Beyond modal idioms and modal harmony: a corpus-based analysis of gradient idiomaticity in MOD + ADV collocations[1]

SUSANNE FLACH ⓘ
Université de Neuchâtel
(Received 12 February 2018; revised 18 June 2020)

How do we know that *would rather* and *may well* are more idiomatic than *would well* or *will really*? Can this intuition be measured systematically in usage data? Traditionally, modal idioms such *had/'d better*, *would/'d rather* or *might (as) well* are seen as distinct from more compositional collocations, which may be modally harmonic (*could possibly*, *will probably*) or not (*could also*, *might even*). Yet the collocation of modal auxiliaries + adverbs (MOD + ADV) is more complex than suggested by a binary classification into idioms and non-idioms. This article uses data from COCA and the method of collostructional analysis to show that the difference between qualitatively distinct types of MOD + ADV is a matter of degree. Modal idiomaticity should be seen as gradient along a continuum from strong association (*would rather*) to strong dissociation (*would well*). The results support assumptions that statistical information about the collocational behavior of modal auxiliaries is a cue for the scope of adverbial modification and is thus an important aspect of speakers' knowledge of modal meaning. The study contributes to recent approaches to modality from a 'combinatorial' perspective, which recognizes the importance of the lexical environment in core areas of grammar.

Keywords: modal auxiliaries, idioms, adverbial modification, collostructional analysis, corpus linguistics

1 Introduction

What distinguishes collocations of modal auxiliaries and an adverb (MOD + ADV) such as *would rather*, *may well* or *could possibly* from *would well*, *can likely* or *may rather*? After all, the bigrams are structurally identical and all are possible and attested in language use. Yet there is a notable asymmetry in how we judge them intuitively: modal idioms such as *would rather* or *might as well* are more idiosyncratic than *could possibly*, and all of them sound more natural than *can well*. These examples give a first impression of the complexity of post-modal adverbial modification which goes beyond a binary distinction into idioms (*would rather*) and non-idioms (*could possibly* or *would well*). Rather, MOD + ADV bigrams seem to lie on a continuum from strong association to strong dissociation (see Wulff 2008). In other words, some modal auxiliaries and adverbs share a closer bond than others, which may explain our intuition about the examples above. From a corpus-linguistic perspective, this raises the question of how

we can approach shades of MOD + ADV idiomaticity in a bottom-up fashion and whether patterns of (dis)preference in usage data tell us something about the scope of adverbial modification.

A collocational study of MOD + ADV bigrams offers a quantitative perspective on the interaction of modality and adverbial modification, which are two highly polyfunctional and context-dependent areas of grammar that are notoriously difficult to describe. We will focus on adverbs in post-modal position, as in example (1), which is more specific than general adverb placement, illustrated in example (2):

(1)   (a) You**'d better** get this sorted out soon.   'modal idiom'
      (b) This **might possibly** be the best show we've seen.   'modal harmony'
      (c) He **would later** become president.   'compositional modification'
(2)   (a) **Perhaps** it was invented by the Greeks.
      (b) They are **probably** home by now.

For the purpose of this article, we distinguish three broad types of MOD + ADV bigrams that we will refine below. They can take the form of highly idiosyncratic modal idioms (1a); they can be modally harmonic, that is, the modal auxiliary and the adverb agree in their modal value (1b); or they can combine with general-purpose adverbs (1c). These types form a reasonably coherent class; the differences, as we will see below, are gradual rather than categorical. Crucially, they are sufficiently distinct from general forms of (sentential) adverb placement in (2).

The heterogeneity of adverbial modification has been widely studied (Greenbaum 1970, 1974; Jacobson 1975; Simon-Vandenbergen & Aijmer 2007; Simon-Vandenbergen 2008; Celle 2009; Aijmer 2013; to mention only a few). But while adverb placement appears unsystematic and free, it is not random (see the overview in Nuyts 2001: chapter 2). Hence, we assume that post-modal adverbial modification is also non-random.

MOD + ADV collocations are understudied in the modality literature, although they are occasionally mentioned in passing (e.g. Coates 1983; Perkins 1983; an exception is Hoye 1997). Some types receive considerably more attention than others, especially in cases of high idiosyncracy. For instance, the modal idioms *'d/would rather, 'd/had better* or *may (as) well* are comparatively well studied (e.g. Jacobsson 1980; Mitchell 2003; van der Auwera & De Wit 2010; Denison & Cort 2010; van der Auwera, Noël & Van linden 2013; Traugott 2016). In addition, collocations with epistemic adverbs have been discussed in a number of theoretical or applied-descriptive contexts (e.g. Coates 1983; Hoye 1997; Xiao 2009 on 'modal harmony'; for 'modal concord' in formal frameworks, see Geurts & Huitink 2006; Zeijlstra 2007; Grosz 2010). By contrast, collocations with general-purpose adverbs (e.g. *still*, *also* or *just*) only attract occasional comments (see Hoye 1997).

The divergence in the amount of attention can be explained with reference to qualitative properties, particularly to the higher degree of idiosyncracy of modal idioms. At the same time, however, 8.5 percent of modal auxiliaries in COCA are directly followed by an

adverb, even if we exclude the negation particles *not/n't* (see section 3). Hence, MOD + ADV is a frequent phenomenon that deserves a closer look.

As we assume no a priori distinction between the types in (1), a number of questions can be addressed. For instance, can we distinguish modal idioms and other forms of post-modal adverbial modification on the basis of a quantitative corpus analysis? How can corpus data be used to measure cohesion between modal auxiliaries and adverbs? Does a probabilistic approach reflect our intuition that *may well* and *could only* are more idiomatic or 'natural' than *would well* and *can rather*? And do distributional patterns inform hypotheses on the scope of adverbial modification?

