common such word pairs are. None of this, of course, invalidates the basic tenets of distributional linguistics – it just means that more care in the interpretation of the data is needed. Instead of positing that similarity in textual distribution indicates similarity in meaning, it is perhaps safer to say that it hints at semantic relatedness of which similarity is one common expression.

3.1.3 Evidence for usage-base models

The usage-based approach is very well supported by evidence, both through the success of distributional models of meaning in computational linguistics and through experimental evidence in psycholinguistics.

There is a preponderance of evidence for speakers’ sensitivity to frequency effects and distributional information at many levels of linguistic structure. The phonological effects of frequency are well-documented. One example is the reduction effect, meaning that highly frequent linguistic items undergo phonetic reduction faster than items with lower frequencies. This effect is found both for specific high-frequency phrases (e.g. Bybee and Scheibman (1999) on the reduction of “don’t” as in “I don’t know” to “I dunno”) and in gradual processes of sound change which operate on all
words but to different degrees depending on their frequencies (e.g. Bybee (2000) on the word-final t/d-deletion in American English). Gahl (2008) studies the effect of frequency on duration and shows that highly frequent words (like ‘time’) are shorter than their less frequent homophones (like ‘thyme’). Empirical evidence also shows that acceptability judgments are shaped by frequencies: Frisch et al. (2000) show that participants judge nonce words (constructed words that do not exist but that could exist because they do not violate any phonotactic rules) as more ‘wordlike’ when they instantiate highly common phonotactic patterns.

The fact that highly frequent words are privileged in word retrieval has been well-documented in psycholinguistics for several decades. In written language, high frequency improves speed in both word naming tasks (Forster and Chambers, 1973) and in lexical decision tasks (Forster, 1976). In spoken language, Savin (1963) shows that word frequency facilitates the accurate identification of words in noise. Inhoff and Rayner (1986) and Rayner and Duffy (1986) study word recognition in eye-tracking experiments and that high-frequency words are recognized more quickly.

Frequency effects are also extremely well documented beyond the word level, showing that speakers keep track of co-occurrences between words and grammatical structures. Arnon and Snider (2010) showed participants four-word sequences and asked them whether they were legal combinations in English. They found that the token frequencies of the sequences were predictive of reaction times (and this effect could not be reduced to the frequencies of the individual words). The authors concluded that
the mental lexicon includes multi-word phrases along with information about their frequencies of occurrence. In an eye-tracking study, Garnsey et al. (1997) showed that speakers use knowledge about the probabilistic argument structure preferences of verbs to resolve temporary ambiguities. Reali and Christiansen (2007) conducted a self-pacing reading task to show that frequent combinations of relative clause type and pronoun type were processed more quickly than less common combinations.

Another effect of token frequency that has been termed the *conserving effect* provides support for the notion of entrenchment: We know that highly frequent forms are more entrenched in their morpho-syntactic structure because they tend to resist change; e.g. high frequency English verbs tend to be irregular (e.g. Hooper, 1976). Langacker’s argument about the rule/list fallacy and his idea that entrenched linguistic forms acquire unit status is also supported by empirical data: Stemberger and MacWhinney (1986) and Stemberger and MacWhinney (1988) study speech errors both in spontaneous speech and in experimental settings and show that even fully regularly inflected word forms are likely stored as separate entries. Hay (2001) investigates under which conditions morphologically complex words are either decomposed into their component parts or treated as non-compositional units. She finds that the best predictor is not absolute frequency but rather the relative frequency of the derived form and the base form. Entrenchment is also supported by research in construction grammar (see 3.2). Linguistic structure is learned through lexically instantiated constructions that are first acquired as complete sequences and then
generalized and further analyzed into their component parts. This supports the idea of the co-existence of redundantly stored linguistic information on different levels of analysis.

Another source of evidence for the usage-based approach to grammar comes from studies that describe how grammar can be learned from ambient language. There are two basic categories of studies here, those that show that language use contains sufficient data to derive structure and those that investigate how children acquire language. In the first category, there are numerous corpus studies that attest the informational richness that is found in the distributional data of spoken language. For example, Redington et al. (1998) conduct several corpus experiments to test whether syntactic categories can be entirely derived from language use without any a priori information (i.e. through some sort of innate grammar). They study the distribution of the 1,000 most frequent words in the CHILDES database (a database for studying child language acquisition, consisting of transcripts of interactions between young children and their caretakers or friends, cf. MacWhinney, 2000) and used cluster analysis to categorize them. The clusters that emerged were very similar to commonly acknowledged word classes such as nouns, verbs, adjectives and prepositions. Studies by Mintz et al. (2002) and Monaghan et al. (2005) arrive at similar results. These studies show that even a particularly noisy corpus that is full of false starts and ungrammatical sentences provides enough distributional information for an unsupervised learning algorithm without any additional input to be able to determine
grammatical structure.

Corpus studies show that raw language is rich in information and that this information is sufficient to derive linguistic structure, thus refuting any nativist arguments about a *poverty of the stimulus* argument. However, they can of course not show if and how speakers actually derive linguistic structure from language use (machine learning may be very different from human learning). There are, however, psycholinguistic studies that show speakers’ sensitivity to frequency effects and distributional patterns. Saffran et al. (1996) show that 8-month old infants use statistical learning to accomplish word segmentation in continuous speech through transitional probabilities between speech sounds. Levelt et al. (2000) find frequency to be an important factor in the acquisition of syllable structures in Dutch children. Zamuner et al. (2004) demonstrate children’s production of nonce words is sensitive to phonotactic frequencies. There is also evidence that both frequencies and context play a role in the generalizations children make. Brooks et al. (1999) found that children are less likely to use verbs they had only encountered in one construction in other constructions if those verbs were frequent. In other words, when frequent verbs were limited to one construction children used this as negative evidence that they could not be extended to other constructions. Less frequent verbs were more readily used in constructions they were not attested in.
3.2 Construction Grammar

Construction grammar is a family of linguistic theories that posit that all linguistic knowledge can be represented through a system of form-meaning pairings, i.e. constructions. It considers constructions to be the primitives or building blocks of language and is thus radically different from and fundamentally opposed to more traditional views of language that see it as a system of words and grammatical rules.

Most constructional approaches are not competing theories to usage-based linguistics. They could be considered two complimentary and closely intertwined approaches or two facets of the same approach. Goldberg’s *Construction Grammar* (Goldberg, 1995; 2006) and Croft’s *Radical Construction Grammar* (Croft, 2001) make explicit commitments to a usage-based account of language. Langacker’s *Cognitive Grammar* (Langacker, 1987; 1991) is an inherently usage-based and constructionist framework. At the same time, research within usage-based linguistics often arrives at the notion of the construction (Tomasello, 1998; Bybee, 2010, inter alia). A usage-based approach is, however, not a prerequisite to construction grammar and, in fact, not all constructional theories, are usage-based. Diessel (2015) describes the relationship between usage-based linguistics and construction grammar as follows:

“Construction grammar has played an important role in the development of the usage-based approach. In fact, in the literature construction grammar is often described as an integral part of the usage-based approach to the study of grammar […] but the notion of construction grammar
refers to a whole family of theories which are not all usage-based […]. Indeed, one of the earliest and most influential construction-based theories, i.e. the sign-based theory of construction grammar developed by Fillmore and Kay (1999), adopts the generative division between competence and performance and disregards usage and development […].” (Diessel, 2015, p. 3)

Usage-based linguistics and most construction grammars are non-nativist, i.e. they do not posit a detailed innate grammar but instead recognize the bottom-up manner in which linguistic structure emerges from language use. According to Langacker (2008, p. 220), all linguistic units are derived from usage events through the processes of *schematization* and *categorization*: “The relationship between units and the usage events that spawn them is tightly constrained by the content requirement, [according to which] units are limited to structures that arise from usage events through two basic cognitive processes: schematization and categorization.” Schematization is an abstraction across multiple experiences, forming mental representations that filter out difference while retaining similarities. As Langacker (2008, p. 17) puts it, schematization is “the process of extracting the commonality inherent in multiple experiences to arrive at a conception representing a higher level of abstraction.” Categorization is a process of classification based on comparing novel information against existing knowledge: “Categorization is most broadly describable as the interpretation of experience with respect to previously existing structures. A category is a set of elements
judged equivalent for some purpose; for example, the alternate senses of a lexical item constitute a category, equivalent in having the same phonological realization” (ibid).

Through these two processes, language structure is derived from usage events. As Perek (2015, p. 6) puts it: “In usage-based approaches, grammar is commonly seen as a vast inventory of symbolic conventions that are extracted from full-fledged utterances through a gradual process of schematization, retaining the syntactic and semantic commonalities across different usage events.” Constructions emerge in a bottom-up fashion when speakers encounter repeated forms of similar linguistic material and then abstract away from them. The noun phrase construction emerges as a generalization over different nominal constructions; the involved word classes are not assumed to be primitives or innate categories but also arise through schematization and categorization. Other categories emerge in a similar way such as the English category of subject which is an abstraction over the agentive roles that occur in various clause-level constructions such as the transitive construction and the ditransitive construction as well as the sole role in intransitive clauses.

In Goldberg’s highly influential Construction Grammar\(^1\), constructions are originally defined as any linguistic form that is not “strictly predictable” from knowledge of other constructions (Goldberg, 1995, p.4). Later versions expand this definition to include fully predictable patterns if they occur with ‘sufficient frequency’ – an idea

\(^1\)I use the term construction grammar to refer to any kind of constructional approach to the study of language. I reserve the term Construction Grammar (with both words capitalized) to refer to Goldberg’s framework specifically.
clearly related to Langacker’s notion of the entrenched linguistic unit as a ‘prepackaged assembly’ and the rule/list fallacy:

“Any linguistic pattern is recognized as a construction as long as some aspect of its form or function is not strictly predictable from its component parts or from other constructions recognized to exist. In addition, patterns are stored as constructions even if they are fully predictable as long as they occur with sufficient frequency.” (Goldberg, 2006, p.5)

Construction grammar aims to account for linguistic knowledge in its entirety, so constructions are posited at any level of complexity and abstraction. This ranges from relatively small and simple linguistic items such as morphemes and words to idioms (both completely lexicalized and partially filled) and even further to argument structure constructions such as the ditransitive construction (cf. Goldberg, 2006, p. 6; Croft, 2001, pp. 17f). As Goldberg (2006, p. 18) puts it, “it’s constructions all the way down.”

Both usage-based linguistics and construction grammar are maximalist approaches that require speakers to store a great amount of item-specific knowledge about the linguistic structures they encounter, such as their frequencies of usage and their contexts of occurrence. As Gurevich et al. (2010) point out, this seems to contradict the conventional wisdom that people are not able to remember sentences verbatim but instead only recall the ‘gist’ of what is said. Sachs (1967, p. 437) for example claims, “the original form of the sentence is stored only for the short time necessary for
comprehension to occur.” Gurevich et al. (2010) criticize earlier studies that found a lack of verbatim memory on the basis that they compared *memory for meaning* with *memory for structure*. Memory for meaning is always superior but this does not indicate a lack of structural or verbatim memory. Quite on the contrary, Gurevich et al. (2010) prove the existence of explicit verbatim memory for language in a series of experiments and conclude that “learners retain remarkably detailed memories for language, over which generalizations and abstractions may readily be formed” (ibid, p. 72). Similarly, Sosa and MacFarlane (2002) study collocations involving the preposition ‘of’ and find that more frequent collocations are more quickly identified by study participants, indicating that for frequent collocations the entire sequence is ‘holistically’ stored in the mental lexicon.

It should be noted that a verbatim memory of language does not seem to be a necessary assumption for the usage-based paradigm. Instead, exemplars could be updated when linguistic items are encountered and then stored as implicit knowledge. Langacker (2008, p. 220) emphasizes that the process of schematization only retains a small portion of the information contained in usage events: “Since only recurring features are reinforced, the units that emerge are far less comprehensive and detailed than the usage events giving rise to them.” Nonetheless, the existence of explicit memory for structure does lend further credence to the idea that the mental lexicon contains rich information about linguistic structures.
3.2.1 The conceptual basis of grammar

While construction grammar rejects the existence of innate grammatical categories, it is certainly not an anti-cognitivist or behaviorist paradigm that considers language as a disembodied, formal system. Instead, many constructional approaches emphasize the importance of a psychologically plausible account of language that is informed both by fundamental processes of human cognition and by the goal of language to conceptualize and communicate common human experiences. A core tenet of Langacker’s work is that grammatical structure is directly derived from a conceptual basis (e.g. Langacker (2006), but throughout his work). A similar concept is found in Fillmore’s theories of Case Grammar (Fillmore, 1968; 1977a) and Frame Semantics (Fillmore, 1977b; 1985), which can be seen as precursors to construction grammar:

“The case notions comprise a set of universal, presumably innate, concepts which identify certain types of judgments human beings are capable of making about the events that are going on around them, judgments about such matters as who did it, who it happened to, and what got changed” (Fillmore, 1968, p.24).

Fillmore (1977a, p.59) distills this definition into the slogan: “Meanings are relativized to scenes.” Drawing on Fillmore’s ideas, Goldberg claims that basic clause-level constructions reflect typical “humanly relevant scenes” (Goldberg, 1995, p. 39). She summarizes this concept in her scene encoding hypothesis:
“Constructions which correspond to basic scene types encode as their central senses event types that are basic in human experience.” (Goldberg, 1995, p. 39)

Goldberg goes on to list some examples of events or scenes:

“Languages are expected to draw on a finite set of possible event types, such as that of someone causing something, someone experiencing something, something moving, something being in a state, someone possessing something, something causing a change of state or location, something undergoing a change of state or location, and something having an effect on someone” (Goldberg, 1995, p. 39).

According to Goldberg, these basic and abstract concepts are so fundamental to the human experience that they are linguistically privileged through their expression in basic clausal constructions, whereas concepts like oversleeping do not possess their own distinct sentence types. A superficially similar concept is also found in Sinclair’s work who believed that the existence of idiomatic, semi-preconstructed phrases “may reflect the recurrence of similar situations in human affairs” (Sinclair, 1991, p. 10). There is however one big difference – Sinclair does not operate on the level of grammatical constructions. He is talking about idioms being used to describe recurrent scenes, not basic clause-level constructions.
3.2.2 The meaningfulness of constructions

An important claim of construction grammar that sets it apart from most other approaches to language is that constructions themselves carry meaning. In other linguistic theories, language is accounted for through a lexicon that contains the meaning component, and a set of grammatical rules that provide syntactic structure, i.e. that syntagmatically relate the lexical items to each other. I will refer to this view which essentially corresponds to Sinclair’s open-slot principle as the lexicon-and-grammar view. In construction grammar, this strict dichotomy between lexicon and grammatical rules is abandoned in favor of a view that regards lexis and grammar as a continuum. In the lexicon-and-grammar view, the only constructions that meaning can be subscribed to are highly idiomatic expressions that violate the principle of compositionality, e.g. ‘kick the bucket’, which means that they need their own entry in the lexicon. An early example for this view is Bloomfield (1933), who states: “The lexicon is really an appendix of the grammar, a list of basic irregularities” (1933, p. 274). In this view, idioms are typically considered to be a marginal phenomenon, relegated to the periphery of language and as such not of interest to the ‘core’ study of language (e.g. Di Sciullo and Williams, 1987). In construction grammar, there is, of course, no such distinction between core and periphery and evidence from various venues of research provide support for this position. Corpus studies show that language is permeated with idiomatic expressions on all levels (e.g. Erman and Warren, 2000) and psycholinguistic experiments and work in second language learning attest to
the importance of recurring formulaic phrases in developing fluency in speech (Syder and Pawley, 1983; Raupach, 1984; Foster, 2001; Wood, 2006).

The meaningfulness of ordinary constructions is an important point because it allows construction grammar to explain why speakers readily understand sentences like 1 and 2 (both taken from Goldberg (1995), p. 9). Sentence 1 instantiates the ditransitive construction. The prototypical meaning of the ditransitive construction has been described as the transfer of an object from one person to another (initiated by the giver), which can be paraphrased as ‘X causes Y to receive Z’. The meaning component of a transfer from X to Y is not coded in any of the words; in the constructional view, this part of the sentence meaning is instead supplied by the construction. For the lexicon-and-grammar view, on the other hand, it would be challenging to explain how ‘bake’, can have a ditransitive reading. If syntactic rules cannot contribute to meaning, one would have to postulate an additional sense of ‘bake’ that includes an optional recipient. The famous caused-motion construction in 2 exemplifies a similar point. If no constructional meaning is assumed, ‘sneeze’, ordinarily a monovalent verb, would require an additional lexical entry that specifies a ditransitive sense with a patient and a path role.

(1) She baked him a cake.

(2) He sneezed the napkin off the table.

Examples like 1 and 2 are not unusual or peripheral cases either; it is entirely common for English intransitive or transitive verbs to have their valency increased
via the inclusion of a patient or recipient/benefactive. An ordinarily intransitive verb like ‘sneeze’ can be part of a multitude of constructions beyond the caused-motion construction: You can sneeze a loud sneeze, sneeze onto something, sneeze yourself silly, or sneeze your way through the cafeteria. ‘Sneeze’ is by no means special; the same is true of many other verbs. These examples are problematic for the lexicon-and-grammar approach because they would require postulating a tremendous number of implausible verb senses. As such, it is often cited as an important advantage of constructional accounts.

Accounts of constructional meaning broadly fall into one of two categories. The first category deals with the characterization of the semantics of canonical syntactic constructions. The most well-known example is perhaps Goldberg’s work on *argument structure constructions* (1995). This line of research aims to explain the contribution constructions make towards meaning that allows speakers to understand sentences such as 1 and 2. It also tries to understand what factors influence the compatibility between words and constructions, e.g. why sentences like ‘she purchased him a book’ or ‘he shouted them the news’ are usually considered ungrammatical or at least infelicitous. Factors limiting the productivity of constructions are often explained with semantic restrictions which are based on the logic behind their conceptual structure. The problem with this is that these arguments are usually based on grammaticality judgments. Goldberg’s early work, in particular, is full of examples of allegedly ungrammatical expressions that are actually well-attested in corpora. Using introspec-
tion and grammaticality judgments about artificial sentences – not just for illustrative purposes but as linguistic evidence that forms the basis of arguments about linguistic theory – is perhaps the biggest problem with some foundational work in construction grammar. Explanations of linguistic structure based on conjecture about conceptual motivations run the risk of becoming *just so stories.* Despite their commitment to usage-based linguistics, many leading theorists within the field tend to use invented examples instead of studying authentic instances of natural language. Fortunately, this appears to be changing with several recent publications being explicitly dedicated to moving the field into a more empirical direction and combining construction grammar with an empiricist approach to usage-based linguistics (Gries and Stefanowitsch, 2007; Gonzalez-Marquez et al., 2007; Glynn and Fischer, 2010; Diessel, 2015; Perek, 2015; Noël, 2016, *inter alia*).

The second category of research into constructional meaning deals with idiomatic constructions. Early construction grammar, in particular, was driven by the need to explain unusual expressions that seemed to fall outside of ordinary language by virtue of their idiomatic meaning, unusual grammar and/or limited productivity. Examples include the *let alone* construction, e.g. “I barely got up in time to eat lunch, let alone cook breakfast” (Fillmore et al., 1988, p. 512), the comparative correlative construction “The bigger they come the harder they fall”, the incredulity construction “Him wear a tuxedo!??” (Lambrecht, 1990, p. 215), the *time-away* construction and “Bill slept the afternoon away” (Jackendoff, 1997, p. 534).
Few studies, however, have investigated the role of frequency in *partially filled constructions*, i.e. the co-selection between words and constructions. Some important forays have been made into this area such as *behavioral profile analysis* (Hanks, 1996) and *collostructional analysis* (Stefanowitsch and Gries, 2003; 2005; Gries and Stefanowitsch, 2004) but there is still little research into the way words co-occur with grammatical constructions and how they form entrenched patterns with distinct meanings. As I will argue here, this is an important facet of constructional meaning that gives rise to the distributional phenomenon of semantic prosody.

### 3.2.3 Evidence for constructions

So far, I have described some analytical advantages of a constructional account, such as the elegant solutions that the idea of meaningful constructions provide when describing valency-increasing operations (e.g. “she threw him a ball”). However, I consider empirical evidence to be a far more convincing argument than elegance, parsimony or austerity of theoretical description. Fortunately, there are several studies that attest the psychological reality of constructions and explain the mechanisms of how speakers acquire constructions through exposure to ambient language.

Goldberg et al. (2004) study the acquisition of argument structure constructions. They find that it is typically a single prototypical exemplar verb with a particularly high token frequency that drives the generalization and the acquisition of constructional meaning. They also discuss how their model disproves the Chomskyan *poverty*
of the stimulus argument and obviates the need for positing the innateness of argument structures. Bencini and Goldberg (2000), Eddington and Ruiz de Mendoza (2010), Kirsch (2018) all use various experimental methods (priming, visual world eye tracking, and a sorting paradigm respectively) to show that argument structure constructions have meaning independent of their lexical instantiations. Work by Bowerman (1973), Braine and Bowerman (1976) and Tomasello (2006) emphasizes the role of constructions in child language acquisition. They show how children gradually acquire grammar through learning to abstract away from fixed patterns – first through the use of highly local patterns with one open slot (pivot schemas and item-based constructions) and later through generalizing more complex constructions. Tomasello (2014, p. 225) summarizes the findings of this line of research: “Ontogenetically, children hear individual utterances and then (re-) construct the abstract constructions of a language. All of this is done with general cognitive processes […]”

3.3 Evaluation and stance

Language is rarely if ever, used to dispassionately convey a set of neutral facts. Instead, it is subjective and used to express point-of-views or attitudes; in discourse speakers constantly position themselves relative to the context of the conversation and one another’s stances.

The importance and ubiquity of evaluative language have been described by numerous linguists. According to Du Bois (2007), “One of the most important things
we do with words is take a stance” (p. 139). In their seminal study of oral narratives, Labov and Waletzky (1967) conclude that a narrative without evaluation is not complete and has “no point” (1967, p. 33). Partington (2015, p. 280) emphasizes how common it is in language use: “Evaluation is pervasive in practically all forms of linguistic communication.” Stubbs stresses both the universality and the diverse nature of evaluation:

“[W]henever speakers (or writers) say anything, they encode their point of view towards it: whether they think it is a reasonable thing to say, or might be found to be obvious, questionable, tentative, provisional, controversial, contractionary, irrelevant, impolite, or whatever. The expression of such speakers’ attitudes is pervasive in all uses of language.” (Stubbs, 1986, p. 1)

Much work on semantic prosody emphasizes its communicative function, more specifically the way it is used to convey evaluative meaning in discourse. Before I discuss the evaluative or attitudinal nature of semantic prosody I will first clarify the terminology used here and approach basic definitions of concepts such as subjectivity, evaluation, attitude and stance-taking.

3.3.1 A working definition of evaluative terminology

A good starting point for a discussion of evaluation is perhaps the notion of subjectivity. Wierzbicka (1988) argues against truth-conditional semantics and emphasizes
the role of a human factor in creating meaning:

“[M]eaning cannot be defined in terms of a relationship between linguistic units and elements of extra-linguistic reality […] meaning consists in human interpretation of the world. It is subjective, it is anthropocentric, it reflects predominant cultural concerns and culture-specific modes of social interaction […]” (Wierzbicka, 1988, p. 2)

Subjectivity is not a linguistic strategy or an optional feature of utterances but instead a fundamental part of how meaning is created. Evaluation is closely related to subjectivity and can be seen as its subcategory. Englebretson (2007) derives evaluation from subjectivity, describing it as subjectivity focused on the content of the conversation:

“[E]valuation can be roughly summed up as subjectivity with a focus. In other words, while subjectivity refers broadly to ‘self-expression’ (Lyons, 1994, p. 13), evaluation implies self-expression that is focused toward a narrow purview – self-expression about the ‘entities or propositions’ (Thompson and Hunston, 2000, p. 5) present in the very language that the speaker or writer is currently producing.” (Englebretson, 2007, p. 16)

Langacker uses different terminology, but he also sees evaluation as a form of subjectivity. In his work, both terms are captured through his concept of construal, which is an obligatory part of the linguistic encoding of any conceptual content:
“A fundamental notion of cognitive semantics is that a predication does not reside in conceptual content alone but necessarily incorporates a particular way of construing and portraying that content. Our capacity to construe the same content in alternate ways is referred to as imagery; expressions describing the same conceived situation may nonetheless be semantically quite distinct by virtue of the contrasting images they impose on it.” (Langacker, 1991, p. 4)

Langacker then goes on to say that “Numerous dimensions of imagery can be discerned” (ibid) and one such “dimension of imagery is construal relative to different background assumptions and expectations (consider stingy vs. thrifty)” (ibid).

Thompson and Hunston (2000) give a useful working definition of evaluation that summarizes all major points discussed so far:

“[E]valuation is the broad cover term for the expression of the speaker or writer’s attitude or stance towards, viewpoint on, or feelings about the entities or propositions that he or she is talking about. That attitude may relate to certainty or obligation or desirability or any of a number of other sets of values.” Thompson and Hunston (2000, p. 5)

To recapitulate, evaluation is an expression of speakers’ (subjective) viewpoints about the propositions they are currently talking about. It is broadly defined to encompass a wide set of categories and is by no means limited to a system of positive or
negative judgments. I refer to items that are classified as either negative or positive as having an *evaluative polarity* or valence (not to be confused with grammatical valency); both are terms occasionally used in the literature (Channell (2000); Ellis et al. (2007); Hauser and Schwarz (2016) inter alia). It is essentially used interchangeably with terms such as attitude or stance.

Stance is a deeply heterogenous concept and many definitions that vary greatly are found across academic disciplines (cf. Englebretson, 2007, pp. 1ff; Kockelman, 2004, pp. 129ff). My use of the term is limited to some form of evaluative or attitudinal stance. Not included in this definition are concepts like the authorial stance or Dennett’s system of physical stance, design stance, and intentional stance.

### 3.3.2 Evaluation and semantic prosody

The context a word habitually occurs in becomes part of our knowledge of that word. In the case of semantic prosody, this knowledge concerns a word’s distribution in evaluative statements. A word like ‘utterly’ for example does not only mean ‘absolutely’ or ‘completely’: Through its common co-occurrence with evaluations like ‘utterly boring’ or ‘utterly pointless’ this evaluative meaning becomes part of its meaning (cf. Louw, 1993, pp. 159-161). As such, there is a difference between calling someone ‘totally unconventional’ and ‘utterly unconventional’ (cf. Hauser and Schwarz, 2018).

This underlines the two different levels of semantic prosody described in the previous chapter. Semantic prosody is not just a matter of co-occurrences, it also serves
a communicative function. It is a tool for stance-taking. The utterance ‘reality set in’ is not a disinterested description of the situation but also clearly conveys the viewpoint that this reality is an unfavorable state of affairs. An example of employing different construals through semantic prosody is the contrast between ‘fraught with’ and ‘brimming with’. Both roughly mean ‘full of’ or ‘replete with’, but through their different collocational profiles they have very different evaluative meanings: ‘Fraught with’ tends to co-occur with words such as ‘danger’, ‘conflict’ or ‘difficulty’, whereas ‘brimming with’ collocates with ‘confidence’, ‘energy’ and ‘enthusiasm’ (cf. Morley and Partington, 2009, p. 147). It is not only in combination with these collocates that the two phraseological units show a distinct evaluative polarity; they also retain it in other contexts: There is a marked difference in evaluation between ‘brimming with emotion’ and ‘fraught with emotion’ (both of which are attested in COCA).

3.3.3 Is evaluative meaning connotational and peripheral?

What kind of meaning does semantic prosody have? Can it be described as a connotational meaning component? Whether or not the notion of connotation is appropriate here might seem like a terminological and definitional question, but still, it is a relevant one. Much work in semantic prosody describes it as connotational. According to Berber Sardinha (2000, p. 93) “semantic prosody is the connotation conveyed by the regular co-occurrence of lexical items.” Similar descriptions are found in Stubbs (2001) and Hunston (2002). Traditionally, evaluative language has been described as
peripheral or connotative. According to Lyons (1977, p.176), “the connotation of a word is thought of as an emotive or affective component additional to its central meaning” and we may say that “a word has a pleasant or desirable connotation” (ibid). Palmer (1981, p. 92) agrees that connotation is “emotive or evaluative meaning.”

Partington (1998, p. 68) describes semantic prosody as “the spreading of connotational coloring beyond single word boundaries.” He is, however, careful to point out that connotation should not be considered to be a secondary or peripheral meaning component: “a word like pig-headed only exists because it has an expressive connotation of disapproval, there is nothing ‘secondary’ about this implication” (ibid, p. 66). Sinclair (2000, p. 200) emphasizes that evaluation is a major factor for our lexical choices: “the reason why we choose to express ourselves in one way rather than another is coded in the prosody.” Additionally, it has been pointed out that some near-synonyms differ in the attitude the speaker expresses towards the entity they describe. Clark (1993, p. 73) contrasts ‘statesman’ and ‘politician’ noting “the former is laudatory and the latter is not.” A similar point can be made about Langacker’s comparison of ‘thrifty’ and ‘stingy’. Evolving his position from his previous definition, Partington (2004, p. 154) points out that it might be preferable to avoid the notion of connotation altogether:

“One of the reasons Martin, Thompson and Hunston on the one hand and the Sinclairians on the other adopt terminology such as evaluative or attitudinal meaning is because of the huge problems associated with the
denotation-connotation distinction. The term ‘connotation’ is made to do an immense amount of work, covering concepts as varied as social connotation (consider *awfully clever* and *dead clever*), cultural connotation (*whisky* in Glasgow or in Riyadh) and expressive connotation, this latter being close to evaluative meaning.” (Partington, 2004, p. 154)

I will avoid the term connotation (unless I am describing an account that uses this terminology) for a number of reasons: Even though Partington (2004) cautions against it, for many researchers connotation does indeed evoke a sense of peripherality, which is an association that should be avoided. Additionally, the distinction between denotation and connotation is far from clear and perhaps not particularly meaningful. Finally, the term connotation is too general and imprecise. Instead, I adopt terms such as evaluative or attitudinal meaning as described above.
Chapter 4
Methodology

This chapter describes the main methodologies used throughout this dissertation. In section 4.1, I summarize the approach of quantitative corpus linguistics and contrast it with ordinary corpus linguistics and other quantitative approaches within linguistics. In section 4.2, I give a practical description of the processes of data retrieval and pre-processing. In section 4.3, I outline collostructional analysis, the main method of quantitative analysis used in this research. In 4.4, I discuss using additional quantitative methods to further analyze the data.

4.1 Quantitative corpus linguistics

The main methodological stance adopted here is the framework of quantitative corpus linguistics. The basic idea is that frequency effects (as described in the section on usage-based linguistics) can be studied by retrieving relevant linguistic items from large collections of electronically searchable text and evaluating them statistically. Stefanowitsch and Gries (2005, pp. 4ff) describe quantitative corpus linguistics as the combination of two approaches to the study of language: First, it is based in traditional corpus linguistics with its interest in a wide range of linguistic phenomena and a strong foundation in linguistic expertise. Second, it commits to a rigorous
quantitative methodology as typically found in computational linguistics or statistical language processing. Stefanowitsch and Gries (2005) list three important conditions for quantitative corpus linguistics. First, the corpus needs to be representative and balanced. Second, it has to be retrieved exhaustively. Third, the data need to be analyzed statistically. I will discuss these points one by one.

First, what does it mean for a corpus to be balanced and representative? The underlying assumption of usage-based corpus linguistics is that the data in a well-balanced linguistic corpus is a rough but useful approximation of speakers’ experience with language. For discussions of balance and representativeness in corpus design, see Atkins et al. (1992) or Biber (1993) inter alia. For more modern approaches, see Temnikova et al. (2014) or Rapp (2014).

While the exact relationship between frequency in text and entrenchment in the conceptual system is far from clear (cf. Schmid, 2010; Blumenthal-Dramé, 2012), quantitative corpus linguistics and the usage-based paradigm have been successfully employed in the study of numerous linguistic phenomena and are supported by a preponderance of data, both in the form of experimental and observational evidence. As mentioned in chapter 3, frequency effects have been found on all levels of linguistic organization from phonology to syntax, and in all areas of study, such as language acquisition, language production, language perception and language change (for more extensive literature reviews see Diessel, 2007; Ellis, 2002).

The second issue is the exhaustive retrieval of the relevant information. No matter
how good a corpus is, if we only retrieve a portion of the linguistic items we want to study, any frequency-based arguments are moot. The same is true if we retrieve false positives and include them in our analysis. This means that both recall and precision should be maximized. In information retrieval, there is always a tradeoff between recall and precision. Automatic approaches to natural language processing bias their algorithms towards one or the other based on the requirements of the specific application. In quantitative corpus linguistics, this means that the researcher often needs to perform tremendously labor- and time-intensive manual work. Typically, this means first optimizing recall, i.e. retrieving as much potentially relevant information as possible without paying much attention to precision. In the second step, the researcher then combs through the data manually and weeds out all irrelevant instances in order to maximize precision. Very good treebanks with manually parsed or corrected syntactic trees theoretically allow the retrieval of linguistic items with great precision and would greatly reduce the need for manual data processing, but these corpora are typically far too small for the linguistic phenomena investigated here.

The third point – the statistical analysis of the data – is one of the points that sets quantitative corpus linguistics apart from traditional corpus linguistics, which often reports raw counts. The two biggest issues with using raw count are that (a) without inferential statistics it is impossible to know if effects are statistically significant and (b) counts that are not normalized are not particularly informative for most types of analysis because they ignore the overall distributions of words in text. Beyond the
need for representative corpora, precise data retrieval, and quantitative methodology, I would add a fourth point: The work does not stop with the statistical analysis. As in traditional corpus linguistics, the data and their implications need to be evaluated qualitatively. This point will become clear throughout the next chapter.

4.2 Data retrieval and pre-processing

Stubbs (1995a) employs two different approaches in his corpus study of semantic prosody. In the first approach, a relatively limited number of concordance lines is retrieved from a corpus. These lines are then inspected and analyzed manually. The advantage of this approach is that it allows a reasonably detailed analysis of the data by the researcher (depending on how much context is retrieved for each token); the downside of this method is that it is slow and limited to how much data a researcher can manually inspect. The second approach described by Stubbs relies on a collocational display. A computer program retrieves all the collocates of a node word within a defined span (typically 3 or 4 words to each side of the node word), counts their occurrences and displays them in the form of a ranked table. This method allows the researcher to take a huge number of tokens into account, but if used alone, it runs the danger of overlooking important details. Stubbs argues that this second approach is a viable alternative to the first one in that it allows for a summary of the data when the corpus is too large to analyze manually. Stubbs (1995a) however also notes that the collocates need to be interpreted with care. He finds ‘great’ to be
one of the most frequent adjectival collocates of ‘cause’, but that is because of strings such as “cause for great concern” (Stubbs, 1995a, p. 15). Collocational displays are a simple and useful but also very coarse tool.

I use an approach that includes both types of methods, the manual analysis of concordance lines and the semi-automatic overview of the most frequently co-occurring items. While the first step is quite similar to what Stubbs describes, the second step is far more involved: What sets the methodology employed here apart from most work in corpus linguistics is that it is used to study constructional phenomena whereas most corpus linguistic research is conducted from a lexical perspective. In traditional collocational studies, the researcher typically retrieves all words within a window around the node word without taking grammatical information into account. This dissertation, on the other hand, deals with the co-selection between words and constructions. The exact data retrieval process differs depending on whether the investigation starts with the word or with the construction. If the starting point is the word, a two-step approach is adopted. In the first step, a relatively small data sample of 200 sentences is manually analyzed. This serves two purposes – it yields important insights into the usage of the word and it allows for the identification of the relevant constructions the word takes part in. Second, once these constructions have been established, they can be searched for in a large corpus with the help of regular expressions so that the words that occur in their open slots can be retrieved. If the investigation starts with a construction, the first step is not necessary; instead,
search queries are formulated and sentences containing the construction are retrieved from the corpus.

This description is a tremendously idealized description of the process. As described in 4.1, a great deal of manual and semi-manual preprocessing is still needed. Experience shows that using what may seem to be a well-designed and very specific search query usually leads to too many false positives. Typically, items are missed because there are some fine details and complexities that are difficult to account for when designing search queries. Instead, the constructions need to be retrieved by using a “maximally underspecified search string” (Stefanowitsch and Gries, 2005, p. 36) to ensure that virtually all relevant examples are caught (alongside many false positives). The process of whittling the data set down to only the true positives is semi-manual and iterative. First, false negatives are manually identified, and systematic generalizations are made on how to exclude them. These generalizations are then translated into regular expressions to find all of these examples. The examples marked for deletion are then inspected to ensure that the regular expressions were not too broad or (less common) too narrow. They are then revised, and the process is repeated as necessary. This procedure continues until the data set is both high in recall and precision. How time-intensive this process is, varies greatly, mostly depending on how common and how productive the investigated construction is and whether there are homonyms or polysemes.¹

¹Take Sinclair’s example of ‘set in’ meaning ‘happen’. The fact that it is a phrasal verb is not
I used *The Corpus of Contemporary American English* or COCA (Davies, 2008) as my main data source. The reason is that it is the largest well-balanced and openly accessible corpus of English, currently consisting of more than 550 million words. It is true that larger does not necessarily mean better when it comes to corpora. Instead, the required size of a corpus depends on the application. For the linguistic phenomena I investigate, very large text sizes are necessary for two reasons. First, some of the linguistic items studied here are relatively uncommon and include extended sequences meaning they have very low token counts. Earlier research was often limited by the small corpora available at the time. Bublitz (1996) for example studies the word ‘prevail’ in the 1 million word *Brown Corpus* (Nelson and Kucera, 1979), the 1 million word *Lancaster-Oslo/Bergen Corpus, LOB* (Johansson et al., 1978; Garside and Leech, 1986) and the 500,000 word *London-Lund Corpus of Spoken English, LLC* (Svartvik and Quirk, 1980; Svartvik, 1990). Even though he only studies a single word (as opposed to a longer semi-idiomatic phrase), he only finds 61 relevant tokens in all three corpora combined. If language structure emerges from language use and we need very large amounts of text to reliably find and describe linguistic patterns, how much language have speakers typically been exposed to? How do corpora compare to speakers’ exposure to language? Stubbs (1995a, p. 19-20) roughly estimates that an average speaker hears, reads and produces somewhere around 500-600 million coded in corpora that only contain words and their part-of-speech tags. As such it is not possible to automatically disambiguate it from the verb ‘set’ meaning something like ‘be situated in a setting’, e.g. ‘the novel is set in Tuscany’, which is actually the most common instantiation of the string.
words over a span of 30 years. He also cites other calculations that put the number at about 300 million over 30 years. This means the amount of data contained in a corpus like COCA is similar to a speakers’ entire experience with language, arguably putting it “within the right range for certain kinds of cognitive modeling” (Stubbs, 1995a, p. 20). It has to be noted however that a similar size does not mean that its other qualities also necessarily correspond to speakers’ experience with language.

A second issue is that relatively large numbers are necessary for the reliable use of quantitative methods. Corpora that are even larger than COCA do exist, but they have often been compiled with the goal of being as large as possible without taking representativeness into account. Additionally, they often are not available as full-text corpora but have been transformed into n-gram counts, which makes them entirely unsuitable for any qualitative analysis.

Once the data set is ready, retrieving only the relevant words from the sentences for further analysis is relatively unproblematic. Finally, the association strength between words and constructions is computed with the statistical methodology of *collostructional analysis*.

### 4.3 Collostructional analysis

Collostructional analysis is a family of quantitative methods developed by Stefanowitsch and Gries with the aim of quantifying the relationships between words and constructions. There are three different types of collostructional methods, (1) *collex-*. 
eme analysis which measures the association of words to particular constructions, (2) distinct collexeme analysis which compares the attraction of words to different (typically related) constructions, and (3) covarying collexeme analysis which investigates the relationships between words within different slots of the same construction (cf. Stefanowitsch and Gries, 2003; 2005; Gries and Stefanowitsch, 2004).

Collostructional analysis comes with its own terminology (cf. Stefanowitsch and Gries, 2003, p. 215): Lexemes that are associated with a construction are referred to as collexemes of this construction. Constructions associated with lexemes are called collostructs. In other words, collostructs are not different from constructions; they are just constructions in the context of collostructional analysis. The combination of collexeme and collostruct, i.e. the entire instantiated sequence is referred to as a collostruction. Collostructional strength (or attraction) is a bidirectional statistical measure of the association between lexemes and constructions, i.e. of a lexeme’s tendency to fill a slot in a construction. Stefanowitsch and Gries (2003) argue against the use of raw frequencies (as used in many older corpus studies) on the basis of both practical and theoretical considerations. Most obviously, raw counts do not take the overall frequencies of a word in the entire corpus into account, often leading to strong associations with extremely frequent words, i.e. function words.

All this terminology is perhaps best illustrated with the help of an example. Let us assume we wanted to study the ditransitive dative construction. Schematically, it can be described as ‘NP_{SUBJ} V NP_{IO} NP_{DO}’ or, in terms of its semantic roles as
Agent Transfer-Action Recipient Theme. In the context of collostrucional analysis, this can be referred to as the collostruct. Examples for the construction include ‘she gave him a book’ or ‘he threw her a ball’. These two examples are fully specified sequences (i.e. all slots are occupied), so they are collostructions.