These questions touch upon idiomaticity in a way that goes beyond the notions idiom, idiosyncrasy or harmony. First, they force us to critically examine the implicit assumption that the difference between idioms and non-idioms is categorical. From a usage-based perspective, which is well suited to capture gradual differences in the idiomaticity of semi-fixed multi-word expressions (Wulff 2008), the argument is that MOD + ADV bigrams are situated along a linguistically meaningful continuum. Second, the questions address the probabilistic nature of the cohesion between a modal auxiliary and an adverb: stronger semantic cohesion should correlate with stronger statistical association at one end of the continuum. Third, they permit a discussion of the implications of an *absence* of cohesion at the other end of the continuum. That is, statistical dissociation provides cues as to which unit of meaning an adverb 'refers to'. This information taps into general questions about the scope of adverbial modification.

This article presents quantitative evidence for an idiomaticity continuum by means of a collostructional analysis of all MOD + ADV observations in the *Corpus of Contemporary American English* (COCA; Davies 2008; see section 3). Although Construction Grammar (CxG) is not a necessary framework for this type of analysis, it is well suited conceptually and methodologically. On the one hand, constructionist approaches have long abandoned the idea of categorical distinctions in lexis and syntax; rather, they assume gradience on many levels of linguistic representation (Langacker 1987, 2000; Goldberg 1995, 2006; Wulff 2008; Hilpert 2014; for a methodological perspective, see Stefanowitsch & Gries 2003). On the other hand, while modal idioms are clearly form–meaning pairings in the CxG sense, the constructional status of modal auxiliaries is less clear, since they resist a straightforward integration into a framework that puts greater emphasis on slot–filler constructions (Hilpert 2016; see the overview in Cappelle & Depraetere 2016a). Thus, the aim of this study is to complement recent usage-based approaches to modality from a distributional perspective, which highlight the modal auxiliaries' associative connections to other items in the network. This article argues that probabilistic information about combinatorial patterns is part of speakers' knowledge in general and of modal constructions in particular (see Hilpert 2014, 2016).

The article is organized as follows. Section 2 discusses the phenomenon in more detail; it pays particular attention to the strengths and limits of categorical distinctions. Section 3 describes data and method. Section 4 presents the results; the implications of association and dissociation are discussed in section 5. Section 6 concludes with some general

remarks and argues (i) that idiomaticity in MOD + ADV is gradient across qualitatively different types and (ii) that speakers make use of this distributional information as part of their constructional knowledge.

2 Post-modal adverbial modification

Extending the overview in (1), we can distinguish four types of MOD + ADV collocation that range from highly idiosyncratic, traditional idioms to seemingly free, predictable combinations. While the types appear to be conceptually discrete, many of their exemplars defy a categorical classification.

The 'modal idioms' in (1a) can be further subdivided into two types, depending on their internal structure. They include *had/'d better* or *had/'d sooner* on the one hand, as in (3), and *would rather*, *may well* or *might as well* on the other, as in (4):[2]

(3) (a) "Then maybe you**'d better** come out . . . and attend to the job yourself." (FIC, 1994)
     (b) I said, 'Let them shoot. I**'d sooner** die by a bullet as die by an explosion.' (NEWS, 1990)
(4) (a) In a city where people **would rather** get wet than admit that it is raining outside, we permitted ourselves the luxury of beginning from truth, not politics. (NEWS, 2004)
     (b) . . . we **may well** be past the point where hearts and minds are winnable. (SPOK, 2004)
     (c) If you are going to be a Mennonite you **might as well** join with me and be a Cub fan. (ACAD, 1992)

The so-called 'comparative modals' *had/'d better*, *had/'d best*, *had/'d sooner*, *had/'d rather*, as in (3), are most idiom-like (Jacobsson 1980; van der Auwera, Noël & Van linden 2013). Unless *'d* is analyzed as *would*, they are the only type without a modal auxiliary. Their classification as modal follows from their functional-semantic properties (e.g. 'peripheral modals', van der Auwera, Noël & Van linden 2013; 'marginal modals', Traugott 2016; see also 'modal indicative', Declerck 2009). They satisfy the textbook definition of idioms, since the deontic or optative import is a property of the combination rather than the additive sum of the parts (Denison & Cort 2010; van der Auwera, Noël & Van linden 2013; Traugott 2016). They share non-compositionality with the examples in (4), which do contain a modal auxiliary. Yet the examples in (4) are clearly idiosyncratic and therefore idiom-like: the meaning of *may well* or *might as well* is non-predictable from the parts: omitting the adverb leads to a considerable change in meaning (*you might as well join me* vs *you might join me*).

The idiosyncrasy of *would rather*, *may/might (as) well* or *should sooner* motivates the treatment as modal idioms (Mitchell 2003). Their classification is sometimes based on the adverb (as BETTER, SOONER or RATHER idioms; see van der Auwera & De Wit 2010; Traugott 2016) and reflects functional-semantic overlaps between the two groups (see overview in van der Auwera, Noël & Van linden 2013). Depending on the aim of an analysis, it makes sense to single out or collapse certain types. That is, it is a matter of focus whether *would*

---

[2] All examples, except paraphrases below, are cited from COCA (Davies 2008).

rather is on par with *had/'d rather* in a class of RATHER idioms. Yet this can be taken as a symptom of gradience, since it illustrates the difficulty in delimiting idioms, which in turn provides evidence for degrees of idiomhood along an idiomaticity cline (see Wulff 2008: chapter 1 for discussion).

The two other groups appear compositional by comparison. The first are 'modal harmony' bigrams (*could possibly*, *will probably* or *might conceivably*). A defining property of modal harmony is the agreement in the modal value between modal auxiliary and adverb, such as POSSIBILITY in *could possibly* or PREDICTION for *will inevitably*:

(5)   (a)   One wonders how they **could possibly** proceed in such an ambivalent manner. (ACAD, 1990)

(b)   When officials have absolute power, they **will inevitably** become corrupt. (ACAD, 1990)

Modal harmony illustrates that idiomaticity and (relative) compositionality are not mutually exclusive. For instance, *could* and *possibly* share a strong bond with unit-like status. Yet the modal and the adverb are in a reinforcing relationship, so that the meaning of *could possibly* is closer to the sum of its parts than the meaning of the traditional idioms. In other words, *could possibly* is idiomatic (or conventional), but less 'idiom-like' (or idiosyncratic) compared to *would rather* or *'d better*.

Harmony effects are sometimes referred to as 'synergism' or 'concord' (Halliday 1970: 331; Lyons 1977: 807; Coates 1983: 45–6; Hoye 1997; for formal analyses of 'concord', see Geurts & Huitink 2006; Zeijlstra 2007; Grosz 2010). The adverbs are described as 'modal adjuncts' (Halliday 1970: 330), 'epistemic adjuncts' (Huddleston & Pullum *et al.* 2002: 173, 767) or 'modal satellites' (Hoye 1997; Xiao 2009). The metaphor of a satellite reflects the special relationship between adverbs and modals; note that many studies in the context of modal–adverb collocation include adverbs in slots other than the post-modal position (Coates 1983; Hoye 1997; Xiao 2009). Since epistemic adverbs form a reasonably closed class (Nuyts 2001: 55), it appears straightforward to delimit this group. In actual practice, however, it is difficult to devise a comprehensive list of adverbs that fall under the concept of harmony.

The classificational problems are also evident in the final group, where modal auxiliaries are followed by general adverbs. In isolation, *only*, *also* or *always* have less or no modal import. But they attain considerable epistemic flavor, since, in combination with a modal, they evaluate the extent of PREDICTION (*will*) or HYPOTHETICAL PREDICTION (*would*):

(6)   (a)   I **can only** get more by denying an equal amount to others. (ACAD, 1990)

(b)   Still, people **will always** generalize from specific encounters. (MAG, 2009)

(c)   Petrov **would never** have described her as beautiful … (FIC, 1997)

In contrast to the types we discussed above, these collocations are more compositional. They add a temporal or interval qualification (*also*, *only*, *always*, *never*) and are clearly not modally harmonic.

This type is rarely discussed (but see comments in Hoye 1997). One reason may be that general adverbs are perceived as non-modal in isolation. They are implicitly assumed to

form compositional sequences, which may not be immediately relevant for a discussion of modality. On the other hand, they occur in idiom-like fixed expressions, where the adverb is an essential element, such as *I **can hardly** wait*, *I **would never** dream of it* or *we**'ll just** have to wait and see*. This cohesion, as well as their high frequency, warrants the inclusion of general adverbs in a study of MOD + ADV bigrams alongside idioms and modally harmonic collocations.

In sum, the four groups appear sufficiently distinct on conceptual grounds, yet their boundaries overlap considerably. In the sense that we refrain from drawing a sharp distinction between 'idioms' and 'collocation' (as in, e.g., Hoye 1997), the discussion is empirically more inclusive. At the same time, the restriction to the post-modal position is more *ex*clusive compared to studies with a much wider definition of a satellite (see Hoye 1997; Xiao 2009). However, the results from the more restrictive analysis permit informed inferences about the scope of adverbial modification beyond the post-modal position (see section 5).

The restriction has a methodological and a linguistic motivation. Methodologically, adjacent items are much easier to extract from usage data and require minimal manual intervention. This means that we can comprehensively exploit large corpora like COCA, where the number of MOD + ADV tokens runs into the hundreds of thousands. The linguistic motivation follows from the assumption that idiomaticity is gradient, covering a continuum from highly idiosyncratic idioms to idiom-like expressions to more freely combining sequences. Hence, a greater inclusiveness pays heed to the observation that neither idiom, nor harmony, nor compositionality or predictability capture the full extent of MOD + ADV collocation on their own.

From a distributional perspective, which assumes that statistical patterns are linguistically meaningful, we can formulate the following expectations. First, greater cohesion, or association, between a modal and an adverb coincides with higher degrees of idiomaticity and/or unit-status. While this is the case for modal idioms almost by definition, it is probabilistic and gradual for the remaining types. Second, conversely, the greater the dissociation, the less likely a combination has unit-status. In other words, in addition to (statistically significant) patterns of mutual preference, the method described below identifies (statistically significant) patterns of *dis*preference. Greater dissociation, i.e. the absence of a statistical relationship between the modal auxiliary and the adverb, has implications for the scope of adverbial modification.

Analyses of MOD + ADV in a collocational context are not new, but the phenomenon remains understudied. Hoye's (1997) study is the most relevant in this respect; it is concerned with descriptive frequency profiles of individual modal auxiliaries. The current study extends on Hoye's concept of 'collocability', but takes a paradigmatic and contingent perspective: it measures (dis)preferences *relative to all other (potential) combinations* in the same pattern. This can be seen as an operationalization of the 'collocational range' of a modal auxiliary (see Greenbaum 1974; Hoye 1997). The insights on combinatorial properties help to address questions of speaker knowledge of modal constructions (see Hilpert 2016).

3 Data and method

This section describes the data set and the method used to investigate it, i.e. Collostructional Analysis (Stefanowitsch & Gries 2003; Stefanowitsch & Gries 2005). CA is a well-established method for the investigation of lexis–construction interaction, but has not featured prominently in the modality literature (but see Hilpert 2011, 2016; Cappelle & Depraetere 2016b; for deontic modality more generally, see, e.g., Van linden 2010a, 2010b, 2012; for German, see Stefanowitsch 2009). The logic of CA is identical to collocational analyses in lexical semantics (Evert 2004), except that it focuses on the co-occurrence of two items *within* a pattern. Like other contingency-based methods, CA goes beyond raw frequencies and identifies positive and negative association, which cannot be inferred from raw frequency profiles.