Collexeme analysis measures the attraction between a construction and a single collexeme, so in our example, we could use it to calculate which collexemes show the strongest degree of collostrucional strength to the verb slot. To calculate the collostrucional strength we need the token counts of all collexemes both within the construction and within the entire corpus, the token counts of the construction in the corpus and the token counts of all words in the corpus. The collostrucional strength can be calculated with various association measures such as Fisher’s’ Exact Test, Mutual Information Score, Odds Ratio and Log Likelihood, the latter of which was chosen here.\(^2\) It is used to rank the collexemes of a construction in order of

\(^2\)There is a great amount of discussion about these measures, (Wiechmann, 2008; Bybee, 2010; Schmid and Küchenhoff, 2013; Gries, 2012; 2015; 2015). It should be noted though that the rankings the different approaches yield are quite similar in most cases. Fisher’s exact test which is recommended by Gries and Stefanowitsch can be extremely computationally demanding, especially considering the type of data analyzed here: Frequent and productive constructions can be associated with dozens, if not hundreds of collexemes. Most of my analyses were finished within a few minutes, but in some cases, the model did not converge in over 48 hours (I tried this both on a modern PC and remotely through two different cloud computing services). The reason that I chose the log likelihood statistic is that it produced rankings similar to Fisher’s exact test. Ultimately, as will become evident in chapter 5, the exact ranking of the collexemes is not of great importance.
their relative degree of attraction (or repulsion) to that construction. These ranked lists are then interpreted by a researcher, which involves some degree of qualitative analysis.

The second collostructional method, distinct collexeme analysis compares the attraction of collexemes to different constructions/collostructs. Where ‘plain’ collexeme analysis compared the frequencies of words within a construction to its frequencies in the rest of the corpus, distinct collexeme analysis compares the frequencies of words within one construction to their frequencies in another construction. In our example, it could be used to test whether personal pronouns are more likely to occur in the ditransitive dative or the prepositional dative construction (‘NP_{SUBJ} V NP_{DO} to NP_{DO}’, e.g. ‘he sent a letter to his father’). As its input, this method takes the token counts for each collexeme within the two constructions.

The third method, covarying collexeme analysis measures the tendencies of collexemes within different slots of the same construction, so in the case of the ditransitive dative construction, it could be employed to investigate which collexemes in the theme slot tend to co-occur with which collexemes in the verb slot. It takes the frequencies of co-occurrence of pairs of collexemes as its input.

4.4 Further analysis

An important point is that collostructional analysis is not a purely quantitative approach but involves the interpretation of the data by a human researcher. The output
from all collostructional methods is a table of collexemes, ranked by their collostructional strength. The collexemes are manually inspected, and often generalizations are made about their semantics and about how they could be grouped together into various semantic fields.

The manual analysis of data is sometimes criticized in corpus linguistics (particularly within large-scale computational approaches), either because of concerns that it introduces a biased and subjective human perspective or because of the argument that as much as possible should be automated in order to allow for the processing of larger amounts of data. There are two ways in which semantic relationships between words can be computed, either through comparing their distributions in text or through using hand-curated databases that include this kind of information. I will summarize both approaches here and describe how they can be used to automatically cluster collexemes into lexical fields. I argue why they may be a useful addition to the linguist’s toolbox, but not a replacement for the linguist themselves.

The first approach is purely bottom-up and relies on raw linguistic data, typically in the form of texts. The idea of using distributional information to describe word similarity is the concept that underlies frequentist approaches in the usage-based paradigm: Similar words have a tendency to occur in similar contexts in language use. In natural language processing, words are often described through their word embeddings, which contain distributional information of words in the form of their co-occurrences with other words and skip-grams (a type of n-gram that can be dis-
continuous. Mathematically they are expressed through a multidimensional vector space where each unique word is represented by a vector. An advantage of vector representations is that relationships between words can be computed through vector operations, e.g. vector addition and subtraction. A classic example goes as follows: \textbf{King - man + woman = queen}. If we subtract the vector for the word ‘man’ from the vector for ‘king’ and then add the vector for ‘woman’ to it, we arrive at a vector that is most similar to the vector for ‘queen’ (again, this is solely based on distributional data without any additional information curated by human researchers). Similarities between words can be calculated by measuring the cosine of the angle between their vectors.

A much more useful application of vectors, however, is calculating similarities between words or to automatically group them together. Similar words should have similar distributions in text and consequently similar vectors. Collexemes can be grouped together based on their word vectors with the help of clustering algorithms. Clustering is an unsupervised machine learning application, meaning the algorithms try to find some kind of natural structure in the data.

I implemented the clustering algorithm for the collexeme analysis with the help of the \textit{spaCy} (Honnibal, 2017) library, a natural language processing module for the \textit{Python} programming language (van Rossum, 1995), which includes pre-trained word vector models based on Stanford’s \textit{GloVe} or Global Vectors for Word Representation algorithm (Pennington et al., 2014). I used the \textit{scikit-learn} machine learning library.
(Pedregosa et al., 2011) to implement a hierarchical clustering algorithm with Ward’s method. One of the advantages of this approach is that it creates a hierarchical tree of clusters that can be represented as a dendrogram. While this does not obviate the need for a human researcher to study the collexemes manually, this visualization is a very convenient way to peruse the data.

An issue with word vectors is that their distributional data are gathered with a ‘bag-of-words approach’: Words are collected, counted and put into the figurative bag without taking any kind of word order information into account (beyond what emerges from n-grams). As such, it is somewhat counter to the methodological approach expressed here, viz., that detailed linguistic analysis requires rich data that is analyzed within its grammatical or constructional context. The counter-argument to this is that while the models that create word vectors are not given explicit grammatical categories, a tremendous amount of information is ‘baked into the cake’, meaning that machine learning models are able to derive this kind of information from their raw linguistic input data.

Gries and Stefanowitsch (2010) propose an approach to clustering collexemes that avoids these issues but that also makes it limited in its application. Instead of clustering words based on their distributions in text, they cluster them based on their covarying collexemes. For example, in a construction with two slots, words in the first slot could be grouped together based on what words they tend to co-occur with in the second slot.
The second approach to automatically clustering collexemes does not depend on their distributional similarities but instead uses lexical databases that contain hand-coded information about lexical relationships. *WordNet* (Miller, 1995; Fellbaum, 1998) is such a database. It consists of sets of cognitive synonyms, called *synsets*, which are interlinked through lexical relations. The most commonly encoded relation is the *super-subordinate relation* also known as *is-a relationship*, which means that the database contains a hierarchical hypernym/hyponym structure. All noun hierarchies terminate in the shared root node ‘entity’. The word ‘table’ for example with the meaning ‘a piece of furniture having a smooth flat top that is usually supported by one or more vertical legs’ has the following hypernym path: ‘table’, ‘furniture’, ‘furnishing’, ‘instrumentality’, ‘artifact’, ‘whole’, ‘object’, ‘physical_entity’ and finally ‘entity’. Words that are not in an *is-a* relationship can be compared through their lowest common hypernyms (e.g. ‘carnivore’ for ‘cat’ and ‘dog’). Their similarity can be calculated by how close they are within this taxonomy, e.g. through a measure that counts the length of the shortest path (number of nodes passed) between them.

I used the following process to use WordNet data to classify collexemes: First, the synsets for all collexemes are retrieved from WordNet. I used the *Natural Language Toolkit* or *NLTK* (Bird et al., 2009) which is a natural language processing module for the Python programming language to access the WordNet information. Then the full paths between all possible pairs of synsets are found. This is not a function available in NLTK, so I implemented it myself. These paths are used to grow a network of
the collexemes and all the synsets that connect them within the WordNet hierarchy. This network was created with the NetworkX Python module (Hagberg et al., 2008). Within this network, clusters can be determined, which are referred to as communities within the field of network studies. Newman (2010, p. 357) describes the problem of community detection as follows: “Loosely stated, it is the problem of finding the natural divisions of a network into groups of vertices such that there are many edges within groups and few edges between groups.”

Since this is neither a central point in my dissertation nor of particular importance to linguistic theory, I will only provide a very brief and greatly simplified description. The WordNet synsets in the network are referred to as nodes and the connections between them are referred to as edges. Community detection typically works by finding groups of nodes that are more closely connected to each other than to the rest of the graph (this is referred to as modularity). Different algorithms are available; I am using the Girvan-Newman algorithm (Newman, 2006). In short, the Girvan-Newman algorithm is a divisive hierarchical clustering algorithm for networks. It creates communities by progressively removing edges from a network. Edges are removed based on their betweenness centrality, i.e. how many nodes they connect in the shortest path throughout the entire network. The idea is that edges connecting separate communities have high betweenness because many shortest paths from one community to another need to travel through them. The output of the algorithm is a hierarchy of clusters that can be visualized as a dendrogram or as a network. It
is a useful tool, but I found the quality of the models it produces to be below the performance of the vector-based models.

In conclusion, complex and *more quantitative* does not necessarily mean better. As I have described, there are limitations and drawbacks. For those reasons, I only use them in addition to qualitative analyses, never instead of them. They are useful to support manual analyses, to check and possibly corroborate intuitions but not to replace them – at least not yet.
Chapter 5

The Case for a Constructional Account

In this chapter, I argue that evaluative meaning should not be studied through words in isolation, but instead within a constructional context. This chapter is organized as follows. In section 5.1, I outline what prompted me to study semantic prosody within a constructional framework. I provide a brief recapitulation of a few prior studies that have analyzed semantic prosody as a word-level phenomenon even though the data hint at the relevance of construction-specific meaning. I also review research that has studied evaluation from a phraseological perspective and discuss to what extent it overlaps with constructional approaches.

In section 5.2, I explore the different uses and semantics that words take on within different constructions. I provide examples for adjectives, verbs, and nouns. Some of the studies involve a reanalysis of classic examples from the prior literature, others are novel examples. All examples are investigated with the help of collostructional analysis; in some cases, the resulting collexeme profiles are further examined through various forms of cluster analysis.

In section 5.3, I adopt a different starting point for my investigation. Instead of starting at the level of the word and exploring the constructions it participates in, I start with the construction. In section 5.4, I summarize the findings of this chapter,
contextualize them within the theoretical frameworks of construction grammar and usage-based linguistics and draw further theoretical conclusions.

5.1 Lexical meaning and constructional meaning

Research on semantic prosody is virtually always done on a purely collocational level: Studies describe the way individual words co-occur with each other without taking grammatical structures into account. Occasionally, some researcher will describe data that hint at the relevance of different grammatical contexts, but ultimately, they still argue on the lexical level. One example of this is the phrasal verb ‘build up’ mentioned in chapter 2. To recapitulate, Louw (1993) discovered that ‘build up’ when used in intransitive constructions often describes bad things while it describes good things when used in transitive constructions. Intransitive ‘build up’ typically refers to the accumulation of things that are often unfavorable, such as ‘stress’, ‘frustration’ or ‘plaque’, whereas in transitive constructions it describes scenes where people build up generally favorable things like ‘savings’, ‘muscles’ or ‘confidence’. Louw’s solution is simply to call this a mixed prosody and attribute the semantic prosody to the lexical level, not to the two patterns that include the verb despite the fact that their semantics are quite different.

Louw (2000) and Louw and Chateau (2010) revisit this example, concluding that semantic prosody has “some relationship with transitivity” (Louw, 2000, p. 53). This sounds as if they may be getting closer to a concept of constructional meaning (or
at least context-dependent meaning) but they never expound upon their statement or explain whether or not this is a systematic relationship, nor do they provide any additional data beyond ‘build up’. Instead, they argue in terms of the real-world entities involved in the described scenes (similar to my description above):

“[W]here human beings are in control of their own destiny and are shaping it transitively for themselves, the semantic prosody is positive, but where people are at the mercy of forces beyond their control, the things which build up intransitively are negative and uniformly threatening.” Louw (2000, p. 53)

This explanation may or may not be accurate for ‘build up’, but it does not become clear if and how it can be generalized to ‘some relationship with transitivity’. Perhaps the authors mean to indicate that sentences with humans in object slots tend to describe undesirable events, although this is conjecture. Ultimately, Louw (2000) makes a general claim about syntactic structure based on the description of one specific semantic relationship.

Another example is an early and small study of the word ‘prevail’ by Bublitz (1996) where he notes that it can have different semantic prosodies depending on its inflectional form. He states that when all forms of the lemma are conflated, 19 occur with a positive semantic prosody, 14 with a negative semantic prosody and another 28 are ‘vague’ (Bublitz, 1996, p. 21). Based on this data, ‘prevail’ does not seem to have a particularly pronounced semantic prosody in either direction. When
the different inflectional forms are considered individually, however, a different picture emerges: The word-forms ‘prevail’ and ‘prevails’ are found in more positive than negative contexts; ‘prevailing’ has more negative than positive examples but also far more neutral ones. The form ‘prevailed’ is split almost evenly between positive and neutral evaluations. Bublitz (1996) concludes that there is an “asymmetrical distribution of semantic prosody, i.e. of different word-forms having different collocates” [ibid, p. 22]. A crucial point he does not mention is that the different word forms occur in different constructions. Most notably, in its form as a present participle, it can be used attributively as in ‘the prevailing opinions.’ A quick inspection of corpus data show that its attributive usage is by far the most common (accounting for about 90% of all instances in a sample of 200 sentences including the verb, taken from COCA), making it a likely reason why Bublitz found it to differ from the other forms.

These two examples illustrate that even though researchers have encountered evidence that hints at the importance of construction-dependent meaning, they do not discuss the data in these terms. Instead, they solely argue along the lines of purely collocational relationships, i.e. word-to-word co-occurrences. This reveals a word-centered perspective of meaning that neatly separates the lexicon and grammar and only assigns semantic value to the former. This lack of awareness of constructional meaning which is found throughout the research literature on semantic prosody needs to be situated within an understanding of the schools of thought guiding this research.
Much of its foundational work was born out of the neo-Firthian (or Sinclairian) tradition of British corpus linguistics, which emphasizes the importance of idiomatic, partially prefabricated linguistic units, e.g. as seen in Sinclair’s idiom principle.

Some of Sinclair’s idiom principle seems compatible with the tenets of construction grammar: The importance of semi-preconstructed phrases in Sinclair’s theory bears some resemblance to the centrality of constructions in construction grammar. The idea that commonly recurring linguistic segments can have unit status even when they can be further analyzed into their component parts is captured through the rule-list fallacy noted in Langacker’s Cognitive Grammar (Langacker, 1987, p. 29).

All this being said, the Sinclairian tradition is not a constructional approach: Even though Sinclair’s notion of the lexical item is complex and goes beyond the level of individual words to also include phraseological units (see Pope (2015) for an in-depth discussion), his work lacks a clear and systematic concept of constructions beyond the notion of the idiom. Furthermore, even though he describes the idiom principle as more important than the open-slot principle, this solely relates to its frequency of instantiation in text – there is no indication that he considers it to be privileged in terms of the emergence of grammatical structures.

Most studies of semantic prosody that go beyond the purely lexical level are found in the field of phraseology. Researchers who are primarily interested in phraseological units, or any kind of partially fixed units of language, whether they are called constructions, formulaic sequences, idiomatic sequences, prefabs, pattern, multi-word
units or pre-packaged assemblies naturally realize the importance of meaning within larger linguistic units and are often interested in their evaluative semantics. It has long been noted that idioms seem to have a tendency to be used for emotive and evaluative language. As Nunberg et al. (1994, p. 493) puts it: “A language doesn’t ordinarily use idioms to describe situations that are regarded neutrally – buying tickets, reading a book […].” A few papers in phraseology mention the concept of semantic prosody: Conklin and Schmitt (2008, p. 71) discusses the dimensions of meaning that formulaic sequences carry and includes semantic prosody. Schmitt (2005) describes that ‘bordering on’ has a negative prosody which is not the case for the verb ‘border’ per se. Conversely, some articles dedicated to semantic prosody discuss idiomatic expressions or extended lexical units: Louw (1993) mentions ‘symptomatic of’, Channell (2000) discusses ‘par for the course’ and ‘roam the streets’, and Morley and Partington (2009) compare the semantic prosodies of ‘brimming with’ and ‘fraught with’.

These studies deal with constructions that have a high degree of idiomaticity. The constructions are often lexically constrained, meaning they have no or few open slots which in turn are often limited to a small number of lexical items. Their meanings tend to be non-compositional and/or their syntax deviates from the usual patterns of grammar. As such they are quite different from the open constructional patterns that are discussed below.

The bottom line is that while there are some similarities between Sinclarian text
linguistics and construction grammar, they are essentially two different schools with different areas of interest and different theoretical foci. Corpus linguists in the Sinclairian tradition who have produced most of the seminal theoretical work on the topic of semantic prosody lack a background in construction grammar. At the same time, linguists working on construction grammar often make an explicit commitment to usage-based linguistics but unfortunately, some of its foundational work is lacking in empiricism. The result is that few researchers conduct corpus research and investigate co-occurrence phenomena and distributional properties of linguistic units within a constructional approach. For the study of semantic prosody, this has meant that so far no one has considered it within a constructional framework; some research acknowledges Sinclair’s idiom-principle (or more generally the importance of linguistic units beyond the level of the word) to some extent but ultimately the data are always analyzed and discussed as a word-level phenomenon.

My dissertation aims to address exactly this issue. I will argue that what has previously been described as semantic prosody can be explained through the interactions of constructions and their words. This view is in stark contrast to previous research in this area, but against the background of the theoretical foundations described in chapter 3, this is by no means a radical proposal. It follows directly from an emergentist, usage-based account, where lexical meaning does not possess ontological primacy but is instead the result of abstraction over context-dependent meaning.

Anecdotal data from previous studies such as the examples by Louw and Bublitz
mentioned above indicate that a constructional account may be valid and relevant for the study of semantic prosody. In the remainder of this chapter, I investigate this question in a more systematic and rigorous way by analyzing the interplay between words and constructions with the help of collostructional analysis. The research question that guides this investigation is to what extent evaluative meaning can be attributed to the word alone, and to what extent it belongs to the construction. More specifically, this means that when we identify the different constructions a given word takes part in, we would expect to find semantic differences, including different semantic prosodies. This is the research hypothesis: There will be systematic semantic differences between different constructions a given word takes part in as expressed through their different collexeme profiles.

5.2 Words and their constructions

5.2.1 ‘Persistent’

When reading the dissertation of a student who was not a native speaker of English, Hunston (2007a) stumbled over the following sentence: “I would like to thank my supervisor for his persistent help and advice” (2007a, p. 259). She notes that this use of the word ‘persistent’ seemed infelicitous to her because it conflicted with her intuitions about its collocational characteristics. Based on KWIC (Key Word In Context) concordance lines, she provides a brief, impressionistic account of the kind of evaluations ‘persistent’ seems to be part of and concludes: “Whether ‘persistence’
in a given instance is interpreted as good or bad seems to depend entirely on the other items in the environment” [ibid, p. 254]. Hunston also remarks that it is consistently negative when followed by a noun.

I decided to conduct a more systematic investigation of ‘persistent’ and the two constructions it occurs in. First, I collected a random sample of 200 sentences containing the word from COCA. I identified which constructions it occurs in, which are the attributive constructions (as part of a noun phrase) and the predicative construction. The sentences were classified as to whether they were used in the attributive or predicative construction and categorized for their evaluative dimensions and semantic preferences (i.e. what kind of lexical fields they tend to occur with). A cross-tabulation of the constructions by the evaluative dimensions is given in Table 5.1.

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Attributive tokens</th>
<th>Predicative tokens</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>25</td>
<td>17</td>
<td>42</td>
</tr>
<tr>
<td>Neutral</td>
<td>17</td>
<td>8</td>
<td>25</td>
</tr>
<tr>
<td>Negative</td>
<td>114</td>
<td>19</td>
<td>133</td>
</tr>
<tr>
<td>Total</td>
<td>156</td>
<td>44</td>
<td>200</td>
</tr>
</tbody>
</table>

Table 5.1: Attitudinal stance found in 200 examples with the adjective ‘persistent’.

‘Persistent’ is much more commonly found in the attributive than in the predicative construction, with about 78% of all instances in the sample falling into the
former category. Overall, the analyzed sentences mostly express negative evaluations. In the attributive constructions, 114 out of 156 tokens or 70% of all occurrences are negative. Typical noun phrases include ‘persistent problems’, ‘persistent pain’, ‘persistent pollutants’ and the fixed expression ‘persistent vegetative state’. It describes medical, social and environmental issues that have become so entrenched within their systems that they have become difficult to rectify.

When used predicatively, however, no clear-cut semantic prosody can be found: There are almost as many sentences in the sample with positive as with negative evaluations (17 and 19 respectively). This is especially true for human entities in the subject position, where ‘persistent’ indicates a sense of hard work or tenacity (often in a struggle against injustice). The difference in evaluations between attributive and predicative use as shown in 5.2.1 is statistically significant (Fisher’s Exact Test, \( p < .001 \)). It has to be noted that predicative items are more difficult to interpret than attributive ones: When the adjective describes a noun, the lexical meaning of the noun often provides a strong hint about the evaluative prosody expressed in the noun phrase. For predicative uses, a clear indicator of the speaker’s attitudinal stance is not always present. ‘Persistent health problems’ or ‘a dramatic and persistent decline in per capita GNP’ describe obviously unfavorable states of affairs while sentences like ‘he is persistent’ do not carry such obvious attitudes. The evaluations expressed in examples 1 and 2 are easily understood, but example 3 is more problematic in this respect. At first, the juxtaposition with ‘aggressive’ may seem to indicate that
it is used in an unfavorable sense, but on the other hand, the meaning of ‘aggressive’ also needs to be considered in the context of ‘persistent’. The meaning could be that ACORN (an advocacy group) is being rightfully assertive and persevering in difficult situations. As a closer look at the example within its larger context shows, the speaker takes a nuanced stance towards the organization, pointing out both positive and negative aspects.

(1) I was both friendly and persistent and eventually got the information I sought.

(2) Dr. Renshaw says one reason lack of desire is so persistent is that people just don’t want to talk about it.

(3) From its inception, ACORN was ‘aggressive and persistent,’ he says

The attributive construction was further analyzed through collexeme analysis. All occurrences were retrieved from COCA. A collexeme analysis was performed to calculate the attraction between the construction and the words in its noun phrase slot. Table 5.2 shows the lemmas of the 20 collexemes with the greatest collostructional strength. There is a preference for lexical fields that describe medical problems, chemical pollution, social, political or economic problems and (less frequently) annoying people.

1Unless indicated otherwise, all numbered examples are authentic sentences taken from COCA. In some cases, they have been slightly shortened. For examples in running text, double quotes are used to indicate authentic data, and single quotes signify a constructed sentence that is purely provided to illustrate a pattern. Constructed sentences are idealized versions of typical usage scenarios.
<table>
<thead>
<tr>
<th>Rank</th>
<th>Collexeme</th>
<th>Tokens</th>
<th>Attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(vegetative) state</td>
<td>117</td>
<td>2,538</td>
</tr>
<tr>
<td>2</td>
<td>problem</td>
<td>228</td>
<td>2,217</td>
</tr>
<tr>
<td>3</td>
<td>pollutant</td>
<td>92</td>
<td>1,470</td>
</tr>
<tr>
<td>4</td>
<td>rumor</td>
<td>86</td>
<td>1,122</td>
</tr>
<tr>
<td>5</td>
<td>pain</td>
<td>100</td>
<td>1,024</td>
</tr>
<tr>
<td>6</td>
<td>cough</td>
<td>60</td>
<td>866</td>
</tr>
<tr>
<td>7</td>
<td>infection</td>
<td>61</td>
<td>675</td>
</tr>
<tr>
<td>8</td>
<td>poverty</td>
<td>54</td>
<td>585</td>
</tr>
<tr>
<td>9</td>
<td>contaminant</td>
<td>35</td>
<td>526</td>
</tr>
<tr>
<td>10</td>
<td>asthma</td>
<td>36</td>
<td>463</td>
</tr>
<tr>
<td>11</td>
<td>symptom</td>
<td>42</td>
<td>438</td>
</tr>
<tr>
<td>12</td>
<td>effort</td>
<td>48</td>
<td>416</td>
</tr>
<tr>
<td>13</td>
<td>OME</td>
<td>21</td>
<td>396</td>
</tr>
<tr>
<td>14</td>
<td>pattern</td>
<td>45</td>
<td>385</td>
</tr>
<tr>
<td>15</td>
<td>difficulty</td>
<td>40</td>
<td>380</td>
</tr>
<tr>
<td>16</td>
<td>critic</td>
<td>35</td>
<td>310</td>
</tr>
<tr>
<td>17</td>
<td>threat</td>
<td>35</td>
<td>285</td>
</tr>
<tr>
<td>18</td>
<td>myth</td>
<td>27</td>
<td>280</td>
</tr>
<tr>
<td>19</td>
<td>delay</td>
<td>27</td>
<td>278</td>
</tr>
</tbody>
</table>
We can explore the collexeme profile further through cluster analysis. Figure 5.1 shows a network representation of how the collexemes are clustered according to their representations in WordNet. This visualization is just meant to give a rough overview of what this kind of network looks like. Each dot is a node in the network, representing a synset and each line is a direct connection, representing a semantic relationship of hypernymy or hyponymy between the synsets. The different communities are color coded. The positions of the nodes are not meaningful (the function that creates the visualization just tries to arrange them in some way). We see that it is a fairly sparse network, with many nodes just being connected to two or three other nodes. Next, we can look at the individual communities in more detail.

One of the communities is shown in figure 5.2. Here, I have assigned a different meaning to the colors of the dots. Red dots depict collexemes of persistent NP, blue dots are not and were just retrieved from WordNet to connect the collexemes. We see that several collexemes belonging to the lexical field of illness have been correctly identified by the clustering algorithm, e.g. ‘pain’, ‘fever’, ‘cough’ and ‘symptom’. Oddly enough, the word ‘worry’ is included in this cluster too, albeit as a peripheral member. The reason seems to be that all words have to be assigned to some cluster and worry did not fit anywhere else any better.
Figure 5.1: Collexeme clusters for ‘persistent’ in WordNet.
Figure 5.2: Collexemes of *persistent NP* mostly belonging to the lexical field of illness.

Figure 5.3 shows another community. It includes internal states, such as ‘concern’, ‘depression’ and ‘anxiety’. It seems that ‘worry’ should have been assigned to this group. The reason why this did not happen is that ‘worry’ is not classified as a psychological state but as a negative stimulus. It should be noted that WordNet does not aim to be a fully consistent ontology of all existing entities, and I am using it in a way it was not technically intended for. Another subgroup that is part of this cluster includes words like ‘poverty’, ‘lack’ and ‘drought’, but also the very general term ‘problem’.

An entirely different way of finding structure within lists of words uses their distributional similarities. Figure 5.4 depicts the hierarchical clusters found among all collexemes for ‘persistent’ in the attributive construction based on their word vectors.
Figure 5.3: Collexemes of *persistent NP* mostly belonging to the lexical fields negative mental states and scarcity.

The clustering algorithm was able to group together words from the semantic field of illness with almost perfect precision. The only words that are not included even though they should have are ‘arthralgia’ and ‘OME’, both of which are very rare and technical terms. Additionally, for the most part, the subdivisions in this category also make sense. ‘Depression’ and ‘anxiety’, both mental issues, form one cluster and on the next level of the hierarchy, they are connected to clusters containing ‘ache’, ‘pain’ and ‘headache’ and to clusters containing ‘fatigue’, ‘symptom’ and ‘weakness’. Another node up in the hierarchy those clusters connect to ‘fever’, ‘cough’ and ‘diarrhea’. Many of the other groupings also correspond to lexical fields. There is the field of social issues, which is subdivided into the more personal terms ‘bias’ and ‘discrimination’ on the one hand and the wider societal problems ‘inequality’ and ‘poverty’ on the
Figure 5.4: Collexeme clusters for persistent NP found through hierarchical cluster analysis of their word vectors.
other hand. It is closely connected to clusters describing economic issues: ‘decline’, ‘unemployment’, ‘inflation’ and ‘deficit’. Other clusters include lexical fields of physical conflict (‘conflict’, ‘struggle’, ‘violence’, ‘attack’ and ‘threat’), questionable beliefs (‘myth’, ‘rumor’, ‘faith’, ‘believe’) or environmental pollution (‘pesticide’, ‘herbicide’, ‘contaminant’, ‘pollutant’). While there is no such thing as a definite ontology that perfectly captures reality, these groupings generally match my own intuitions very closely. Additionally, the way the dendrogram visually arranges the information is convenient and makes it easy to find lexical fields within the data.

Discussion

‘Persistent NP’ and ‘be persistent’ have taken on different meanings through their different distributions in language use. The prototypical meaning of the former could perhaps be paraphrased as ‘of an undesirable state of affairs (typically health issues, pollutants or social, political or economic problems): having the tendency to remain unchanged.’ The latter means something like ‘having a tendency to continue an action in the face of difficulty.’ These two different meanings are not absolute but instead describe probabilistic tendencies. It is noteworthy that there is a clear split between the two forms in terms of the arguments they take. The attributive use occurs largely with states of affairs while predication is more common with humans. This asymmetry may well have played a part in the emergence of the different meanings. At this point, however, the two meanings have become entrenched and are not dependent on the
animacy of the arguments they take. ‘Persistent’, when used attributively with a person, more often than not, displays a negative attitude.

There are two main takeaways from this case study. First, it provides evidence that semantic prosody cannot solely be assigned to the word-level as a sort of context-independent lexical meaning. Instead, there is also construction-specific meaning. If there is such a thing as semantic transfer, it does not permeate to the level of the word. Constructional information is retained as meaning is shaped through usage events. This meaning is not conflated across different grammatical patterns. The attributive use of ‘persistent’ does not carry over to its predicative use. I discuss the implications of these data and the data from the following studies for the notion of semantic transfer in the conclusion to this chapter. Second, we see that evaluative meaning cannot be broken down into a simple positive vs negative polarity. Instead, it encompasses far more complex facets of meaning. This is a point that will be revisited in the next chapter in more detail.

5.2.2 ‘Cause’ as a verb

The verb ‘cause’ is perhaps the single most-discussed word in the literature on semantic prosody. It has been examined in numerous studies and has become a go-to example for researchers that has been discussed in numerous context, from its implications for the theory of semantic prosody to diachronous studies, language learning and contrastive linguistics (Bublitz, 1996; Hunston, 2007a; Smith and Nordquist,
What makes ‘cause’ interesting is that it occurs in various grammatical constructions. I hypothesize that some of the different constructions ‘cause’ takes part in will differ in their collexeme profiles and in their evaluative semantics.

The first systematic investigation is found in Stubbs (1995a). His study of ‘cause’ in the LOB corpus (the Lancaster-Oslo-Bergen corpus, a one million word corpus of British English that replicates the design of the Brown Corpus) finds that ‘cause’ occurs in 80% negative, 18% neutral, 2% positive evaluations. He concludes: “CAUSE acquires guilt by association. At some point, the word itself acquires unpleasant connotations and affects” (Stubbs, 1995a, p. 51). Larger studies by Xiao and McEnery (2006) and Dam-Jensen and Zethsen (2008) have since found similar results for ‘cause’. A simple collocational analysis seems to confirm the findings of these results: Table 5.3 shows the ten most common noun collocates for the verb ‘cause’ within a 4:4 search window in COCA. The counts are given for the lemma of the collocates, i.e. singular and plural forms are combined. When we conflate across all forms of the verb, the collocates do indeed paint an overwhelmingly negative picture. All the top collocates either describe bodily injury (‘disease’, ‘harm’, ‘pain’) or generally bad things (‘problem’, ‘trouble’, ‘loss’).
Table 5.3: The most common noun collocates for the
verb ‘cause’ in COCA.

<table>
<thead>
<tr>
<th>Collocate</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>problem</td>
<td>3,276</td>
</tr>
<tr>
<td>damage</td>
<td>3,105</td>
</tr>
<tr>
<td>disease</td>
<td>1,822</td>
</tr>
<tr>
<td>pain</td>
<td>1,654</td>
</tr>
<tr>
<td>cancer</td>
<td>1,232</td>
</tr>
<tr>
<td>trouble</td>
<td>1,203</td>
</tr>
<tr>
<td>harm</td>
<td>1,142</td>
</tr>
<tr>
<td>injury</td>
<td>1,109</td>
</tr>
<tr>
<td>loss</td>
<td>998</td>
</tr>
<tr>
<td>infection</td>
<td>799</td>
</tr>
</tbody>
</table>

However, as indicated before, a simple collocational study as in 5.3 does not reveal the entire picture. In order to study the types and the semantics of the constructions that ‘cause’ takes part in, I adopt a two-step approach as described in chapter 4: I first perform a manual investigation of a manageable sample of sentences including ‘cause’ through which all its constructions are identified. In a second step, the relevant constructions are then retrieved from the entire corpus and the association between the constructions and their collexemes (i.e. the arguments the constructions take) are
calculated with the help of collostructional analysis.

A sample of 200 occurrences of the lemma ‘cause’ was retrieved COCA. The sentences were classified based on their grammatical constructions, and – despite the difficulties with this approach – an attempt was made to classify them according to their evaluative polarity. Due to the relative rarity of unequivocally positive statements and the difficulty in distinguishing between positive and neutral evaluations, the two were combined and a distinction was only made between negative and non-negative evaluations. The results are summarized in Table 5.4. Five different constructions were identified: The most common is the transitive construction (example 4), where the caused event occurs in the direct object slot of the construction. Next is the infinitival complement construction (5) which is slightly less common. This is the only construction where the caused event is not encoded with a noun phrase but instead with a complement clause with the to infinitive. The passive use of ‘cause’ (6) is also attested and relatively common. ‘Cause’ occurs in both constructions of the English dative alternation, the ditransitive (or double object) construction (7) and the prepositional dative (8) but both are rare.

<table>
<thead>
<tr>
<th>Construction</th>
<th>Tokens</th>
<th># Negative</th>
<th>% Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transitive</td>
<td>86</td>
<td>70</td>
<td>0.81</td>
</tr>
<tr>
<td>Infinitival complement</td>
<td>78</td>
<td>33</td>
<td>0.42</td>
</tr>
<tr>
<td>Passive</td>
<td>26</td>
<td>13</td>
<td>0.50</td>
</tr>
<tr>
<td>Construction Type</td>
<td>Count</td>
<td>Collexeme</td>
<td>Number of Uses</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>-------</td>
<td>-----------</td>
<td>----------------</td>
</tr>
<tr>
<td>Ditransitive</td>
<td>6</td>
<td>6</td>
<td>1.00</td>
</tr>
<tr>
<td>Prepositional dative</td>
<td>4</td>
<td>4</td>
<td>1.00</td>
</tr>
<tr>
<td>Total</td>
<td>200</td>
<td>126</td>
<td>0.63</td>
</tr>
</tbody>
</table>

Table 5.4: Constructions used with verbal ‘cause’.

(4) The conservation leader said other mistakes in spraying had **caused** serious damage in Ohio and Wyoming (*transitive*)

(5) Developments over the next two months, however, **caused** the President to reconsider the question of the timing. (*infinitival complement*)

(6) There will be losses **caused** by emergencies that arise while he is away at his off-farm job. (*passive*)

(7) In any other man this reassurance to the electorate would have **caused** us a profound moral shock. (*ditransitive*)

(8) [...] he knew that retaliation could **cause** only violent warfare and disaster to business [...] (*prepositional dative*)

All five constructions were retrieved from COCA and were analyzed with the help of collexeme analysis. I discuss each construction in its own section.
Transitive construction

Table 5.5 shows the 20 nouns with the strongest attraction to the direct object slot of the transitive construction with ‘cause’ as its head verb. The noun phrases have been reduced to their head nouns (determiners and adjectives have been removed) and the counts are based on the lemmatized forms of both ‘cause’ and the relevant collexeme, e.g. the token count for ‘problem’ includes expression like ‘caused a problem’ or ‘causing great problems’. Not included are expressions that are part of other constructions like ‘caused him problems’ which belongs to the ditransitive construction.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Collexeme</th>
<th>Tokens</th>
<th>Attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>problem</td>
<td>1,579</td>
<td>10,891</td>
</tr>
<tr>
<td>2</td>
<td>damage</td>
<td>835</td>
<td>8,408</td>
</tr>
<tr>
<td>3</td>
<td>harm</td>
<td>478</td>
<td>5,356</td>
</tr>
<tr>
<td>4</td>
<td>trouble</td>
<td>619</td>
<td>5,201</td>
</tr>
<tr>
<td>5</td>
<td>death</td>
<td>622</td>
<td>4,180</td>
</tr>
<tr>
<td>6</td>
<td>cancer</td>
<td>575</td>
<td>3,916</td>
</tr>
<tr>
<td>7</td>
<td>stir</td>
<td>253</td>
<td>3,653</td>
</tr>
<tr>
<td>8</td>
<td>pain</td>
<td>395</td>
<td>2,911</td>
</tr>
<tr>
<td>9</td>
<td>disease</td>
<td>332</td>
<td>1,813</td>
</tr>
<tr>
<td>10</td>
<td>injury</td>
<td>198</td>
<td>1,408</td>
</tr>
<tr>
<td>11</td>
<td>change</td>
<td>321</td>
<td>1,361</td>
</tr>
</tbody>
</table>
Table 5.5: Collexemes with the strongest attraction to the direct object slot in the transitive causative construction.

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>confusion</td>
<td>141</td>
</tr>
<tr>
<td>13</td>
<td>accident</td>
<td>155</td>
</tr>
<tr>
<td>14</td>
<td>havoc</td>
<td>84</td>
</tr>
<tr>
<td>15</td>
<td>AIDS</td>
<td>126</td>
</tr>
<tr>
<td>16</td>
<td>concern</td>
<td>184</td>
</tr>
<tr>
<td>17</td>
<td>illness</td>
<td>126</td>
</tr>
<tr>
<td>18</td>
<td>explosion</td>
<td>95</td>
</tr>
<tr>
<td>19</td>
<td>crash</td>
<td>112</td>
</tr>
<tr>
<td>20</td>
<td>inflammation</td>
<td>97</td>
</tr>
</tbody>
</table>

The object slot in the transitive construction shows a strong association with words from lexical fields of health issues or diseases (e.g. ‘pain’, ‘harm’, ‘AIDS’ etc). Another semantic field that is commonly encountered in this construction concerns destruction (‘damage’, ‘explosion’, ‘havoc’). Other very strongly associated collexemes include the more general terms ‘problem’ and ‘trouble’.

Due to the productivity of *cause* in the transitive constructions, it may be useful to further explore its collexemes through cluster analysis. Figure 5.5 is a dendrogram of the hierarchical clusters for the 40 collexemes with the strongest attraction to the construction, based on their word vectors. This analysis largely confirms the
lexical fields described in the previous paragraph but also indicates that there are further lexical fields. Many of the collexemes belong to a lexical field of illness of physical discomfort. Another field that is less obvious within the 20 top collocates presented in 5.5 concerns words like ‘consternation’, ‘uproar’ or ‘controversy’. Also, according to their word vectors, collexemes expressing destruction (‘damage’, ‘crash’) are connected to terms of physical harm, like ‘injury’ or ‘death’.

The only two collexemes with a high collostructional strength that are not inherently negative are ‘change’ and ‘stir’. The word ‘change’ by itself does not evoke any negative connotations. It could be used to denote positive, negative or neutral change. When change is *caused*, however, it occurs in mostly negative contexts. In a random 200 sentence sample of ‘cause change’ retrieved from COCA, 75% of the sentences were negative, 12% were neutral or unclear in their evaluation and only 4% were positive. The changes described in the sample sentences often concern environments issues or illnesses, e.g. “cause changes inside of cells that lead to cancer”. The picture for ‘stir’ is different: Causing a stir means something like bringing out excitement in people through one’s actions. This can be positive excitement as in example (9) or it can be bad and scandalous as in (10). Finally, it can be devoid of any obvious positive or negative evaluation (11). It is not entirely clear, why ‘stir’ is protected from being coerced into carrying a negative evaluation in the same way as ‘change’. A possible answer might be that it has to do with the fact that ‘cause a stir’ is a fixed idiom.
Figure 5.5: Collexeme clusters for *cause* in the transitive construction, found through a hierarchical cluster analysis of their word vectors.
(9) A new crop of young, multicultural, female hip-hop acts is causing a stir on the
Internet and in indie-label conference rooms

(10) The incident caused a stir, but Eugene declined to press charges

(11) His arrival in the village caused a stir.

The data for the agentive slot of the transitive construction are not provided here in tabular format but can be summarized as follows: The most common collexemes in the agentive slot are disease agents (‘virus’, ‘bacteria’) and unhealthy substances or behaviors (‘chemicals’, ‘smoking’, ‘drugs’), which makes sense considering that many of the caused events are health issues. The vast majority of agents associated with ‘trouble’ are humans. For ‘problem’ no generalizations seem possible; it co-occurs with all sorts of entities, from animates to abstract states of affairs.