3.1 Data: source and retrieval

Three queries were run on the mid-2015 offline version of the *Corpus of Contemporary American English* (COCA; Davies 2008).[3] The first query extracted the core modals as strings (unambiguously tagged 'vm', i.e. *can*, *could*, *may*, *might*, *must*, *should*, *shall*, *will*, *would*, *'d*, *'ll*). The adverb slot contained any number of adverbs (retrieving, e.g. *could also*, *might as well*, *will almost certainly*); negation in *not/n't* was excluded. The second query retrieved *had/'d better/best* and their modifications. In order to avoid duplicate matches with *'d* from the first query as well as ambiguity with a contracted *'d* in perfect auxiliaries, the second search excluded the modal tag (vm) and required an infinitive after *better/best*. The third query retrieved the string *had rather* with zero or more intervening adverbs; these results were manually cleaned. In the displays below, *'d* is represented as *'d(wd)* if retrieved from the first query (likely *would*) and as *'d(hd)* if retrieved from the second (likely *had*). The data and the query documentation are available at https://osf.io/f6azk/

The final data set contains 441,608 MOD + ADVobservations (436,436 for query 1; 5,147 for query 2; 25 for query 3). This amounts to about 8.5 percent of all modal auxiliaries in COCA (5,173,007). Since contractions are treated as distinct from their full forms, there are 13 types in slot A (*can*, *could*, *may*, *might*, *must*, *should*, *shall*, *will*, *would*, *'d(wd)*, *'ll*, *had*, *'d(hd)*). A total of 5,012 types occur in slot B: multi-token adverbs account for 65 percent of types, but only for 7.5 percent of tokens (the bulk of these are *no longer*, *as well*, *at least*, *very well*, *most likely* and *almost certainly*).

Table 1 provides an overview by frequency, from which it is clear that raw occurrence is not the most comprehensive indicator of cohesion, let alone of idiomhood. The top ten do not include a single modal idiom and only one modally harmonic bigram (*will probably*). Frequency lists mask strong association between low-frequency items. For instance, it is implausible to assume that the frequency of *can also* (12,756) reflects a higher degree of

---

[3] Excluding punctuation and missing tokens due to copyright restrictions, this offline COCA version contains roughly 445m tokens.

Table 1. *Top ten MOD + ADV sequences by raw frequency; COCA*

MOD + ADV collocation	Frequency
can also	12,756
would never	12,545
can only	9,986
will never	9,015
may also	7,304
'll never	7,088
will also	6,995
can still	5,467
would also	5,136
will probably	4,960

cohesion than *'d(hd) better* (3,561). Since both *can* and *also* are very frequent on their own, their co-occurrence is expected to be high by chance alone. Thus, in order to measure cohesion, the frequency of co-occurrence must be controlled for the frequency of its parts.

3.2 Method: Collostructional Analysis

Collostructional Analysis (CA) normalizes raw co-occurrence by taking into account overall frequencies of the units involves. For the current purpose, I used Co-Varying Collexeme Analysis (CCA). CCA is a variant of CA, which quantifies attraction or repulsion between the items in two slots of a pattern (Gries & Stefanowitsch 2004; Stefanowitsch & Gries 2005). In the context of MOD + ADV, we assume that each type in slot A (the modal) can in principle co-occur with each type in slot B (the adverb). If the combination of A and B were free, each combination would occur roughly as often as expected given the individual frequencies of A and B in the pattern. (C)CA calculates whether an observed value deviates from expectation; the corresponding test statistic is interpreted as a measure of attraction or repulsion between A and B.

To illustrate, we assess the level of attraction between the parts of the most frequent combination *can also* in the 441,608 MOD + ADV observations. We need four frequencies (see table 2): in COCA, *can* and *also* co-occur 12,756 times, *can* occurs 73,594 times without *also*, *also* occurs 34,788 times without *can*, and 320,470 MOD + ADV observations involve neither *can* nor *also*. Note that CCA ignores the *corpus* frequencies of *can* (1,054,081) and *also* (508,808), as well as the corpus total (this version of COCA: 444,797,856). In contrast to collocational methods (see Evert 2004), where measures are calculated across total corpus frequencies, CCA is based on the frequencies *within* a pre-defined pattern, such as MOD + ADV.

Based on the four conditions, the expected frequency of *can also* is 9,297.[4] Since this is lower than the observed value, *can* and *also* are positively associated. The log likelihood

---

[4] Multiplying the column total by the row total divided by the table total, i.e. $E = \frac{47544 \times 86350}{441608} = 9296.5$.

Table 2. *2 × 2 contingency table for* can also

	also	*¬also*	Row sum
can	12,756	73,594	86,350
¬can	34,788	320,470	355,258
Column sum	47,544	394,064	441,608

value (G^2)[5] for this table is 1,670, which is statistically significant at $p < .001$. This means that *can* and *also* are *significantly positively associated*. In other words, their co-occurrence is not merely a function of their individual frequencies, but suggests a linguistically relevant level of mutual cohesion.

To take the example of a traditional modal idiom, the association for *'d(hd) better* is influenced by two types of skew (see table 3): slot B is largely restricted to *better* and *better* is skewed toward *'d(hd)*. The expected frequency of 52 is not only much lower than the observed frequency of 3,561, the association score ($G^2 = 31,412$) vastly exceeds the value for *can also* ($G^2 = 1,670$). This is mathematical confirmation for the intuition that *'d(hd) better* is a more cohesive unit than *can also* despite its lower raw frequency.

Two remarks illustrate the additional advantage of association over raw frequency. CCA can be used to identify negative association and the absence of an association. As an example for the former, *would well* occurs only 14 times out of an expected 1,463 and is thus significantly dissociated ($G^2 = 3,113$, neg). For the latter, *must really* occurs 157 times out of an expected 172; since this difference is not statistically significant, there is no association in either direction ($G^2 = 1$; $p = .25$). Non-associated items are usually not discussed in the CA literature, but we will return to potential implications of this phenomenon in the context of MOD + ADV in section 5.

The procedure is repeated for each (potential) MOD + ADV type using appropriate software. All calculations reported below were performed using the collex.covar() function in the R package {collostructions} (Flach 2017). The output of CCA is a ranked list of MOD + ADV bigrams in descending order of attraction, which represents a continuum of (waning) idiomaticity. Note that dissociation is represented in a graph below as a negative value; since it is incorrect to represent a squared value (G^2) as a negative, this should be read as 'directed association strength' (see figure 1).

---

[5] In large data sets, G^2 is better suited for ranking than the p-value of the Fisher–Yates Exact Test, which is traditionally used in CA (Stefanowitsch & Gries 2003): p_{FYE} is 0 for the 33 most attracted and the 21 most repelled MOD + ADV in COCA, which prohibits rankings at both ends of the continuum (see table 5). G^2 is not subject to this problem.

Table 3. *2 × 2 contingency table for* 'd(hd) better

	better	¬*better*	Row sum
'd(hd)	3,561	234	3,795
¬*'d(hd)*	2,524	435,289	437,813
Column sum	6,085	435,523	441,608

Figure 1. Modal idiomaticity continuum (CCA); COCA

Table 4. *Sample input for Relative Entropy (H_{rel})*

	'd(hd)	*'d(wd)*	*might*	*can*	…
better	3,561	4	165	483	
best	234	0	169	753	
also	0	541	2,145	12,756	
never	0	3,714	1,068	3,466	
…	…	…	…	…	
H_{rel}	**.029**	**.348**	**.413**	**.477**	

3.3   Relative entropy

A related point for idiomaticity concerns the propensity of modals to combine with adverbs. To use an extreme comparison, *would* and *will* combine with vastly more adverb types than *'d(hd)* or *had* (by definition). This skew is a rough operationalization of Hoye's (1997) notion of the 'collocational range' of a modal auxiliary.

Relative Entropy (H_{rel}) measures the level of skew in a distribution. Imagine that modal A combines with six adverbs in a set $A = \{5, 4, 3, 2, 1, 0\}$ and modal B with the same set of adverbs as $B = \{7, 6, 1, 0, 0, 0\}$. Modal A is more evenly distributed, indicated by a higher Relative Entropy ($H_{rel} = .83$) than B ($H_{rel} = .50$). A maximally skewed modal $C = \{10, 0, 0, 0, 0, 0\}$ has a H_{rel} of 0. Thus, the higher a modal's H_{rel}, the more varied its collocational range. Overall frequency is not relevant for H_{rel}, although H_{rel} and frequency are correlated in the present context (see below).

Table 4 shows a snippet of the table for the calculation of H_{rel} (Gries 2010: 273). The rows represent all adverbs that have at least one significant positive relationship with one of the modal auxiliaries ($G^2 > 3.84$, $p < .05$).

4   Modal idiomaticity from a collostructional perspective

This section presents the results from two angles. First, it describes patterns in MOD + ADV and the continuum of idiomaticity, as determined by CCA. It is followed by a brief discussion of the modal auxiliaries' collocational ranges.

4.1   MOD + ADV *idiomaticity*

Of 65,156 possible MOD + ADV combinations, 11,374 are attested and 1,546 of these are significantly associated *or* dissociated at $p < .001$ ($G^2 > 10.83$).[6] Table 5 lists the top thirty attracted (left) and repelled (right) MOD + ADV bigrams, which we can interpret linguistically as the two ends of the idiomaticity continuum. Figure 1 visualizes the continuum and includes a random selection for non-associated types.

---

[6] This is a more conservative threshold given the size of the data set. At $p < .05$, roughly half of attested types are significantly associated or dissociated.

Table 5. *Top thirty most strongly associated (left) and dissociated (right)* MOD + ADV
combinations

	Associated				Dissociated			
	MOD + ADV	Obs	(Exp)	G^2	MOD + ADV	Obs	(Exp)	G^2
1	'd(hd) better	3,561	(52)	31,412	can never	3,466	(9,102)	5,890
2	'd(wd) rather	4,363	(344)	17,802	can rather	3	(1,524)	3,380
3	might as well	3,347	(195)	17,434	can probably	783	(3,191)	3,119
4	may well	4,532	(592)	13,856	would well	14	(1,463)	3,112
5	had better	1,221	(19)	9,753	can soon	35	(1,513)	3,051
6	can only	9,986	(4,432)	7,489	will rather	12	(1,320)	2,778
7	'll just	4,617	(1,355)	5,802	will well	17	(1,300)	2,687
8	'll never	7,088	(2,906)	5,492	would also	5,136	(9,062)	2,657
9	could possibly	2,197	(452)	4,909	may probably	6	(1,260)	2,597
10	will likely	3,186	(922)	4,848	will just	1,313	(3,675)	2,412
11	could easily	3,014	(907)	4,002	can well	150	(1,500)	2,313
12	might well	1,958	(383)	3,733	could rather	2	(1,061)	2,274
13	might even	2,338	(552)	3,660	can likely	30	(1,065)	2,098
14	may also	7,304	(3,671)	3,582	would easily	146	(1,269)	1,852
15	could hardly	2,326	(650)	3,437	'd(wd) also	541	(2,098)	1,841
16	could barely	1,775	(432)	3,170	will easily	123	(1,128)	1,653
17	would later	1,463	(387)	2,702	might probably	6	(814)	1,630
18	will soon	3,195	(1,310)	2,600	would better	155	(1,160)	1,593
19	'll probably	2,768	(1,019)	2,369	will even	523	(1,874)	1,587
20	must also	2,887	(1,122)	2,364	may always	88	(1,037)	1,564
21	would rather	3,304	(1,485)	2,280	would as well	13	(744)	1,522
22	might otherw.	828	(108)	2,238	can as well	23	(763)	1,489
23	can easily	2,990	(1,302)	2,237	will better	127	(1,030)	1,441
24	will always	4,377	(2,274)	2,024	may ever	51	(860)	1,414
25	would never	12,545	(8,872)	1,940	can eventually	178	(1,103)	1,390
26	may even	2,288	(855)	1,919	could also	4,048	(6,475)	1,317
27	will probably	4,960	(2,764)	1,867	might only	186	(1,130)	1,302
28	should also	3,626	(1,765)	1,832	could likely	40	(742)	1,279
29	will eventually	2,296	(955)	1,802	can ever	854	(2,179)	1,263
30	could almost	1,250	(355)	1,790	will as well	23	(661)	1,246

CCA identifies the traditional modal idioms *'d(hd) better*, *'d(wd) rather*, *might as well*, *may well* and *had better* as the top five. The most frequent combination *can also* is ranked only thirty-second by association. Conversely, the clearest idiom by intuition and statistical cohesion, *'d(hd) better*, is ranked only thirty-second by frequency. This is a plausible result, since we would expect idioms with higher degrees of idiosyncracy to have the strongest levels of statistical cohesion.