Prepositional dative construction

Table 5.6 shows the 20 nouns with the strongest attraction to the direct object slot in the prepositional dative construction. Like the data found in the 200 sentence sample, the frequency of the causative prepositional dative is relatively low compared to the transitive or passive construction. The top four collexemes which make up more than 85% of all occurrences of the construction, all express a sense of violation of bodily integrity. Other strongly attracted collexemes share a sense of irritation, which includes both physical and related psychological effects (‘distress’, ‘embarrassment’, ‘discomfort’, ‘inconvenience’). The recipient (not given in the table) is mostly either
a person, a human body part, an organization or (less frequently) the environment.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Collexeme</th>
<th>Tokens</th>
<th>Attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>damage</td>
<td>414</td>
<td>6,925</td>
</tr>
<tr>
<td>2</td>
<td>harm</td>
<td>187</td>
<td>3,185</td>
</tr>
<tr>
<td>3</td>
<td>injury</td>
<td>70</td>
<td>882</td>
</tr>
<tr>
<td>4</td>
<td>pain</td>
<td>39</td>
<td>400</td>
</tr>
<tr>
<td>5</td>
<td>distress</td>
<td>12</td>
<td>145</td>
</tr>
<tr>
<td>6</td>
<td>concern</td>
<td>14</td>
<td>108</td>
</tr>
<tr>
<td>7</td>
<td>irritation</td>
<td>7</td>
<td>96</td>
</tr>
<tr>
<td>8</td>
<td>stress</td>
<td>7</td>
<td>57</td>
</tr>
<tr>
<td>9</td>
<td>inflation</td>
<td>6</td>
<td>56</td>
</tr>
<tr>
<td>10</td>
<td>suffering</td>
<td>5</td>
<td>49</td>
</tr>
<tr>
<td>11</td>
<td>resistance</td>
<td>6</td>
<td>46</td>
</tr>
<tr>
<td>12</td>
<td>problem</td>
<td>9</td>
<td>37</td>
</tr>
<tr>
<td>13</td>
<td>change</td>
<td>8</td>
<td>35</td>
</tr>
<tr>
<td>14</td>
<td>danger</td>
<td>4</td>
<td>34</td>
</tr>
<tr>
<td>15</td>
<td>disruption</td>
<td>3</td>
<td>26</td>
</tr>
<tr>
<td>16</td>
<td>blow</td>
<td>3</td>
<td>26</td>
</tr>
<tr>
<td>17</td>
<td>exposure</td>
<td>4</td>
<td>24</td>
</tr>
<tr>
<td>18</td>
<td>loss</td>
<td>4</td>
<td>22</td>
</tr>
</tbody>
</table>
Table 5.6: Collexemes with the strongest attraction to the *theme slot* (= caused event) in the prepositional dative causative construction.

**Ditransitive dative construction**

The ditransitive construction summarized in Table 5.7 shows a focus on internal states. The three collexemes most strongly attracted to the construction are ‘pain’, ‘problem’ and ‘trouble’. Otherwise, most words designate negative emotions (‘grief’, ‘distress’, ‘anxiety’, ‘embarrassment’).

<table>
<thead>
<tr>
<th>Rank</th>
<th>Collexeme</th>
<th>Tokens</th>
<th>Attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>pain</td>
<td>216</td>
<td>2,837</td>
</tr>
<tr>
<td>2</td>
<td>trouble</td>
<td>141</td>
<td>1,740</td>
</tr>
<tr>
<td>3</td>
<td>problem</td>
<td>182</td>
<td>1,726</td>
</tr>
<tr>
<td>4</td>
<td>grief</td>
<td>48</td>
<td>653</td>
</tr>
<tr>
<td>5</td>
<td>distress</td>
<td>46</td>
<td>650</td>
</tr>
<tr>
<td>6</td>
<td>harm</td>
<td>48</td>
<td>643</td>
</tr>
<tr>
<td>7</td>
<td>discomfort</td>
<td>38</td>
<td>560</td>
</tr>
<tr>
<td>8</td>
<td>concern</td>
<td>45</td>
<td>421</td>
</tr>
<tr>
<td>9</td>
<td>anxiety</td>
<td>35</td>
<td>405</td>
</tr>
</tbody>
</table>
An interesting point concerns the difference between the two dative constructions. The prototypical meaning of the English dative constructions has been described as the successful transfer of a theme from an agent to a recipient (cf. Goldberg, 1995, p. 36). When both the ditransitive and the prepositional dative are instantiated with ‘cause’ in their verb slot, there is a successful metaphorical transfer of an abstract theme to a human experiencer. The semantic difference between the two constructions can be further investigated through the method of distinctive collexeme analysis.
Table 5.8 shows the 20 collexemes with the greatest degree of attraction to one construction over the other. The data corroborate the impression that the ditransitive is associated with internal states while the prepositional dative tends to describe physical damage and harm. There is, in fact, a very clearly delineated separation between the two constructions along the lines of the semantic fields of their collexemes. Even though both constructional patterns have low occurrence frequencies, they have become specialized towards expressing only internal or external states respectively. The semantic differences in the two constructions of the dative alternation are an example of the Principle of No Synonymy of Grammatical Forms. A famous version of the principle of no synonymy in its general form was formulated by Bolinger, who noted that differences in form always implied (possibly) subtle differences in meaning: “A difference in syntactic form always spells a difference in meaning” (Bolinger, 1968, p. 127). Construction grammarians emphasize that the principle of no synonymy holds not only for lexical items but also for constructions (including argument structure constructions). As Goldberg puts it, “if two constructions are syntactically distinct, they must be semantically or pragmatically distinct” (Goldberg, 1995, p. 67).

Ramonda (2014, p. 60) cites two examples from Fillmore (1968, p. 49) to illustrate this principle:

(12) “Bees are swarming in the garden.”

(13) “The garden is swarming with bees.”

Ramonda (2014) argues that the difference between the two sentences consists
in whether the bees occupy the entire garden or just a part of it: Example 13 is a statement about the garden (meaning the entire garden). The sentence indicates that the bees are flying around the whole garden. Example 12 focusses on the bees, indicating that they are limited to one area in the garden. While this explanation sounds intuitive, there is really no evidence for it. The distinctive collexeme analysis of the dative alternation, on the other hand, provides concrete data that support the principle of no synonymy: There are two related constructions that are similar but not the same, both in form and function. They become specialized in a quantifiable way, with one expressing the transfer of mental states and the other describing the transfer of physical states.  

---

2It is not entirely clear how this distinction is motivated. It could be iconic, i.e. there might something about the grammatical structure of the constructions that make them conceptually more suitable for either semantic field. There is a tremendous amount of literature on the motivation behind the dative alternation. Stefanowitsch and Gries (2003) argue that in the ditransitive, the theme is metaphorically transferred to the recipient; the metaphorical recipient is interpreted as an experiencer of this result. The inclusion of the experiencer makes the ditransitive suitable for encoding mental states and experiences. Alternatively, a motivation could be found in linguistic interaction. Finally, it may just have been some kind of random drift: It may just have happened because some relatively frequent collexemes became attached to one of the constructions more or less randomly; once they were associated with the construction, they then attracted more semantically similar lexemes to the construction.
<table>
<thead>
<tr>
<th>Collexeme</th>
<th>Ditransitive</th>
<th>Prep. Dat.</th>
<th>Preference</th>
<th>Attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>damage</td>
<td>14</td>
<td>414</td>
<td>Prep. Dat.</td>
<td>261</td>
</tr>
<tr>
<td>trouble</td>
<td>141</td>
<td>0</td>
<td>Ditransitive</td>
<td>68</td>
</tr>
<tr>
<td>problem</td>
<td>182</td>
<td>9</td>
<td>Ditransitive</td>
<td>64</td>
</tr>
<tr>
<td>harm</td>
<td>48</td>
<td>187</td>
<td>Prep. Dat.</td>
<td>60</td>
</tr>
<tr>
<td>pain</td>
<td>216</td>
<td>39</td>
<td>Ditransitive</td>
<td>40</td>
</tr>
<tr>
<td>injury</td>
<td>10</td>
<td>70</td>
<td>Prep. Dat.</td>
<td>32</td>
</tr>
<tr>
<td>grief</td>
<td>48</td>
<td>0</td>
<td>Ditransitive</td>
<td>23</td>
</tr>
<tr>
<td>discomfort</td>
<td>38</td>
<td>0</td>
<td>Ditransitive</td>
<td>18</td>
</tr>
<tr>
<td>anxiety</td>
<td>35</td>
<td>0</td>
<td>Ditransitive</td>
<td>16</td>
</tr>
<tr>
<td>difficulty</td>
<td>32</td>
<td>0</td>
<td>Ditransitive</td>
<td>15</td>
</tr>
<tr>
<td>embarrassment</td>
<td>21</td>
<td>0</td>
<td>Ditransitive</td>
<td>10</td>
</tr>
<tr>
<td>deal</td>
<td>19</td>
<td>0</td>
<td>Ditransitive</td>
<td>9</td>
</tr>
<tr>
<td>anguish</td>
<td>18</td>
<td>0</td>
<td>Ditransitive</td>
<td>8</td>
</tr>
<tr>
<td>weight</td>
<td>18</td>
<td>0</td>
<td>Ditransitive</td>
<td>8</td>
</tr>
<tr>
<td>distress</td>
<td>46</td>
<td>12</td>
<td>Ditransitive</td>
<td>6</td>
</tr>
<tr>
<td>inconvenience</td>
<td>14</td>
<td>0</td>
<td>Ditransitive</td>
<td>6.</td>
</tr>
<tr>
<td>change</td>
<td>0</td>
<td>8</td>
<td>Prep. Dat.</td>
<td>5</td>
</tr>
<tr>
<td>concern</td>
<td>45</td>
<td>14</td>
<td>Ditransitive</td>
<td>5</td>
</tr>
</tbody>
</table>
Another observation is that the prepositional dative accounts for only 4 out of the 20 collexemes, but 3 out of 4 of these occur in the top half of the table. This is both because there is little overlap between the two sets of collexemes and because the prepositional construction shows an even stronger lexical bias than the ditransitive. The collostructional profiles of both constructions have long-tailed distributions, i.e. a very small number of collexeme types accounts for the majority of all tokens while the rest of the collexemes occur only very few times (including a long tail of types that occurs only once or twice). This tendency is notably stronger in the prepositional dative than in the ditransitive dative.

**Passive construction**

The corpus data showed that passivized ‘cause’ without an explicitly encoded agentive argument is extremely rare. I will, therefore, limit my discussion to passives that include the agent of the causation. Tables 5.9 shows the 20 collexemes with the strongest attraction to the patient role (the result of the causation). Table 5.10 shows the 20 collexemes with the strongest attraction to the agent role.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Collexeme</th>
<th>Tokens</th>
<th>Attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>damage</td>
<td>557</td>
<td>7,282</td>
</tr>
<tr>
<td></td>
<td>Collexeme</td>
<td>Count</td>
<td>Attraction</td>
</tr>
<tr>
<td>---</td>
<td>----------------</td>
<td>-------</td>
<td>------------</td>
</tr>
<tr>
<td>2</td>
<td>problem</td>
<td>513</td>
<td>4,324</td>
</tr>
<tr>
<td>3</td>
<td>disease</td>
<td>384</td>
<td>3,654</td>
</tr>
<tr>
<td>4</td>
<td>death</td>
<td>267</td>
<td>2,346</td>
</tr>
<tr>
<td>5</td>
<td>injury</td>
<td>195</td>
<td>2,113</td>
</tr>
<tr>
<td>6</td>
<td>infection</td>
<td>201</td>
<td>1,911</td>
</tr>
<tr>
<td>7</td>
<td>loss</td>
<td>191</td>
<td>1,799</td>
</tr>
<tr>
<td>8</td>
<td>illness</td>
<td>125</td>
<td>1,396</td>
</tr>
<tr>
<td>9</td>
<td>harm</td>
<td>108</td>
<td>1,290</td>
</tr>
<tr>
<td>10</td>
<td>pain</td>
<td>139</td>
<td>1,255</td>
</tr>
<tr>
<td>11</td>
<td>accident</td>
<td>99</td>
<td>1,033</td>
</tr>
<tr>
<td>12</td>
<td>delay</td>
<td>89</td>
<td>953</td>
</tr>
<tr>
<td>13</td>
<td>suffering</td>
<td>81</td>
<td>944</td>
</tr>
<tr>
<td>14</td>
<td>change</td>
<td>144</td>
<td>913</td>
</tr>
<tr>
<td>15</td>
<td>outbreak</td>
<td>72</td>
<td>841</td>
</tr>
<tr>
<td>16</td>
<td>disruption</td>
<td>73</td>
<td>829</td>
</tr>
<tr>
<td>17</td>
<td>pollution</td>
<td>77</td>
<td>786</td>
</tr>
<tr>
<td>18</td>
<td>cancer</td>
<td>101</td>
<td>714</td>
</tr>
<tr>
<td>19</td>
<td>effect</td>
<td>122</td>
<td>700</td>
</tr>
<tr>
<td>20</td>
<td>devastation</td>
<td>46</td>
<td>680</td>
</tr>
</tbody>
</table>

Table 5.9: Collexemes with the strongest attraction to the *theme slot* (= caused event) in the passive construction.
The relationship between caused and causing event has not been investigated quantitatively but a qualitative inspection of the retrieved sentences shows some clear tendencies in the data. Patients that belong to the category of illness and injury are typically partnered with agents from the semantic fields of disease agents. Expression of (environmental) destruction collocate with catastrophic events. The collexemes of the passive construction are relatively widely spread out, i.e. they show a lesser lexical bias than, the two dative constructions. There are numerous other collexemes that are both still relatively common and that do not fall into the lexical fields described here.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Collexeme</th>
<th>Tokens</th>
<th>Attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>bacterium</td>
<td>146</td>
<td>2,292</td>
</tr>
<tr>
<td>2</td>
<td>virus</td>
<td>179</td>
<td>1,955</td>
</tr>
<tr>
<td>3</td>
<td>lack</td>
<td>102</td>
<td>955</td>
</tr>
<tr>
<td>4</td>
<td>exposure</td>
<td>100</td>
<td>947</td>
</tr>
<tr>
<td>5</td>
<td>change</td>
<td>106</td>
<td>705</td>
</tr>
<tr>
<td>6</td>
<td>infection</td>
<td>79</td>
<td>675</td>
</tr>
<tr>
<td>7</td>
<td>disease</td>
<td>88</td>
<td>656</td>
</tr>
<tr>
<td>8</td>
<td>stress</td>
<td>61</td>
<td>588</td>
</tr>
<tr>
<td>9</td>
<td>failure</td>
<td>60</td>
<td>521</td>
</tr>
<tr>
<td>10</td>
<td>water</td>
<td>71</td>
<td>477</td>
</tr>
<tr>
<td>11</td>
<td>hurricane</td>
<td>38</td>
<td>467</td>
</tr>
</tbody>
</table>
Table 5.10: Collexemes with the strongest attraction to the agent slot (= causer) in the passive construction.

<table>
<thead>
<tr>
<th></th>
<th>Collexeme</th>
<th>Rank</th>
<th>Attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>factor</td>
<td>70</td>
<td>460</td>
</tr>
<tr>
<td>13</td>
<td>mutation</td>
<td>60</td>
<td>449</td>
</tr>
<tr>
<td>14</td>
<td>trauma</td>
<td>38</td>
<td>432</td>
</tr>
<tr>
<td>15</td>
<td>fungus</td>
<td>30</td>
<td>427</td>
</tr>
<tr>
<td>16</td>
<td>activity</td>
<td>70</td>
<td>424</td>
</tr>
<tr>
<td>17</td>
<td>earthquake</td>
<td>37</td>
<td>419</td>
</tr>
<tr>
<td>18</td>
<td>pollution</td>
<td>42</td>
<td>416</td>
</tr>
<tr>
<td>19</td>
<td>condition</td>
<td>57</td>
<td>387</td>
</tr>
<tr>
<td>20</td>
<td>air</td>
<td>55</td>
<td>383</td>
</tr>
</tbody>
</table>

**Infinitival complement construction**

The infinitival complement construction is fundamentally different from the other constructions discussed so far. The transitive, passive, ditransitive and prepositional dative construction are all monoclausal and encode the result of the causation in a noun phrase. The infinitival complement, on the other hand, is a periphrastic causative where the caused micro-event is expressed with a *to-infinitive* complement. This makes the application of collexeme analysis less straightforward. Unlike noun phrases which can often be reduced to nouns while retaining the relevant information, complement clauses cannot be pared down to a single unit that can be treated as the
collexeme. One option is to consider only the verb of the complement clause although this makes for a poor representation of the clausal meaning. Nonetheless, it can be used as a starting point of the investigation. The data for the collexeme analysis on the verb of the complement clause are given in Table 5.11.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Collexeme</th>
<th>Tokens</th>
<th>Attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>lose</td>
<td>543</td>
<td>5,191</td>
</tr>
<tr>
<td>2</td>
<td>fall</td>
<td>295</td>
<td>2,913</td>
</tr>
<tr>
<td>3</td>
<td>become</td>
<td>350</td>
<td>2,464</td>
</tr>
<tr>
<td>4</td>
<td>say</td>
<td>342</td>
<td>1,407</td>
</tr>
<tr>
<td>5</td>
<td>rise</td>
<td>155</td>
<td>1,396</td>
</tr>
<tr>
<td>6</td>
<td>feel</td>
<td>200</td>
<td>1,393</td>
</tr>
<tr>
<td>7</td>
<td>drop</td>
<td>146</td>
<td>1,320</td>
</tr>
<tr>
<td>8</td>
<td>miss</td>
<td>155</td>
<td>1,252</td>
</tr>
<tr>
<td>9</td>
<td>break</td>
<td>155</td>
<td>1,248</td>
</tr>
<tr>
<td>10</td>
<td>swell</td>
<td>98</td>
<td>1,235</td>
</tr>
<tr>
<td>11</td>
<td>change</td>
<td>175</td>
<td>1,213</td>
</tr>
<tr>
<td>12</td>
<td>turn</td>
<td>177</td>
<td>1,206</td>
</tr>
<tr>
<td>13</td>
<td>look</td>
<td>204</td>
<td>1,159</td>
</tr>
<tr>
<td>14</td>
<td>collapse</td>
<td>99</td>
<td>1,106</td>
</tr>
<tr>
<td>15</td>
<td>stop</td>
<td>139</td>
<td>1,059</td>
</tr>
</tbody>
</table>
Table 5.11: Collexemes with the strongest attraction to the infinitival complement construction.

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>go</td>
<td>307</td>
</tr>
<tr>
<td>17</td>
<td>vomit</td>
<td>47</td>
</tr>
<tr>
<td>18</td>
<td>rethink</td>
<td>65</td>
</tr>
<tr>
<td>19</td>
<td>question</td>
<td>94</td>
</tr>
<tr>
<td>20</td>
<td>move</td>
<td>129</td>
</tr>
</tbody>
</table>

Some of the verbs seem like they would mostly be used to express negatively evaluated events (e.g. ‘lose’, ‘fall’ or break’) but others appear to be entirely neutral (‘say’, ‘become’). There are also words like ‘feel’, which clearly have an emotional and evaluative meaning but without looking at the data, it is not obvious which way it would go. Clearly, a more in-depth analysis of the distribution of the verbs is required. For the 10 verbs with the strongest attraction to the construction, I looked at a random sample of 200 sentences of the verb in the construction ‘cause NP to V’. If less than 200 sentences were available, all examples were analyzed. Where possible, an attempt was made at classifying the examples with respect to their evaluative polarity.

‘Lose’: In most examples, the causee is a person and the caused event is about a loss concerning internal states like consciousness, sleep, confidence, focus, interest or faith. The loss of more concrete things is also relatively common and typically relates to losing ownership and as such is expressed with possessives, e.g. someone
lost their job, their house, their license. The great majority of the examples carry a negative evaluation. 97% of the sentences describe a negative event, 3% are neutral and none are positive. ‘Fall’: By far the most common case is about people physically falling, often resulting in injury. It is also used in the metaphorical sense of decreasing numbers (e.g. prices, stocks). 83% of the sentences were negative, 2% positive and 16% neutral. ‘Become’: Here the picture is less clear. 56% of the sentences in the sample were classified as negative, 19% as positive and 25% were neutral or carried no clear evaluation. No particular semantic generalization can be made. ‘Say’: This seems completely neutral. The sentence does not express any evaluative attitude. ‘Rise’: The most common causees include ‘wages’, ‘stocks’, ‘insurance rates’, ‘costs’. Other examples are ‘sea level’ or ‘popularity’. About 80% are negative, 10% are positive and 10% are neutral.

‘Feel’: The feelings described here are overwhelmingly negative, including the causee being ‘embarrassed’, ‘alarmed’, ‘unsafe’ or ‘confused’. 79% of the sampled sentences were negative, 6% were positive and 15% were neutral. ‘Drop’: Some examples are transitive and describe the physical dropping of things. Others are intransitive include the dropping of ‘credibility’, ‘grades’ or ‘funding’. Evaluations are 82% negative, 5% positive and 13% neutral. ‘Miss’: Almost all examples are about missing in the sense of not being able to participate in something (‘causing it to miss opportunities’, caused him to miss the final nine games). A few examples are about missing people. Every single sentence is negative. ‘Break’: A lot of examples
are about physical breaking. There are also many examples where ‘break’ is not a verb of its own but part of a phrasal verb such as ‘break out/up/even/down’. 68% of the examples are negative, 2% are positive and 30% are neutral. ‘Swell’: Includes literal and less often metaphorical swelling, e.g. ‘these veins can also cause feet to swell’ or ‘caused unemployment to swell’. Virtually all examples are negative.

In summary, the complement clauses that occur with the most common collexemes of the to-infinitive construction describe a great variety of different scenes, most of which are unfavorable. Beyond that, few generalizations can be made. A common type of causee is a person who is negatively affected by the caused event over which she or he typically has little control.

Discussion

If we just look at verbal ‘cause’ without further qualification it is true that it tends to collocate with words that could be described as negative. But this statement misses a tremendous amount of both differences between grammatical patterns and semantic detail of constructional meaning. None of the constructions can be characterized as merely negative. For most of them, there is a tendency to co-occur with two different sets of words. The first group consists of a small number of highly frequent collexemes such as ‘problem’ and ‘trouble’ that are found in most other constructions as well. In the second group are words from particular semantic sets which are specific to that construction. For example, transitive ‘cause’ typically expresses health problems
while ditransitive ‘cause’ is associated with negative emotional states. Based on these findings, it seems questionable that semantic prosody is a two-term system of negative or positive evaluations. The next chapter deals with this question in more detail.

5.2.3 ‘Cause’ as a noun

The previous examples presented here have shown constructional differences in evaluative meaning for both adjectives and verbs. Nouns are often thought to be more stable in meaning than verbs or adjectives. In this section, I show that nouns are also subject to construction-dependent variation in their semantics.

While ‘cause’ is perhaps the most-studied word in all research of semantic prosody, most work on it is limited to its verbal form. A few studies make mention of it: Stubbs (1996) mentions that the negative prosody he found for verbal ‘cause’ likely does not extend to its nominal use. Louw and Chateau (2010, p.762) state that compared to its verbal use, cause as a noun is “less monolithic in its prosody, but even so the most frequent expression in the COCA corpus is ‘cause for concern.’” Smith and Nordquist (2012) find that the noun ‘cause’ is associated with different semantic prosodies, depending on its context: The ‘cause of NP’ construction has a strongly negative semantic prosody, while the ‘cause for NP’ construction often occurs as part of positive evaluations. While nominal ‘cause’ occurs in several different constructions, I focus my investigation on ‘cause for NP’ and ‘cause of NP’ because this comparison provides sufficient data to draw the relevant theoretical conclusions.
Cause for NP

The ‘cause for NP’ construction is relatively productive, i.e. it can be instantiated by a large number of different collexemes. It is not limited to expressing only positive or negative evaluations. Instead, examples such as sentences 14 and 15 are commonly encountered in corpus data:

(14) Developments in reproductive technologies that are a cause for concern

(15) The rebirth of flowers and leaves on the trees was cause for celebration.

The results of the collexeme analysis for ‘cause for NP’ are given in table 5.12. A cursory examination of the ranked collexemes shows that collexemes with negative connotations (e.g. ‘concern’, ‘alarm’, ‘worry’), and positively evaluated collexemes (e.g. ‘celebration’, ‘optimism’, ‘hope’) are both relatively common. The overall evaluative tendency seems to skew negative though because negative collexemes are more common, both in terms of types and tokens. ‘Cause for concern’ which strikes me as an idiomatic and entrenched sequence makes up for more than one-third of all tokens attested for the construction.

If we assign an evaluative valence to each collexeme in isolation and count all types, 58% of all types are negative, 28% are positive and 14% are neutral. These numbers need to be put into perspective: Type frequencies mean that each unique collexeme is counted. If we were to make type counts across all collexemes, this would mean that extremely frequent items and items that only occur a single time would
both be counted once. I limit the type counts to items that are attracted to the
construction, i.e. that show a statistically significant tendency to occur within the
construction compared to the rest of the entire corpus. Very frequent words that
rarely occur within a given construction would show repulsion for that construction.
This cut-off can have a decisive impact on what type frequencies look like; token
frequencies, on the other hand, are much less affected. In terms of token frequencies
(i.e. counting total occurrences), 70% of the attracted collexemes are negative, 26%
are positive and 4% are neutral.

The causing event (the cause in ‘cause for NP’) is linguistically encoded in too
many ways to be retrieved automatically and analyzed quantitatively. A manual
inspection of the data shows that no matter whether the collexemes are positive or
negative, most sentences describe human reactions that are often emotional, typically
to large-scale events that are out of their control, e.g. economic, political or societal
developments and environmental issues (see example sentences 14 and 15).

<table>
<thead>
<tr>
<th>Rank</th>
<th>Collexeme</th>
<th>Tokens</th>
<th>Attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>concern</td>
<td>595</td>
<td>8,158</td>
</tr>
<tr>
<td>2</td>
<td>celebration</td>
<td>190</td>
<td>2,698</td>
</tr>
<tr>
<td>3</td>
<td>alarm</td>
<td>152</td>
<td>2,280</td>
</tr>
<tr>
<td>4</td>
<td>optimism</td>
<td>70</td>
<td>1,004</td>
</tr>
<tr>
<td>5</td>
<td>hope</td>
<td>45</td>
<td>484</td>
</tr>
<tr>
<td></td>
<td>Collexemes</td>
<td>Frequency</td>
<td>Count</td>
</tr>
<tr>
<td>---</td>
<td>-------------------------</td>
<td>-----------</td>
<td>-------</td>
</tr>
<tr>
<td>6</td>
<td>worry</td>
<td>36</td>
<td>449</td>
</tr>
<tr>
<td>7</td>
<td>complaint</td>
<td>20</td>
<td>184</td>
</tr>
<tr>
<td>8</td>
<td>dismissal</td>
<td>13</td>
<td>148</td>
</tr>
<tr>
<td>9</td>
<td>decline</td>
<td>18</td>
<td>146</td>
</tr>
<tr>
<td>10</td>
<td>behavior</td>
<td>21</td>
<td>144</td>
</tr>
<tr>
<td>11</td>
<td>panic</td>
<td>13</td>
<td>137</td>
</tr>
<tr>
<td>12</td>
<td>anxiety</td>
<td>13</td>
<td>126</td>
</tr>
<tr>
<td>13</td>
<td>shame</td>
<td>12</td>
<td>111</td>
</tr>
<tr>
<td>14</td>
<td>suspicion</td>
<td>11</td>
<td>102</td>
</tr>
<tr>
<td>15</td>
<td>joy</td>
<td>11</td>
<td>101</td>
</tr>
<tr>
<td>16</td>
<td>war</td>
<td>20</td>
<td>92</td>
</tr>
<tr>
<td>17</td>
<td>excitement</td>
<td>9</td>
<td>88</td>
</tr>
<tr>
<td>18</td>
<td>failure</td>
<td>14</td>
<td>87</td>
</tr>
<tr>
<td>19</td>
<td>despair</td>
<td>8</td>
<td>87</td>
</tr>
</tbody>
</table>

Table 5.12: Collexemes with the strongest attraction to the noun slot in *cause for NP*.

There is however more to the story: The construction is often negated, as in 16 which changes the attitudinal stance conveyed in the overall utterance from negative to positive. To quantify how often this happens, I analyzed a sample of 200 sentences. 146, or 73% of the sentences included a collexeme could be described as negative and 54 or 27% of the sentences had a positive collexeme. Out of the 146 sentences with
negative collexemes, 41 sentences were negated, flipping the utterance-level evaluation to positive. Negated sentences with positive collexemes such as example 17, on the other hand, are exceedingly rare: The sample included only two examples, making up less than 4% of the tokens in this category. This means that the attitudes expressed in sentences including *cause for* are almost split evenly between positive and negative (93 and 107 examples respectively).

(16) This usually isn’t a cause for concern.

(17) The result was not cause for celebration.

What then is the semantic prosody of *cause for NP*? The answer depends on whether we consider the construction by itself or within its entire utterance. The local constructional meaning alone carries a negative evaluation in 73% of the 200 analyzed examples, which is very similar to the 70% of negative tokens found through collexeme analysis. If we consider the meaning of the wider utterance, only 54% of all examples are negative, meaning that there is virtually no evaluative preference one way or another.

**Cause of NP**

The result of the collexeme analysis for *cause of* are given in table 5.13. The most common and most strongly attracted collexeme by far is ‘death’, which appears to have the status of a fixed expression. In terms of their lexical fields, most collexemes fall within three main categories, physical harm (‘injury’, ‘death’), illness (‘cancer’,
‘disease’) and conflict (‘war’, ‘violence’, ‘conflict’). More general terms like ‘problem’ or ‘failure’ are common as well. Beyond this, there is the legal idiom ‘cause of action’ and two uses of ‘cause’ in a different sense: ‘cause of peace’ and ‘cause of freedom’. In this sense it expresses an ideal or goal to strive for, a sense of the word most often found in ADJ N constructions, e.g. ‘good cause’ or ‘charitable cause’.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Collexeme</th>
<th>Tokens</th>
<th>Attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>death</td>
<td>2,647</td>
<td>29,838</td>
</tr>
<tr>
<td>2</td>
<td>action</td>
<td>331</td>
<td>2,456</td>
</tr>
<tr>
<td>3</td>
<td>problem</td>
<td>408</td>
<td>2,449</td>
</tr>
<tr>
<td>4</td>
<td>disease</td>
<td>308</td>
<td>2,134</td>
</tr>
<tr>
<td>5</td>
<td>cancer</td>
<td>245</td>
<td>1,650</td>
</tr>
<tr>
<td>6</td>
<td>failure</td>
<td>178</td>
<td>1,394</td>
</tr>
<tr>
<td>7</td>
<td>poverty</td>
<td>144</td>
<td>1,379</td>
</tr>
<tr>
<td>8</td>
<td>crash</td>
<td>141</td>
<td>1,361</td>
</tr>
<tr>
<td>9</td>
<td>accident</td>
<td>140</td>
<td>1,262</td>
</tr>
<tr>
<td>10</td>
<td>blindness</td>
<td>87</td>
<td>1,223</td>
</tr>
<tr>
<td>11</td>
<td>illness</td>
<td>131</td>
<td>1,198</td>
</tr>
<tr>
<td>12</td>
<td>injury</td>
<td>135</td>
<td>1,184</td>
</tr>
<tr>
<td>13</td>
<td>violence</td>
<td>134</td>
<td>1,101</td>
</tr>
<tr>
<td>14</td>
<td>pain</td>
<td>123</td>
<td>1,008</td>
</tr>
</tbody>
</table>
Table 5.13: Collexemes with the strongest attraction to the noun slot in *cause of NP*.

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>conflict</td>
<td>116</td>
</tr>
<tr>
<td>16</td>
<td>freedom</td>
<td>117</td>
</tr>
<tr>
<td>17</td>
<td>mortality</td>
<td>113</td>
</tr>
<tr>
<td>18</td>
<td>war</td>
<td>151</td>
</tr>
<tr>
<td>19</td>
<td>crisis</td>
<td>102</td>
</tr>
<tr>
<td>20</td>
<td>peace</td>
<td>109</td>
</tr>
</tbody>
</table>

Even though the two constructions are structurally alike and may seem like they could be used in similar contexts, their collexeme profiles are very different. *Cause for* is largely associated with mental states, which can be positive or negative. *Cause of* is mostly associated with external states. Apart from the sense of ‘the cause of freedom’, it is virtually exclusively used to express negative events.

The findings of this investigation can be summarized with three main points. First, the data show that nouns can also be context-dependent in their meanings. It would not make sense to assign semantic prosody to the noun ‘cause’. Instead, there is an evaluative dimension of meaning that is inherent to the lexically instantiated constructions *cause of NP* and *cause for NP*. Second, at least in this case, semantic prosody cannot be described in terms of a positive/negative polarity. Such a polarity cannot be assigned to *cause for* because it habitually occurs with both positive and negative internal states. This is a theme found with many constructions: They are
highly evaluative, but they exist at both poles of a positive-negative spectrum. Additionally, what sets cause of NP and cause for NP apart from each other is not their evaluative polarities, but their semantic preferences. There is little overlap between the lexical fields of their collexeme profiles.

The third point is methodological. We need to be careful to characterize words or constructions through their collexeme profiles, because some collexemes tend to occur in negated clauses. In 2, I mentioned this issue in the context of words like ‘alleviate’. Here however, we cannot predict a relationship of reversal from the lexical meaning of the collexeme. I return to this issue in the discussion section of this chapter, alongside related evidence from another construction.

5.3 Constructions and their words

The previous case studies have demonstrated that the same lexical word can be associated with different attitudinal meanings depending on what construction it occurs in. I argued that this shows that it is the overall pattern, not the lexical word alone that possesses the semantic prosody. In this section, the argument is taken one step further: There are constructions that are associated with particular evaluations, some of this meaning is independent of the lexical words in their open slots and some of it emerges through their interplay.
5.3.1 ‘There BE something ADJ about’

In this section, I describe the constructional semantics of the there BE something ADJ about construction. It is briefly mentioned by Hunston and Sinclair (2000, pp. 85f) as an example of a pattern commonly used in evaluation: “Its function is to give a subjective judgement about something and this is typically a judgement of good or bad” (p. 85). Syntactically, the evaluated entity directly follows the construction and is either encoded as a noun phrase or a gerund clause (cf. examples 18 and 19 below).

I conducted a collexeme analysis to study the words with the strongest attraction to the adjective slot in the construction. The results are given in table 5.14. The semantics of the construction are highly evaluative and often include an emotional reaction to something unusual. The evaluations have a strong good/bad component and include both ends of the continuum (though rarely the middle). The negative collexemes all focus on the danger of something different and unknown. Strongly attracted collexemes include ‘odd’, ‘strange’, ‘weird’, ‘unnatural’, ‘creepy’, ‘unreal’ and ‘fishy’. While the word ‘different’ is not necessarily bad on its own, within this construction, it tends to occur in context with negative evaluations, e.g. “But as many award ceremonies as I have attended over the years, there was something different about this one – like an overlying sense of dread.” Still, there are several examples where it is not clear whether the speaker intends to convey a neutral or negative attitude – or if he or she is being purposely vague and perhaps hints at a negative interpretation without explicitly committing to it. The positive collexemes focus on

At this point of analysis, there BE something ADJ about seems to be a relatively idiomatic construction, with strong evaluative semantics that are specified for a meaning of unusualness, but that are open in terms of the evaluative valence. A more detailed reading of the examples in context shows that there is more to the story. The evaluative polarity varies depending on the entity that is described by the construction. Intuitively, it seems that when a person is described, the meaning is overwhelmingly negative (“There’s something weird about her”) and when an event or place is described it is more likely to be positive than negative (“There is something magical about flying from summer to winter”).

<table>
<thead>
<tr>
<th>Rank</th>
<th>Collexeme</th>
<th>Tokens</th>
<th>Attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>special</td>
<td>117</td>
<td>1,405</td>
</tr>
<tr>
<td>2</td>
<td>different</td>
<td>122</td>
<td>1,119</td>
</tr>
<tr>
<td>3</td>
<td>odd</td>
<td>54</td>
<td>730</td>
</tr>
<tr>
<td>4</td>
<td>strange</td>
<td>56</td>
<td>724</td>
</tr>
<tr>
<td>5</td>
<td>familiar</td>
<td>55</td>
<td>667</td>
</tr>
<tr>
<td>6</td>
<td>magical</td>
<td>26</td>
<td>378</td>
</tr>
<tr>
<td>7</td>
<td>comforting</td>
<td>22</td>
<td>366</td>
</tr>
<tr>
<td>8</td>
<td>wrong</td>
<td>33</td>
<td>320</td>
</tr>
</tbody>
</table>
In order to test these impressions, I took a 200 sentence sample from COCA; I classified all examples for the type entity described and encoded the attitudinal stance expressed in the construction. Overall, 46% of all sentences in the sample express a negative evaluation, 38% are positive and 16% are neutral. While the classification included some subjective judgment calls, I grouped the themes described by the construction into five categories. Most common were descriptions of personal

<table>
<thead>
<tr>
<th></th>
<th>Collexemes</th>
<th>Frequency</th>
<th>Attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>unreal</td>
<td>17</td>
<td>297</td>
</tr>
<tr>
<td>10</td>
<td>fishy</td>
<td>16</td>
<td>279</td>
</tr>
<tr>
<td>11</td>
<td>funny</td>
<td>24</td>
<td>239</td>
</tr>
<tr>
<td>12</td>
<td>wonderful</td>
<td>22</td>
<td>210</td>
</tr>
<tr>
<td>13</td>
<td>weird</td>
<td>17</td>
<td>199</td>
</tr>
<tr>
<td>14</td>
<td>unnatural</td>
<td>11</td>
<td>192</td>
</tr>
<tr>
<td>15</td>
<td>creepy</td>
<td>13</td>
<td>172</td>
</tr>
<tr>
<td>16</td>
<td>exciting</td>
<td>15</td>
<td>170</td>
</tr>
<tr>
<td>17</td>
<td>unusual</td>
<td>16</td>
<td>169</td>
</tr>
<tr>
<td>18</td>
<td>unique</td>
<td>14</td>
<td>125</td>
</tr>
<tr>
<td>19</td>
<td>sad</td>
<td>11</td>
<td>109</td>
</tr>
<tr>
<td>20</td>
<td>endearing</td>
<td>7</td>
<td>104</td>
</tr>
</tbody>
</table>

Table 5.14: Collexemes with the strongest attraction to the adjective slot in the *there BE something ADJ about* construction.
characteristics as in example (18), followed by human actions (19). The other categories are less common and include places, often in nature (20), verbal statements (21) and art (22). 9 examples could not be classified and are not included in the rest of the analysis.

(18) There’s something creepy about Leonard. (person)

(19) There’s something really beautiful about cooking for someone and feeding them. (action)

(20) There’s something very calming about the building and the verdant and rocky hills surrounding the winery. (place)

(21) there was something fishy about Paddy’s stories. (statement)

(22) There’s something magical about electronics artwork from the’ 90s. (art)

For each category, I calculated a valence score, a simple descriptive statistic that summarizes the evaluative valence of each category with a single number, ranging from -1 (completely negative) to 1 (completely positive). To calculate the score, negative tokens are subtracted from positive tokens and the result is divided by all tokens (including neutral tokens). The results are given in Table 5.15.

<table>
<thead>
<tr>
<th>Category</th>
<th>Tokens</th>
<th>Positive</th>
<th>Negative</th>
<th>Neutral</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal</td>
<td>75</td>
<td>19</td>
<td>50</td>
<td>6</td>
<td>-0.39</td>
</tr>
<tr>
<td>Action</td>
<td>61</td>
<td>32</td>
<td>21</td>
<td>8</td>
<td>0.13</td>
</tr>
<tr>
<td>Place</td>
<td>26</td>
<td>15</td>
<td>5</td>
<td>6</td>
<td>0.45</td>
</tr>
</tbody>
</table>
Table 5.15: Attitudinal stance found for different semantic fields in 200 examples of the *there BE something ADJ about* construction.

<table>
<thead>
<tr>
<th>Category</th>
<th>Negative</th>
<th>Neutral</th>
<th>Positive</th>
<th>Total</th>
<th>Cadence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbal</td>
<td>18</td>
<td>0</td>
<td>15</td>
<td>3</td>
<td>-0.81</td>
</tr>
<tr>
<td>Art</td>
<td>11</td>
<td>9</td>
<td>2</td>
<td>0</td>
<td>0.60</td>
</tr>
</tbody>
</table>

Most categories are in some way related to humans. Descriptions of human personal characteristics are more often negative than positive. Examples for negatively evaluated collexemes include ‘odd’ or ‘creepy’, positively evaluated collexemes span a wide range and include words such as ‘magnetic’, ‘unique’ or ‘endearing’. ‘Familiar’, which is also relatively common is one of the few more neutral words in this otherwise highly attitudinally charged construction. Descriptions of human actions vary in their evaluations but are slightly more positive than negative. Collexemes are varied, including words such as ‘exhilarating’ or ‘satisfying’ but also ‘dishonorable’, ‘sinister’ or just ‘different’.