There are at least four additional patterns. First, while the modal idioms dominate the top of the list, one – *would rather* – is ranked rather low (rank 21) and much lower than the

contracted variant *'d(wd) rather*. Generally, contractions rank higher than their full forms, which is indicative of a stronger bias toward idiosyncrasy (see below).

Second, general adverbs not only outnumber epistemic adverbs, they are also relatively evenly spread across *only*, *just*, *never*, *easily*, *even*, *also*, *later*, *soon*, *otherwise*, *always* and *almost*. There is a tendency for historically and/or semantically related modal auxiliaries (*can*/*could*, *may*/*might*, *'ll*/*will*, etc.) to attract similar adverbs, but each auxiliary has at least one very salient 'satellite' in the top list that it doesn't share with its 'relative'. Examples of similar patterning of pairs include the positive association between *even* and *may*/*might* or *can*/*could* (all other modals disprefer *even*) or *easily* with *can*/*could* (dispreferred by all others others). An example for dissimilar patterning of pairs is *never*, which is positively associated with *'ll*, but not with *will*, while *also* is strongly attracted to *may*, *must*, *should* and *could* and much less strongly attracted to *might*, *shall* or *can*. These tendencies indicate that many bigrams are conventionalized chunks: if co-occurrence were random, we would expect general adverbs to be evenly distributed across (pairs of) modal auxiliaries. That said, it is a matter of definition in how far *easily*, *only*, *even* or *always* are actually non-modal – they *do* have epistemic, speaker-based import in combination with modals (e.g. in *this could easily be done* or *one might even argue*).

Third, epistemic adverbs are underrepresented among the top thirty (*could possibly*, *will likely* or *'ll*/*will probably*). Their relative absence from the top of the list is partly due to the lower frequencies of epistemic adverbs (CA favors frequent items). However, their systematic absence is also noteworthy with regard to modal harmony (see below).

Finally, as alluded to above, there is a conspicuous pattern with contractions: *'ll* and *'d(wd)* appear to have 'a life of their own'. Specifically, *will* and *'ll* combine with very different adverbs at the top or even have 'contradictory' associations. For example, *'ll just* is associated, but *will just* is dissociated. While this reflects oral discourse to some extent, it points to a broader pattern (which is also evident in the COCA-spoken data). The adverbs attracted to *'ll* cluster around expressions of intention and proximity (*just*, *never*, *both*, *ever*, *even*), while *will* attracts adverbs which signal the prediction of results (*likely*, *probably*, *eventually*, *undoubtedly*, *inevitably*, *ultimately*). For *would* and *'d(wd)*, the picture is less clear, but it is interesting to note that variations of MOD *rather* (*much rather*, *just rather*, *still rather* or *really rather*) are associated predominantly with *'d(wd)*. On the one hand, it points toward greater flexibility within the *'d(wd) rather* idiom; on the other, it parallels the contractions' stronger propensity to form idiom-like combinations.

There are three types of negative association at the other end of the continuum, which we will return to in detail in section 5. Some bigrams are dissociated for mathematical reasons. The obvious cases involve the adverbs *well*, *rather* and *better*, which are largely restricted to *may*/*might*, *'d*/*would* and *'d*/*had*, respectively. By the logic of CA, all other modals show a strong dispreference; they also sound decidedly odd (*can rather*, *will well* or *would better*). The second type concerns high-frequency combinations with general adverbs that are similarly affected by skews and also

repelled for mathematical reasons (*also*, *still*, *yet*, *easily*). However, these are *not* unidiomatic (*would easily*). The final type includes those we expect to be repelled by conflicting (modal) values: adverbs of prediction (e.g. *probably* and *likely*) are unlikely to combine systematically with modal auxiliaries of ability, permission or inference (*can*, *may*, *might*). Despite these differences, the three types of repulsion share a linguistic interpretation, which we will also discuss in section 5.

4.2   *Propensity toward adverbial modification*

Recall from section 3.3 that we can use Relative Entropy (H_{rel}) to measure the 'collocational range' of a modal auxiliary (Hoye 1997). A more varied modal auxiliary will have a higher H_{rel}, as it indicates greater distributional spread. Conversely, a lower H_{rel} indicates less variability and thus a higher propensity toward idiomhood. The dotplot in figure 2 shows the H_{rel} for each of the thirteen slot A types (see table 4).

Four clusters emerge: *'d(hd)* and *had* have by far the lowest entropies (almost by definition): *'d(hd)* is restricted to *better* and *best*, while *had* is slightly more productive and combines with *rather* and *better*/*best* and their modifications (*just better*, *far better*, *damn well better*). Contracted *'ll* and *'d(wd)* form a second cluster. The remaining two clusters are the mid-frequency (*may*, *might*, *should*, *must*, *shall*) and the high-frequency modal auxiliaries (*would*, *will*, *could*, *can*), respectively.

One contributing factor is frequency, which is correlated with Relative Entropy ($r = .66$, $p < .05$). At the same time, a modal's variability is indicative of semantic generality, and semantically general items tend to be more frequent. The stronger skew of *'ll* and *'d(wd)* indicates that limited generality accounts for lower Relative Entropy more than frequency. Figure 3 illustrates this effect. The relationship between variability and frequency for the full forms is linear and near-perfectly correlated ($r = .97$, $p < .001$). However, while *'ll* and *'d(wd)* are in the same frequency band as *must*, *should*, *might* and *may*, they are much less variable. Thus, frequency alone does not account for fluctuation in the collocational range.[7, 8]

In summary, two distributional patterns illustrate how MOD + ADV bigrams go beyond modal idioms or modal harmony. First, the four groups of post-modal adverbial modification that were distinguished on qualitative grounds in section 2 are quantitatively different only as a matter of degree. While idioms tend to cluster at the top end of the continuum, there is considerable overlap between and across groups along the cline of association. Second, modal auxiliaries differ in variability, with a higher tendency toward idiom-like behavior for those that are also functionally

---

[7] A potential objection concerns the strong bias of contracted forms *'ll* and *'d(wd)* toward spoken language. However, the patterns in the COCA-spoken data are essentially identical.

[8] It is interesting to note at this juncture that a threshold greater than $G^2 > 3.84$ ($p < .05$) to determine the set of adverbs over which H_{rel} is calculated (see section 3.3) does not affect the clusters. A higher threshold increases H_{rel} for all modals (because it removes low-frequency adverbs in the long tail). The notable exception is *shall*, which moves toward *'d(wd)* and *'ll* for higher thresholds (due to an increasingly lower H_{rel}), highlighting the oddness of *shall* in the paradigm of full form modals.
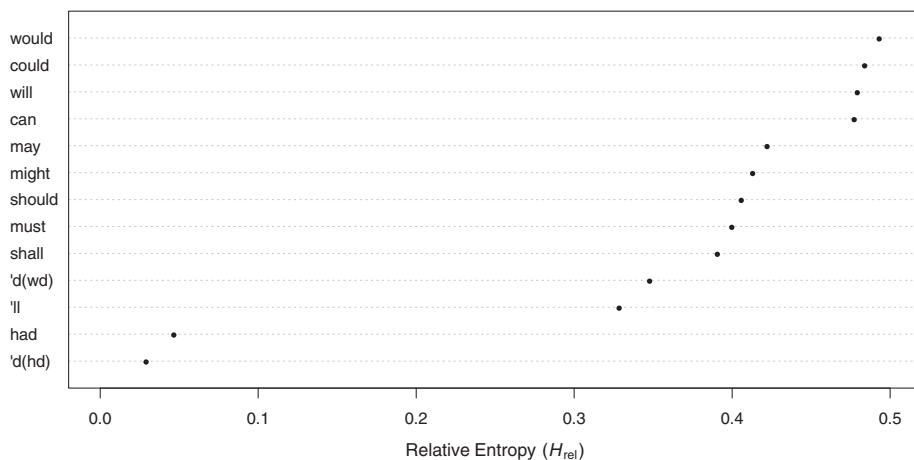
Figure 2. Diversity of post-modal adverbial modification (H_{rel})
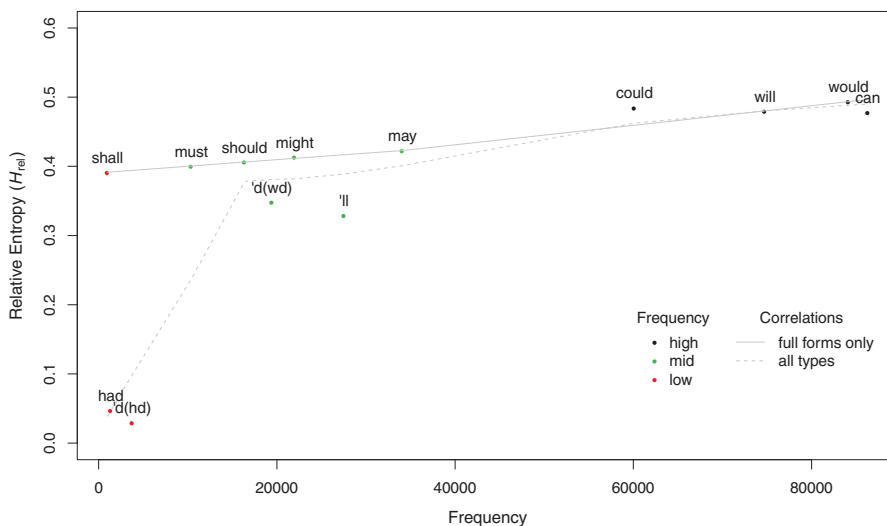

Figure 3. Relative Entropy (H_{rel}) by frequency (full COCA)

restricted. The question is how to interpret presence and absence of association in MOD + ADV, to which we now turn.

5 Discussion

The numerical results confirm an idiomaticity continuum for MOD + ADV (see figure 1). Higher-ranked bigrams correlate with greater idiomaticity, regardless of the qualitative

type. On the other hand, there are several types of relevant non-associations. There is true repulsion, mostly with bigrams that have conflicting modal values; for others, the repulsion is only apparent (e.g. *would easily*). This section discusses the implications of association and dissociation and argues that distributional information of MOD + ADV potentially provides cues to the scope of adverbial modification.

Since the results at the top end of the continuum can straightforwardly be interpreted as higher degrees of idiomhood and idiomaticity, more emphasis will be on the absence of an association and on repulsion. That is, we focus on the mid- and end-sections of the continuum, respectively. Despite some key differences, all types of dispreference are based on a simple assumption that follows from the logic of statistical association. Trivially, the stronger the attraction between two items, the more likely they are to form a cohesive, indivisible unit. This is most intuitive for modal idioms (*You'd better be sorry* vs **Better, you'd be sorry*). Conversely, the stronger the repulsion, the more likely the adverb has wider scope and modifies a unit of meaning beyond the modal auxiliary.

In the case of a strong positive association, the adverb modifies 'backward', qualifying modality. To illustrate, consider *barely* and *hardly*, which are positively associated only with *can* ($G^2 = 511$ and $G^2 = 695$) and *could* ($G^2 = 3,170$ and $G^2 = 3,437$):

(7)   (a)   He **could barely** keep his thoughts straight, but he knew things had gone wrong. (FIC, 2007)
       (b)   We can't see her and we **can barely** hear what they're talking about . . . (FIC, 2006)
       (c)   My husband has such bad road rage that we **can barely** stay in the car together for an hour. (MAG, 2010)
       (d)   We **can hardly** wait for it to come in and take Ryan home with us. (SPOK, 2006)

Here, *barely* and *hardly* qualify the agents' ability of keeping, hearing, staying or waiting, rather than the actions themselves. The adverbs have scope over the modal auxiliary, although both may be part of a larger chunk (*can hardly wait*). A similar connection between cohesion and modification of modality is assumed to be at work for all strongly associated MOD + ADV bigrams.