Verbal statements are overwhelmingly characterized as negative (not a single positive example is attested in the sample). The semantic field of the collexemes is much more specific than in the other categories. Most common are words like ‘fishy’ or ‘wrong’ that question the veracity of the statements. Places show generally positive evaluations with collexemes such as ‘magical’, ‘comforting’; but they can also be ‘weird’. Works of art are the category with the highest percentage of positive statements (although also the smallest sample size). Art can be ‘magical’ or ‘beautiful’,
although ‘sinister’ is also found.

Hunston and Sinclair (2000) generalize the construction to also include phrases like *there BE nothing ADJ about* or *there BE something ADJ with*. The data show that these slight differences in their structure lead to very different collexeme profiles and constructional meanings. Let us first contrast constructions with ‘something’ and ‘nothing’: The collexemes associated with *there BE nothing ADJ about* can be categorized into three types. The first category consists of just one entry, the highly attracted collexeme ‘illegal’. About a third of the instances include some sense of contrast, e.g. ‘*while* there is nothing illegal about’ or ‘there is nothing illegal about X *but/except*’. The second type is the word ‘funny’; the speaker conveys their opinion that there is a bad situation that is not taken seriously enough (‘there’s nothing funny about what’s happening in his life’). The third type includes collexemes such as ‘special’, ‘new’, ‘unusual’ or ‘remarkable’. It describes the unremarkable nature of the described entity and no further generalizations seem possible.

If we exchange the word ‘about’ for the word ‘with’ in the last slot of the construction, its use and meaning changes dramatically as well. Out of roughly 1,000 occurrences attested in COCA for the *there BE something ADJ with* construction, virtually all contain the collexeme ‘wrong’ in the adjective slot. There is no semantic component of unusualness as was the case with the construction with ‘about’; instead it is a simple judgment that something is not right. This can be a moral judgment (‘there is something wrong with killing children’) or just a statement that something
does not function correctly (‘there is something wrong with your car’). The only other examples, e.g. ‘weird’ or ‘odd’ are attested less than a handful of times. The picture for there BE nothing ADJ with is very similar. The pattern is more common, with over 1,600 occurrences in COCA and adjective slot is virtually always occupied by the word ‘wrong’.

Discussion

The meaning of there BE something ADJ about can be summarized as follows: A speaker chooses this construction to express that the entity they are describing is in some way unusual. This description is almost always strongly evaluative and can be positive or negative. Whether this unusualness is seen as positive or negative depends on the type of entity that is being evaluated: When personal characteristics or verbal statements are described, the adjective usually conveys a negative stance. Unusual places and works of art on the other hand tend to be positively evaluated.

The following conclusions can be made from this case study: First, the there BE something ADJ about construction is an example of a construction that is meaningful in itself – there are consistent semantic patterns across its uses, independent of lexical instantiations. Second, this meaning is complex and evaluative; it does not have an overall tendency towards a positive or negative polarity. Third, if we dig deeper we can make out some further distinctions, i.e. evaluation varies based on the type of entity described. Fourth, this attests to the context-dependent nature of word meaning, with
the collexeme ‘different’ taking on a sense of oddness in this context. This semantic component is imparted upon the word by the construction. The construction in turn has been loaded with this meaning through collexemes, such as ‘odd’ or ‘strange’.

Fifth and finally, closely related patterns such as *there BE nothing ADJ about* or *there BE something ADJ with* are associated with entirely different use patterns and semantics.

There may be an issue with teasing out fine distinctions in constructional meaning through the semantic categories of its collexemes: At what level of granularity is the constructional meaning a psychologically real mental representation and at what point is it just an epiphenomenon, i.e. a meaning that is created ad hoc, derived from conceptual structures and not part of the mental lexicon? It seems uncontroversial to say that the meaning of the *there BE something ADJ about* construction includes a component of unusualness. This is a generalization that has become conventionalized. Word like ‘normal’, ‘mundane’ or ‘boring’ are not attested as collexemes of this construction in the entire 550 million word COCA corpus.

The question is whether the same argument can be made about the skew in meaning when comparing the different categories of entities described by the construction. Here, conceptual motivations that drive the semantics of the different lexical fields can readily be found. It makes sense that something unusual could be good or bad. It also makes sense that people who behave in an unusual manner are suspicious and may be dangerous. Similarly, when people’s claims and statements are unusual, they
are likely untrue. The point is that this semantic component could be derived *on the fly*, i.e. speakers could choose this construction because it makes sense conceptually – which means that it may not be part of their mental lexicon. There are several open questions: How fine are the distinctions people store? How deep can we reasonably go to separate out meaning components? Which categories are linguistically real and which are *analytic fictions*? In other words, just the because we split the data into infinitely fine categories, this does not mean that there is any real linguistic value to it. These issues are revisited in chapter 9.

5.3.2 The ‘into Ving’ causative

Research on stance-taking and in the field of phraseology has long shown that longer patterns are often used to convey highly evaluative meaning. Hunston (2007b), for example, shows the stances taken with patterns such as “if . . . so be it”, “to the point of” or “the wrong end of the stick.” These sequences are typically highly idiomatic; they have little grammatical and semantic productivity and often have only one open collexeme slot. The construction I am discussing in this section is different – it is an abstract schema with many open slots and very little pre-defined lexical content.

The *into Ving* causative construction has the schema $NP_{Subj} \ V \ NP_{Obj} \ into \ Ving$. The basic structure of this causative construction is as follows: There is a subject ($NP_{Subj}$, the causer) who acts ($V$, the causing event) upon an object ($NP_{Obj}$, the causee) so that the object performs another action ($Ving$, the resulting event). Ex-
amples of the construction include “she talked me into buying a flatbed truck” or “they pressured him into eating pork.” The into Ving causative has been discussed in the research literature on causation. Wierzbicka (1998), for example, describes its semantics as follows:

“[T]he causee originally didn’t want to do what he or she did, […] the causee’s action is ‘triggered’ by the causer’s will, not by the causee’s own will, […] the causee is unaware of what is happening (namely, that his or her action is ‘triggered’ by the causer’s will).” (Wierzbicka, 1998, pp. 125f)

All occurrences of this construction were retrieved from COCA. First, I analyzed a sample of 200 sentences. The results show overwhelmingly negative attitudes towards the expressed propositions. Typical examples include ”fool ourselves into thinking,’ ‘pressured him into going’ or ‘tricked me into doing.’ The meaning of these utterances is always that one person coerces or deceives another person to perform an action that will typically yield unfavorable results for that second person.

A collexeme analysis was performed for the first verbal slot of the V NP into Ving construction. The 20 most strongly attracted collexemes are presented in Table 5.16. Almost all verbs have a transparently negative meaning by themselves and encompass a sense of either deception or force. Perhaps the only exception is ‘talk’. While there is no negativity attached to the verb itself, within this construction virtually all sentences carry this evaluation. This example shows again that the analysis should not end with the collocate analysis. We cannot determine the meaning of
a construction by looking at its collexemes alone. This would, in fact, be circular reasoning, given the context-dependent nature of words. Out of all the collexemes, ‘talk’ has the second strongest attraction to the construction. This does not mean that the construction is commonly devoid of attitudinal semantics. Instead, if we further investigate ‘talk’ in this context we see that it is coerced into taking on the same aura of deception as many of its fellow collexemes.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Collexeme</th>
<th>Tokens</th>
<th>Attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>trick</td>
<td>673</td>
<td>12,560</td>
</tr>
<tr>
<td>2</td>
<td>talk</td>
<td>1,035</td>
<td>10,360</td>
</tr>
<tr>
<td>3</td>
<td>fool</td>
<td>379</td>
<td>6,012</td>
</tr>
<tr>
<td>4</td>
<td>coerce</td>
<td>252</td>
<td>4,469</td>
</tr>
<tr>
<td>5</td>
<td>coax</td>
<td>181</td>
<td>2,897</td>
</tr>
<tr>
<td>6</td>
<td>force</td>
<td>269</td>
<td>2,816</td>
</tr>
<tr>
<td>7</td>
<td>pressure</td>
<td>180</td>
<td>2,765</td>
</tr>
<tr>
<td>8</td>
<td>scare</td>
<td>146</td>
<td>1,863</td>
</tr>
<tr>
<td>9</td>
<td>lure</td>
<td>121</td>
<td>1,631</td>
</tr>
<tr>
<td>10</td>
<td>delude</td>
<td>88</td>
<td>1,533</td>
</tr>
<tr>
<td>11</td>
<td>bully</td>
<td>101</td>
<td>1,453</td>
</tr>
<tr>
<td>12</td>
<td>manipulate</td>
<td>111</td>
<td>1,403</td>
</tr>
<tr>
<td>13</td>
<td>goad</td>
<td>82</td>
<td>1,403</td>
</tr>
</tbody>
</table>
We can further investigate the semantics of the into Ving causative construction with the help of covarying collexeme analysis. This allows us to explore which collexemes within two slots of the same construction show a preference of co-selection.

Table 5.17 shows the output of this analysis, comparing the two verbal slots, i.e. which causation event tends to co-occur with which caused event.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Word 1</th>
<th>Word 2</th>
<th>Tokens</th>
<th>Attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>fool</td>
<td>thinking</td>
<td>221</td>
<td>448</td>
</tr>
<tr>
<td>2</td>
<td>delude</td>
<td>thinking</td>
<td>66</td>
<td>153</td>
</tr>
<tr>
<td>3</td>
<td>mislead</td>
<td>thinking</td>
<td>96</td>
<td>129</td>
</tr>
<tr>
<td>4</td>
<td>force</td>
<td>hiding</td>
<td>24</td>
<td>110</td>
</tr>
<tr>
<td>5</td>
<td>talk</td>
<td>letting</td>
<td>60</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>verb</td>
<td>object</td>
<td>count</td>
<td>probability</td>
</tr>
<tr>
<td>---</td>
<td>------</td>
<td>----------</td>
<td>-------</td>
<td>-------------</td>
</tr>
<tr>
<td>6</td>
<td>mislead</td>
<td>believing</td>
<td>52</td>
<td>68</td>
</tr>
<tr>
<td>7</td>
<td>talk</td>
<td>staying</td>
<td>36</td>
<td>60</td>
</tr>
<tr>
<td>8</td>
<td>talk</td>
<td>going</td>
<td>73</td>
<td>55</td>
</tr>
<tr>
<td>9</td>
<td>talk</td>
<td>coming</td>
<td>48</td>
<td>54</td>
</tr>
<tr>
<td>10</td>
<td>fool</td>
<td>believing</td>
<td>66</td>
<td>49</td>
</tr>
<tr>
<td>11</td>
<td>mislead</td>
<td>assuming</td>
<td>10</td>
<td>46</td>
</tr>
<tr>
<td>12</td>
<td>coerce</td>
<td>confessing</td>
<td>10</td>
<td>44</td>
</tr>
<tr>
<td>13</td>
<td>deceive</td>
<td>thinking</td>
<td>33</td>
<td>34</td>
</tr>
<tr>
<td>14</td>
<td>coerce</td>
<td>having</td>
<td>20</td>
<td>31</td>
</tr>
<tr>
<td>15</td>
<td>deceive</td>
<td>believing</td>
<td>22</td>
<td>30</td>
</tr>
<tr>
<td>16</td>
<td>lull</td>
<td>believing</td>
<td>21</td>
<td>28</td>
</tr>
<tr>
<td>17</td>
<td>lull</td>
<td>forgetting</td>
<td>5</td>
<td>27</td>
</tr>
<tr>
<td>18</td>
<td>talk</td>
<td>moving</td>
<td>18</td>
<td>25</td>
</tr>
<tr>
<td>19</td>
<td>lull</td>
<td>thinking</td>
<td>29</td>
<td>24</td>
</tr>
<tr>
<td>20</td>
<td>coerce</td>
<td>signing</td>
<td>16</td>
<td>22</td>
</tr>
<tr>
<td>21</td>
<td>trick</td>
<td>admitting</td>
<td>8</td>
<td>22</td>
</tr>
<tr>
<td>22</td>
<td>scare</td>
<td>stopping</td>
<td>4</td>
<td>22</td>
</tr>
<tr>
<td>23</td>
<td>coerce</td>
<td>participating</td>
<td>8</td>
<td>21</td>
</tr>
<tr>
<td>24</td>
<td>talk</td>
<td>taking</td>
<td>52</td>
<td>21</td>
</tr>
<tr>
<td>25</td>
<td>force</td>
<td>being</td>
<td>16</td>
<td>20</td>
</tr>
<tr>
<td>26</td>
<td>delude</td>
<td>believing</td>
<td>20</td>
<td>19</td>
</tr>
</tbody>
</table>
The strongest attraction between collexemes is found between some verbs of deception (the causing event) and psych verbs (the caused event). Further inspection of the data explains why this is the case. Within the V into Ving construction, some verbs of deception, such as fool, delude, mislead or deceive are solely used to describe changing someone’s beliefs about something. As such, these verbs overwhelmingly occur with words such as ‘think’ or ‘believe’. Other verbs of deception, and also verbs of coercion, on the other hand, are about causing someone to perform some kind of action. The verbs include ‘talk’, ‘trick’, ‘pressure’ and ‘coax’. They describe these actions as not constrained to any particular semantic field.

In order to quantify whether my impression was correct, I calculated the entropy of the words in the first slot of the construction using the SciPy Python library (Jones et al., 01 ). The basic idea here is based on Kilgarriff’s concept of collocationality: “A word that is very ‘collocational’ is one which has a strong tendency to appear with particular words, rather than appearing freely with large numbers of words. This theme is captured mathematically by ‘entropy’ [ . . . ]” (Kilgarriff, 2006,
Collocationality is essentially an operationalization of the concept of linguistic productivity. The idea of collocationality can be applied to collexemes by simply limiting the analysis to items occurring within a construction. And indeed, words like *pressure, coax, talk* and *trick* have a much higher degree of entropy than words like *deceive, fool, mislead* and *delude*, indicating that the distribution of collexemes of the former group is wider and more varied than that of the latter group, which confirms my intuitions.

**Discussion**

The prototypical meaning of the *V NP into V-ing* construction can be described as follows: Person *A* manipulates person *B* through action *X* to perform action *Y*. *X* (the first verb slot in the construction) can either entail deception or coercion. *Y* (the second verb slot) can be any action but is most often a change of beliefs. *B* would not have performed *Y* if it had not been for *X*. The outcome of *Y* is often favorable for *A* and undesirable for *B*.

Wierzbicka’s (1998) definition of the construction mostly holds up, but the collexeme analysis reveals some details that her introspective approach missed. *B* is not necessarily unaware that he or she is doing something that is not in their interest. This is the case if *X* is a verb of deception and *Y* is a psych verb (‘fooled me into thinking’), but it is not the case if *X* is a verb of coercion (‘bullied me into leaving’). In the latter case, *B* can be fully aware of what is happening and still give in. Addi-
tionally, $A$ and $B$ can be identical. In this case, $B$ is expressed as a reflexive pronoun and the construction almost almost expresses a meaning of self-delusion (‘we deluded ourselves into believing it’).

The constructional meaning has become so strongly entrenched that it cannot be overridden by the words that fill the slots. This is evidenced by the verb ‘talk’ in the first verbal slot. The semantic component of deception or coercion is not part of the invariant, lexical meaning of ‘talk’ because it does not appear in many of the contexts the verb occurs in. Within the context of the into-causative construction, however, this meaning is clearly present. Similar to ‘different’ in the there BE something ADJ about construction, this meaning is imparted upon the collexeme through the construction. The construction in turn acquires this part of its semantics through its other strongly attracted collexemes, virtually all of which either express deception or coercion.

There are other constructions in English that have somewhat similar meanings but lack the same attitudinal evaluation. One example is the $V$ NP to $V$ construction can (among other things) be used for this purpose, e.g. ‘convinced me to go’. Another possible pattern is $[get] \ NP \ to \ V$, as in ‘he got people to come’. These two constructions also encode the meaning ‘person $A$ gets person $B$ to do $X$’; the attitude the speaker displays towards the situation, however, is completely different. The into-causative construction necessarily signals the disapproval of the speaker, meaning that this evaluative component is part of the constructional semantics.
Obviously, such a strong association with a semantic prosody is not found in all constructions. Basic grammatical schemas that are highly productive and versatile, such as the transitive construction or the noun phrase construction cannot be said to carry a semantic prosody. Instead, they are semantically open and can be employed in a wide range of expressions. It seems that semantic prosodies are particularly likely to be relevant in somewhat idiomatic constructions, i.e. constructions that are both limited in their functional scope and constrained in their constructional form. The ‘V Obj into Ving’ construction fulfills both of these criteria: It has a fairly rigid grammatical structure and is almost exclusively used to express situations of deception or coercion.

5.4 Discussion

The main aim of this chapter was to show that semantic prosody cannot be reduced to the word level. The data show that words have context-dependent meanings that depend on the constructions they occur in. Differences are found both in terms of the evaluations that are expressed and in terms of the types collexemes that are attracted to different constructions. Lexically instantiated constructional patterns carry meaning that cannot be reduced to the lexical semantics of words alone. Methodologically, collexemes can be most easily identified by combining two tools. First, a table of the output from the collexeme analysis is created. This includes the ranked collexemes alongside with their collostructional strength, which makes it easy to determine which
individual collexemes are most strongly attracted to the construction. Second, the lexical fields can be visualized with the help of a clustering algorithm, either in the form of a network representation or through a dendrogram.

The study of ‘persistent’ shows that even when we consider semantic prosody in the sense of a simple positive/negative polarity, it is not a word-level phenomenon but differs across constructional contexts. One construction has negative connotations while the other one does not exhibit this tendency. The various constructions verbal ‘cause’ takes part in are differentiated through their collexeme distributions. This is particularly evident in the case of the dative alternation, where there is a clear separation between two constructions that would traditionally be considered synonymous. For verbal ‘cause’, all its constructions share a tendency towards negative evaluations but the same is not true for nominal ‘cause’. The two constructions ‘cause for’ and ‘cause of’ are used in entirely different contexts. Not only do they occur with collexemes from different lexical fields, but they also differ in the attitudes they tend to express, with ‘cause of’ being the far more negative of the two.

Finally, there are constructions that can be directly associated with evaluative meaning. The *there BE something ADJ about* construction is used to convey a sense that something is highly unusual. This evaluation can be negative or positive. The *V NP into Ving* construction is another highly evaluative construction, but it is quite different from the *there BE something ADJ about* construction. First, it is less idiomatic, with its only lexically fixed element being the preposition ‘into’. Second,
its semantics are virtually always negative.

Beyond this basic summary of the data, further conclusions can be made. One important point is that constructions enter into local relations with their collexemes and create idiosyncratic meaning. Construction grammar accounts of argument structure call attention to the meaningfulness of constructions. For example, the ditransitive construction conveys a type of transfer. Importantly, it is the construction itself that possesses this meaning, regardless of its lexical instantiation. The studies presented in this chapter add another dimension to this idea, viz., the ways in which the semantics of instantiated constructional patterns (or collostructions) are created. Let us consider the example of attributive ‘persistent’ again. There is some disagreement within construction grammar whether or not meaningless constructions exist. Goldberg (1995) claims that all constructions are meaningful while Fillmore posits semantically null constructions. According to Fillmore et al. (2012), there are three types of constructions for which no meaning should be posited and one of them is modifier-head. The attributive use of ‘persistent NP’, which instantiates this construction is a meaningful pattern: Persistent has a different meaning in this context than in other contexts. This is of course not to say that Fillmore et al. (2012) was wrong and that the incredibly productive and frequent modifier-head construction carries meaning per se (I am taking no position on this particular issue). Instead, it means that words and their constructions interact to create meaning that is partially idiomatic and not reducible to formal rules.
Another finding concerns the concept of semantic transfer, i.e. the proposed transfer of meaning between frequently co-occurring linguistic items. The data in this chapter show that there may be a way of redeeming Louw’s analogy of prosody. As described in chapter 2, it does not make sense to use the analogy of speech prosody in the Firthian sense to explain a semantic transfer from collocates to a node word because no such transfer exists in phonetic coarticulation. What makes more sense however is to define this idea within a constructional context. As the studies in this chapter have shown, there is no such thing as a semantic transfer that colors words in such a way that they also retain this meaning in other contexts. Instead, words have context-dependent semantics, meaning that some components of their meanings only exist within a specific context. Examples of these contexts include the *there BE something ADJ about* construction and the *V into Ving* causative. ‘Different’ does not generally mean ‘odd’, but it acquires this meaning within the *there BE something ADJ* construction. ‘Talk’ does not typically involve trickery or coercion, but it does so in the context of ‘talking somebody into doing something’. In this sense, it is not entirely different from coarticulation.

This is not only true for the larger, phraseological constructions but is also found in the interplay between ‘ordinary’ grammatical constructions and their collexemes, such as argument structure constructions or the noun phrase construction. The expression ‘cause change’ is almost never used to describe positive change. Through the strong attraction of transitive ‘cause’ to collexemes like ‘trouble’, ‘damage’ or ‘problems’, a
negative evaluation is part of its meaning. This association is so strong that the ‘cause 
NP’ pattern is almost never used to express a change for the better. In the few attested 
examples in my sample that do describe positive change, the speakers apparently 
felt the need to explicitly mark the change as good with an adjective, e.g. ‘cause 
positive change’. The far more common choice when talking about positive change 
is to use a different pattern, e.g. ‘bring about change’ or ‘make a change’. When 
speakers plan their sentences, they seem to account for the evaluative dimensions of 
the constructions and words at their disposal and make choices that are aligned with 
the attitudinal stance they are taking towards the content of their statement.

A similar point can be made about ‘persistent’ in the noun phrase construction. 
In this context, it co-occurs with lexical fields like illnesses and social, economic or 
environmental issues. It seems that when people hear an utterance like ‘thank you for 
your persistent help and advice’, they derive the contextual meaning of ‘persistent’ 
through comparison with its collexeme profile. Again, this semantic component only 
arises in this context and does not transfer over to the predicative use of ‘persistent’.  

Not all words show differentiated context-dependent meanings, for the simple 
reason that some words are limited to occurring in only one or two constructional

---

3It needs to be emphasized that words do not get *extra* meaning in context that is added onto their lexical meaning. In usage-based theory, context-dependent meaning is primary. A word only exists in its contextual instantiations in language usage. Lexical meaning exists only as an abstraction over all these contexts, and it is questionable to what extent it is a real-world category or whether it should only be treated as a descriptive heuristic.
contexts. The verb ‘attain’ for example, which is often cited as a rare example for
a word with positive semantic prosody is solely used transitively and does not occur
in intransitive or ditransitive clauses. As such, there is little construction-dependent
meaning – the verb only appears in transitive and passive contexts, both of which are
quite similar. There is however some measure of context-dependent meaning based
on the nature of the collexemes (which were again established through collexeme
analysis based on sentences retrieved from COCA). They can be broadly classified
into four groups: First, words like ‘status’, ‘rank’ and ‘power’, second ‘goal’ and
‘objective’, third ‘knowledge’, ‘education’ and fourth ‘age’. The verb ‘attain’ means
different things in attaining an age than in attaining knowledge or in attaining a goal.
Additionally, attaining power is the only collocation that often conveys a negative
attitude (e.g. ‘Immediately upon attaining power each dictator has suppressed all free
speech except his own’).

A similar point can be made about the verb ‘alleviate’, which is also solely a tran-
sitive verb. Its usage in COCA shows that the literal sense of alleviating symptoms
of illness is relatively common, as in the collexemes ‘pain’, ‘suffering’ and ‘symptom’.
Even more frequent are collexemes such as ‘problem’, ‘poverty’, ‘concern’, ‘stress’,
‘anxiety’, ‘fear’ or ‘pressure’. There is a difference between ‘alleviating poverty’ and
‘alleviating pain’. In the former example, ‘alleviate’ implies a sociopolitical action,
in the latter it is a medical treatment. Unlike in the example of ‘attain’, there is no
difference in the evaluative dimension here.
The same can, of course, be said of all sorts of words and contexts and it also includes words that already display constructional differences. ‘Causing death’ is different from ‘causing concerns’. The former implies the direct causation of a physical change and the latter describes the more indirect and complex causation of a mental state. ‘Persistent questioning’ describes a constantly repeated action, ‘persistent poverty’ is a continuous state, ‘persistent pollutants’ describe the chemical property of a type of matter and ‘a persistent critic’ describes the character and behavior of a person.

None of this is strictly speaking limited to evaluative meaning, which is an important point. These are completely ordinary processes of context-dependent meaning. Evaluation is an essential part of meaning and as such often shows context-dependent variation as well. This can include a dimension of negative or positive evaluation, but it does not need to. As I will argue throughout the next chapter and in my final conclusions in chapter {conclusions\_ch}, semantic prosody does not exist as a privileged type of meaning but can instead simply be accounted for through ordinary semantic processes.

Once we are able to study semantic prosodies within a constructional framework, we are able to distinguish much finer categories and facets of meaning. This opens up the possibility of investigating another central claim about semantic prosody: that it is a simple dichotomous category, consisting of good and bad evaluations. This idea is discussed in the next chapter.
Chapter 6

Beyond a Polarity of Evaluative Meaning

In the previous chapter, I investigated a fundamental notion about the structure of semantic prosody: the idea that it is part of the lexical semantics of words. I argued that it has been studied at the wrong level and that it should be posited at the constructional level, not the word-level. In this chapter, I explore a key claim about the function of semantic prosody: the idea that it is used to express either a negative or a positive evaluation.

The meaning of semantic prosody is typically considered to be a dichotomous system of evaluation, sometimes referred to as *prosodic valence* or *evaluative polarity*. It is often described through pairs of terms like good and bad, positive and negative, or favorable and unfavorable. When analyzing individual sentences, it is treated as a binary category. The overall semantic prosody of a word is described as a continuum between good and bad, depending on the frequencies of its positive and negative instances, e.g. ‘set in’ may be described as slightly more negative than ‘happen’ (cf. Partington, 2004). Crucially, this is a unidimensional continuum without any room for other meaning components.

As cited in chapter 3, evaluation can be described as “subjectivity with a focus” (Englebretson, 2007, p. 16). It is not limited to a negative-positive dichotomy but
encompasses all sorts of attitudinal stances, e.g. whether something is “obvious, questionable, tentative, provisional, controversial, contractionary, irrelevant, impolite, or whatever” (Stubbs, 1986, p. 1). Some researchers however still propose an account of evaluation that assigns a privileged role to a positive-negative polarity. Thompson and Hunston (2000) posit a system of evaluation that consists of four parameters but consider the good-bad parameter to be the most fundamental one:

“We have, […] identified four parameters of evaluation: good-bad, certainty, expectedness, and importance. We would wish to argue, however, that evaluation is essentially one phenomenon rather than several, and […] the most basic parameter, the one to which the others can be seen to relate, is the good-bad parameter.” Thompson and Hunston (2000, p. 25)

The meaning of semantic prosody is almost universally described as a type of evaluation that is limited to the good-bad parameter. An example of this view is Hunston (2002):

[Semantic prosody] accounts for ‘connotation’: The sense that a word carries a meaning in addition to its ‘real’ meaning. The connotation is usually one of evaluation, that is, the semantic prosody is usually negative or, less frequently, positive.” Hunston (2002, p. 142)

In this chapter, I argue against this view. The data show that semantic prosody is not limited to a positive-negative dimension but can instead be associated with
various types of evaluative meaning. From this standpoint, there is nothing special about semantic prosody; it conveys the same kinds of attitudinal stances that are found all throughout language.

The remainder of this chapter is organized as follows: Section 6.1 deals with the difficulties of separating semantic prosody from semantic preference. In section 6.3, I argue that no empirical evidence has been produced to corroborate the existence of evaluative polarity as its own semantic category, much less as the most fundamental phenomenon that underlies all other types of evaluation.

In sections 6.4 – 6.6, I make the case that the evaluative semantics that emerge through lexical co-selection cannot be broken down into only a negative/positive polarity. First, the evaluations are often more complex, second, a negative/positive dimension is not always present, and, third, the same linguistic unit can be associated with opposite evaluations.

Section 6.7 deals with the issue that assigning a good/bad polarity is highly subjective, often problematic and it is not clear at what level it should be implemented. This is not just a practical but also a theoretical issue, as it introduces unwarranted assumptions that neither aid the analysis of the data nor are supported by them. Section 6.7 summarizes the results and discusses their theoretical implications.
6.1 Semantic preference and semantic prosody

The idea that a word is associated with negatively or positively evaluated collocates is an artifact of a lexical perspective of meaning that conflates all contextual occurrences of words. As shown in chapter 5, the different constructions verbal ‘cause’ takes part in display distinct and well-delineated collexeme profiles. In the transitive and passive constructions, it often occurs with illnesses, in the prepositional dative construction it occurs with external states and in the ditransitive dative with internal states, while the infinitival complement construction shows a very varied and productive collexeme profile. A word-centered approach that only investigates the collocates of words without considering constructional differences misses these crucial details and may well conclude that the only possible generalization is that the collocates of ‘cause’ are mostly negative. For ‘cause’ as a noun the constructional differences are even more pronounced. A word-centered, collocational study would find all sorts of unrelated words and would entirely overlook the differences between ‘cause for’ and ‘cause of’.

Some prior studies have noted the relevance of semantic preference, typically in cases where constructional differences are less relevant, e.g. in Sinclair’s original remarks on the phrasal verb ‘set in’. As a result, semantic prosody has been described as a phenomenon that is related to semantic preference but ultimately separate from it. Partington (2004, p. 150) regards semantic prosody as an abstraction from semantic preference: “prosody is at a further stage of abstraction than preference. In
fact, semantic preference generally remains relatively closely tied to the phenomenon of collocation.”

This is an abstraction, researchers can posit as part of their analysis, but it is not clear that this is an abstraction that speakers make. The fact that semantic preference allows us to describe the data in more detail does not necessarily mean that semantic prosody does not exist. It does, however, mean that the patterns found in prior research can be accounted for in a different way. This leaves us with the question what the evidence for semantic prosody as a two-term system of evaluation is. I explore this question in the next three sections.

6.2 Lack of evidence

Although it is often claimed that there is a polarity of good and bad at the core of evaluation, these claims are never corroborated by data. An example is the following statement by Morley and Partington (2009, p. 141):

“Semantic prosody is an expression of the innate human need and desire to evaluate entities in the world they inhabit as essentially good or bad. Different terms are used in the literature: positive and negative, favourable and unfavourable, desirable and undesirable, but evaluation at its most basic is a two-term system. We argue that the drive to evaluate is innate because human beings, and probably any biological organism capable of decision making, are goal-driven, and it is essential (phylogenetically, for
survival) to judge whether the outcome of any decision, or of any scenario confronted, will be beneficial or otherwise to the organism. All the other factors influencing decision-making (possibility, willingness, importance, and so on) follow upon this fundamental binary divide.”

Morley and Partington (2009) argue within a line of reasoning similar to evolutionary psychology, claiming that the need for evaluation in human survival drives the creation of a dichotomous system of good and bad. They do however not provide any empirical evidence whatsoever, linguistic or otherwise, nor do they provide citations to previous research to corroborate their claims. This does not mean that they are necessarily wrong. Intuitively, good and bad seem to be psychologically salient categories. However, any positive claim carries a burden of proof, which Morley and Partington (2009) do not meet. Any claim that is presented without evidence can be dismissed without evidence.

Morley and Partington (2009) also make fairly strong claims. Even if we concede that there is a psychological reality to the concepts of good and bad, their other statements do not necessarily follow, i.e. it does not prove that these concepts would be expressed linguistically in the way they describe. Some approaches in construction grammar posit that humanly relevant experiences are reflected through privileged structures in language (e.g. Goldberg’s scene encoding hypothesis), but this is not proof that good/bad evaluations would be expressed through habitual co-occurrence patterns of words.
6.3 Evaluation is often more complex than evaluative polarity

In the previous chapter I discussed the complex semantics of *there BE something ADJ about*. This construction is by no means an outlier; there are many examples of semi-idiomatic constructions that possess strong evaluative semantics that cannot simply be broken down in terms of a positive or negative polarity. One such example is the *get oneself Ved* construction. The prototypical use of this construction evaluates someone as getting into a bad situation due to their own fault. The most common collexemes in the verbal slot are words like ‘shot’, ‘hurt’, ‘fired’, ‘incarcerated’, ‘mixed up’, ‘tangled up’. It is used to describe bodily injury (‘he always gets himself hurt’), career-related or legal problems (‘he managed to get himself fired twice’) or inserting oneself into complicated and unfavorable situations (‘the old coot got himself tangled up with a long-running feud’). Positive verbs are found too (‘cleaned up’, ‘straightened up’, ‘turned around’, ‘calmed down’) but many of their uses are in irrealis mode or negated (‘he needs to get himself straightened up’ or ‘he just couldn’t get himself calmed down’). The meaning of this construction has a strong negative component, but it is not the only component. It also includes the notion that the person brought the negative consequences upon themselves through their own action and their own fault. The construction can be generalized to *get oneself X*, because the result does not necessarily need to be expressed through a past participle verb. Another pattern is *get oneself in(to) NP* and the most common collexeme in the NP slot is ‘trouble’.
Only to be Ved describes a human (or human-like) entity performing an action which is then rendered moot by another entity’s counteraction, as in examples 1 - 3. The action performed by the first entity is typically difficult or requires a great deal of effort or time. The second entity tends to be in some way superior to the first or exert authority over it. As such it tends to overcome or overrule the first entity, even through a much less involved counteraction (especially in 2 & 3). While the evaluation expressed in this construction could be described as negative, it goes well beyond it. It also conveys a sense of futility and perhaps points out the absurdity of how great efforts can be rendered void by trivial counteractions.

(1) Nicea took advantage of those defeats to capture Adrianople and threaten Constantinople, only to be repulsed by Theodore’s forces in a major battle that ended the war.

(2) Corrupted intellectuals proclaimed their undying loyalty to General Morales Bermudez, only to be removed at the whim of even secondary power brokers.

(3) Some businessmen spend a month on applications only to be told that they used wrong-colored ink.

Wind up Ving (examples 4 & 5) is another highly evaluative construction. It describes an action that is often (though not always) undesirable and that the actor has little control over or at least did not expect. End up Ving (examples 6 & 7) is a very similar construction with comparable semantics.

(4) Many drone strikes wind up killing innocent civilians and fuel terrorism.
(5) It’s just ironic that I wound up coaching with him.

(6) Mike joked that he had no idea how he ended up teaching.

(7) These kids end up running a little bit too wild.

The semantics of these patterns can be summarized as follows. They carry negative evaluations that are usually relatively mild in their unfavorability (however, cf. example 4 for an exception). They typically describe a situation that is the result of a prior action or event that often remains undefined or unexplained. This situation tends to be unwanted, unexpected and was arrived at in a manner that lacked planning or control. The constructional semantics are partially compositional and conceptually motivated. The fact that they describe results or endpoints is driven by the semantics of the first verb (‘end up’ and ‘wind up’). The lack of volitionality and control appears to be driven by the fact that they are influenced by an underexplained prior event. Their tendency to occur with negative evaluations is a direct consequence of this lack of volitionality. One important point that this example perhaps demonstrates more clearly than the previous examples is that there is no primacy of evaluative polarity. The negative evaluation of wind/end up Ving is not the base or the starting point for further evaluative semantics; instead it results from them.
6.4 Some evaluations lack the positive/negative dimension

While semantic prosody is most often described as a dichotomous system of evaluation, not all researchers explicitly commit to this view. Sinclair’s view on the issue, for example, is not entirely clear. While he often describes semantic prosody as good or bad, he never defines it in these terms. In other words, he never states that it is more than a convenient label to summarize the data. More importantly, he sometimes includes other semantic components in his descriptions of it. In his analysis of the phrase ‘the naked eye’, he mentions that it often occurs with limiting adverbs such as ‘barely’ or ‘rarely’, creating a prosody of ‘difficulty’ (Sinclair, 1996, p. 87). For ‘true feelings’, he finds a prosody of ‘reluctance’ or ‘inability’ [ibid p. 89]. As the examples in the previous section, Sinclair’s examples display complex evaluative semantics. Not only do they go beyond a negative/positive polarity but they relegate it to a minor position. It could be argued, however, that they still include a dimension of good and bad in addition to their other semantics. This is however not always the case; there are linguistic items that are strongly evaluative but not necessarily positive or negative.

One example of a linguistic item that is used to evaluate but not to make a good-bad judgment is striking NP. The ten collexemes with the strongest collostructional attraction to this pattern are listed in table 6.1. The majority of its collexemes can be categorized into one of three groups, describing (i) a trait of something (feature, aspect, characteristic), (ii) similarity or difference (resemblance, difference, contrast,
similarity, parallel, change) or (iii) an example or result (example, thing, image, result, pattern, finding). Sentences 8 – 10 illustrate the usage.

(8) A third striking feature of the conference was the variety of issues addressed by participants.

(9) I suppose I currently bear a striking resemblance to that Buddha under the Bodhi Tree.

(10) The striking fact about the cultural diversity of ethnic groups in China is that there are so many forms of intangible culture still alive.

Striking NP evaluates something as being particularly salient, standing out in obviousness, importance or both. Most of the examples do not convey a sense that this is a good or a bad thing. At the very least, no such thing as a systematic evaluative polarity is found here.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Collexeme</th>
<th>Tokens</th>
<th>Attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>resemblance</td>
<td>179</td>
<td>3,075</td>
</tr>
<tr>
<td>2</td>
<td>feature</td>
<td>182</td>
<td>2,149</td>
</tr>
<tr>
<td>3</td>
<td>similarity</td>
<td>142</td>
<td>2,114</td>
</tr>
<tr>
<td>4</td>
<td>contrast</td>
<td>146</td>
<td>1,790</td>
</tr>
<tr>
<td>5</td>
<td>difference</td>
<td>156</td>
<td>1,511</td>
</tr>
<tr>
<td>6</td>
<td>thing</td>
<td>137</td>
<td>857</td>
</tr>
</tbody>
</table>
Other constructions that forms of ‘strike’ occur in, such as *strike NP as ADJ* and *strikingly ADJ* show no clear tendencies in this regard. The construction *strike NP as ADJ* (e.g. ‘the phone call struck me as odd’) is mostly used to express that something strikes a person as ‘odd’, ‘funny’, ‘strange’, ‘unusual’, ‘absurd’ or ‘curious’. The collexemes show that the construction is used to evaluate something as unusual. This can be positive, negative or devoid of any clear value judgment. The collexemes for *strikingly ADJ* (e.g. “but this is far from the only difference between these two strikingly different brothers”) can be categorized into three semantic groups: (i) likeness (‘similar’, ‘different’), (ii) attractiveness (‘beautiful’, ‘handsome’, ‘attractive’, ‘pretty’) and (iii) modernity (‘modern’, ‘contemporary’). The second and perhaps the third group show an evaluative valence, but the items in the first group are largely neutral in this regard.

Ultimately, there is no strictly principled relationship between polarity and other forms of evaluation. It can be present, but it does not need to be. Some linguistic items such as ‘set in’ are associated with collocates that are overwhelmingly nega-

<table>
<thead>
<tr>
<th>Rank</th>
<th>Collexeme</th>
<th>Count 1</th>
<th>Count 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>parallel</td>
<td>46</td>
<td>620</td>
</tr>
<tr>
<td>8</td>
<td>aspect</td>
<td>53</td>
<td>508</td>
</tr>
<tr>
<td>9</td>
<td>image</td>
<td>61</td>
<td>491</td>
</tr>
<tr>
<td>10</td>
<td>pattern</td>
<td>39</td>
<td>324</td>
</tr>
</tbody>
</table>

Table 6.1: The ten collexemes most attracted to the NP slot in the *striking NP* construction.
tive. Other linguistic items such as *striking NP* convey evaluations that do not have a strong valence. Yet others still are associated with different lexical fields, some of which do and some of which do not show this valence (e.g. *strikingly ADJ*). Additionally, as I argue in the next section, it is not uncommon for items that tend to express strong value judgments to be associated with both positive and negative evaluations.

### 6.5 Some words are associated with opposite evaluations

The distributional hypothesis states that semantically similar words tend to occur within similar distributional contexts. We find that this is not limited to relationships of synonymy but also includes antonymy. For example, ‘alleviate’ and ‘exacerbate’ tend to take similar object arguments. As such, the same linguistic unit can habitually occur in statements that express either positive or negative evaluations.

Some studies from the previous chapter have already shown this. The collexemes with the strongest attraction to the NP slot in the *cause for NP* construction are ‘concern’ and ‘alarm’ but also ‘celebration’ and ‘optimism’. The *there be something about NP* construction describes something as unusual which can be good or bad. A similar case can be made about the nominal extraposition construction, *it be ADJ the NP* (‘it is amazing the things people will do’). Both constructions are associated with collexemes that tend to occupy extreme positions on either end of the positive-negative continuum.

A straightforward example for a lexeme that is found in diametrically opposed
evaluations is ‘absolutely’. For the construction absolutely ADJ, four antonymous word groups are found: (i) ‘right’, ‘correct’ and ‘true’ compared to ‘wrong’ and ‘false’; (ii) ‘necessary’, ‘essential’, ‘crucial’, ‘indispensable’ and ‘vital’ contrasted with ‘unnecessary’; (iii) ‘sure’ and ‘certain’ versus ‘unbelievable’; (iv) ‘beautiful’, ‘amazing’, ‘wonderful’ and ‘perfect’ but also ‘horrible’, ‘terrible’, ‘absurd’. Most of the highly attracted collexemes tend to be positive (see table 6.2), but their negative counterparts are frequent as well. No argument could be made that ‘absolutely wrong’ sounds infelicitous because ‘absolutely right’ is so frequent. A similar point can be made about absolutely V. The verb slot in the construction is commonly filled with both ‘agree’ and ‘believe’ but also ‘disagree’ and ‘refuse’.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Collexeme</th>
<th>Tokens</th>
<th>Attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>right</td>
<td>2,055</td>
<td>21,963</td>
</tr>
<tr>
<td>2</td>
<td>necessary</td>
<td>599</td>
<td>5,954</td>
</tr>
<tr>
<td>3</td>
<td>convinced</td>
<td>240</td>
<td>5,442</td>
</tr>
<tr>
<td>4</td>
<td>true</td>
<td>596</td>
<td>4,780</td>
</tr>
<tr>
<td>5</td>
<td>correct</td>
<td>420</td>
<td>4,522</td>
</tr>
<tr>
<td>6</td>
<td>essential</td>
<td>433</td>
<td>4,503</td>
</tr>
<tr>
<td>7</td>
<td>certain</td>
<td>380</td>
<td>3,096</td>
</tr>
<tr>
<td>8</td>
<td>sure</td>
<td>408</td>
<td>2,849</td>
</tr>
<tr>
<td>9</td>
<td>wrong</td>
<td>292</td>
<td>2,460</td>
</tr>
</tbody>
</table>
Another example is *sheer NP*, which occurs with a number of different semantic fields some of which are not related to each other in any obvious way. The most common collexemes express vast magnitude (‘number’, ‘size’, ‘volume’) or physical power (‘strength’, ‘force’, ‘momentum’). Also very common are intense emotions, both positive (‘joy’, ‘pleasure’, ‘bliss’, ‘excitement’) and negative (‘terror’, ‘desperation’, ‘boredom’, ‘frustration’). Collexemes conveying perseverance (‘grit’, ‘willpower’, ‘persistence’, ‘determination’, ‘stubbornness’) are highly frequent as well; these collexemes are related to terms of boldness (‘bravado’ or ‘audacity’) which are in turn connected to similar but more negative characteristics (‘folly’, ‘stupidity’ ‘arrogance’). In summary, ‘sheer NP’ occurs with all sorts of collexemes. Some of the lexical fields do not necessarily involve a value judgment (‘sheer number’, ‘sheer momentum’). Others are either located at the extreme poles of a positive-negative continuum (‘bliss’ vs ‘terror’) or span across it (‘grit’, ‘audacity’, ‘folly’).

The nominal extraposition construction, *it be ADJ the NP* (Michaelis and Lambrecht, 1996) is another example of a construction that occurs both with strongly positive and negative evaluations. It is used to take the stance that something is remarkable or extreme (‘it’s incredible the things ordinary people will do in wartime’). These evaluations can be positive (‘it’s wonderful the way they’re balancing work and
family’) or negative (‘it was horrible the way they died’). While its prosodic valence is certainly a salient aspect of its meaning, it is not at its core. Instead, it is the judgment that something is remarkable that constitutes its invariant core and that gives rise to the positive or negative evaluations. Again, this contradicts the idea that evaluation is primarily a matter of a positive/negative polarity.

In summary, the idea of semantic prosody that words have a tendency to occur with either negative or positive evaluations cannot be maintained. It is entirely normal for lexical items and constructions to be associated with both positive and negative collexemes. None of this is surprising. Semantic preference is an important principle in lexical co-selection, but lexical fields are not organized by synonymy alone. Instead, it is common for words within the same lexical fields to be antonyms (or near-antonyms).

### 6.6 The difficulty of assigning evaluative polarity

A basic problem with semantic prosody is its inherent vagueness and subjectivity. So far, nobody has been able to provide an operationalized definition of what a positive or negative evaluation is. Instead, this issue is usually glossed over and apparently taken as intuitively obvious. Dilts and Newman (2006) make a similar point:

“[T]he researcher is required to make evaluative judgments in the absence of a set of principled criteria to guide the evaluation. Terms such as ‘good’ or ‘bad’, ‘positive’ or ‘negative’ etc. are introduced at will and without
much care taken to explain the basis for the judgment. [...] In fact, the issue is rarely even addressed in the typical prosody study.” (Dilts and Newman, 2006, p. 233)

As Stewart (2009b) points out, contrary to what is often said in the literature, semantic prosody is not revealed through corpora. Corpora may reveal distributional patterns, but semantic prosody is a researcher’s subjective interpretation of these patterns. This is not just a theoretical issue but also a problem on a practical level. Many instances of semantic prosody are difficult to judge as positive or negative. One question is at what level the evaluation should be carried out. Consider examples 11 - 13 (all taken from the Brown corpus). ‘Cancer’ in example 11 describes a dangerous illness so it does not seem to be a stretch to call it a bad thing. The sentence, however, talks about curing this illness. ‘Cure’ is one of the verbs that Hunston (2007a) describes as having a reversing relationship to its collocates (see chapter 2).

The question then is what aspect matters in the case of ‘cure cancer’ – the fact that ‘cure’ co-occurs with a negative word or the fact that the compositional meaning of the larger phrase is favorable? As discussed in chapter 2, there is no consensus on how to deal with this issue; Ellis et al. (2007) espouse the former view while Morley and Partington (2009) are an example of the latter view. How we analyze it may depend on the specific questions we are asking, but even then, things are not entirely clear. If we follow the Firthian maxim of knowing a word ‘by the company it keeps’ we could make the case for either approach: Speakers may well remember a word’s
collocates but also the types of statements it occurs in.

(11) This machine will cure your cancer-ridden body.

(12) But none of this could soothe the exacerbated nerves.

(13) And he caused the fumble that set up our touchdown.

(14) From its inception, ACORN was “aggressive and persistent,” he says. It used “direct action” — rallies, protests and confrontations — to get attention. Sometimes people were arrested, he says.

The same issue is taken a step further in example sentence 12. Again, we have a relationship of reversal between the verb and its argument, but in this case, the sentence is also negated. There are three different ways to evaluate the sentence: First on the level of the collocate, second on the level of the verb phrase (which negates it through the semantics of the verb) and third, on the sentence level (which negates it again, through the negative pronoun ‘none’). Example (13) takes things another step further to an even broader level. Here, the issue is extra-linguistic; there is a problem of perspective, i.e. the question from whose point of view something is favorable or not. The sentence describes a scenario where there are two groups whose interests are directly opposed. A player in a football match makes a mistake which is bad for his team and consequently good for the other (the speaker’s) team. The question is whether this should be treated within a local context (the verb ‘cause’ describes a
direct negative effect) or whether it is the evaluation from the speaker’s perspective in a broader context that counts.

Finally, sometimes speakers do not take a clear positive or negative stance on a topic but instead either leave their own appraisal open or make a nuanced point, highlighting both positive and negative aspects. This is shown in example 14, a slightly expanded version of the sentence mentioned in the previous chapter. It is not obvious what the speaker thinks about the group. He mentions some issues but does not condemn them. Corpus linguistics, particularly in its more quantitative forms tends to treat evaluation as a property of text, e.g. something that can be summarized with a sentiment score. Research in evaluation, however, shows that it is the expression of speakers’ subjectivity, a means by which they position themselves relative to the content of the discourse.

The discussion in this chapter has mostly been focused on semantic prosody as a local distributional phenomenon, which Hunston (2007a) refers to as the lexical priming perspective. If we consider semantic prosody from the discourse perspective instead, it is still unclear what its role could be. Morley and Partington (2009) explains that the two perspectives are inseparably interconnected: The lexical priming perspective gives rise to the discourse perspective which in turn organizes the former in a larger context. If we reject semantic prosody from the lexical priming perspective, the question is what remains of it at all. There is evaluative cohesion, which serves to organize speech and text in an evaluatively consistent manner. Partington
(2017, p. 202) regards this as the function of semantic prosody (which he refers to as evaluative prosody): “Evaluative prosody is the ‘invisible’ non-obvious cohesive glue by which texts cohere evaluatively (and not just propositionally). It creates evaluative harmony, to avoid mixed evaluative messaging.” But again, evaluative cohesion (or clashing) can be accounted for through semantic preference; choices of linguistic items can be explained by comparisons with the lexical fields they habitually co-select with.

6.7 Discussion

The data presented in this chapter do not support the traditional account of the function and meaning of semantic prosody. One important finding is that the notion of evaluative polarity cannot be maintained. The corpus evidence shows that the evaluative meaning of semantic prosody is much more complex. First, many words and constructions have evaluative semantics that go well beyond a simple polarity. Second, evaluations do not have to be defined for a positive/negative meaning. Third, the same linguistic structure can be associated with both positive and negative evaluations. What all the examples discussed in this chapter have in common is that in contrast to what has been described in the literature (Thompson and Hunston, 2000; Hunston, 2002, inter alia), there is no indication that semantic prosody is at its core a matter of a positive/negative dichotomy. Such claims do not explain what exactly is meant by semantic prosody being primarily a matter of evaluative polarity. In other
words, it is not clear what the ontological status of this primacy is supposed to be and where it comes from or is situated. If we examine how semantic prosody arises from usage events, however, we see a positive/negative polarity is a direct result of broader evaluative semantics, not its antecedent. In this light it makes sense that evaluative meaning can have a tendency toward either negative or positive evaluations (e.g. *end up Ving*), that it can be undefined for this dimension (*striking NP*), or that that the same construction can be associated with both (*it be ADJ the NP*).

Another issue concerns the relation between semantic prosody and semantic preference. No patterns exist in corpus data that cannot be accounted for through semantic preference. It is sometimes argued that the linguistic generalizations speakers make include an abstraction away from lexical fields to a two-term system of evaluation (e.g. Bednarek, 2008; Partington, 2004). This seems like a claim that goes well beyond what the data warrant. It shifts the discriptum away from language usage to mental representations and makes psychological claims based on corpus data alone. Additionally, the proposed patterns are not found in corpus data so it seems speculative that they should exist in the mind.

Positing semantic prosody as a separate phenomenon in its own right adds both an unwarranted assumption on the theoretical level and an additional layer of subjectivity on the level of the data analysis. Partington’s (2004) point that semantic preference is “closely tied to the phenomenon of collocation” (p. 150) is well-taken. Patterns of collocates and collexemes arise directly from corpus data. Semantic pref-
ference introduces a relatively minor step of abstraction into the analysis. As results of the cluster analyses in chapter 5 show, lexical fields are natural semantic groupings that emerge from their distributions in data. Semantic prosody, on the other hand, requires a major step of abstraction that is much harder to justify. There is no theoretical reason to assume such a generalization exists and, more importantly, corpus evidence does not support it either. In summary, there are two aspects of evaluation I have discussed in this chapter. The first is that the evaluations linguistic items are associated with can be understood through their habitual co-occurrences with lexical fields. The data do not attest to the existence of semantic prosody as its own category, separable from semantic preference. The second aspect is the principle of evaluative cohesion on the larger level. Introducing the concept of semantic prosody on this level does not add anything to the analysis either.
Chapter 7

Investigating the Claim of Hiddenness

Almost every research article on semantic prosody includes a section where the authors state in one form or another that it is ‘hidden’ from speakers’ explicit awareness, intuition or introspection. To mention just a few examples, Partington (1998, p.77) claims: “Semantic prosody is very frequently hidden from immediate view and therefore difficult to predict out of context [. . .].” Channell (2000, p. 41) states: “evaluative polarity is not usually accessible to intuition” and Berber Sardinha (2000, p. 106) puts it almost the exact same way: “semantic prosody is not directly accessible by intuition.” Adolphs and Carter (2002, p.7) agree: “semantic prosodies are difficult, if not impossible, to determine on the basis of intuition alone.”

Xiao and McEnery (2006, p. 103) find that patterns of co-selection in general and semantic prosody, in particular, are not easily accessed by intuition; “a speaker’s intuition is usually an unreliable guide to patterns of collocation and [. . .] intuition is an even poorer guide to semantic prosody.” McEnery et al. (2006) similarly conclude that semantic prosody “is at least as inaccessible to a speaker’s conscious introspection as collocation is.” Many authors follow such statements up by concluding that semantic prosody can therefore only be revealed through the study of corpus data. Louw and Chateau (2010, 756), for example, put it as follows: “Semantic prosody is
I will argue that the issue is more complicated than this and that claims about the hiddenness of semantic prosody are both problematic and unsubstantiated. The remainder of this chapter is structured as follows. In section 7.1, I critically examine the accounts of hiddenness in the literature and argue why many of the claims are not convincing. In 7.2, I provide a comprehensive review of psycholinguistic studies that investigate the conceptual basis of semantic prosody and discuss what their results mean for the notion of hiddenness. In section 7.3, I provide data from elicited judgments of evaluative valence and investigate how speakers’ intuitions compare to corpus data. Section 7.4 discusses the results of this chapter and their implications.

### 7.1 Issues with claims of hiddenness

Statements about the covertness of semantic prosody are typically put forward with little or no evidence to support them. It seems odd that such a central and much-repeated claim is simply taken for granted and has never been rigorously investigated. Instead, it is usually based on anecdotal information or the researcher’s own introspection. Some researchers, for example, point out that semantic prosody is often not captured in dictionaries. This is taken to be evidence that it is hidden even from professional lexicographers. Sinclair (1998, p. 16) finds ‘budge’ to be a problem for dictionaries. He cites its definition in the Longman Dictionary of Contemporary En-
glish, which misses the tendency of the verb to occur in negated clauses and as such in contexts that convey a sense of difficulty (e.g. “the door won’t budge”). Stubbs (1995a, p. 27) reports finding neutral definitions in various dictionaries for ‘cause’ and Channell (2000) notes that the Oxford Advanced Learners’ Dictionary defines ‘right-on’ and ‘par for the course’ without mentioning any negative connotations.

Out of all the foundational papers that aim to formulate a theory of semantic prosody, Stubbs (1995a) is the only researcher who makes at least some attempt at a survey of speakers’ intuitions. Concerning the word ‘cause’ he states: “Some native speakers (but not all) that I have informally tested do produce one or two examples of such unpleasant collocations, but such native speaker data are very sparse and unreliable indeed. Neither do widely used dictionaries explicitly draw attention to such negative cases.” To my best knowledge, this is the extent of evidence presented to show that semantic prosody is hidden: It is sometimes not included in dictionaries, and a qualitative, informal survey seemed to suggest that speakers have relatively little knowledge of the typically negative collocations of the verb ‘cause’. The overwhelming majority of research articles simply asserts that semantic prosody is hidden without providing any evidence at all.

Researchers seem to be comfortable making unsubstantiated claims about the hiddenness of semantic prosody because it follows widespread conventional wisdom within corpus linguistics: the idea that intuition can generally not be trusted as a guide to frequency-based phenomena. Stubbs (1993, p. 17), for example, claims that
“native speakers have no reliable intuitions about such statistical tendencies.” In the same vein, Bublitz (1996, p. 23) states that “intuitions about frequency and likelihood of co-occurrence are notoriously thin and not always accurate.” Hunston (2002, p. 21) believes: “It is almost impossible to be conscious of the relative frequency of words, phrases and structures except in very general terms.” Typical examples in corpus linguistics in general (i.e. outside of research on semantic prosody) include Biber et al. (1996, p. 120), who asserts: “Intuitions regarding lexical associations are often unreliable and inaccurate” and Hoey (2005, p.133), who declares: “if there is one thing a corpus linguist quickly learns it is that their intuitions almost always simplify the picture or tell outright lies.” These claims are often used to argue in favor of the importance of corpus linguistics and its superiority to introspective methods. Channell (2000, p.39), for example, contends that “analysis of evaluation can be removed from the chancy and unreliable business of linguistic intuitions and based in systematic observation of naturally occurring data.” Unfortunately, this claim, too, is accepted as a general truth without providing evidence or references to other research in order to corroborate these claims. In his review of the literature, McGee (2012, p. 80) makes the same point: “The actual evidence forwarded to support such claims is typically rather thin: being either anecdotal [...] or indirect, for example on the basis of material included or excluded in teaching materials [...].”

There is research outside of corpus linguistics that shows that these assumptions are unwarranted. Research in psychology shows that speakers do in fact possess
reliable intuitions about word frequencies. There is a strong positive relationship between word frequencies in reference corpora on the one hand and speakers’ word familiarity ratings or their subjective word frequency estimates on the other hand. Log-transformed corpus frequencies generally model speakers’ intuition fairly well (a common mathematical relationship between stimulus and human perception, known in psychophysics as the Weber-Fechner law). This is a very robust finding that has been replicated many times through a research tradition spanning several decades (cf. Tryk, 1968; Shapiro, 1969; Carroll, 1971; Balota et al., 2001, inter alia). These findings directly contradict commonly held believes in corpus linguistics.¹

It appears that corpus linguists are generally not aware of this research. Rapp (2005), who works within corpus linguistics and computational linguistics, investigates the relationship between word familiarity ratings and corpus frequencies. He comes to the same conclusion as the seminal studies in psychology, finding a strong correlation between subjects’ familiarity ratings and log-transformed frequencies in corpora but does not cite any of those studies, which, again, indicates a lack of interdisciplinary communication.²

¹This research is also remarkable in that it shows that speakers do not only reliably track frequencies of occurrence (as is the central tenet of exemplar theory and an important assumption in usage-based linguistics), but in this case, this knowledge is also not hidden but instead explicit and directly accessible to them.

²The only exception seems to be research in second language learning, which does, in fact, conduct studies based on relevant previous work in psychology (e.g. Arnaud, 1990; Schmitt and Dunham,
When we move on from single-word frequencies to word co-occurrences or frequencies of larger linguistic units, research studies become sparser and the results are less straightforward. Hoffmann and Lehmann (2000) retrieved two-word collocations (mostly adjective-noun and noun-noun collocates) with low frequencies but high log-likelihood values from the BNC. In other words, these were word-pairs with relatively low overall frequencies of occurrence in the BNC, but strong collocates because their two words tended to co-occur together, e.g. “unsung heroes” (p. 21). They presented the study participants with the first words of the word-pairs and asked them to fill in the second word. L1 speakers performed well, matching the corpus data in 70% of all cases. Siyanova and Schmitt (2008) asked participants to rate English adjective-noun pairs as to whether the constitute frequent or infrequent collocations. They find that L1 speakers perform this task with a high degree of reliability. Hoffmann Thomas (2007) studies preposition placement in English relative clauses and finds a strong overlap between corpus evidence and grammaticality judgments. Wulff Stefanie (2009) models idiomaticity judgments for V NP collocates with various distributional features through regression analysis and is able to account for almost 80% of all variation.

Other studies, however, do not find such straightforward associations between speakers’ elicited linguistic judgments and corpus data. McGee Iain (2009) presented

---

3In this case, transitional probabilities from the first word to the second may have been a more appropriate metric than a bidirectional measure of collocational strength.
participants with high-frequency adjectives and asked them to provide the most frequent noun collocates. The results are mixed, and the author hypothesizes that there may be an availability heuristic for frequency judgments. Mcgee discusses different potential explanations as to why the most available words may not always be the most frequent ones. A number of experiments on verb sense and verb subcategorization probabilities by Roland and Jurafsky (2002) find more differences than similarities between corpora and experimental measures which the authors attribute to the differences between isolated sentences and connected discourse. They conclude that “‘test-tube’ sentences are not the same as ‘wild’ sentences’” (Roland and Jurafsky, 2002, p. 327).

There are a few takeaways. The idea that semantic prosody is hidden from speakers’ conscious awareness is proposed without presenting any substantial evidence. This seems to stem from an attitude within corpus linguistics that generally dismisses the validity of intuitions about frequency-based language phenomena. However, experimental evidence shows that the picture is more complex. The question to what extent speakers possess explicit knowledge of frequency-based language phenomena is both complicated and understudied – at least when we move beyond the level of the individual word. As a result, many complexities remain unknown. Some studies find a relationship between corpus frequencies and speaker judgments while others do not. It is not clear, why this is the case. It may differ depending on the specific linguistic phenomenon under investigation or it may differ based on how the task is
posed to the participants. At the very least, however, it is incorrect to assume that
semantic prosody is hidden because frequency-based phenomena always are.

Another issue concerns the terminology used within the field of semantic prosody
research. It is not entirely clear what researchers mean when they describe semantic
prosody as being hidden from intuition, inaccessible through introspection or not part
of speakers’ (conscious) awareness or knowledge of language. These terms all remain
undefined and are often used interchangeably. Louw (1993, p.173), for example, goes
back and forth between intuition and introspection: “It may well turn out to be the
case that semantic prosodies are less accessible through human intuition than most
other phenomena to do with language [. . .] corpus linguistics reveals a greater and
greater mismatch between the products of introspection about language and those of
direct observation.”

This is not an issue that is limited to the study of semantic prosody, but that is
common throughout linguistics in general. Willems (2012, p. 671) problematizes the
lack of a distinction and finds: “It is currently customary in both Anglo-American
and European linguistics not to distinguish between intuition and introspection.” A
few linguists do attempt to differentiate between the two concepts (Pateman, 1987;
Itkonen, 2008; Schütze, 2016, inter alia) but arrive at different conclusions about
what defines and distinguishes them. None of their descriptions seem to have gained
traction in linguistics.

While the conflation of intuition and introspection is common in linguistics, it
could be particularly problematic for research into semantic prosody because a distinction is possibly relevant in this context. As Stewart (2009b, pp. 126ff) points out, when it comes to speakers’ knowledge and understanding of semantic prosody, there may be a difference between *immediate insight* and *intellectual examination*. Stewart sees linguistic intuition as an “instinctive, immediate ‘feel’ for language” (p. 128). Introspection, on the other hand, is typically considered as the metacognitive operation of carefully examining one’s own thought processes. Speakers may not be able to instantly determine the semantic prosody of a linguistic item; when given some time, however, they may be able to go over some typical usage cases and make deliberate and thoughtful generalizations about what evaluative contexts an item tends to occur in. Hunston (2007a) makes a similar observation, noting that when speakers are made aware of semantic prosodies they had not noticed before, they often find them to agree with their intuitions:

“The semantic prosody of a word is often not accessible from a speaker’s conscious knowledge. Few people, for example, would define *SET IN* as meaning ‘something bad starts to happen,’ but when the negative connotation is pointed out in many cases it accords with intuition (*A spell of fine weather set in* sounds very odd, for example).” (Hunston, 2007a, p. 142)

There is another terminological pitfall. We need to be careful what we mean by speakers having (or lacking) intuitions: There is a difference between having intuitions
about the semantic prosody of a lexical item (that surface in our communicative competencies) and being able to explicitly access our intuition through introspection in order to form an intellectual judgment.

Several researchers have noted that while speakers lack reliable intuitions when asked about words in isolation, they are clearly sensitive to them in language usage. Again, we need to clarify what is meant by intuitions in this context and be careful not to conflate intuition and introspection. What researchers typically mean when they say that speakers lack intuition about semantic prosody, is, that they are unable to retrieve accurate information about semantic prosody by querying their language intuitions by means of introspection. Their intuitions do not rise to the level of explicit knowledge, but this does not mean that speakers do not possess implicit knowledge of semantic prosody. As Partington (2014b, p. 288) puts it, speakers ‘intuitively’ use lexical items in a prosodically coherent way: “Whilst it is true that the average speaker may not be able to articulate the complex prosodic behaviour of, say, *build up* [...], they have no difficulties in intuitively employing it appropriately in conversation.” Partington (2004) expands upon this idea and accounts for implicit knowledge of semantic prosody through priming:

“If the favourable or unfavourable evaluation of an item said to display semantic prosody is not part of its in-built, inherent meaning [...], then how do language users decide to employ such items in the appropriate environment? The answer is that language users have a set of mental
rules derived from the priming process, alongside or integrated with the mental lexicon, of how items should collocate. These rules are not always open to casual introspection.” (Partington, 2004, p.132)

Stewart (2009b, p. 129) agrees that while speakers may lack explicitly accessible knowledge of semantic prosody, they do possess some form of knowledge which becomes evident through their usage in communication: “Knowledge of co-occurrence is not necessarily either conscious or explicitly recollectable but remains part of our communicative competence. Competent speakers’ knowledge of the item set in, then, includes the fact that it is not normally found in a favourable environment.” Louw (1993, p. 157) on the other hand goes as far as to claim that even in usage, speakers might not be consciously aware of the semantic prosodies they employ, which “may mark the speaker’s real attitude even when s/he is at pains to conceal it.” Tognini-Bonelli (2001) following Louw’s line of argument believes that “This lack of control suggests that semantic prosodies may operate mainly subliminally and are not readily available to the speaker as discourse devices at the conscious level” (p. 112).

The idea that speakers store some information about semantic prosody (or frequencies of co-occurrence in general) is almost trivial; otherwise, how would they employ it in language production and comprehension and how would these patterns arise in corpora? Whitsitt (2005, p. 294) makes a similar point: “intuition […] is the very thing that makes a corpus possible, for it is surely the collection of people’s intuitive use of language that makes it possible for a corpus to contain ‘real’
language.”

At this point, it becomes clear that we need to further revise the ideas about intuitions of semantic prosody, at least to be more precise in our terminology. We need to distinguish between *having* intuitions on the one hand and *querying* these intuitions to come to an explicit, intellectual understanding on the other hand. When Stubbs (1993, p. 17) claims that speakers have “no reliable intuition” about co-occurrence frequencies or when Bublitz (1996, p. 23) asserts that these intuitions are “notoriously thin and not always accurate”, they fail to make this distinction. In more recently published articles, authors are starting to be aware of this distinction. Partington (2015, p. 288), for example, states: “what speakers may know via introspection about a language feature – especially when asked about it with no context supplied – is not the same as their intuitive awareness of how to use it.” There is not sufficient evidence to posit claims about whether or speakers have conscious knowledge of semantic prosody, but it seems obvious that they have at least some form of intuitive knowledge.

The next problem with the claim that semantic prosody is hidden is perhaps the most obvious. There are linguistic items for which speakers are in fact explicitly aware of their semantic prosody. First, a definitional question needs to be addressed: Do we consider semantic prosody to be a feature of all linguistic items or do we limit our investigation to seemingly neutral words? As briefly mentioned in chapter 2, the great majority of published articles avoids discussing overtly evaluative words and instead focuses on ostensibly neutral words, such as ‘happen’ or ‘cause’. With
very few exceptions, researchers do not clarify why this is the case. It could either be because seemingly neutral words make for a more interesting research topic or because evaluative words are by definition excluded.

There are only a handful of exceptions. A few articles on semantic prosody include case studies of clearly evaluative words, but do not discuss why they included them: Stubbs (2001), for example, briefly mentions ‘reckless’ and Zhang (2013) describes the diachronic development of words like ‘horribly’ or ‘dreadfully’. Other articles make a distinction between more and less overtly evaluative lexical items. Channell (2000) contrasts lexis with “obviously and intuitively accessible polarity” (p. 44), e.g. “self-important” (p. 43), and “less obvious” examples (p. 44), e.g. “par for the course” (p. 47). It should be noted, however, that Channell does not mention the term semantic prosody, nor reference any research into it. Instead, she talks about “pragmatic meaning” (p. 39) and “evaluative polarity” (p. 41). Hunston (2001) makes a similar distinction, finding that some lexical items “evaluate explicitly in the sense that their evaluative meaning is immediately obvious to a speaker of English” (p. 21). In other cases, “this function is much less obvious” (ibid) and the “evaluation is being expressed prosodically throughout the clause, rather than being restricted to particular lexical items” (ibid). Hunston (2001) does however not take a position on whether the former examples should also be analyzed within their prosodic contexts.

Lewandowska-Tomaszczyk (1996, p. 157) explicitly addresses the issue whether or semantic prosody is a relevant category for overtly evaluative words and concludes:
“I would see no reason [...] not to use the term semantic prosody for the cases when the semantic load of an item is quite explicit.” There is indeed nothing within the traditional definitions of semantic prosody that would exclude obviously evaluative lexical items. Sinclair (1998, pp. 14-15) considers semantic prosody to be one of the ‘obligatory’ components of any lexical item, meaning that it is, in fact, a meaningful category for any word, not only ‘neutral’ ones. This means that semantic prosody would at best sometimes be hidden from speakers’ conscious awareness, more specifically only in the case of lexical items that lack any obvious evaluative semantics in isolation but that are prosodically primed for the evaluative contexts they occur in.

This position is shared by some authors, who distance themselves from Louw’s (1993, p. 173) assertion that “semantic prosodies have in large measure and for thousands of years remained hidden from our perception and inaccessible to our intuition.” Partington (2004), for example, notes that lexicographers did, in fact, document the unfavorable collocations of some words, even prior to using corpus methods:

“There may be a slight overstatement here. That a number of items have good or bad prosodies was clear before the advent of corpus lexicography, e.g. items like commit or perpetrate are usually described, even in pre-corpus dictionaries, as collocating unfavourably.” (Partington, 2004, p. 155)

McGee (2012, p. 171) agrees, finding that Louw’s position is “generally understood to be exaggerated.” A categorical distinction between transparently evaluative lexical
items and potentially hidden lexical items is possibly a false dichotomy. There may well be a continuum from obvious to hidden items. Morley and Partington (2009) make the case for this idea:

“The ‘obviousness’ of evaluative connotation is best considered as a cline. Items such as *callow* and *venerable* seem to express fairly clearly unfavourable and favourable connotation / evaluation respectively. Items such as *peddle, commit* and *fraught with* are somewhere in the middle, whereas *[not] budge, border on* and *dealings with* and the different evaluative polarities of *build up [NP]* and *[NP] builds up* were entirely obscure until assistance came to hand in the form of corpora.” (Morley and Partington, 2009, p. 151)

Morley and Partington’s (2004) claim seems reasonable. It is, however, not based on evidence, but only on the authors’ intuitions (some of which, coincidentally, do not align with mine). Again, experimental evidence would be required to test this idea.

Finally, asking about speakers’ knowledge about semantic prosodies of words strikes me as asking the wrong question. As discussed in the previous chapters, there are two issues with this: For one thing, as I argue in chapter 5, the lexical approach fails to tease out distinctions in contextual meaning. Words do not exist in isolation. An adjective like ‘persistent’ carries an entirely different evaluative meaning in attributive contexts than in predictive usage. Additionally, as argued in chapter 6,
it is doubtful whether an abstraction to the level of a positive or negative evaluative polarity is a real and relevant mental category. Speakers may be able to make an ad hoc generalization about an items’ evaluative polarity when required to do so, but this does not mean that this generalization is at the core of semantic prosody or that it is stored in the long-term memory as its own category.

In summary, the claims about the covertness of semantic prosody need to be much more restricted than they are usually presented in the research literature. First, they only apply to a subset of lexical items: Words and phrases that possess prosodic primings but that do not carry any obvious inherent evaluative meaning in isolation. Second, speakers do not ‘lack intuitions’ about semantic prosody. Instead, they may not be able to access their intuitions in order to retrieve explicit information about semantic prosody. Third, semantic prosodies may be on a continuum of hiddenness, varying in terms of how obvious or non-obvious their evaluative primings are. This would mean that only a small number of lexical items could potentially be considered to be truly hidden. Finally, and most importantly, even this cannot be taken for granted. No strong evidence has been put forward to prove that semantic prosodies are ever hidden.

7.2 Psycholinguistic studies on semantic prosody

In the previous section, I argued that claims about the hiddenness of semantic prosody are unproven and likely overstated. A crucial problem is that authors tend to make
conclusions about psychological phenomena solely on the basis of corpus data. Psycholinguistic studies of semantic prosody do exist, but they are few and far between and do not directly investigate the question of hiddenness. In this section, I provide a review of these studies and discuss what their finding mean for the idea that it is a hidden category.

In a small study that aims to compare corpus data with elicited data, Nordquist (2004) investigates whether speakers are able to access the semantic prosody of the word ‘cause’ when they are presented with them outside of any discoursal context. Participants were asked to use the word in three sentences each. Nordquist analyzed the sentiments expressed in the sentences and found that 70% of all produced utterances carried negative evaluations. While the author does not provide a quantitative comparison between this result and corpus data, there appears to be a relatively good match. In a 200-sentence sample analyzed in chapter 2, 63% of the sentences were found to be negative. It has to be noted that the study design does not determine whether the results are due to knowledge of semantic prosody or simply due to collocational knowledge. Additionally, the study is similar to a free association task – participants were not asked to provide the most common collocates for ‘cause’ but instead simply to produce three sentences that included the word. As McGee Iain (2009) and Siyanova and Schmitt (2008) point out, the way that judgments are elicited from subjects likely affects study results.

Ellis et al. (2007) aim to establish priming for semantic prosody in a lexical de-
cision task. While they show the existence of collocational priming, they are not able to find evidence for priming effects based on semantic prosody alone. In other words, they do not find priming for word pairs that are congruent in terms of their semantic prosodies but that are not collocates. Pairs like ‘attain goals’ (collocational match) are processed faster than pairs like ‘attain problems’ (collocational mismatch), but the same is not true when comparing ‘attain virtue’ (no collocational match but evaluative polarity match) to ‘attain-harm’ (evaluative polarity mismatch). Negative experimental results are technically not interpretable due to the change of committing type II errors, e.g. the statistical power might not be sufficient to detect the effect, or the way the treatment was administered might not have been optimal. Still, the fact remains that there is no evidence to show that semantic prosody affects lexical access and word recognition (Ellis et al., 2007, p. 106). The authors conclude that priming in lexical decision tasks is an example of ‘repetition priming,’ i.e. specific constructions are stored and become quicker to retrieve, but there are no wider generalizations: “It appears then that fluent lexical access is due to memory for particular lexical associations – there are no top-down semantic generalizations upon this level of processing” (ibid).

Ellis and Frey (2009) find that affective priming is influenced by semantic prosody.

---

4These conclusions seem somewhat problematic in that they confound ‘failing to reject the null hypothesis’ with ‘accepting the null hypothesis’ which (as mentioned above) is an invalid argument in hypothesis testing. A more careful conclusion might be: “There is no evidence for top-down semantic generalizations upon this level of processing.”
In the affective priming paradigm (Fazio et al., 1986), subjects are asked to judge a target word as positive or negative as quickly as possible. When the prime and the target have a congruent evaluative polarity, response times are faster than in the incongruent condition. Ellis and Frey (2009) presented participants with 40 verbal primes, half of which had positive semantic prosodies and the other half of which had negative semantic prosodies. Next, the subjects were shown a target noun and asked to indicate whether this word is positive or negative (the polarity of the target is in no way hidden). Primes and targets were designed not to be collocates. The results showed that responses to congruent pairs were quicker and more accurate than for incongruent pairs. A possible issue with this study is that the researchers only focus on the semantic prosodies of their primes, and do not consider whether they possess any transparently evaluative meanings. One of their verbal primes, for example, is ‘suffer’ which strikes me as obviously negative.

In a supplemental task, Ellis and Frey (2009) elicited explicit ratings of evaluative valence from their participants. Subjects were asked to rate the 20 verb primes from the main study on a nine-point scale of pleasantness from most positive (+4) to most negative (-4). The authors then compared these ratings to the corpus data of semantic prosody. Their corpus data are based on type counts of collocates. The authors do not explain why they use type counts rather than token counts. Most frequentist approaches are more interested tokens than in types. The reason is that type frequencies are seen as a feature of the lexicon; they describe the available items for a given category. Token counts, on the other hand, reflect usage. Some studies have found areas where type counts are more important.
ratings had a strong positive relationship with the absolute number of positive and negative collocates ($\beta = 0.57$, $p < .001$, $R^2 = 0.32$).

There are, however, some outliers: ‘Cause’ and ‘commit’ are rated by participants as essentially neutral while corpus data found it to be negative. ‘Cure’, is rated as positive although the authors assigned it a negative semantic prosody. ‘Lack’ is rated as negatively evaluated while its corpus data show a strong positive semantic prosody. The issue with ‘cure’ and ‘lack’ is obvious: They both tend to occur with negative collocates but they reverse their evaluative meanings. Typical examples are ‘cure cancer’ or ‘a lack of support’. As discussed in chapters 2 and 6, many authors would consider the overall stance taken in the entire utterance as the relevant evaluation (e.g. Partington (2004)). Theoretical considerations aside, this explains why the correlation between semantic prosody and elicited valence ratings is not higher.

Guo et al. (2011) explicitly address the question of conscious or unconscious knowledge of semantic prosody in their study on the effect of incidental versus intentional learning on acquisition. The researchers retrieved sentences for six English words from different corpora. The six words were then replaced with pseudo-words in order to eliminate any possible influences of previous knowledge (e.g. the word ‘enhance’ was replaced with the pseudo-word ‘briten’). All study participants were native speakers of Chinese who were highly competent second language learners of English. The

---

Furthermore, Smith (2001) compared token and bigram counts, but these studies typically investigate productivity, which is a different topic from what Ellis and Frey (2009) are studying.
study consisted of a training phase and a test phase. In the training phase of the experiment, the incidental learning group was assigned to the ‘read’ condition, where they were asked to read and try to comprehend the sentences. The intentional learning group was in the ‘rule search’ condition where they were asked to determine the rules governing the use of the pseudo-words. A third group that served as a control group did not participate in the training phase.

In the test phase, acceptability judgments were elicited for new sentences, half of which matched and half of which mismatched the semantic prosodies established in the training phase. Additionally, participants rated their confidence in their own judgment and attributed their answers to either implicit knowledge (‘guessing’, ‘intuition’) or to explicit knowledge (‘memory’, ‘rule’). Overall, the ‘rule search’ group performed the best in the judgment task, followed by the ‘read’ and then the control group. Additionally, the ‘rule search’ group was more confident in their correct answers than the ‘read’ group. There was no reliable difference between the two experimental groups concerning the self-reported knowledge attribution (implicit vs explicit). The authors conclude that both explicit and implicit knowledge of semantic prosodies are acquired through incidental learning, but that intentional learning is more beneficial for conscious acquisition. On a broader level, these results indicate that speakers may possess explicit knowledge of semantic prosody, even if they acquired this knowledge through incidental learning (i.e. through normal language exposure).
McGee (2012) primarily investigates knowledge of semantic prosody in L2 learners of English. What makes this paper relevant in this context however is the inclusion of a native speaker group. Like Nordquist (2004), McGee (2012) tests knowledge of semantic prosody by asking the test subjects to use words in context. He presents them with the following lexical items: ‘bring about’ (described by Xiao and McEnery, 2006, Louw (2000)), ‘cause’ as a verb (introduced by Stubbs, 1995a), ‘completely’ (Partington, 2004), ‘face’ as a verb (Tognini-Bonelli, 2001), ‘potentially’ (Stubbs, 2001), ‘provide’ (Stubbs, 1995b) and ‘regime’ (Channell, 2000). Participants were then asked to write down up to “three dictionary-style example sentences [...] which would exemplify the word or phrase’s use in its typical lexical and grammatical environment” (McGee, 2012, p. 173). Unlike Nordquist’s (2004) design, this task thus emphasizes that participants should use words within typical contexts. The author analyzed the sentiments expressed in the sentences and categorized them as either positive, negative or neutral. The results for the native speaker group are given in 7.1.

<table>
<thead>
<tr>
<th>Lexical item</th>
<th>Polarity in corpora</th>
<th>Positive</th>
<th>Negative</th>
<th>Neutral</th>
</tr>
</thead>
<tbody>
<tr>
<td>bring about</td>
<td>positive</td>
<td>82%</td>
<td>13%</td>
<td>5%</td>
</tr>
<tr>
<td>cause</td>
<td>negative</td>
<td>6%</td>
<td>84%</td>
<td>10%</td>
</tr>
<tr>
<td>completely</td>
<td>negative</td>
<td>31%</td>
<td>57%</td>
<td>12%</td>
</tr>
<tr>
<td>face</td>
<td>negative</td>
<td>0%</td>
<td>94%</td>
<td>6%</td>
</tr>
<tr>
<td>potentially</td>
<td>negative</td>
<td>29%</td>
<td>71%</td>
<td>0%</td>
</tr>
<tr>
<td>Lexical item</td>
<td>Polarity in corpora</td>
<td>Positive</td>
<td>Negative</td>
<td>Neutral</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------</td>
<td>----------</td>
<td>----------</td>
<td>---------</td>
</tr>
<tr>
<td>provide</td>
<td>positive</td>
<td>86%</td>
<td>0%</td>
<td>14%</td>
</tr>
<tr>
<td>regime</td>
<td>negative</td>
<td>5%</td>
<td>57%</td>
<td>39%</td>
</tr>
</tbody>
</table>

Table 7.1: Evaluative polarity for sentences produced by speakers in McGee (2012).

The experimental results generally match the semantic prosodies that were established in previous corpus studies. A severe limitation is that the corpus data are only reported as either positive or negative, making a fine-grained comparison between corpus data and elicited data impossible.

Hauser and Schwarz (2016) conduct a series of experiments into the inferences speakers derive from semantic prosody, both in term of judgments they make and in their production of language. First, they test whether knowledge of semantic prosody would affect the interpretation of ambiguous sentences. Study participants were presented with the sentence “Surprisingly, ingestion of the substance caused/produced endocrination of abdominal lipid tissue” with each group only seeing either the word ‘caused’ or ‘produced. They were then asked to decide whether “endocrination of abdominal lipid tissue” was a good or a bad thing. The sentence was ambiguous in the sense that it was not actually meaningful because the word ‘endocrination’ was made up. Participants were significantly more likely to think that it was was a bad thing when it was caused (72.9%) than when it was produced (48.5%). In a
follow-up study, the authors determine that semantic prosody does not significantly influence judgments of positive or negative evaluation, when the described state is either unambiguously positive or negative: There is no difference between ‘cause’ and ‘produce’ when ‘shrinking cancer tumors’ (in the positive condition) or ‘gall bladder infections’ (in the positive condition) are evaluated.

In another study, the authors investigated the effect of varying the agent of the causation. Participants were given a sentence completion test in the form “[Actor] caused/produced ___.” The actor could be positive (e.g. ‘the aid workers’), negative (‘the workers’) or neutral (‘the terrorists’). Based on word valence norms from Warriner et al. (2013), Hauser and Schwarz (2016) assigned valence ratings from 1 (very negative) to 9 (very positive) to the words produced by the participants. They found significant main effects for the verb and actor. There was also a significant interaction between actor and verb, showing that semantic prosody had a stronger effect if the actor was neutral or ambiguous. Hauser and Schwarz (2016) also show that semantic prosody influenced inferences speakers/listeners make about actors (and not just about outcomes). A political representative who ‘produced additional work for middle class families in his district’ is judged more favorably than a representative who ‘caused additional work for middle class families in his district’.

Hauser and Schwarz (2018) conduct several more experiments in the same vein, although this time they do not include production studies. In one study, they asked participants to rate a person who is described as either ‘an utterly changed man’ or ‘a
totally changed man’ along several dimensions. ‘Utterly’ and ‘totally’ were rated as synonyms in a pilot study but differ in their semantic prosodies, with utterly having a stronger negative tendency. Compared to the ‘utterly’ condition, the man described in the ‘totally’ condition was rated higher in warmth and competence, and his change was judged to be more likely to be positive. In a similar study, participants are asked to rate a new boss who was described as either utterly or totally unconventional. The totally unconventional boss is rated as more favorable, but this effect is smaller than in the previous study, again indicating that the influence of semantic prosody on valence is greater for semantically underspecified utterances.

In another study, participants rated online dating profiles that included either prosodically charged word or neutral synonyms, e.g. “I commit/engage in what I do” in the negative condition and “My friends say I attain/get what I want” in the positive condition. This study exemplifies an issue with much psycholinguistic research into semantic prosody: It is very difficult to design good stimuli because there are numerous variables that are difficult to match between the different experimental conditions. Ideally, the independent variable (in this case semantic prosody) should be the only difference between the two experimental conditions; all other variables should be tightly controlled. In this experiment (and to a lesser degree in other studies as well) the pairs clearly differ along other dimensions as well. For example, ‘get what pronoun want’ is an idiomatic sequence and is attested 1772 times in COCA; ‘attain what pronoun want’ is attested five times. This could mean that the former phrase
is perceived as more felicitous than the latter, which could skew the results. The ‘commit’ versus ‘engage in’ pair is problematic as well. It is not clear if the stimulus is ‘commit’ or ‘commit to’. If the former is true, the sentence seems ungrammatical or at least infelicitous. If the latter is the case, the semantic prosody is nowhere near as negative as they think: While ‘commit NP’ has an overwhelmingly negative semantic prosody (being strongly associated with committing crimes), ‘commit to NP’ is positive (e.g. ‘we are committed to providing the best service to our customers’).

There are more examples: One prompt reads “People say I tend to escalate/heighten any situation.” Setting aside the issue of felicitousness again (‘escalate the situation’ is by the far more common of the two), it seems strange that the two verbs are supposed to be matched in their semantics (i.e. their meaning in isolation, without considering semantic prosody). ‘Escalate’ is transparently negative to the point that this is captured in dictionary definitions. It is understandably difficult to find near-synonyms that only differ in their semantic prosody, especially if we except the idea that a semantic transfer occurs from the habitual context to the node word. This also shows, again, that the distinction between the stable, core meaning of a word and its contextual use is highly problematic.

While the description of their stimulus design is somewhat unclear, it seems that Hauser and Schwarz (2018) drew from different sources and did not pre-test all their stimuli in pilot studies. Unfortunately, the examples described here are not exceptions. Almost all their stimulus pairs in this experiment are problematic, falling into
one the three categories described here: (a) one stimulus results in a much more common and arguably more felicitous utterance, (b) the semantic prosody of a word is not what the authors think it is (c) a word carries an obvious negative or positive evaluation. Despite this, or in some cases perhaps because of this, Hauser and Schwarz (2018) find the expected effect of semantic prosody on favorability ratings.

In summary, the most fundamental finding of these experimental studies is that speakers possess an intuitive understanding of the prosodic evaluative semantics of lexical items; they are clearly sensitive to the types of evaluations words tend to occur in, both in language production and language comprehension. As Nordquist (2004) and McGee (2012) show, when prompted to use prosodically primed items to create or complete sentences, speakers do so in ways that match the items’ distributions in corpora. Hauser and Schwarz (2016) and Hauser and Schwarz (2018) show that speakers additionally use knowledge of semantic prosody to make inferences about the sentiments expressed in statements, leading them to make evaluative judgments about both actions/events and the involved actors. They also show that the evaluations carried by semantic prosody are subtle and can be overruled or at least backgrounded by more explicit evaluative information. Semantic prosody is particularly relevant in utterances that are otherwise underspecified in their evaluative semantics. These results show that semantic prosody does not just surface in language usage in the form of patterns of co-occurrence in corpora but that it is also stored in the long-term memory in some form.
Most of the studies discussed here confirm the ideas about speakers’ intuitive and implicit understanding of semantic prosody proposed in the previous section, but they provide little evidence about the existence of explicit knowledge. Only two studies provide some insights into explicit knowledge. Guo et al. (2011) demonstrate the acquisition of semantic prosodies in an experimental setting. They conclude that even linguistic knowledge that has been acquired through incidental learning can be partially conscious. This indicates that knowledge of semantic prosody may, in fact, be explicit. A methodological caveat, however, is that this claim is solely based on participants’ self-reported knowledge attribution. Ellis and Frey (2009) provide data from a supplemental study that asked participants to rate the evaluative valences of their verbal primes. Even though there are some methodological issues, their data show a correlation between elicited data and corpus data.

7.3 Explicit knowledge of semantic prosody

The studies discussed so far provide important insights into speakers’ knowledge of semantic prosody, but they mostly deal with implicit knowledge. They do not provide any definitive evidence as to whether semantic prosody is retrievable in the form of explicit knowledge. As described in the first section of this chapter, researchers who talk about the covert nature of semantic prosody typically deal with the latter not the former: They mean that it is inaccessible to introspection, not that there is a lack of implicit knowledge.
In order to test this idea, I conducted a survey where subjects were directly asked to explicitly access their knowledge of semantic prosody. In the context of the framework laid out in chapters 5 and 6, this is, to some extent, asking the wrong question. However, the aim of this study is to test if the claims found in the prior research literature are accurate, which requires investigating them on their own terms.

7.3.1 Methodology

The aim of the study was to investigate whether speakers are able to access their knowledge of semantic prosodies. 17 native speakers of American English who were undergraduate students at Rice University and enrolled in an introductory-level linguistics class participated in the study for a small amount of extra credit. The study design was simple: Participants were presented with a total of 55 linguistic items (both words and larger units) and asked to decide whether these items tended to occur in positive or negative evaluative contexts. They rated the items on a Likert scale from 1 to 7, 1 meaning maximally negative, 4 being completely neutral and 7 being maximally positive. All items were taken from the research literature on semantic prosody. The items were presented one at a time, and a forced delay of four seconds was used to encourage the participants to take a moment to think about the semantic prosody. Beyond this, they could take as much or as little additional time as they wanted. After a one-second delay, the next item was presented. Every participant was given an individually randomized list to avoid any potential order effects. Verbs
were indicated with the *to-infinitive* form (e.g. ‘to cause’). The study was conducted on a computer and was programmed in the experiment builder software *OpenSesame* (Mathôt et al., 2012).

How the evaluations contained in corpus data should be quantified in order to compare it to the elicited ratings is a nontrivial question. First, a pretheoretical issue should be addressed: When comparing corpus data and elicited data, researchers from different academic disciplines often show a bias as to which of the two they consider as having the primacy over the other, based on their own field of study. When there is a mismatch between the two, corpus linguists tend to assume that the intuitions are faulty, while psychologists take issue with the design and representativeness of the corpus. Instead of asking to what extent speakers possess knowledge of semantic prosody, the question should perhaps be framed differently: Can we find a way to model speakers’ ratings with corpus data?

It is not clear how speakers arrive at their verdicts and it may depend on the way the question is asked. In this experiment, participants are not explicitly asked if the given item itself was positive or negative; instead they are asked about the utterances it tended to occur in. They could use different strategies to arrive at their judgments – they may try to think of typical usage instances, e.g. what kinds of sentences the items tend to occur in or what their most common collocates are. Alternatively, they simply follow their ‘gut feelings’ without introspectively going over typical uses. For this reason, I created three corpus measures against which the elicited data were
compared. The different measures were retrieved for a portion of the data in order to test which metric best models the elicited data. The best score was then used for the entire data set.

The first measure is a valence score of collexeme types. I first determined the most common construction a word occurred in. For the word ‘happen’, for example, this is the intransitive construction. Other types were excluded from the analysis, e.g. the pattern ‘happen to be’ (‘I happen to be a big fan of his music’). I conducted collexeme analyses for all words to determine their entire collexeme profiles. All collexemes that showed collostructional attraction to the stimulus item were included in the profile. Words that showed collostructional repulsion (e.g. words that only occurred very few times in these contexts and that were not specific enough to the construction) were not included. I then classified all collexemes as either positive, neutral or negative. This classification was only based on the sequence of the two collexemes in isolation, e.g. ‘change happened’ would be regarded as neutral even though the entire utterances may show tendencies towards negative or positive evaluations. The valence score is then calculated as follows: Positive collexemes are assigned a value of 1, negative collexemes a value of -1 and neutral collexemes a value of 0. These scores are added up and divided by the total number of collexeme types. For example, a lexeme with 140 positive, 205 negative and 150 neutral collexeme types would be assigned a value of \((140-205)/(140+205+150) = -65/495 = -.13\). The resulting measure ranges from -1 to 1, -1 being entirely negative and 1 being entirely positive.
The second measure is the valence score of collexeme tokens. The first analytical steps are identical to the type-based valence score. All collexemes for the most common constructional context of a given lexical item are retrieved and classified as positive, negative and neutral. The only difference is that token frequencies are taken into account. For each collexeme, their evaluative polarity (encoded as 1, -1 or 0) is multiplied by their token frequencies. The products are added up and divided by the sum of all collexeme tokens. The resulting token valence score also ranges from -1 to 1. The third measure is based on the evaluations found within entire utterances. For each item, a random 200-word sample was retrieved. The sentiments expressed by the sentences were analyzed and categorized as positive, negative and neutral and assigned scores of 1, -1 and 0 respectively. These values were summed and divided by the total number of sentences, again leading to an overall score ranging from -1 to 1.

A valence score was also created for the experimental data. Responses were originally collected on a scale from 1 to 7. For each item, the mean was calculated across all subjects’ responses. These values were then centered around zero and normalized to range from -1 (indicating a maximally negative rating) to 1 (completely positive). The resulting score is directly commensurable with all the corpus-based measures. For each corpus metric, a simple linear regression analysis was performed to test how well it predicted the elicited data.
7.3.2 Results

Inter-rater reliability

Before any comparisons are made to a corpus as a point of reference, a question that could be asked is how much speakers agree in their judgments. Assessing agreement between raters is far from trivial; various statistical indices are available but there is no definite answer as to what constitutes good (or sufficient) agreement. A commonly used measure of inter-rater reliability (or inter-rater agreement) is Krippendorff’s alpha. An analysis of the dataset yields a value of alpha = .513.

The meaning of this value is not a clear-cut mathematical question but rather a matter of interpretation that depends on a few conceptual questions, such as the purpose of the data. A value of .513 for Krippendorff’s alpha or similar indices of inter-rater reliability would typically be considered to indicate “moderate agreement” (Landis and Koch, 1977). What ‘moderate agreement’ means really depends on the question we are asking. Concerning the assessment of intercoder reliability, Lombard et al. (2002, p. 593) acknowledge that there are no “established standards”, but state that a coefficient of .80 is acceptable in most situations, while the .70 criterion is often found in exploratory research. In principle, intercoder reliability refers to the same basic concept as inter-rater reliability but has a somewhat different goal: It informs us how much we can trust data that have been ‘filtered’ through the (perhaps somewhat subjective) content analysis of human coders (Lombard et al., 2002, p.589).

In other words, if we wanted to operationalize the semantic prosody of a given
word as its score on a Likert scale as rated by a group of native speakers, we would have to conclude that our data are not sufficiently reliable. Basing a study of semantic prosody on the intuitions of speakers would not be a good methodology. But this is of course not the aim of the survey. The question is whether or not speakers have any explicit knowledge of semantic prosodies at all.

**Corpus data and elicited ratings**

Retrieving and analyzing the corpus data required to create all three corpus measures for 55 items is not feasible. Therefore, the performance of the three scores was tested on a smaller portion of the data with the aim to identify the measure that best predicted the elicited data. This measure was then selected to be applied to the entire dat. The smaller portion is based on two sets of words found in Partington (2004). Partington uses semantic prosody to distinguish between near-synonyms. He compares two groups of near-synonyms, ‘happen words’ and amplifying intensifiers.

The so-called happen words are among both the earliest and the most studied examples of semantic prosody. Sinclair (1987) described the tendency of ‘set in’ to co-occur with unfavorable things; Sinclair (1991, p. 112) describes how ‘happen’ is “associated with unpleasant things.” In addition to these two words, Partington (2004) also include the related words ‘occur’, ‘come about’ and ‘take place’. He concludes that “there are different degrees of bad prosody among the items in the group; ‘set in’ has the worst prosody, followed by ‘happen’, followed by ‘occur' and
‘take place’, while ‘come about’ does not seem to evince any particular inclination” (Partington, 2004, p. 144).

The data for all metrics of the happen words are presented in Table 7.2. The column labeled Judgments contains the valence score assigned by speakers. The next three columns contain the three different corpus-based valence scores. The Type column contains the valence scores for collexeme types, the Token column contains the token-based scores and the Utterance column represents the sentiment score of the entire sentence.

<table>
<thead>
<tr>
<th>Item</th>
<th>Judgments</th>
<th>Type</th>
<th>Token</th>
<th>Utterance</th>
</tr>
</thead>
<tbody>
<tr>
<td>happen</td>
<td>0.04</td>
<td>-0.22</td>
<td>-0.11</td>
<td>-0.43</td>
</tr>
<tr>
<td>take place</td>
<td>-0.02</td>
<td>-0.16</td>
<td>-0.14</td>
<td>-0.10</td>
</tr>
<tr>
<td>occur</td>
<td>-0.06</td>
<td>-0.28</td>
<td>-0.32</td>
<td>-0.38</td>
</tr>
<tr>
<td>set in</td>
<td>-0.12</td>
<td>-0.78</td>
<td>-0.58</td>
<td>-0.71</td>
</tr>
<tr>
<td>come about</td>
<td>0.12</td>
<td>0.06</td>
<td>0.13</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Table 7.2 : Speakers’ judgments and three different corpus measures: ‘Happen words’, Partington (2004)

For each corpus metric, a simple linear regression analysis was performed. The type-based valence score significantly predicted speakers’ ratings, $R^2 = .78, p = .047$. The token-based valence score was an even stronger predictor of speakers’ ratings, $R^2 = .941, p = .006$. The valence score based on the utterance-level sentiment failed to
reach significance at $\alpha = .05$ with $R^2 = .639$, $p = .105$. Figure 7.1 shows a scatterplot with a regression line for the strongest predictor, the token valence score, and speakers’ ratings. Two points need to be made here. First, the number of observations is very small, which, in part, explains why the coefficients of determination are relatively high. Second, valences scores are mostly higher for the elicited ratings than the corpus data. In other words, speakers appear to underestimate the degree to which the items are negative (assuming that a direct comparison between the magnitudes of elicited data and corpus data is valid). In the linear regression model, this is reflected in the intercept, $b = .065$ for the token-based valence score. No penalty is applied for the inclusion of the intercept.
Another set of words described in Partington (2004) are the ‘amplifying intensifiers’ absolutely, perfectly, entirely, completely, thoroughly, totally and utterly. Partington (2004) only establishes distinct semantic prosodies for utterly (negative) and perfectly (positive). The valence score for all three corpus metrics and the speakers’ ratings are shown in Table 7.3.

<table>
<thead>
<tr>
<th>Item</th>
<th>Judgments</th>
<th>Type</th>
<th>Token</th>
<th>Utterance</th>
</tr>
</thead>
<tbody>
<tr>
<td>absolutely</td>
<td>0.31</td>
<td>0.15</td>
<td>0.53</td>
<td>-0.09</td>
</tr>
<tr>
<td>perfectly</td>
<td>0.78</td>
<td>0.79</td>
<td>0.79</td>
<td>0.73</td>
</tr>
<tr>
<td>entirely</td>
<td>0.02</td>
<td>0.40</td>
<td>0.18</td>
<td>-0.27</td>
</tr>
<tr>
<td>completely</td>
<td>0.14</td>
<td>-0.04</td>
<td>0.01</td>
<td>-0.27</td>
</tr>
<tr>
<td>thoroughly</td>
<td>0.27</td>
<td>0.30</td>
<td>0.34</td>
<td>0.35</td>
</tr>
<tr>
<td>totally</td>
<td>0.33</td>
<td>0.06</td>
<td>-0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>utterly</td>
<td>-0.31</td>
<td>-0.37</td>
<td>-0.46</td>
<td>-0.24</td>
</tr>
</tbody>
</table>

Table 7.3: Speakers’ judgments and three different corpus measures: ‘Amplifying intensifiers’, Partington (2004)

As with the happen words, a simple linear regression analysis was performed for each corpus measure to compare it to the elicited data. The type-based and the utterance-level valence scores both reached significance at the .05 level and performed almost equally well, with $R^2 = .67$, $p = .024$ and $R^2 = .683$, $p = .022$ respectively.
Again, the token-based valence score was the strongest predictor of speakers’ ratings, $R^2 = .766$, $p = .01$. The relationship is visualized in a scatterplot in figure 7.2.

Based on this relatively small sample of the data, the token-based valence score was adopted as the measure apply to the entire data set. The valence scores based on speakers’ ratings and on token frequencies are given in table 7.4. A simple linear regression analysis finds significant relationship, $R^2 = .61$, $p < .001$, indicating a moderately strong relationship (cf. figure 7.3 for the plot).
<table>
<thead>
<tr>
<th>Item</th>
<th>Judgments</th>
<th>Token</th>
</tr>
</thead>
<tbody>
<tr>
<td>(to) abound</td>
<td>0.24</td>
<td>0.1</td>
</tr>
<tr>
<td>(to) advocate</td>
<td>0.55</td>
<td>0.24</td>
</tr>
<tr>
<td>(to) affect</td>
<td>-0.2</td>
<td>-0.57</td>
</tr>
<tr>
<td>(to) bring about</td>
<td>0.14</td>
<td>0.08</td>
</tr>
<tr>
<td>(to) cause</td>
<td>-0.24</td>
<td>-0.84</td>
</tr>
<tr>
<td>(to) come about</td>
<td>0.12</td>
<td>0.13</td>
</tr>
<tr>
<td>(to) commence</td>
<td>0.31</td>
<td>0.01</td>
</tr>
<tr>
<td>(to) commit</td>
<td>-0.02</td>
<td>-0.94</td>
</tr>
<tr>
<td>(to) create</td>
<td>0.43</td>
<td>0.24</td>
</tr>
<tr>
<td>(to) end up</td>
<td>-0.35</td>
<td>-0.78</td>
</tr>
<tr>
<td>(to) gain</td>
<td>0.51</td>
<td>0.72</td>
</tr>
<tr>
<td>(to) happen</td>
<td>0.04</td>
<td>-0.11</td>
</tr>
<tr>
<td>(to) occur</td>
<td>-0.06</td>
<td>-0.32</td>
</tr>
<tr>
<td>(to) peddle</td>
<td>-0.22</td>
<td>-0.25</td>
</tr>
<tr>
<td>(to) prevail</td>
<td>0.69</td>
<td>0.29</td>
</tr>
<tr>
<td>(to) provide</td>
<td>0.47</td>
<td>0.65</td>
</tr>
<tr>
<td>(to) restore</td>
<td>0.67</td>
<td>0.93</td>
</tr>
<tr>
<td>(to) set in</td>
<td>-0.12</td>
<td>-0.58</td>
</tr>
<tr>
<td>(to) sit through</td>
<td>-0.45</td>
<td>-0.13</td>
</tr>
<tr>
<td>(to) take place</td>
<td>-0.02</td>
<td>-0.14</td>
</tr>
<tr>
<td></td>
<td>Value1</td>
<td>Value2</td>
</tr>
<tr>
<td>--------------------------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>(to) undergo</td>
<td>-0.35</td>
<td>-0.33</td>
</tr>
<tr>
<td>absolutely</td>
<td>0.31</td>
<td>0.53</td>
</tr>
<tr>
<td>bordering on . . .</td>
<td>-0.39</td>
<td>-0.67</td>
</tr>
<tr>
<td>brimming with . . .</td>
<td>0.35</td>
<td>0.45</td>
</tr>
<tr>
<td>bustling with . . .</td>
<td>0.33</td>
<td>0.29</td>
</tr>
<tr>
<td>buzzing with . . .</td>
<td>0.35</td>
<td>0.34</td>
</tr>
<tr>
<td>completely</td>
<td>0.14</td>
<td>0.01</td>
</tr>
<tr>
<td>consequence</td>
<td>-0.67</td>
<td>-0.3</td>
</tr>
<tr>
<td>crawling with . . .</td>
<td>-0.76</td>
<td>-0.68</td>
</tr>
<tr>
<td>dealings</td>
<td>-0.41</td>
<td>-0.28</td>
</tr>
<tr>
<td>effect</td>
<td>-0.02</td>
<td>-0.13</td>
</tr>
<tr>
<td>entirely</td>
<td>0.02</td>
<td>0.18</td>
</tr>
<tr>
<td>ever-present</td>
<td>0.1</td>
<td>-0.38</td>
</tr>
<tr>
<td>fraught with</td>
<td>-0.69</td>
<td>-0.84</td>
</tr>
<tr>
<td>impressive</td>
<td>0.84</td>
<td>0.93</td>
</tr>
<tr>
<td>incredibly</td>
<td>0.51</td>
<td>0.19</td>
</tr>
<tr>
<td>largely</td>
<td>-0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>omnipresent</td>
<td>0.25</td>
<td>-0.12</td>
</tr>
<tr>
<td>perfectly</td>
<td>0.78</td>
<td>0.79</td>
</tr>
<tr>
<td>persistent</td>
<td>0.14</td>
<td>-0.7</td>
</tr>
<tr>
<td>pervasive</td>
<td>-0.49</td>
<td>-0.34</td>
</tr>
</tbody>
</table>
Table 7.4: Speakers’ judgments compared to token frequencies.

<table>
<thead>
<tr>
<th>Expression</th>
<th>Score1</th>
<th>Score2</th>
</tr>
</thead>
<tbody>
<tr>
<td>rife with ...</td>
<td>-0.49</td>
<td>-0.61</td>
</tr>
<tr>
<td>sheer</td>
<td>0.02</td>
<td>0.14</td>
</tr>
<tr>
<td>signs of ...</td>
<td>-0.37</td>
<td>-0.35</td>
</tr>
<tr>
<td>somewhat</td>
<td>-0.24</td>
<td>-0.09</td>
</tr>
<tr>
<td>swarming with</td>
<td>-0.71</td>
<td>-0.3</td>
</tr>
<tr>
<td>teeming with</td>
<td>-0.22</td>
<td>0.28</td>
</tr>
<tr>
<td>tenacious</td>
<td>0.16</td>
<td>0.06</td>
</tr>
<tr>
<td>thoroughly</td>
<td>0.27</td>
<td>-0.02</td>
</tr>
<tr>
<td>to the point of ..</td>
<td>-0.41</td>
<td>-0.86</td>
</tr>
<tr>
<td>totally</td>
<td>0.33</td>
<td>-0.02</td>
</tr>
<tr>
<td>tremendous</td>
<td>0.18</td>
<td>0.16</td>
</tr>
<tr>
<td>ubiquitous</td>
<td>0.16</td>
<td>-0.04</td>
</tr>
<tr>
<td>universal</td>
<td>0.37</td>
<td>0.19</td>
</tr>
<tr>
<td>utterly</td>
<td>-0.31</td>
<td>-0.46</td>
</tr>
</tbody>
</table>

7.3.3 Discussion

How do speakers arrive at their ratings of lexical items’ evaluative valences? In this study, the best predictor was a score based on the token frequencies of their collexemes. One important takeaway of this finding is that the heuristics speakers employ
Figure 7.3: Correlation between corpus data and elicited judgments for all items.

appear to concern relatively local relationships. The overall sentiment, expressed in
the larger utterance, plays a less important role than the proximal sentiment con-
veyed through the lexical item and its collexemes. It appears that speakers go over
typical uses of lexical items through retrieving their collocates or collexemes and are
able to generalize across their evaluative semantics. This is not surprising, given the
prior research on collocational knowledge. Additionally, the fact that the token-based
score was a better predictor than the type-based score shows that the frequency of
the collocates/collexemes matters. Again, this seems reasonable, given that much
research shows that highly frequent words are privileged in their retrieval (cf. chapter
3).
A caveat is that the results may also depend on the elicitation technique. In this study, participants were asked to rate whether the items tended to occur in negative or positive contexts. Additionally, a four-second pause was introduced between the presentation of the stimulus and the rating screen. This was done to encourage participants to take some time to inquire their intuitions, and not hastily categorize words that lacked immediately obvious evaluative semantics as neutral. The idea was not to test whether speakers have an instantaneous understanding of semantic prosody, but if they can retrieve any explicit knowledge of it at all.

Another important point and a serious limitation of this study is that the corpus metrics for type and token scores were solely based on the most common construction a given linguistic item occurred in. This is a simplification of the data that works well in most cases but not in all. It explains why for a few lexical items there was a clear mismatch between corpus data and elicited ratings. One example of this is ‘persistent’. Type and token valence scores were retrieved for its most common usage, the attributive use in the ‘persistent NP’ construction, which resulted in overwhelmingly negative scores. Speakers, on the other hand, rated it as moderately positive, suggesting that they mostly considered it in its predicative form. Another example is ‘commit’. The construction ‘commit NP’ (e.g. ‘commit suicide’) is approximately twice as frequent as ‘commit to NP’ (e.g. ‘commit to peace’). Despite this, on average participants rated it as almost completely neutral with only a very slight negative tendency. The ratings for ‘commit’ were also among the most varied across subjects,
including both strongly negative and positive ratings. This indicates that speakers likely differed as to which sense of the word they were thinking of. As with ‘persistent’, it is not clear why speakers did not show a stronger tendency towards the much more common usage of the word.\footnote{If these two items are removed from the regression analysis, the coefficient of determination improves to $R^2 = .69$.}

For a few lexical items, however, a mismatch between corpus-based scores and elicited judgments cannot be blamed on constructional differences. ‘Cause’, for example, was rated as only moderately negative while its token valence score is overwhelmingly negative. Similarly, speakers rated ‘prevail’ as strongly positive (a result that happens to match my own intuitions) but the corpus data shows that it also commonly occurs with neutral collexemes, e.g. *condition, view, attitude* and *situation*. These examples suggest that the salience of the collocates/collexemes is not solely driven by frequency, but that more complex heuristics are at play. Language is full of idiosyncrasies and perhaps some linguistic items possess particularities that lead otherwise useful heuristic processes astray. This, however, is an issue that goes well beyond the research question at hand and will as such not be further explored.

Despite these issues, the overall message is clear: Speakers have a better handle of semantic prosodies than has been assumed in the prior literature. The token-based valence score accounted for 62% of all variance of speakers’ ratings. For a very simple single predictor, this is not a bad value, especially considering that the items con-
sisted of examples from the literature, most of which had previously been described as hidden from intuition. For some items, there was a mismatch between corpus and intuitions, but this does not necessarily mean that speakers lack intuitions. Other explanations are possible: Maybe ratings should have been elicited differently. Alternatively, a different operational definition of corpus scores with more sophisticated feature engineering could have led to closer alignment.

### 7.4 Discussion

The idea that semantic prosody is hidden from speakers’ intuitions is a claim so commonly repeated that it has become conventional wisdom. In this chapter, I set out to investigate if this idea holds up under empirical scrutiny.

In section 7.1, I argued that there are several issues with the typical descriptions of semantic prosody as hidden from intuition. In summary, little evidence has been put forward to support this claim directly, and the assumption that speakers generally lack knowledge of frequency-based phenomena is contradicted by psycholinguistic studies. Additionally, there are definitional issues, both concerning the conflation of introspection and intuition and the need for a distinction between implicit and explicit knowledge. Finally, semantic prosody is often not hidden; at best it is invisible in a few rare cases. Section 7.2 reviewed the psycholinguistic research literature on semantic prosody and discusses the ramifications for the notion of hiddenness. The findings corroborate the idea that competent speakers possess an intuitive, implicit
understanding of semantic prosody that they skillfully employ both in language production and comprehension. Relatively little can be said about explicit knowledge of semantic prosody.

In section 7.3, I explored how elicited ratings of semantic prosody compare with corpus data. The results showed that speakers generally possess better explicit knowledge of semantic prosody than assumed in the prior literature. A valence score based on relative token frequencies of positive, neutral or negative collexemes was proposed to model the elicited ratings. It seems plausible to me that this score approximates the psychological strategies speakers use to arrive at their ratings, i.e. they may use collocational knowledge to estimate the typical evaluative contexts of lexical items. In some cases, e.g. ‘cause’, speakers underestimated the degree of its negative semantic prosody, but they correctly identify its negative tendency. It should also be noted that it is not clear, how closely corpus data and ratings should line up for individual data points in order to make a decision whether the semantic prosody for this specific item is hidden or not. This is not a question that can be solved by a statistical test.

In conclusion, the claim that semantic prosody is hidden from speakers’ intuitions cannot be maintained, at least not in its commonly stated form. Speakers possess a relatively good explicit understanding of the kinds of evaluative contexts linguistic items tend to occur in. Only for a few words, no close alignment between corpus data and elicited data was found. In some cases, this is very likely due to methodological issues (e.g. ‘commit’ and ‘persistent’). In other cases, it is not clear whether this is
due to methodological issues or due to a true lack of explicit awareness. By and large, though, the degree of knowledge of semantic prosody seems to be about what should be expected of frequency-based language phenomenon. As such, when it comes to speakers' intuitions and introspective knowledge of semantic prosody, there is nothing 'special' about it that makes it stand out from the rest of language.
Chapter 8

Semantic Prosody from a Cross-Linguistic Perspective

In this chapter, I examine semantic prosody from a cross-linguistic viewpoint. The goal is to determine whether or not the findings described in the previous chapters are limited to English. Since I consider semantic prosody to be a normal component of context-dependent meaning, it seems likely that the constructional effects described in chapter 5 are not specific to the English language. If language structure universally emerges in a bottom-up manner as hypothesized in usage-based approaches, we should expect context-dependent schematizations to be an important locus of meaning in other languages as well. This hypothesis will be put to the test in a contrastive study of German and English.

In section 8.1, I discuss prior contrastive studies on semantic prosody. The great majority of research into semantic prosody has been published on English; in other languages, it remains a relatively understudied phenomenon. Almost all work on languages other than English is done from a cross-linguistic perspective, virtually always contrasting that language with English. In this section, I first describe the prior literature on German and English, which is very sparse. Then I provide a summary of cross-linguistic research on semantic prosody in other languages.
In section 8.2, I discuss two original studies contrasting semantic prosody in German and English in a parallel corpus. In the first study, I investigate German equivalents of English verbal ‘cause’ in its various constructional contexts. In the second study, I examine the German intensifier *vollkommen* and its English analogs. Beyond the proximal motivation to explore semantic prosody in German and investigate how similar or different it is to English, this research is guided by two more fundamental, theoretical questions. The first question is whether semantic prosody in German is also constructionally conditioned in a similar way as it is in English. In other words, do we also need to go beyond the level of individual words and consider their constructional context in our description of evaluative meaning? The second question concerns the scope of evaluative meaning of semantic prosody. Does semantic prosody in German go beyond a polarity of positive and negative evaluations, as I suggested it does in English?

Section 8.3 reviews the results, contextualizes them in within the framework laid out in the previous chapters and discusses their theoretical implications.

8.1 Previous Studies

8.1.1 Research on Semantic Prosody in German

There is a marked lack of studies on semantic prosody in German. To my best knowledge, this section, as brief as it is, constitutes a near-exhaustive overview of all the relevant original research. All studies described here are cross-linguistic; there
do not appear to be any monolingual studies on semantic prosody in German. The only articles or books that are not included here only make a very brief mention of semantic prosody in German but do not provide any data and do not conduct any original research.

Kenny (1998) argues that literary translators tend to use a ‘toned down’ vocabulary that lacks the prosodic implications of the original, resulting in ‘sanitized’ translations. She provides one example, a German translation of an English text where the word ‘giro’ is translated with *Scheckheft* (‘checkbook’). The author argues that ‘giro’ carries a negative evaluation and is an informal term for an unemployment benefit check. The German translation is entirely neutral and does not share this prosodic meaning. Following this line of research, Kenny (2001) and (2006) investigates the translation of creative language between German and English. While the author cautions against sweeping generalization based on her limited data sample, she finds that translations tend to be ‘normalized’, i.e. more conventional than the original and lacking the same semantic preferences and semantic prosodies.

In the introduction to a volume on corpora in German studies, Dodd (2006) mentions semantic prosody alongside several other topics and asks: “To what extent are semantic prosodies language-specific? Are they ‘universal’?” (p. 10). He provides some tentative data in the form of 50 KWIC lines for the German verb *verursachen* (‘cause’) and finds that it is similar to English ‘cause’, leading the author to conclude that this “suggests that at least some semantic prosodies do ‘translate’ across language
divides” (ibid).

In what appears to be the only cross-linguistic study of semantic prosody of any kind that does not include English as one of its languages, Oster and van Lawick (2008) examine translations between German and Spanish/Catalan and compare idiomatic phraseological units that are generally considered to be equivalent. They provide corpus data to show that semantic preference and prosody can vary cross-linguistically and need to be accounted for by translators. The authors conclude:

“For translators, semantic preference is important because it tells them something about the semantic fields that the collocates of a lexical item may come from. Semantic prosody imprints a specific argumentative direction on an utterance through positive or negative connotations. In both cases, failure to meet the target language conventions may lead to unintended awkwardness in the translation.” (Oster and van Lawick, 2008, p. 342)

Beyond this there is very little research into semantic prosody in German, be it from a mono-linguistic or a contrastive perspective. There is some further research on related topics: Some authors, for example, criticize that bilingual dictionaries are often not informed by word co-occurrence data without specifically mentioning semantic prosody. Siepmann (2005), for example, compares French, English and German bilingual dictionaries and argues that they are often not informed by distributional data, leading to entries that lack semantic-pragmatic precision. Similarly, there is re-
search on the mistakes non-native speakers make in the realm of lexico-grammar and style, but they also do not make reference to semantic prosody. For example, Lorenz (1998) describes the overuse of adjective intensification by advanced German learners of English, creating “an impression of ‘non-nativeness’ or ‘lack of idiomaticity’” (p. 53).

8.1.2 Other Contrastive Studies

Partington (1998) is an early example for an, albeit small, cross-linguistic study of semantic prosody. The author compares two-word pairs in Italian and English. ‘Impressive’ in English tends to collocate with words such as ‘achievement’ or ‘talent’, while Italian *impressionante* tends to co-occur with neutral or negative items, such as ‘a series of price rises’ or ‘assassination attempts’ (pp. 77-78). Similarly, English ‘incite’ is clearly negative, collocating with unfavorable words such as ‘hatred’ or ‘violence’, whereas Italian *incitare* is generally favorable, meaning something like ‘encourage’. Partington (1998, p. 77) concludes: “[C]ognate, or ’look-alike words’ from two related languages can have very different semantic prosodies.”

Berber Sardinha (2000) compares semantic prosody in English and Portuguese, finding both differences and similarities between the two languages. The author finds that English ‘commit’ and Portuguese *cometer* are similar in their semantic prosodies, as are ‘cause’ and *causar*, while ‘happen’ and *acontecer* exhibit different semantic prosodies. For ‘set in’ Berber Sardinha (2000) discusses several potential transla-
tions but finds no suitable equivalent in Portuguese that truly captures the meaning and collocational behavior of the English phrasal verb. The author summarized the findings as follows:

“The results indicate that semantic prosodies may vary across Portuguese and English. The present study, therefore, corroborates other contrastive studies which have identified discrepancies between seemingly equivalent items in different languages (e.g. Partington 2008).” (Berber Sardinha, 2000, p. 106)

Tognini-Bonelli (2001) describes a methodological approach to finding adequate translation through the use of corpora. Comparing English and Italian, the author finds that words or phrases with seemingly synonymous semantics may differ in their semantic prosodies and other distributional properties. For English ‘real’ it is difficult to find an Italian translation that matches its distribution in language use, and its *prima facie* translation *vero* is not a good option (pp. 144ff). Furthermore, semantic prosody can vary cross-linguistically in subtle and perhaps unexpected ways. The different versions of the English idiom ‘in (the) case (of)’ have different distributional properties and as such different meanings. For ‘in case’, the author posits a semi-ironical stance towards a hypothetical situation that is not shared by ‘in case of’ (p. 142). While the potential Italian translations are generally similar to English in their semantic prosodies, they do not exhibit the same variation. Tognini-Bonelli (2001, p. 144) concludes that this “could generate a trap for the unaware translator
because the correspondence is similar but not as systematic.”

Xiao and McEnery (2006) study semantic prosody in English and Chinese. They aim to answer two main questions: whether Chinese exhibits semantic prosody at all, and, if it does, how similar it is for lexical items with similar denotational meanings across the two languages. The authors summarize their findings as follows:

“Our contrastive analysis shows that semantic prosody and semantic preference are as observable in Chinese as they are in English. As the semantic prosodies of near-synonyms and the semantic preferences of their collocates are different, near-synonyms not interchangeable in either language.” (Xiao and McEnery, 2006, pp. 124f)

There is, however, more detail and complexity to these results. While the authors found that near-synonyms may not be directly interchangeable in terms of their semantic prosodies and semantic preferences, they also point to some systematic similarities between the two languages. In a study of various lexical items describing consequence, they find that the English words ‘outcome/result’, ‘consequence’ and ‘aftermath’ can be arranged on a continuum from positive to negative. A similar continuum exists for the corresponding Chinese terms. They also find a closely corresponding Chinese equivalent for English ‘cause’. Other terms, however, lack such direct interlinguistic synonyms. Overall, the authors emphasize the similarities in semantic prosody between the two languages, especially considering how different the two languages are: “While English and Chinese are distinctly unrelated, the collo-
cational behaviour and semantic prosodies of near synonyms are quite similar in the two languages” (Xiao and McEnery, 2006, pp. 125).

Munday (2011) studies semantic prosody in English and Spanish in the context of using corpora in translation studies. The author notes that some translations, such as English ‘loom large’ and Spanish *cernerse* may match in terms of their semantic prosodies (here defined in the sense of their evaluative polarity); in this case, both ‘loom large’ and *cernerse* tend to describe unfavorable things. In terms of other distributional qualities, such as semantic preference and the wider interpretation of context and stance, however, they clearly differ. ‘Looming large’ is, for example, often followed by a prepositional phrase such as *in his mind* that is used to express mental pressure or preoccupation; this use is absent from Spanish *cernerse*.

Volanschi (2012) contrast semantic prosody in English and French. Some of their findings display similar tendencies as Munday’s (2001) work on English and Spanish. Their comparison of English ‘commit’ and French *commettre* shows that while the two words are similar in terms of their evaluative polarities, they differ in their semantic preferences, e.g. French does not have an equivalent to English ‘to commit suicide’ but instead uses the reflexive verb *se suicider*. Additionally, French *commettre* can be used with *roman*, ‘novel’ and *oeuvre (d’art)*, ‘work (of art)’, using a prosodic clash to create irony and to insinuate that the novel or work of art is so bad that it is a crime. The authors point out that while ‘to commit a novel’ would be unidiomatic
in English, it is possible to use ‘perpetrate’ in this context.\textsuperscript{1} This, again, indicates that while it may be difficult to find cross-linguistic synonyms that match along all dimensions of meaning and use, similar structures and strategies do seem to exist across different languages. The authors also study the effect of genre on semantic prosody and find that, as in English, semantic prosodies in French also appear to be \textit{smoothed over} in scientific usage.

Ebeling (2013) & (2014) compares semantic prosody in English and Norwegian. The author finds that for English ‘commit’ and ‘signs of’, Norwegian possesses functionally equivalent items that match their English counterparts in terms of their distributions in language use. For ‘utterly’ and ‘cause’, on the other hand, no directly equivalent items exist in Norwegian.

Another type of evidence comes from research on how L2 speakers employ semantic prosodies. Several studies have shown that ESL students and even advanced translation students are often not aware of semantic prosodies, and use language that contains infelicitous prosodic clashes. These findings have been established for native speakers of Arabic (McGee, 2012), Chinese (Lu, 2005; Wang and Wang, 2005; Naixing, 2006; Zhang, 2009, 2014; Li and Liu, 2017), Farsi (Ahmadian et al., 2011; Elahi and Rahbar, 2018) and Korean (Lee, 2011). The relevance of these studies in this context is that a lack of prosodic awareness of non-native speakers supports the idea

\textsuperscript{1}A corpus search shows that this use of ‘perpetrate’ is extremely rare. ‘Play’, ‘novel’, ‘book’ or ‘work of art’ are all not attested as collocates of ‘perpetrate’ in COCA. However, when we use the web as a corpus, they are attested a few times.
that semantic prosodies are language-specific. The sole exception to this tendency is Dam-Jensen and Zethsen (2008) whose study found that Danish translation students had a relatively good awareness of semantic prosody in English. Their study participants translated from English to their native Danish and chose words with appropriate semantic prosodies with an accuracy of 75% to 92%, depending on their level of education and on the specific English items.

Many of these articles criticize ESL instruction and bilingual dictionaries for presenting vocabulary in isolation, and focusing on denotative meaning while ignoring co-selection tendencies such as semantic prosody or preference. Some researchers point out that dictionaries often provide translations that are inappropriate in terms of their evaluative polarities. Li and Liu (2017), for example, find that translations in Chinese-English bilingual dictionaries often mismatch with regards to their semantic prosodies.

In summary, research on semantic prosody from a contrastive perspective shows a complicated and nuanced picture. How different or similar semantic prosody is cross-linguistically is a subjective question. It greatly depends on the scope and the level of granularity in the analysis. Some researchers, e.g. Dodd (2006) define semantic prosody narrowly, limiting it to evaluative polarity. These researchers tend to find more cross-linguistic similarities than differences. Other researchers include the study of semantic preference in their scope (either subsuming it under semantic prosody or including it as an additional part of a linguistic item’s distributional properties),
e.g. Munday (2011), resulting in a more fine-grained description that reveals more cross-linguistic differences. Researchers like Tognini-Bonelli (2001) who also include phraseological units in their investigations find that small differences in linguistic form can lead to marked differences in usage. This type of variation is often idiomatic and language-specific, meaning that seemingly equivalent phrasemes in two languages may be more different than is immediately obvious.

What counts as similar or different also depends on the researchers’ goals. Much research on semantic prosody from a cross-linguistic perspective is done within the field of translation studies and aims to improve the training of translation students. As such, it tends to emphasize the idea that translators need to be aware of cross-linguistic differences in semantic prosody (Oster and van Lawick, 2008; Volanschi, 2012, inter alia). Similarly, researchers in second language learning find that L2 speakers tend to lack knowledge of semantic prosody and point out that language instruction needs to include information about cross-linguistic differences in semantic prosody.

Another theoretical consideration about cross-linguistic differences relates to the distinction between specific linguistic items and the linguistic system in general. Concerning specific linguistic items, Stewart (2009a, p. 32) echoes a conclusion commonly found among semantic prosody researchers who have not conducted any contrastive studies themselves: “semantic prosodies of near-synonyms are unpredictable across […] language pairs, in some cases being quite similar and in others quite differ-
ent.” While this may, at first glance seem like a reasonable summary of the research literature, it is a problematic and ultimately trivial statement. There is no reason to assume that they should be predictable. The whole idea is, in fact, not meaningful on a theoretical level and can be dismissed without even having to consider cross-linguistic evidence.

Even within the same language near-synonyms often differ with regards to their evaluative semantics. In fact, languages can have sets of lexical items that are near-synonyms with respect to their basic propositional meanings but that differ in terms of their evaluative semantics and in terms of the types of words they tend to co-occur with. Examples in English include Partington’s (2004) happen words ‘happen’, ‘set in’, ‘occur’, ‘come about’ and ‘take place’, and amplifying intensifiers ‘absolutely’, ‘perfectly’, ‘entirely’, ‘completely’, ‘thoroughly’, ‘totally’ and ‘utterly’ (also see the discussion of Xiao and McEnery (2006) findings on a related topic below). Semantic prosody cannot be predictable across languages, because it is not clear which words from such sets should be matched up across two different languages. Arbitrarily picking two words from these sets and comparing their semantic prosodies does not tell us anything about how universal semantic prosody is.

Two points need to be made here. First, researchers who have conducted cross-linguistic studies are of course aware of this. As such, they typically start with a lexical item in one language and try to find its closest equivalent from a list of items in another language. In some cases, researchers may find cross-linguistic pairs that
are quite similar in their distribution, while in other cases no good options may exist. Ebeling (2014), for example, found a close equivalent of English ‘commit’ in Norwegian, but did not find a good translation for ‘utterly’.

Second, while Stewart’s statement is theoretically void, it can be rephrased into a practically useful concept: The *prima facie* translation of a word will often not match its semantic prosody. In other words, the first translation offered in bilingual dictionaries is often not informed by distributional information. Additionally, cognates between two languages often do not share similar semantic prosodies. Pointing this out can indeed be valuable, because, as Partington (1998, p. 78) puts it: “The pitfalls for translators unaware of such prosodic differences are evident.” Munday (2011, p. 172) makes a similar statement about unexpected prosodic mismatches among cognates in two different languages: “One may describe this as a subtler variation on the old false cognate (*faux amis*) translation chestnut.” These are technically not false cognates but instead genuine cognates that are false friends. Still, the main point stands: Across languages, cognates may not match in terms of semantic prosody.

Moving on from a cross-linguistic comparison of semantic prosody in specific linguistic items to its study in the linguistic system in general, the current research literature offers very little data. On a fundamental level, semantic prosody has been found to exist in every language it has been investigated in. It is unlikely that this is due to the *file-drawer effect* (a publication bias that favors the publication of positive results); it seems that the absence of semantic prosody in a language would be a much
more noteworthy finding than its existence.

Beyond its mere existence, there is little cross-linguistic research into the structure of semantic prosody. Xiao and McEnery (2006) believe that their study of Chinese and English reveals systematic similarities. In both languages, for example, the *happen words* can be arranged on a continuum from less to more favorable (or at least neutral). While it is questionable to interpret this as a profound, shared systematicity, it seems reasonable to suggest that languages may have similar sets of near-synonyms that differ in their evaluative polarities (and likely also in their collocational behavior). In the case of the *happen words*, a speaker may want to take an evaluative stance towards the occurrence of an event they are describing. Having different *words* with different evaluative semantics at their disposal seems like a useful strategy that both English and Chinese appear to implement. It is, of course, impossible to extrapolate how common this is in the languages of the world based on a sample of two languages.

A final point is that none of this seems to be specific to semantic prosody. As much of the research cited in this section shows, it is not only the evaluative polarity that can differ between words and their translation but the entire distributional context. This includes a preference for lexical fields and grammatical constructions. By the same token, a lack of implicit awareness of semantic prosody among L2 speakers is not surprising. On a more general level, studies have shown that non-native speakers lack knowledge of collocations and lexical patterning tendencies (Lorenz, 1998; Altenberg and Granger, 2001; Nesselhauf, 2003, inter alia). As such, there is again no indication
that semantic prosody is a separate phenomenon that cannot be subsumed among
general linguistic processes of co-selection.

There are some limitations to the existing contrastive studies. First, as is the case
with research into semantic prosody in general, most studies focus their investigation
on the level of the individual word. In some cases, larger phraseological units are also
considered, but these are relatively rigid, idiomatic structures, not fully productive
grammatical constructions. Second, few of the studies described here work with par-
allel corpora. Instead, they contrast linguistic items from different languages within
separate, unrelated corpora. The studies presented in the next section aim to address
both of these issues.

8.2 Semantic prosody in German and English

The data for the two studies in this chapter come from the European Parliament
Proceedings Parallel Corpus 1996-2011, or Europarl Corpus, (Koehn, 2005), a par-
allel corpus that consists of translations of the same material into multiple different
languages. Its size is roughly 60 million words per language. All sentences (and
sometimes slightly larger utterance units) are indexed in such a way that the cor-
responding sentences in other languages can be easily retrieved. This allows for a
direct comparison of how the same meaning is expressed across languages. I used

\[2\] The corpus is available for download at http://www.statmt.org/europarl. The data from release
v7 was used (https://www.statmt.org/europarl/v7/).
spaCy (Honnibal, 2017), a natural language processing library for the programming language Python to tokenize, part-of-speech tag, lemmatize and parse the raw Europarl data in order to prepare it for further linguistic analysis.\(^3\)

While the Europarl corpus is an excellent resource for contrastive studies, two caveats need to be addressed. First, the corpus does not include any information as to what the original language of any given utterance is. This means that when I, for example, talk about an English phrase and its German translation, this does not necessarily mean that the German version is a direct translation of an English original. The opposite could just as well be the case, or, both could be translations of a third language. The term ‘translation’ is simply a shorthand here to denote that there are two linguistic structures in two different languages that aim to convey the same contents.

Second, while the translations are all done by professional translators, this does not necessarily mean that they are particularly authoritative. They are not definite translations because no such thing exists. A similar point could, of course, be made about the quality of evidence in any unilingual corpus. In other words, the data need to be taken for what they are: language in a natural context that exemplifies the common usage of the linguistic structures we wish to study. While it is valid to point out the difficulty of translating semantic prosody across languages, as some

\(^3\)The Europarl corpus can also be searched online through various web interfaces, but I found that the tools provided by SpaCy to perform with superior accuracy.
of the research described in the previous section does, the patterns that emerge in contrastive studies of parallel corpora are still valuable.

8.2.1 Study 1: English ‘cause’ and its German counterparts

In this study, I investigate how English verbal ‘cause’ with its various constructions is translated into German. I first provide a general overview of expressions of causality in German that correspond to English sentences with ‘cause’. Next, I distinguish between the different constructions of ‘cause’ and investigate what German structures they are aligned with. In chapter 5, I made the argument for constructional meaning, so I expect the different constructional variations of ‘cause’ to have different analogs in German.

Overview

In order to get a general overview of the data, I first retrieved a random sample of 200 sentences containing a form of the lemma ‘cause’, tagged as a verb from the English portion of the Europarl corpus. I then retrieved the corresponding German sentences and analyzed the way causation was expressed in both languages. The most common verbs in German corresponding to English verbal ‘cause’ are given in table 8.1.

<table>
<thead>
<tr>
<th>Rank</th>
<th>German expression</th>
<th>English translation</th>
<th>Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>verursachen</td>
<td>cause</td>
<td>53</td>
</tr>
<tr>
<td>2</td>
<td>führen zu</td>
<td>lead to</td>
<td>13</td>
</tr>
</tbody>
</table>
There is a multitude of expressions used in German that correspond to English ‘cause.’ This does, of course, not mean that German possesses a more complex or varied system of causatives than English. Instead, it is a direct result of the way the data were retrieved: I retrieved occurrences of the English verb ‘cause’, effectively limiting the sample of English sentences to this causative, while imposing no such restrictions on the German sentences. Searching for a German causative verb such as verursachen also results in various expressions of causation in the corresponding English sentences. Sentences (1) – (10) illustrate the use of the ten most common German verbs that are used as analogs of English ‘cause’. Unless otherwise indicated, all examples provided in this chapter are authentic examples from the Europarl cor-

<table>
<thead>
<tr>
<th></th>
<th>English</th>
<th>German</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>emerge</td>
<td>entstehen</td>
<td>11</td>
</tr>
<tr>
<td>4</td>
<td>trigger</td>
<td>auslösen</td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>evoke</td>
<td>hervorrufen</td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td>bring</td>
<td>bringen</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>cause</td>
<td>anrichten</td>
<td>5</td>
</tr>
<tr>
<td>8</td>
<td>create</td>
<td>erzeugen</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>let</td>
<td>lassen</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>give cause</td>
<td>Anlass geben</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 8.1: The most common German verbs or structures corresponding to verbal ‘cause’ in the Europarl corpus.
pus. Some examples have been simplified, typically by omitting material that is not relevant to the points they illustrate.

(1) *Die Wirtschafts- und Sozialkrise wird zweifelsohne immer mehr soziale Unruhen verursachen.*

‘The economic and social crisis is undoubtedly going to cause increased social unrest.’

(2) *Seine Bestimmungen sind unpräzise und führen nicht nur zu Kontroversen, sondern bereits zu einer Teilung in der EU-Führung.*

‘Its provisions are not precise, and not only are they causing controversy, but they are already causing a division among the EU’s leaders.’

(3) *Wie groß ist der bisher schon entstandene Schaden für die europäische Wirtschaft?*

‘How great is the damage which has already been caused to the European economy?’

(4) *Dieser letztenannte Punkt scheint im Rat Kontroversen ausgelöst zu haben.*

‘The latter appears to have caused controversy in the Council.’

(5) *China hat diese Gespräche abgebrochen und den Dalai Lama beschuldigt, Gewalt hervorzurufen.*

‘China broke up these talks, accusing the Dalai Lama of... causing violence...’
(6) Das unsägliche Leid, das diese Kriminellen über die Welt gebracht haben...
‘... the unspeakable suffering these criminals have caused in the world’

(7) Dies hat sehr viel Unheil angerichtet...
‘This has caused a great deal of harm...’

(8) Einige Redner haben schon darauf hingewiesen, ... dass Pestizide nachweislich Krebs erzeugen können.
‘A number of speakers have already pointed out ... that pesticides can cause cancer.’

(9) die neuerliche Ölpest... [lässt] heute die Fischer in Galicien verzweifeln...
‘the latest oil spill,... has today caused despair amongst the fishermen of Galicia.’

(10) ... wir [sehen] viele spektakuläre Errungenschaften..., jedoch auch Entwicklungen, die zur Besorgnis Anlass geben.
‘...we witness many spectacular achievements but also developments that cause concern.’

Causation is not always expressed through causative verbs in German. Another common approach is to encode the caused event in an intransitive clause with the causer expressed in a prepositional phrase. The verbs describing the caused action are often change-of-state verbs or verbs of inception. The most common verb in this context is entstehen (‘arise’, ‘emerge’) as exemplified in sentence (3).
Another common strategy in German is to use lexical causatives such as *verzögern* (‘cause a delay’) as in example (11) or *schädigen* (‘cause damage’) as in example (12). Other lexical causatives include *anheizen* (‘cause (a conflict) to escalate’, lit. ‘stoke’) and *beunruhigen* (‘cause anxiety’, lit. ‘upset’, ‘agitate’). Lexical causatives are relatively common, making up a little of 10% of the analyzed sample.

(11) *Von daher wollen wir jetzt nicht verzögern...*

‘Therefore we do not want to cause a delay now...’

(12) *Wird also die Kommission... Ernährungsweisen, die nachweislich die Gesundheit der Kinder schädigen, verbieten?*

‘So, will the Commission... ban the use of diets proven to cause damage to children’s health...?’

Phraseological expressions with verbs of causation that only collocate with a very restricted set of nouns encoding the caused event are also relatively common. In some cases, they are heavily idiomatic, and their meaning cannot be reduced to their components. Examples include *Bedenken erwecken* (‘cause concerns’, lit. ‘awaken concerns’) and *Verwirrung stiften* (‘cause confusion’, lit. ‘donate/endow confusion’). None of these individual examples are particularly common but the general principle is prevalent.

As described in chapter 5, verbal ‘cause’ is highly productive and is used to express at least three different types of causation: (i) The transitive and passive constructions
encode the direct causation of an event or state-of-affairs (‘caused an accident’). (ii) The two dative constructions describe the metaphorical transfer of (ii, a) an internal state (‘caused him trouble’) or (ii, b) an external state (‘cause damage to them’). (iii) The infinitival complement describes the effect on an actor who retains some degree of volitionality and control over their actions (‘caused me to reevaluate my position’).

Based on the parsed data from the Europarl corpus, I formulated syntactic rules to automatically segment all English sentences containing verbal ‘cause’ into the five different constructions. A little more than half of the sentences included the transitive construction, about a third used the passive construction, 10% were the to-infinitive construction and the ditransitive and prepositional dative both accounted for a few percent each. This breakdown happens to be almost the same as in the sample of 200 sentences I analyzed manually. When these data are compared to the data described in chapter 5, there is one notable difference: For reasons that are not immediately obvious (although register and genre may play a role), Infinitival complements are much less common. Other than that, the data are relatively similar: Transitive and passive constructions are both highly common, while the ditransitive and the prepositional constructions are exceedingly rare.

The German expressions of causation are retrieved through their alignment with English sentences containing verbal instances of ‘cause’, but they are not only discussed within this context. Their usage in other contexts, i.e. within the entire corpus is also described. This includes both descriptions of their collexemes profiles and of
their English analogs. Since this study includes several words that are examined from multiple contexts, the results are summarized in a maximally concise manner that aims to convey the relevant findings while avoiding an unnecessary level of detail.

**Transitive construction**

The English transitive construction with ‘cause’ is most commonly translated with a very similar grammatical structure in German: a transitive construction with a causative verb. The most frequent German verbs of causation found in sentences corresponding to English sentences with ‘cause’ in the transitive construction are, in descending order of frequency: *verursachen* (‘cause’), *führen (zu)* (‘lead to’), *hervorrufen* (‘evoke’), *bringen* (‘bring’), *auslösen* (‘trigger’), *machen* (‘make’), *kommen (zu)* (‘come to’) and *anrichten* (‘cause’).\(^4\) This list is relatively similar to the counts from the overall sample of 200 sentences with English ‘cause’ (see table 8.1), which is not surprising considering that the transitive construction has the highest token count in the data.

*Verursachen* is not only by far the most common translation of transitive ‘cause’; the two words also have fairly similar collexeme profiles. Whether we investigate ‘verursachen’ only in this context (i.e. only in German sentences that correspond to English sentences with transitive ‘cause’) or within the entire Europarl corpus,

\(^4\) Note that for cross-linguistic comparisons throughout this chapter I am only using raw counts. The reason is that I simply aim to describe the most common translations; I am not proposing some kind of cross-linguistic measure of association that would require word counts to be normalized.
it tends to co-occur with the same vocabulary. In both cases, the most strongly attracted collexemes to *verursachen* are words like *Problem* (‘problem’), *Kosten* (‘costs’), *Schaden* (‘damage’), *Katastrophe* (‘catastrophe’), *Leiden* (‘suffering’), *Krise* (‘crisis’), *Schwierigkeit* (‘difficulty’) and *Krankheit* (‘illness’). Additionally, if we search for all German sentences that include *verursachen* and then analyze the corresponding English sentences, it turns out that over half of them use ‘cause’. Other common translations include ‘create’, ‘lead to’ and ‘produce’.

Despite their similarities, there are two noteworthy differences between the two verbs. First, *verursachen* is not as productive as ‘cause’. Roughly two-thirds of all tokens with *verursachen* are in the transitive construction and a third is part of the passive construction. The ditransitive dative and prepositional dative are attested but are extremely rare, accounting for less than 1% of all sentences with *verursachen* in the Europarl corpus. Using it with an infinitival complement is ungrammatical and not attested.

Second, there are some collexemes of ‘cause’ in English whose German analogs do not tend to co-occur with *verursachen*. One example is ‘cause a stir’. Common German translations include *Aufsehen erregen* (lit. ‘arouse a sensation’), *Aufregung geben* (lit. ‘give excitement’) and the idiom *Staub aufwirbeln* (lit. ‘whirl up dust’, similar to the English idiom ‘kick up dust’). Another example is ‘cause confusion’, which is typically translated as *Verwirrung hervorufen* (lit. ‘evoke confusion’) oder *Verwirrung stiften* (lit. ‘donate confusion’). A third example is ‘cause
concern’. Analog expressions in German include *Sorgen bereiten* (‘cause concern’) and the noun-adjective compound *besorgniserregend sein* (‘be concerning’, lit. ‘be concern-evoking’).

Another frequently encountered translation of transitive ‘cause’ is *führen zu* (‘lead to’). Collexemes are highly varied, and few generalizations seem possible, except that they tend to denote abstract concepts. When considered in the context of the entire Europarl corpus, the most frequent collexeme is *nichts* (‘nothing’). Other frequent collexemes include *Problem* (‘problem’), *Verspätungen* (‘delays’) and *Störung* (‘disturbance’) but also *Verbesserung* (‘improvement’), *Wachstum* (‘(economic) growth’) and *Rechte* (‘civil liberties’). Unlike was the case with *verursachen* which strongly tended to convey negative evaluations, there is no strong evaluative polarity here, with 61% of the collexemes being negative, 31% positive and 9% neutral. The most common English analog of *führen zu* is ‘lead to’ which also happens to be its literal translation.

The collexeme profile of the causative verb *hervorrufen* (‘evoke’) can be categorized into three groups, listed here in descending order of centrality to the construction. The first group consists of words from the semantic field of illnesses, e.g. *Symptom* (‘symptom’), *Verletzung* (‘injury’), *Allergie* (‘allergies’) and *Krankheit* (‘illness’). The second group contains terms denoting conflict and disagreement, such as *Konflikt* (‘conflict’), *Spannung* (‘tension’) *Widerstand* (‘resistance’) and *Reaktion* (‘reaction’). The third semantic field is about concerns, including words like *Besorgnis* (‘concern’).
and Bedenken (‘misgivings’). Here, the most common vocabulary in English is ‘cause’, ‘create’, ‘make’, ‘lead to’ and ‘generate’.

Auslösen (‘trigger’) mostly occurs with collexemes describing confrontational human interactions, such as Diskussion (‘discussion’), Debatte (‘debate’), Reaktion (‘reaction’) and Konflikt (‘conflict’). Words describing emergencies such as Krise (‘crisis’) and Katastrophe (‘catastrophe’) are also common. English analogs include ‘cause’, ‘trigger’, ‘create’, ‘provoke’.

While less common than the other verbs described here, anrichten is noteworthy because it is far more semantically specialized than any of the other verbs described so far. It is solely used to describe destruction. Its top collexemes are Schaden (‘damage’), Zerstörung (‘destruction’) and Unheil (roughly ‘harm’, ‘misery’). This appears to be a hard semantic restriction, not a statistical tendency. Additionally, its collexemes are frequently intensified by adjectives, e.g. enormer Schaden (‘enormous damage’) or massive Zerstörung (‘massive destruction’). The English verb ‘wreak’ behaves in a similar way but is even more restricted is that is almost exclusively occurs in the idiom ‘wreak havok’, whereas anrichten occurs with several different words, albeit from a very limited lexical field.

Passive

By and large, the causative verbs found in German equivalents of English sentences with ‘cause’ in the passive construction are very similar to the ones in the transitive
construction and much of what has been described in the previous section applies here as well. As was the case with the transitive construction in English, \textit{verursachen} is also the most commonly used verb to translate passive ‘cause’.

One difference is that the second most common verb in German sentences corresponding to passive ‘cause’ is non-causative verb \textit{entstehen} (‘emerge’, ‘arise’), which is relatively uncommon in the transitive construction. It is an example of causation being expressed with an inchoative verb, typically combined with a causer encoded by a prepositional phrase, in this case \textit{durch} (lit. ‘through’, here ‘by’, ‘due to’). The collexeme profile of \textit{entstehen} consists of abstract nouns and is very varied. It includes words like \textit{Kosten} (‘costs’), \textit{Problem} (‘problem’), \textit{Hindernis} (‘obstacle’) but also \textit{Möglichkeit} (‘possibility’) and \textit{Arbeitsplatz} (‘job’). The corresponding English sentences include similar structures with inchoative verbs, such as ‘arise’, ‘emerge’ or ‘develop’ but also causative verbs like ‘create’, ‘make’ or ‘cause’.

\textbf{Infinitival complement}

\textit{Führen} (‘lead to’) is by far the most common German verb in sentences corresponding to English ‘cause’ with the infinitival complement. Various types of complement clauses are used; the most common of which are introduced by the subordinating conjunction \textit{dass}, as in example (13). In the German data, infinitival complements are very uncommon in this context.

(13) \textit{Angst um die eigene Sicherheit führt oft dazu, dass wir Strategien anwenden},
die uns in falscher Sicherheit wiegen.

‘Fear for our own security often causes us to adopt strategies that create a false sense of security.’

Führen is not only used with complement clauses, however. Another common strategy is illustrated in example (14). The caused event is expressed without a verb phrase and instead only through a noun phrase, while the causee is encoded with a prepositional phrase (in this case with the preposition bei).

(14) Die Krise..., führt bei vielen Jugendlichen zu Teilnahmslosigkeit und Gewaltbereitschaft.

‘The crisis... is causing many young people to adopt apathetic or violent attitudes.’ (lit. ‘The crisis is leading to apathy and propensity towards violence among young people’)

Other German verbs commonly encountered in this context include bringen zu (‘bring to’, as in ‘bring s.o. to do sth.’), lassen (‘let’), veranlassen (roughly ‘prompt’). Like führen, they can be used with various types of complement clauses. Unlike führen, however, they cannot be used to express causation in simple transitive clauses. In these situations, the causing event and the caused event are usually clearly separated, both conceptually and linguistically. The causee tends to retain a fair degree of control and volition over the caused action. When all occurrences of these verbs are considered, we see that they are not particularly often aligned with ‘cause’ in English.
There are numerous expressions in English, including verbs such as ‘compel’, ‘prompt’ and ‘lead to’.

**Prepositional dative**

As in chapter 5, the Europarl data for English ‘cause’ in the prepositional dative show that it most commonly occurs with external states such as ‘damage’, ‘harm’, ‘disruption’, ‘suffering’ or ‘injury’. In German, the ditransitive dative is the preferred construction to express any kind of transfer. As its English equivalent, it can express causation through the metaphorical transferal of a state upon a causee. Unlike the English ditransitive, it is not specialized in describing the causation of internal states; instead, the causation of external states is just as common. In the Europarl data, English sentences that use the prepositional dative are commonly conveyed through the ditransitive dative in German.

The reason for this seems to be that German does not possess a propositional dative in the same way that English does. Similar grammatical structures do exist, but there are some important differences. First, while the recipient role can be expressed through a prepositional phrase in German, there is no single preferred preposition like English ‘to’; instead the prepositional phrase can contain one of various different prepositions, such as *in* (‘in’), *für* (‘for’), *an* (roughly: ‘at’), *mit* (‘with’) or *bei* (roughly: ‘by’, ‘at’). Second, none of these prepositions is fully productive. Instead, they tend to be restricted to specific sets of verbs. For example, the German equivalent
of ‘it caused damage to the forests’ is *es verursachte Schäden in den Wäldern* (‘it caused damage in the forests’); the translation of ‘it caused damage to the budget’ is *es hat Einbußen für den Haushalt gebracht* (‘it brought losses for the budget’). Third, these prepositional constructions are far less common in German than in English.

Prepositional phrases in German are, however, regularly used in a different role within trivalent causative scenes: They commonly encode the role of the causer, as in example (15). English dative constructions are not always conveyed through equivalent structures in German. In this case, the clause is passivized; the causee is the subject of the clause and the predicate is a lexical causative verb.

(15) *Durch die Explosion wurde... das Parlamentsgebäude stark beschädigt.*

‘The explosion also caused very considerable damage to... the Parliament building.’ (lit: ‘through the explosion the Parliament building was damaged considerably)

The verb most frequently found in German analogs of ‘cause’ in the prepositional dative is *zufügen* (roughly ‘inflict’). In the German portion of the Europarl corpus, it tends to occur with collexemes such as *Schaden* (‘damage’), *Leid* (‘suffering’), *Niederlage* (‘defeat’), *Verletzungen* (‘injuries’).

The use of compounds is another common strategy in German. Instead of using a dative construction as in the English sentence ‘cause damage to their health’, the metaphorical transfer of damage is instead conveyed through the noun-noun compound in *Gesundheitsschäden verursachen* (lit: ‘cause health damage’) in German.
Alternatively, use of a predicative adjective derived from a noun-adjective compound, e.g. *gesundheitsschädlich sein* (lit. ‘be health-damaging’) is also attested several times in the Europarl data.

**Ditransitive**

The English data in the Europarl corpus show a very similar distribution to the data discussed in chapter 5. The most strongly attracted collexemes express internal states, e.g. ‘concern’, ‘anxiety’ and ‘pain’. Other than that, ‘problem’ and ‘difficulty’ are also relatively common.

The German verb *bereiten* (lit. ‘prepare’) is a common translation of ‘cause’ in the ditransitive construction. Like its English equivalent, it can be used in the ditransitive construction to express causation. ‘Bereiten’ occurs almost exclusively with internal states such as *Sorge* (‘concern’, ‘anxiety’) and *Schmerzen* (‘pain’), or with *Problem* (‘problem’) and *Schwierigkeiten* (‘difficulties’). When uncoupled from the English ditransitive construction with ‘cause’ it also expresses the causation of positive mental states, *Freude* (‘joy’) being its strongest collocate. In English, the causation of ‘joy’ is most often expressed through a ditransitive with ‘give’ or ‘bring’.

The phrasal verb *erfüllen mit* (lit. ‘fill someone with’) is a somewhat similar example. It solely occurs in the ditransitive construction and exclusively expresses the causation of mental states (this is not a statistical tendency, but instead a categorical rule). As is the case with *bereiten*, in this context it occurs with negative emotions,
but in general, it also occurs with positive emotions.

Another common strategy in German is to use lexical causatives. The transitive verb *beunruhigen* (with the experiencer expressed as the direct object) is commonly used as an analog to ditransitive ‘cause’ with collexemes such as ‘anxiety’, ‘anguish’ or ‘concern’.

When causation is expressed through the dative construction in German, the verb does not necessarily need to be causative. The verb *entstehen* (‘emerge’, ‘arise’), for example, would traditionally be described as a monovalent verb that occurs in intransitive sentences. As previously mentioned, it can convey quasi-causative scenes when the causer is encoded in a prepositional phrase. An additional recipient can be included through the dative construction. Example (16), a constructed example that is based on attested patterns from authentic sentences from Europarl, illustrates this usage:

(16) *Ihm entsteh-t dadurch ein Problem.*

3SG.DAT arise-3SG through this a problem.

‘This causes him a problem’ (lit. ‘to him a problem arises due to this’)

The recipient does not need to be an indirect object (marked through the dative case in German) but can also be an oblique object, encoded as a prepositional phrase. This is one of the more common used of the otherwise rare prepositional dative in German. Sentence (17), another constructed example illustrates this in the context of an impersonal construction.
In order to avoid unnecessary repetition, I will forgo a summary of the results of this study here. Instead, its most important findings are summarized in 8.3 when they are discussed in their wider, theoretical implications.

### 8.2.2 Study 2: *vollkommen*

This study deals with the German intensifier *vollkommen*, usually translated as ‘completely’, ‘totally’ or ‘perfectly’ in the Europarl corpus. English intensifiers tend to exhibit constructional differences in their semantics and also defy the notion of semantic prosody as a good/bad polarity (e.g. see ‘absolutely’ in chapter 6). This investigation aims to determine whether the same can be established for German intensifiers.

In theory, *vollkommen* can function as an adverb and as an adjective. The adjectival usage, however, is exceedingly rare in the Europarl corpus (and I suspect in the language overall) as will not be considered in this study. The two uses that are considered in this study are modifying adjectives and modifying verbs.
Vollkommen + adjective

_Vollkommen + adjective_ is its most frequent construction in the Europarl corpus, accounting for roughly three-quarters of all occurrences of the lemma. It most commonly occurs in the form of adjective predicates, typically forming part of an evaluative statement, where the speaker takes a stance in relation to an inanimate and often abstract subject. It is also the most versatile construction in terms of its collexemes, the 20 strongest of which are given in table 8.2. Five different lexical sets can be identified that account for a large portion of all collexemes. They are listed in descending order of their centrality to the construction.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Collexeme</th>
<th>Translation</th>
<th>Tokens</th>
<th>Attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>inakzeptabel</td>
<td>unacceptable</td>
<td>78</td>
<td>323</td>
</tr>
<tr>
<td>2</td>
<td>klar</td>
<td>clear</td>
<td>127</td>
<td>161</td>
</tr>
<tr>
<td>3</td>
<td>transparent</td>
<td>transparent</td>
<td>35</td>
<td>119</td>
</tr>
<tr>
<td>4</td>
<td>unzureichend</td>
<td>insufficient</td>
<td>37</td>
<td>116</td>
</tr>
<tr>
<td>5</td>
<td>richtig</td>
<td>correct</td>
<td>81</td>
<td>110</td>
</tr>
<tr>
<td>6</td>
<td>bewusst</td>
<td>aware</td>
<td>71</td>
<td>105</td>
</tr>
<tr>
<td>7</td>
<td>unangemessen</td>
<td>inappropriate</td>
<td>21</td>
<td>86</td>
</tr>
<tr>
<td>8</td>
<td>unangebracht</td>
<td>inappropriate</td>
<td>17</td>
<td>82</td>
</tr>
<tr>
<td>9</td>
<td>unverständlich</td>
<td>incomprehensible</td>
<td>19</td>
<td>74</td>
</tr>
<tr>
<td>10</td>
<td>falsch</td>
<td>wrong</td>
<td>33</td>
<td>63</td>
</tr>
</tbody>
</table>
Table 8.2: Collexemes with the strongest attraction to the adjective slot in the vollkommen + adjective construction.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Collexeme</th>
<th>Translation</th>
<th>Tokens</th>
<th>Attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>unterschiedlich</td>
<td>different</td>
<td>26</td>
<td>63</td>
</tr>
<tr>
<td>12</td>
<td>überflüssig</td>
<td>superfluous</td>
<td>17</td>
<td>56</td>
</tr>
<tr>
<td>13</td>
<td>absurd</td>
<td>absurd</td>
<td>17</td>
<td>52</td>
</tr>
<tr>
<td>14</td>
<td>unrealistisch</td>
<td>unrealistic</td>
<td>13</td>
<td>51</td>
</tr>
<tr>
<td>15</td>
<td>anders</td>
<td>different</td>
<td>36</td>
<td>51</td>
</tr>
<tr>
<td>16</td>
<td>unabhängig</td>
<td>independent</td>
<td>33</td>
<td>49</td>
</tr>
<tr>
<td>17</td>
<td>unannehmbar</td>
<td>unacceptable</td>
<td>17</td>
<td>48</td>
</tr>
<tr>
<td>18</td>
<td>legitim</td>
<td>legitimate</td>
<td>12</td>
<td>38</td>
</tr>
<tr>
<td>19</td>
<td>zufrieden</td>
<td>content</td>
<td>28</td>
<td>35</td>
</tr>
<tr>
<td>20</td>
<td>unnötig</td>
<td>unnecessary</td>
<td>12</td>
<td>30</td>
</tr>
</tbody>
</table>

First, a typical use of vollkommen adjective is to describe something as unacceptable, insufficient or inappropriate. The vocabulary includes inakzeptabel (‘unacceptable’), unzureichend (‘inadequate’), unangemessen (‘inappropriate’), unangebracht (‘inappropriate’, ‘unwarranted’), absurd (‘absurd’) and unannehmbar (‘unacceptable’). Collexemes like angemessen (‘appropriate’) or akzeptabel (‘acceptable’) are attested but are much less common than their negative equivalents. The most common English intensifiers in this context are ‘utterly’, ‘totally’ and ‘completely’. It
is, however, not unusual for the English translations to entirely omit the intensifier, especially if the adjective is categorical in its meaning, such as ‘unacceptable’.

Second, it commonly evaluates something more generally as either right or wrong. The two main collexemes are richtig (‘right’) and falsch (‘wrong’). While the type count of this category is very small, it has a high token count. The English sentences corresponding to vollkommen richtig use adverbs such as ‘absolutely’ or ‘perfectly’, e.g. ‘absolutely right’, ‘perfectly correct’. The English versions of sentences with vollkommen falsch in German typically include intensifiers such as ‘utterly’, ‘totally’, ‘completely’ or ‘seriously’ (e.g. ‘utterly false’, ‘totally wrong’, ‘completely wrong’, ‘seriously wrong’).

Third, vollkommen adjective marks evidentiality. The speaker either indicates that something is obvious or that someone (typically the speaker themselves) are aware of something. In the former case, this is usually expressed with an impersonal construction as in example (18). The most common collexeme here by far is klar (‘clear’, ‘obvious’). Other examples include transparent (‘transparent’), eindeutig (‘unambiguous’) and offensichtlich (‘obvious’). In English, the equivalent is typically expressed through ‘perfectly’ or ‘absolutely’ and ‘clear’ or ‘obvious’.

(18) Es ist vollkommen klar, dass wir in zwei bis drei Jahren in diesen Branchen das gleiche Problem haben werden.

‘It is perfectly clear that we will have the same problem in these industries in two or three years’ time.’
The dative construction is used to express that someone is aware or certain of something. The most important collexemes are *klar* and *bewusst* (‘aware’). In most cases, the speaker talks about their own awareness or certainty. The typical sequence is *mir ist vollkommen klar* (‘it is entirely clear to me’). The preferred corresponding English expressions are ‘fully aware’ and the more specialized ‘well aware’ as in (19):

(19) *Mir ist vollkommen klar, dass dieser Vorschlag komplex und heikel ist.*

‘I am well aware that this proposal is complex and sensitive.’

The fourth category of collexemes is used to express difference or novelty. Examples include *anders* (‘different’), *unterschiedlich* (‘different’) and *neu* (‘new’). Here the most common intensifier in English is ‘completely’, followed by ‘entirely’. There are also specialized intensifiers such as ‘essentially’ or ‘altogether’ that occur in relatively few other contexts and are strong collocates of ‘different’ in English. In contrast to the other three collexeme categories which all tended to form part of predicate adjectives, this category is commonly part of noun phrases (e.g. ‘a completely / entirely different matter / question / situation / opinion’). The fifth group of collexemes describes something as unnecessary and includes collexemes such as *überflüssig* (‘superfluous’) or *unnötig* (‘unnecessary’). The most frequent English intensifiers are ‘completely’, ‘utterly’ and ‘totally’.
Vollkommen + Verb

The 20 collexemes with the strongest attraction to *vollkommen* + verb are listed in table 8.3. When used to modify verbs, *vollkommen* is strongly attracted to verbs that express agreement, approval or support. Its strongest collocate by far is *zustimmen* (‘agree’, ‘approve’). Other verbs in this category include *unterstützen* (‘support’) and *verstehen* (‘understand’). In the great majority of cases, the speaker expresses their stance towards a position. The intensifiers most frequently used in the corresponding English sentences are ‘fully’, ‘wholly’ and ‘completely’.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Collexeme</th>
<th>Translation</th>
<th>Token</th>
<th>Attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>zustimmen</td>
<td>agree</td>
<td>145</td>
<td>1,453</td>
</tr>
<tr>
<td>2</td>
<td>verstehen</td>
<td>understand</td>
<td>38</td>
<td>362</td>
</tr>
<tr>
<td>3</td>
<td>ignorieren</td>
<td>ignore</td>
<td>24</td>
<td>272</td>
</tr>
<tr>
<td>4</td>
<td>unterstützen</td>
<td>support</td>
<td>24</td>
<td>132</td>
</tr>
<tr>
<td>5</td>
<td>entsprechen</td>
<td>correspond</td>
<td>16</td>
<td>101</td>
</tr>
<tr>
<td>6</td>
<td>widersprechen</td>
<td>disagree</td>
<td>8</td>
<td>89</td>
</tr>
<tr>
<td>7</td>
<td>vergessen</td>
<td>forget</td>
<td>11</td>
<td>84</td>
</tr>
<tr>
<td>8</td>
<td>zuwiderlaufen</td>
<td>go against</td>
<td>6</td>
<td>78</td>
</tr>
<tr>
<td>9</td>
<td>verändern</td>
<td>change</td>
<td>7</td>
<td>66</td>
</tr>
<tr>
<td>10</td>
<td>übergehen</td>
<td>skip</td>
<td>7</td>
<td>64</td>
</tr>
<tr>
<td>11</td>
<td>ausschöpfen</td>
<td>fully utilize</td>
<td>5</td>
<td>59</td>
</tr>
</tbody>
</table>
Verbs conveying disagreement such as *widersprechen* (‘disagree’, ‘contradict’) or *zuwiderlaufen* (‘go against’, ‘run contrary to’) are also attested but are much less common. They are not direct opposites of the agreement verbs: They do not only convey people’s disapproval but also the idea that different facts or positions contradict each other. The most common English intensifier in this context is ‘totally’, as in ‘totally contradict’ and ‘be totally contrary to’ or ‘run totally counter to’.

Another identifiable lexical field includes verbs that deal with ignoring a person or a topic. The most important collexemes include *ignorieren* (‘ignore’), *übergehen*
(‘skip’, ‘bypass’, ‘ignore’) and perhaps also missachten (‘disregard’) and vegessen (‘forget’). The most common English intensifier across these sentences with these verbs is ‘completely’, followed by ‘totally’.

Beyond these categories, there is a long tail in the collexeme distribution that includes many words that only occur a few times. By and large, few semantic generalizations can be made. One observation that may be noteworthy is that despite the fact that most of its strongest collexemes could be classified as positive, many of its relatively low-frequency collexemes are neutral or negative.

As with the first study, a summary of these data is provided as part of the discussion in section 8.3.

8.3 Discussion

In the first study, the Europarl data revealed a multitude of causative expressions in German, including analytical and lexical causatives. None of them are anywhere near as productive as ‘cause’. Above all, what this study shows is that there is no single ideal translation for English ‘cause’. This is not surprising, given how productive the verb is and how much fine detail the collexeme analyses of its different constructions reveal. As such, it is clear we need to find German causative verbs that correspond not to ‘cause’ per se, but to the different constructional uses of ‘cause’.

The types of causation expressed through ‘cause’ in the transitive or passive in English is translated with verbs such as verursachen, führen zu, entstehen, hervorrufen,
auslösen and anrichten in German. The most common one them by far is verursachen, which Dodd (2006) mentioned as an example of semantic prosody translating well across languages. The two words are indeed relatively similar in their collexeme profiles and are frequently chosen in corresponding sentences by the translators in the European parliament. A closer examination however also reveals differences. Verursachen is almost exclusively used in the active and passive constructions, while ‘cause’ is more versatile. Not only does verursachen not commonly occur in other constructions in German, it is also very rarely used for causative events that would be expressed through the infinitival complement, the prepositional dative and the ditransitive construction in English. Additionally, it tends not to be used for expressions like ‘cause a stir’.

The infinitival complement is translated with verbs like veranlassen, bringen and lassen, often also part of complement clauses. Generalization about constructional meaning are less salient here, because complement clauses cannot be reduced so single-word collexemes.

Both the prepositional dative and the ditransitive dative in English are typically expressed through the ditransitive dative in German. Like English, German also distinguishes between the causation of internal and external states. Unlike English, German does not use different constructions but instead different verbs of causation, e.g. zufügen for external states and bereiten for internal states. At first glance, the German ditransitive with bereiten appears to be very similar to the English ditran-
sitive with ‘cause’, both in its grammatical structure and its semantics. They both
express the causation of internal states. There is however one significant difference
in that bereiten is equally used with positive and negative mental states. The same
is true for erfüllen mit; again, the verb occurs in the ditransitive and conveys the
causation of internal states but can be used for both positive and negative mental
states.

The ditransitive with zufügen is a relatively good match for English ‘cause’ in
the prepositional dative. The collexeme profiles are relatively similar, although there
is one important difference: Zufügen also occurs with a few collexemes describing
internals states, including Leiden ‘suffering’ which is one its most strongly attracted
collocates.

On a theoretical note, there is one caveat that has so far been ignored (both in
this chapter and in the relevant previous literature): Comparing collexemes profiles
across languages is problematic because it involves working with translations, the
pitfalls of which much of this chapter has dealt with. Every word comes with its own
distribution in language, which in turn consists of words that are also defined by their
distributions and so on. This kind of reasoning quickly leads to an infinitely recursive
circular argument. At this point, it seems reasonable to decide that we need to work
with approximations and that pairs like German Problem and English ‘problem’ are
similar enough for our purposes.

In summary, the finer the granularity of our analysis, the more details we find and
the more opportunity for cross-linguistic differences we encounter. In both languages, verbs of causation follow complex and detailed patterns of co-selection, making it nearly hopeless to find exact matches between two languages. Within specific constructional contexts, e.g. transitive ‘cause’ compared with transitive *verursachen*, we can find patterns with similar distributions, but even in those cases, it seems doubtful that they would ever be perfectly congruent.

The most important theoretical implication of this study concerns the context-dependent nature of meaning. As argued in chapter 5, the different collostructions ‘cause’ takes part in have different meanings, going beyond what the lexical perspective can account for. As such, we would expect them to be conveyed through different linguistic structures in other languages. The German data in this chapter bear out this hypothesis. German puts them into separate categories, less so by using different constructions, but by using different verbs of causation.

On a different note, the cross-linguistic data also provide some, albeit tentative, answers to the question to what extent semantic prosody can be explained through conceptual motivations. Hunston (2007a) surmised that the negative evaluative polarity so commonly conveyed in causation is driven by the fact that a human actor taking on the role of causee indicates a lack of volition, which is often unfavorable. Xiao and McEnery (2006) share this view and believe that similarities in semantic prosody between Chinese and English are partly due to conceptual motivations:

“[A]t least part of the explanation, in our view, can be found in the
common basis of natural language semantics - ‘the conceptual system that emerges from everyday human experience’ (Sweetser 1990: 1). However, as different languages can have different ranges of near synonyms, near synonyms and their close translation equivalents in different languages may also demonstrate, to some extent, different collocational behaviour and semantic prosody.” (Xiao & McEnery, 2006, p. 125)

The data in this chapter lend further support to this theory. In both English and German, causation tends to describe negative states-of-affair. More importantly, however, in both languages, this effect is weaker for verbs where the causee retains some control or volition over the caused action. German führen zu and its English ‘equivalent ‘lead to’ have weaker negative polarities than verursachen and ‘cause’. This is not only true when human causees are involved but also carries over to other contexts.

Moving on to the second study, there are three main takeaways from the investigation of vollkommen. First, it shows that semantic prosody in German is also dependent on grammatical context. As the data presented in this chapter show, there is a marked difference between vollkommen + adjective and vollkommen + verb, with both of them displaying distinct collexeme profiles. The shortest possible summary is that when used to modify adjectives, vollkommen most commonly conveys disagreement, labelling something as unacceptable or insufficient; when used with verbs on the other hand, its most important use expresses agreement. For both collostruc-
tions, additional lexical fields can be identified among their collexemes, none of which are related, further attesting to their distinctness and supporting the importance of context-dependent meaning.

The second point is a direct corollary of the first. Similar to English ‘cause’, there is no single best English translation of ‘vollkommen’. Its meaning depends on its constructional context and this context-dependent semantic variation is not congruent with that of any intensifier in English. As such, the best English translation varies with context. This concerns both the grammatical context (modifying adjectives or verbs) and, more locally, the class of adjectives it modifies. When used with adjectives that express unacceptability, the most common English translations include the intensifiers ‘utterly’ and ‘totally’, whereas when used with verbs of agreement, the most common translations are ‘fully’, ‘wholly’ and ‘completely’.

Third, as in English, semantic prosody in German cannot be broken down into a simple polarity of negative and positive evaluations. The semantics found in the collexeme profile of ‘vollkommen + adjective’ go well beyond a generally unfavorable evaluative polarity. They are specifically associated with a sense of unacceptability and inappropriateness. Additionally, not all of its collexemes are negative – one of the most strongly attracted collexemes is *richtig* (‘right’, ‘correct’). There are other examples that support the same point: German *bereiten* and *erfüllen mit* from the study on ‘cause’ are further example of words that defy a polarity of positive and negative evaluations. They commonly occur with both, positive (*Freude*, ‘joy’) and

On a final note, as with any contrastive study, care must be taken not to make sweeping typological generalizations based on a direct comparison of two languages. It is not clear how universal constructional meaning is. Xiao and McEnery (2006) briefly mention that there may be differences in semantic prosody in English and Chinese due to the fact that the latter is a strictly isolating language: “A more general difference between English and Chinese is that collocation and semantic prosody may be affected by morphological variations in English but not in Chinese, which lacks such variation” (Xiao and McEnery, 2006, p. 125). Evidently, more work is required in other languages.
Chapter 9

Conclusions

In this chapter, I discuss the conclusions of my research and come to a final verdict about semantic prosody. The chapter is organized into three parts.

In section 9.1, I recapitulate the most important findings of my reanalysis of semantic prosody and distill the evidence from the previous chapters into a comprehensive description. Among other things, I describe the importance of constructional context and the nature of evaluative meaning. Additionally, I reexamine open questions and areas of disagreement commonly found in the research literature within this approach and argue that it allows for semantic prosody to be accounted for as a natural part of ordinary semantic processes in language.

Section 9.2 is a discussion of the theoretical and methodological implications of my research. I situate the findings within a broader theoretical context, problematize some of their basic assumptions and provide a conclusion about semantic prosody.

Section 9.3 identifies avenues of further research and points out open questions that are beyond the purview of the present work.
9.1 A re-evaluation of semantic prosody within a constructional framework

9.1.1 Central findings

One of the most important findings of the present research is that the constructional level needs to be considered in the study of semantic prosody. As demonstrated in chapter 5, words do not always have the same meaning but instead vary in their semantics based on their constructional contexts. The adjective ‘persistent’, for example, conveys a different evaluative stance depending on whether it is used attributively or predicatively. In the former case, through its association with collexemes from lexical fields describing medical problems, environmental pollution and socio-economic issues, it adopts a meaning of describing an unfavorable state that is difficult to overcome. In its predicative usage, no such meaning exists. In this case, while it is not inextricably linked to any specific lexical sets, it generally describes people as persevering valiantly in the face of difficult opposition.

The study of verbal ‘cause’ identified different construction-dependent semantics, a finding which was corroborated through the contrastive study in chapter 8. I argued that the different lexical fields the different constructions of ‘cause’ occurred with indicated different meanings. The comparison with German showed that they were indeed different because they were expressed through different means in German (both different verbs and different constructions). Additionally, context-dependent meaning
was also established in German through the investigation of the adverb *vollkommen* modifying either adjectives or verbs.

Once we move past a lexical approach that conflates all context-dependent variants of a word, the data allow us to understand the use and meaning of linguistic items in much finer detail. This leads to the second central finding of this work: The meaning of semantic prosody cannot be described as a dichotomous system of positive and negative evaluations. Several examples presented in chapters 5 & 6 do not fit into such a two-term system. These examples can be loosely classified into three categories: First, some patterns have meanings that go well beyond a positive/negative polarity. One example for this is the construction *get oneself Ved* (e.g. ‘he got himself fired’). While a negative evaluation features prominently in the semantics of the construction, it also includes other components, such as the notion that the person is at fault for the unfavorable situation they are in. The second category concerns constructions that are evaluative but not defined for the positive/negative dimension. *Striking NP* (e.g. ‘a striking resemblance’), for example, evaluates something as being particularly noteworthy or salient without making a good/bad judgment. The third category of constructions that defy the notion of semantic prosody as evaluative polarity consists of patterns that are associated with both, positive and negative evaluations. An example for this is *absolutely ADJ*, which occurs with pairs of antonyms, such as ‘right’ and ‘wrong’, ‘essential’ and ‘unnecessary’ or ‘wonderful’ and ‘horrible’.

The differences in constructional semantics are often partially compositional and
conceptually motivated through their components. An example for this is *there BE something ADJ about*. The invariant core of its meaning is that the theme that it describes is odd or unusual. The type of unusualness is derived from the type of theme. People who are described as unusual are often (although certainly not always) found to be strange or creepy (e.g. ‘there is something weird about this guy’). Unusual statements are often suspicious and likely untrue (‘there is something fishy about their claims’). Places, especially locations in nature, on the other hand, can often be unusual in a good way (‘there is something magical about the forest in winter’). Conceptually speaking, it makes intuitive sense that unusualness can be a more positive concept for some entities than for others.

This leads to another important point about evaluative meaning in general. Thompson and Hunston (2000, p. 25) identify four evaluative dimensions: “good-bad, certainty, expectedness, and importance.” The authors then go on to say that “evaluation is essentially one phenomenon rather than several, and […] the most basic parameter, the one to which the others can be seen to relate, is the good-bad parameter” (ibid). This second point is not supported by the data presented in chapters 5 & 6. In many cases, other semantic parameters are fundamental, and the good/bad parameter secondarily results from them. This is particularly obvious for constructions that vary in their evaluative polarity but remain consistent in other semantic components. In the example of *there BE something ADJ about*, the invariant meaning concerns a sense of unusualness. The good/bad parameter is
compositionally derived from the combination of this sense of unusualness and the type of theme described. As such, the positive or negative evaluations emerge from other evaluative meaning components; they are not their antecedent.

This construction is not an outlier. As mentioned in chapter 6, the nominal extraposition construction *it BE ADJ the NP* is another example. It is used to evaluate something as extraordinary, which can either be positive (“It’s incredible the things that these young women are doing”) or negative (“It was despicable the way Pop treated the kids”). Further examples are intensifiers, which often occur with extreme evaluations on both sides of the positive/negative spectrum. This is not only true for English, as the example of ‘absolutely’ shows, but also holds up in German, as the study of *vollkommen* in chapter 8 demonstrates.

This does not mean that a negative or positive stance is never an important or salient part of meaning. For words like ‘horrible’, it may be reasonable to posit the expression of negativity at the core of its meaning. As the data described throughout this work, however, show this is not always the case. In some cases, it is an essential meaning component, in others, it is not present and in others yet, it is a consequence of the interactions of other meaning components. In any case, the good/bad parameter cannot be maintained as the basis or fundamental unit of evaluative meaning.

These insights give rise to a third central tenet of the present research. Semantic prosody is not a separate or even problematic and enigmatic phenomenon. Instead, it can be subsumed among *ordinary* context-dependent meaning and such does not
pose a challenge to a usage-based model of language. On a related note, the data also show that evaluation should not be described as connotational or peripheral. It is just one part of meaning and in many cases, it can absolutely be central to the semantics of a linguistic unit. An important theoretical implication of all this is that it casts further doubt on the idea that semantic prosody is a separate phenomenon that exists outside of and in addition to ‘normal’ meaning. On a global level, the good/bad parameter surfaces as a statistical tendency in language use and is an epiphenomenon of the interaction of the semantic processes described here.

One of the most commonly repeated claims about semantic prosody is that it is hidden from speakers’ intuitions or introspections. As I argued in chapter 7, there are many issues with this idea. First, the terminology is unclear and often remains undefined; terms like introspection and intuition are often conflated. Second, little to no evidence is put forward to support these claims. Third, there is an implicit assumption of a primacy of corpus data over speakers’ intuitions or introspective knowledge. This is an untenable position, based on the fact alone that corpus patterns never exist outside of our analysis. They are not real-world phenomena but are instead always filtered through our theoretical categories. There is no such thing as theory-independent evidence. In this case, the crucial point is the level of abstraction: The corpus data in the previous chapters did not establish semantic prosody as a category in its own right, so it is highly questionable that speakers would have mental representations of it (more on the question of mental representations in the next
Despite these issues, a comparison of elicited judgments and corpus data showed that speakers were, in fact, able to arrive at valence scores for linguistic items that correlated relatively well with corpus data. It is, however, possible that these judgments were made ad hoc and derived from general knowledge of distributional properties. Similarly, a review of the psycholinguistic literature on semantic prosody showed that, just as the corpus data, experimental data are not able to establish semantic prosody as its own category, separate from collocational phenomena such as semantic preference.

The idea that semantic prosody is a hidden category has always been considered one of its main selling points, setting it apart from other linguistic phenomena and presenting it as a particularly intriguing riddle. The data in chapter 7, however, suggest that there is nothing special about semantic prosody in this regard. For one thing, it is clearly not completely hidden. For another thing, as a review of studies comparing corpus data and psycholinguistic experiments shows, speakers’ knowledge of distributional phenomena varies widely, and it is not clear how these results should be interpreted.

9.1.2 Common questions about semantic prosody

Within the framework laid out here, we can now propose answers to specific questions about semantic prosody, including open and controversial issues. The question what
*mixed prosodies* are, for instance, easily solved within a constructional account of meaning. The basic idea of mixed prosodies is that the same word can be associated with different semantic prosodies. Schmitt (2005) found that *bordering on* tends to occur as part of negative evaluations (‘bordering on ridiculous’), a tendency does not apply to the verb ‘border’ per se. Louw (1993) noticed that ‘build up’ could occur in both, positive (‘build up confidence’) and negative contexts (‘frustration built up in him’). Within a constructional approach, these differences are simply examples of context-dependent meaning.

*Bordering on* is an idiom that describes something being taken to an extreme; this extreme is often negative (‘obsession’, ‘contempt’, harassment’) but can also be positive (‘reverence’). *Build up* in transitive constructions describes people acquiring (mostly positive) traits through hard work (‘strength’, ‘confidence’, ‘muscles’), whereas in intransitive constructions it describes the accumulation or increase in intensity of inanimate objects or states. When it describes purely physical processes (‘pressure’, ‘heat’, ‘tension’), the evaluative dimension is not necessarily clear; when it describes human affairs or psychological states, they are often unfavorable (‘debt’, ‘stress’, ‘anger’). The semantics of *bordering on* and *build up* both exceed simple evaluative polarities. In both cases, the positive/negative dimension results from more complex evaluative semantics, which in turn, are derived from and are part of their overall meanings. Transitive and intransitive *build up*, in particular, are quite different in their overall semantics; the former prototypically describes the controlled
and purposeful creation of human characteristics while the latter deals with a typically uncontrolled accrual or escalation. This is another example that illustrates that semantic prosody can simply be accounted for as an ordinary component of meaning.

These two examples also introduce another question into the discussion: To what extent can semantic prosody be explained through conceptual structure? Several researchers have pointed out that ‘cause’ seems to lose its negative sense in scientific texts. As Hunston (2007a) notes, there is likely a conceptual motivation: “cause implies something undesirable only when human beings, or at least animate beings, are clearly involved” (p. 253). This explanation is supported by evidence but needs to be further explored and generalized. Transitive ‘cause’ tends to be negative when humans are either directly (‘cause cancer’) or indirectly (‘cause an inflation’) affected. The key factor appears to be volitionality and control; something acts upon people in a way that they cannot control and may not want. When humans are unaffected (‘cause a chemical reaction’) there is no clear evaluative meaning.

This pattern holds up throughout different examples. ‘Cause’ tends to be less negative in the infinitival complement construction than in the transitive construction and the causee retains more control in the former than in the latter. Transitive build up tends to occur with personal improvements because it describes a person purposely and willingly working towards a goal. Intransitive build up tends to describe undesirable states when a person is affected because it expresses and uncontrolled process. When it describes physical processes that do not affect people, there is no
clear evaluative stance. Other conceptual motivations exists as well, e.g. the *there BE something ADJ about* construction where a sense of unusualness can be construed as weird and threatening or alternatively as intriguing and enjoyable, depending on the theme it describes.

Another fundamental question is where semantic prosody comes from, or through which mechanisms it arises. The most prominent explanation in the research literature is *semantic transfer*, the idea that through frequent co-occurrence, collocates *rub off* on a node word, or *imbue* it with an *aura* of positive or negative meaning. While much work in usage-based linguistics support the notion that distributional information is part of our encyclopedic knowledge of words, the concept of semantic transfer is clearly too simple. First, it is within the lexical perspective and fails to account for constructional differences. Second, it commits to an evaluative dichotomy which is not supported by the data. Third, within a constructional context, semantic prosody is better understood as a paradigmatic rather than a syntagmatic relationship. To understand this, we need to consider the interplay between construction and words.

Words are context-dependent, and constructions impart meaning upon their collexemes. Constructions are often partially idiomatic, but they are also partially compositional, i.e. they receive part of their meaning through their collexemes. The mechanism with which constructions create context-dependent meaning can be understood through the paradigmatic relation between fellow *co-collexemes* (words that occur in the same slot of a construction): A word’s contextual meaning can
arise through a paradigmatic comparison to its co-collexemes, especially when the word’s invariant meaning is not clearly defined for the evaluative dimension. Consider, for example, the verb ‘talk’ in the V into Ving causative construction. By itself, ‘talk’ does not possess strong evaluative semantics. Its invariant lexical meaning is undefined or open for this semantic dimension. When placed in the V into Ving construction (e.g. ‘he talked him into buying that car’), the verb takes on additional meaning: It is not only about verbal communication anymore but also includes a sense of malicious deception and is used to bring about a situation that is unfavorable. This sense can be derived through a paradigmatic comparison to the collexemes in this slot, e.g. ‘trick’ or ‘deceive’.

Let us consider another example. Gabrovšek (2007, p. 13) points out that “to make an *unforgettable mistake“ is an example of an inappropriate semantic prosody. While the idea of a mistake that is impossible to forget seems perfectly reasonable in theory, a native speaker of English would not express it in such a way, because ‘unforgettable’ is typically used to describe happy memories, collocating with words such as ‘experience’, ‘moment’, ‘day’ or ‘memory’. A competent speaker would compare ‘mistake’ to other nouns that occur in this slot and conclude that it does not fit into this paradigm; it is highly marked and either infelicitous or constitutes creative language use with a specific and novel rhetorical purpose. Positing a prosodic clash, i.e. a mismatch in evaluative polarity is not sufficient – the lexical field of the collexemes is relevant as well.
This also helps us understand the relationship between semantic preference and prosody. While there is a fair amount of discussion about this relationship in the literature, it always remains theoretical and terminological. The orthodox position is that they are related but different. Their exact relationship remains ill-defined and their distinction is based on a priori assumptions, not on empirical data. Researchers claim that the two should not be conflated because by definition they mean different things – but they do not test if this distinction is actually mirrored in the data. In other words, semantic prosody is different from semantic preference because that is how they are defined. The former deals with dichotomous evaluative stances and the latter deals with the co-occurrence of words with semantic fields.

Some researchers make the argument that semantic preference and semantic prosody exist at different levels and have different functions. Partington (2004) describes semantic preference as an essentially collocational phenomenon, while semantic prosody relates to the attitudinal or evaluative component of what is said. He concludes:

“[S]emantic preference is a ‘narrower’ phenomenon – relating the node item to another item from a particular semantic set – than prosody which can affect wider stretches of text. They […] interact: the former, preference, contributes powerfully to building the latter, prosody; conversely, the latter dictates the general environment which constrains the preferential choices of the node item.” (Partington, 2004, p. 151)
It is not entirely clear what Partington means by this. In this context, he seems to understand semantic prosody as something like evaluative cohesion. Sinclair (1998) makes a similar point but goes a step further. In his account, it is semantic prosody that drives lexical choices: “the semantic prosody of an item is the reason why it is chosen, over and above the semantic preferences that also characterise it” (p. 20).

In both, Partington’s and Sinclair’s description, semantic prosody is seen as functionally different from semantic preference and the latter contributes to the former in some vague, undefined way. The problem is that defining semantic prosody through its function does not obviate the need for providing evidence in favor of it. If semantic prosody is real it needs to have a quasi-material linguistic basis. The present investigation of evaluative meaning beyond prosodic polarity, however, shows that semantic prosody is a result, or, an epiphenomenon of semantic preference. Across all the data discussed throughout this dissertation, both in my original research and in other works that I cited, there is no evidence that separates semantic prosody from semantic preference. Evaluative meaning is created as part of contextual meaning, which can be described through semantic preference. Semantic prosody results from these evaluative semantics.

All these findings dovetail into a simple question: What is the final verdict on semantic prosody? Should it be adapted to include constructional meaning and complex evaluative semantics, or has it been explained away and lost its merit as a linguistic category? The next section deals with this question.
9.2 Semantic prosody in a broader theoretical context

9.2.1 The primacy of context-dependent meaning

Throughout this dissertation, the descriptions of evaluative semantics assign a great deal of meaning to the context-dependent level. Within the usage-based model, this is not a radical proposal. In a usage-based account of language, linguistic structure emerges from language use. Speakers learn words in their contexts and with their context, retaining distributional information as part of their meaning. Meaning varies across contexts and the lexical meaning of a word is often considered to be its invariant core meaning, consisting of the commonalities that are left over when all context is eliminated. As Langacker (2000, p. 35) puts it, “a lexeme is not precisely the same in all its environments. Since elements are always shaped by the contexts in which they occur, it is only by abstraction away from contextual variation that a constant representation emerges.” He then goes on to say:

“[A] variant enters into a kind of ‘ecological system’ with its structural context and does not necessarily exist outside that habitat. I am suggesting that these context-dependent variants may be more fundamental than the context-neutral schematizations we regard as primary.” (Langacker, 2000, p. 35)

A result of the primacy of context-dependent meaning is also that when lexical meaning is considered in isolation, words often show variation in meaning. According
to Langacker, this is the reason why it is so problematic to determine “whether an expression is truly ‘ambiguous’, so that we must indeed posit two senses, or whether it is only ‘vague’, in which case there may be just one” (Langacker, 2000, p. 36).

Partington (2004, p. 154) makes a similar point to argue that evaluative semantics are not peripheral or connotative: “evaluative meaning is just as ‘core’ as denotation, perhaps more so given the protean nature of the denotation of most words, i.e. its high dependency on use in context, as underlined in post-Firthian lexical studies.” As corpus evidence shows, evaluative meaning can also be ‘protean’ in nature; a word’s evaluative meaning can vary tremendously across different grammatical contexts. This is not the case for a word like ‘pig-headed’ but it is certainly the case for word like ‘persistent’ or ‘border(ing on)’.

Langacker and Partington are not alone in their accounts of the context-dependent nature of meaning. In an article titled “I don’t believe in word senses”, Kilgarriff (1997) criticizes traditional accounts of word senses. He reviews data from corpus linguistics and psycholinguistics and contrasts them with introspective arguments (e.g. grammaticality tests), concluding that word senses only exist as “abstraction from clusters of corpus citations” (Kilgarriff, 1997, p. 91). Much research comes to similar conclusions, emphasizing the importance of context and the impossibility of considering words in isolation, (Kintsch, 1988; McClelland et al., 1989; Berkum et al., 2003; Hauser, 2017, inter alia).

A particularly radical take on mental representations involved in word meaning is
the *ad hoc cognition* (AHC) framework put forward by Casasanto and Lupyan (2015). The notion of ad hoc categories was coined by Barsalou (1983), who describes them as categories that are constructed ad hoc, in situations where no *common categories* (categories that are well-established in memory) are available. While there has been some previous research on the role of ad hoc categories in lexical meaning (e.g., Allott and Textor, 2012), Casasanto and Lupyan (2015) claim that all, not just some *concepts*, *categories*, and *word meanings* (CC&Ms) are constructed ad hoc. According to this view, words are infinitely polysemous, there is no such thing as context-independent meaning and the notion of an invariant core meaning is an illusion. Word meanings are never accessed, but instead constructed on the basis of various contextual retrieval cues:

> “According to the AHC framework, all concepts, categories, and word meanings are constructed ad hoc and differ from one instantiation to the next, within and between individuals and groups. CC&Ms are inseparable from their contexts, and are shaped by the contexts in which they are instantiated on every time scale, from the millisecond to the lifetime.”

(Casasanto and Lupyan, 2015, p. 561)

Casasanto and Lupyan (2015) note that similarities across different instantiations of CC&Ms may lead to the emergence of some stability in meaning but there is no actual reality to them:

> “[C]oncepts, categories, and word meanings only exist as ‘analytic fictions
created by those who study them’ (Barsalou, 1987, p. 119), rooted in folk-psychological notions (e.g., concepts are entries in a mental encyclopedia, word meanings in a mental dictionary).” (Casasanto and Lupyan, 2015, p. 544)

There are two problems with AHC as it is presented in its current form. First, while the authors talk at length about the physical, social, linguistic and biological cues that activate concepts, there is no description of what kind information is stored in the long-term memory. We find statements such as “instantiating a concept is always a process of activating an ad hoc network of stored information in response to cues in context” (Casasanto and Lupyan, 2015, p. 546), but this ‘stored information’ is never described at any point in the paper.

This leads to the second issue: The authors only discuss word meaning. They reject the notion of conventionalized form-meaning pairings, explicitly disagreeing with Langacker and demonstrating a lack of awareness of the emphasis on context-dependent meaning in Cognitive Grammar and other constructional approaches. These two issues, the lack of any description of linguistic knowledge and the emphasis on word meaning raise an important question: If words are not stored as isolated items, what is being stored: statistical information about language use events, abstractions from language sequences, exemplars of linguistic items on various levels along with their contexts? Some approaches to these questions have been formulated (e.g. cf. Elman (2011) for a proposal about how lexical knowledge
could be stored without a dictionary-like mental lexicon) but there are certainly no
definitive answers. A version of AHC that discusses this question would perhaps still
be at odds with much of linguistic orthodoxy but may find much common ground
with construction grammar.

9.2.2 Do speakers have mental representation of semantic prosody?

The discussion of the context-dependent nature of meaning in the previous section
naturally segues into another question: What exactly do we mean when we talk about
mental representations or the mental lexicon?

The notion of mental representations is used by many linguists, but it is rarely ever
discussed in depth. It is undoubtedly a tremendously complex concept that is posited
at a high level of abstraction, far removed hard empirical evidence. Even though there
is a great amount of academic inquiry into the mental lexicon (cf. Aitchison, 2012, for
an overview), much remains unknown. As such, more care should perhaps be exercised
when making strong statements about mental representations. This is a criticism that
could be leveled at much work in linguistics, not just the study of semantic prosody.
It is, however, by no means a rejection of a cognitivist approach to the study of
language. Language is not an abstract, disembodied symbolic system that is best
described through formal logic, such as lambda calculus or phrase structure rules. It
is a human endeavor and should be studied as such. Numerous cognitive processes
have been found to be relevant to language, e.g. automatization, priming, analogy,
selective attention, exemplars, metaphor, metonymy, schematization, categorization, figure-ground segregation and social cognition. Knowledge of processes such as these should be used in linguistics because they deepen our understanding of language and, in return, because language is often an important source of evidence for them. The problem arises when linguists make psychological claims that go beyond what their data warrant.

The question is then, how far can we go in our interpretation of the data and in what way do our theoretical generalization correspond to the generalizations speakers make? Corpus evidence shows that many constructions have probabilistic tendencies to fill their open slots with particular words, i.e. there are relationships of attraction or rejection between constructs and collexemes. This is a proximal description of the data – all we do is use a statistical measure to discuss the actual raw textual data. The next step is to characterize constructions by the types of collexemes they take, which involves grouping collexemes into semantic classes. While this analysis is slightly more removed from the data, it is still unproblematic. We know that these classes are real because they can be derived from the data by statistical methods in a purely bottom-up manner, as demonstrated by the cluster analysis in chapter 5. A similar approach is also found in Gries and Stefanowitsch (2010), where the authors use a hierarchical clustering algorithm to group together collexemes they previously identified within constructions. Again, they arrive at clusters that closely correspond to the semantic fields a human researcher would identify. This information can be
used to provide a qualitative description of the constructional semantics.

The next analytical step that is made by many researchers is to claim that speakers retain mental representations of the evaluative valences they encounter in language, giving rise to semantic prosody as a psychologically real and relevant category. As Partington (2004, p. 150) concedes: “prosody is at a further stage of abstraction than preference. In fact, semantic preference generally remains relatively closely tied to the phenomenon of collocation.” This step may be more problematic than many proponents of semantic prosody seem to think since it involves numerous strong and untested assumptions about the human mind. Even if we do not go as far as Casasanto and Lupyan (2015), it seems reasonable to ask whether such categories are real or indeed just “analytic fictions created by those who study them” (Barsalou, 1987, p. 119).

Based on all the evidence and arguments presented throughout this work, there is no basis to postulate semantic prosody as its own distinct mental category. It is not supported by corpus evidence or psycholinguistic experiments and is theoretically problematic. The next section deals with the question whether it may still be useful to main semantic prosody as linguistic category for practical reasons.

9.2.3 Is semantic prosody a useful concept?

There is a famous aphorism in science that all models are wrong, but some are useful (cf. Box, 1979). Scientific models are by definition not exact representations of reality
but necessarily involve some measure of simplification and idealization. A related concept that is sometimes evoked by linguists is that of the *useful fiction* – essentially a simplification that is technically false but that is made because it is practical and convenient (cf. Hall, 1965). Déchaine (2005), for example, describes the boundaries between linguistic subfields such as phonology, morphology, syntax and semantics as a useful fiction. Wardhaugh (1968) believes that phonemes not only do not exist but that they are also not always “a particularly useful fiction” (ibid, p. 434).

As these two examples show, the merit of linguistic descriptors is sometimes not only considered based on their veracity but also on purely utilitarian terms. The question then is whether semantic prosody is a useful fiction. In other words, even we do not accept it as a psychologically real category, should it still be retained as a useful linguistic concept? Since this question does not deal with veracity but instead with usefulness, it is the usefulness of semantic prosody that needs to be evaluated.

Usefulness is always a matter of context of application. Let us first consider the context of academic discourse. Can semantic prosody simply be used as a *heuristic device*? This term requires a brief explanation because I am using it in a different sense than in chapter 7, where it described a rule-of-thumb, a cognitive shortcut used in problem-solving that is not optimal but often practical and sufficient (though it can introduce biases). In this chapter, the term ‘heuristic’ is meant to be understood in the context of philosophy of science. It is a descriptive concept, proposed on the level of analysis and with the sole purpose of aiding the analysis. It does not claim
to reflect underlying structures of reality or explain causal mechanisms. In the case
of semantic prosody, using it as a descriptive heuristic is very similar to Sinclair’s
first remarks on the phenomenon: Some words tend to collocate with other words
that appear to carry negative evaluations. No further claims are made as to whether
this is a psychologically real category or just an epiphenomenon of other linguistic
mechanisms that simply surfaces in the corpus data.

Sinclair’s (1987) remarks were certainly useful. They opened a new field of research
in pointing to a linguistic phenomenon that had previously not been studied in a
systematic manner. But how long do heuristics or useful fictions remain useful?
Haugen and Dil (1972) makes this point about another useful fiction in language:

“The concept of language as a rigid, monolithic structure is false, even if
it has proved to be a useful fiction in the development of linguistics. It is
the kind of simplification that is necessary at a certain stage of a science,
but which can now be replaced by more sophisticated models.” (Haugen
and Dil, 1972, p. 325)

Simplification is a natural part of science, but so is progress. As our understanding
of an issue becomes more mature, concepts that were once useful are superseded by
more mature theories. I believe this point has been reached for semantic prosody, at
least in the way it is usually defined. It is a heuristic that describes patterns that arise
in text. It has little explanatory power because it does not identify any underlying
mechanisms. Therefore, semantic prosody has outlived its usefulness.
The next application of semantic prosody that deserves consideration is second language learning. Numerous studies have concluded that L2 speakers are often not aware of semantic prosody. However, L2 speakers often lack any kind of distributional knowledge, be it collocational, constructional or related to semantic preference. A lexicon-and-grammar approach that lists dictionary definitions of words and teaches grammatical rules without providing sufficient exposure to language in usage is clearly suboptimal. Fortunately, there are approaches to second language teaching that include the use of corpora in the classroom (cf. Wichmann and Fligelstone, 2014, for an overview). The various methods range from data-driven learning that is guided by teachers (Johns, 2002) to corpus-based discovery learning that emphasizes the autonomy of students (Bernardini, 2002). Other approaches promote the use of parallel corpora, especially in translation studies, e.g. Oster and van Lawick (2008), Pearson (2014) and Aston (1999), who notes that parallel corpora “may help learners to distinguish between different contexts of use, and reduce their tendency to think in terms of one-to-one equivalence” (p. 300).

Milojkovic (2013) reports on a method that specifically aims to teach students understanding of distributional knowledge such as collocations or semantic prosody through the use of corpus data. Based on Louw’s Contextual Prosodic Theory (Louw, 2000; Louw and Milojkovic, 2016) students are taught basic corpus linguistic methodology and are presented with concordance lines in the KWIC format to help them learn contextual meaning of the node words. Stewart (2009a) describes a similar
program that is taught to final year translation students. These approaches seem promising, but it is not clear that a specific focus on semantic prosody is needed, especially within an approach that carries the abstraction of evaluative meaning to the level of evaluative polarity. Instead, it may be preferable to allow learners to naturally acquire full evaluative semantics in a similar way to native speakers.

9.3 Further avenues of research

The aim of this dissertation was to provide a relatively comprehensive reanalysis of semantic prosody. Nevertheless, there are some interesting issues that were outside the purview of my investigation and have not been addressed here so far. In this section, I briefly point out these desiderata and then discuss some recent developments in methodological approaches.

One issue concerns the diachronic processes that give rise to evaluative meaning. Much of the foundational work into semantic prosody describes the concept of semantic transfer as a diachronic process: Through frequent co-occurrence with evaluative words over time, linguistic items become loaded with evaluative meaning (cf. Louw, 1993; Stubbs, 1995a; Bublitz, 1996; Channell, 2000, inter alia). As Whitsitt (2005) points out, these claims are made without presenting any diachronic data. There is still astonishingly little research into the diachrony of semantic prosody (Smith and Nordquist, 2012, is a rare exception). Even if we reject the traditional notion of semantic prosody, as I have suggested, the evaluative meaning that emerges through
patterns of co-selection is still subject to diachronic developments. The inclusion of the constructional level adds an additional layer of complexity to this, making it a promising subject for future research.

A similar case can be made about the lack of in-depth discourse-analytical studies on semantic prosody. Many researchers emphasize the communicative role of semantic prosody in discourse. Nonetheless, most of its research is either done within literary studies or corpus linguistics. There do not appear to be any studies that investigate how semantic prosody plays out in linguistic interactions. The approach I have suggested here considers semantic prosody to be an ordinary part of meaning, but this does not obviate the need to study it within an interactional framework.

Other future developments concern improvements in research methodology, both in term of data retrieval and data analysis. In recent years, proposals have been made on how to automate linguistic data analysis processes, both relating to construction grammar and semantic prosody. These efforts are driven by two different motives.

First, there is the idea that algorithms may become better at analyzing linguistic data than people. The goal is to remove the element of human subjectivity. Dilts and Newman (2006) suggest a lexicon and rule-based approach to classifying semantic prosody with the aim to “explore prosody with less dependence on a researcher’s subjective, evaluative judgments” (p. 234). The authors formulate an algorithm based on earlier work by Osgood et al. (1957) who aimed to establish a scientific way of measuring meaning. Osgood et al. (1957) assembled 76 antonymous adjective pairs
to exhaustively describe the semantic space of adjectives and studied how subjects used them to describe various concepts. Dilts and Newman (2006) use the evaluations established for these adjectives to calculate the valence of the nouns they co-occur with.

While this is an interesting approach, there are several issues with it. First, it only describes lexical meaning, the insufficiency of which I have argued throughout this dissertation. This is particularly astonishing because the authors use collexeme analysis to rank the attraction of adjectives to the attributive slots of the noun phrase. The results of the collexeme analysis are then applied to the entire lexical meaning of the word, even though they do not study the word in any other contexts. Second, the analysis is limited to nouns and is not applicable to any other word classes (or larger linguistic items). Third, solely using collocates to describe a word’s distribution is problematic. The authors include the example of the word ‘joke’ which tends to co-occur with negative adjectives, such as ‘sick’ or ‘cruel’ resulting in a negative valence score. Fourth, as this example shows, a description of evaluative meaning in terms of a positive/negative polarity is not sufficient.

The second goal in improving corpus methodology concerns the efficiency of the analysis. Computers can perform analytical steps in seconds that would take human researchers hours, days or even longer. This is not a simple matter of convenience but instead enables research on a much larger scale, covering more information by orders of magnitude. Much of the research presented here makes extensive use of computa-
tional tools but still requires a great amount of manual post-processing. In the case of semantic prosody and construction grammar, a fully automatized methodology would allow researchers to ‘map the field’, i.e create a comprehensive dictionary of all relevant linguistic items in the entire language. Louw and Chateau (2010) regrets that so far, no means of exhaustively identifying all semantic prosodies have been put forward:

“An area of neglect (caused by genuine puzzlement over a number of years) in the work on semantic prosodies has been the failure of investigators, this author included, to propose some means of finding all of the semantic prosodies in the language, and that, preferably by computational means.”

(Louw and Chateau, 2010, p. 59)

This appears to be largely a result of the common definition of semantic prosody. All linguistic items have distributional contexts and many of them are evaluative. Far fewer of them may express their evaluative semantics in such a that they would give rise to what is traditionally regarded as semantic prosody, i.e. in the form of an evaluative polarity on the lexical level.

A promising new approach that aims to improve both the speed and the quality of constructional analyses is put forward by Tang (2017), who introduces a new methodology for retrieving constructions, called ‘Lexeme-based collexeme analysis’. The author describes a machine-learning-based tool that automatically retrieves the typical constructions of any given lexeme and performs collexeme analyses for all of
them. The author claims that his method is not only more efficient than a manual approach but also superior in its accuracy. A caveat is that the quality of machine learning models not only depends on the sophistication of their algorithms but to a great degree also upon the quality of their input. While models are improving at an impressive rate, skepticism about the accuracy of automatically parsed syntactic trees is currently still warranted. Nonetheless, Tang’s (2017) results are encouraging, and the coverage of data that his method promises may open new avenues of research within quantitative corpus linguistics.
Bibliography


Arnaud, P. J. (1990). Subjective word frequency estimates in l1 and l2. In Meet-
ing of the World Congress of Applied Linguistics sponsored by the International Association of Applied Linguistics. ERIC.


Halliday, M. A. (1993). Quantitative studies and probabilities in grammar. In Hoey,


Perek, F. (2015). *Argument Structure in Usage-Based Construction Grammar: Ex-


Schmid, H.-J. (2010). Does frequency in text instantiate entrenchment in the cognitive