By contrast, if the relationship between a modal and an adverb is a strong dissociation, modification tends to work *forward*, that is, the adverb modifies the infinitival group. Consider the examples in (8) for *would barely* ($G^2 = 731$, neg), *will barely* ($G^2 = 831$, neg) or *may hardly* ($G^2 = 670$, neg):

(8)   (a)   Furthermore, NASA sought a substance that would **barely expand** or contract as it passed through extremes of temperature . . . (ACAD, 1990)
       (b)   He would **barely talk** to any of them, except Mat . . . (FIC, 1991)
       (c)   Now it looks as though the company will **barely earn** $1.48 this year. (MAG, 1993)
       (d)   That may **hardly seem likely** in the current political environment. (MAG, 2005)

Here, *barely* qualifies expansion, the intention to talk or one's earning of money, not (past) prediction. If the adverbs are part of a conventional collocation with the rightward

infinitival group, they pertain to the situation (*barely talk*, *hardly likely*), not to modality (cf. *could hardly wait* above).[9]

The same argument applies to adverbs that are not associated with any of the modal auxiliaries. A few random examples with at most mild (dis)preferences include *will seriously* ($G^2 = 5$, neg), *would utterly* ($G^2 = 13$, neg) or *can forcefully* ($G^2 = 9$, neg):

(9)   (a) If continued, they will **seriously threaten** the quality of research and education UC can provide, to the detriment of California and the nation. (NEWS, 1993)

      (b) Told that an order to advance would **utterly crush** the retreating Rebels, Meade hesitated. (MAG, 1993)

      (c) The church can **forcefully stand up** in the public arena and say, 'Look, we've got to think about these people as human beings.' (NEWS, 2006)

In brief, the absence of statistical cohesion increases the likelihood that the adverb is part of the infinitival group. This holds irrespective of whether the adverbs are otherwise distinctive for other modal auxiliaries (e.g. *barely*, *hardly*) or infrequent with few or no statistically significant relationships (e.g. *utterly*, *seriously*).

Another case of wider scope is sentential modification, where the adverb qualifies the proposition, not the modal or the infinitival group. To illustrate, consider *understandably*, which is mildly associated only with *might* ($G^2 = 31$, pos):

(10) (a) Nations **will understandably** resist imposing taxes on their own industries to provide a global benefit if this simply causes production or future investment to move to other nations without such taxes. (ACAD, 1997)

     (b) **Understandably**, nations **will** resist imposing taxes on their own industries.

(11) (a) However, parents **may understandably** feel that decisions about inheritances are theirs alone to make. (news, 2005)

     (b) **Understandably**, however, parents **may** feel that decisions about inheritances . . .

Similarly, *evidently* is only (very weakly) associated with *will* ($G^2 = 8$, pos):

(12) (a) But gold was still gold, money was still money, money **could evidently** buy anything, and he was going to be rich enough to start over. (FIC, 1998)

     (b) **evidently**, money **could** buy anything . . .

     (c) money **could** buy anything, **evidently** . . .

The absence of an association indicates a weaker connection, which in turn reflects greater positional variability. In other words, a lower unit-like status increases the likelihood that the adverb has scope over the proposition. In these cases, the adverb often has no

---

[9] As an anonymous reviewer points out, the difference between (7) and (8) is special in that it relates to the difference between external negation (of modality) vs internal negation (of the situation) (Palmer 1990, 1995; Depraetere & Reed 2006). Negation is beyond the current discussion, but the special status of *barely* and *hardly* as near-negative adverbs is not in contradiction to the claim that cohesion is a decisive factor in determining the scope of modification. We would expect similar statistical behaviour at work with the negation of modality (i.e. higher cohesion) vs the negation of the proposition (i.e. lower cohesion). However, this might be very difficult to measure empirically (see the discussion below on masking modal subsenses).

particular relationship in either direction: neither *could evidently* nor *evidently buy* are intuitively very strong collocates.

The phenomenon of modal harmony likely falls into this area. On the one hand, we could expect harmony to correlate with strong cohesion between an auxiliary and an adverb of the same modal value. This connection appears trivial: two items are more likely to co-occur to form a unit when they are also semantically compatible. Indeed, many MOD + ADV combinations discussed as harmonic or synergetic (Hoye 1997: 80f. 216; Geurts & Huitink 2006) have positive associations. However, the association strength is often rather low, for instance in *must inevitably* ($G^2 = 68$) or *might possibly* ($G^2 = 7$). Two exceptions with strong unit-status include *will probably* ($G^2 = 1,867$) and *could possibly* ($G^2 = 2,197$). On the other hand, there are numerous counterexamples with no or even a negative association, including *must certainly* ($G^2 = 15$, neg), *must really* ($G^2 = 1$, ns) or *may possibly* ($G^2 = 117$, neg).

Hence, the evidence on modal harmony is inconclusive from a distributional view. This adds weight to the (implicit) idea that harmony is a wider phenomenon (Lyons 1977: 807; Hoye 1997: 81f.). With the exception of a few highly frequent collocations (*could possibly*, *will likely*, *will probably*), modal harmony seems to be, on the whole, a phenomenon of sentential modification, based on the logic of the absence of cohesion and given only mild association in either direction for the majority of obvious candidates.

Finally, two types of dissociation deserve a brief discussion. We may call the first 'true repulsion', illustrated by *well* and *better*, which are only associated with *may/might* and *'d (hd)/had*, respectively. For reasons that follow from the non-prototypicality of *well* and *better* as MOD + ADV adverbs, they are strongly repelled by all other modal auxiliaries. Where they do co-occur, they are found in contexts that are reminiscent of split infinitives (*to deal better with X > to better deal with X*) and thus signal forward modification:

(13) (a) I will **well and faithfully discharge** the duties of the office on which I am about to enter. So help me God. (ACAD, 2002)
       (b) so she could **better deal** with performance problems. (ACAD, 2010)

The other repulsion type is only apparent and mainly affects modals in their non-central sense(s). An example is *easily*, which is positively associated only with *could* ($G^2 = 4,002$) and *can* ($G^2 = 2,236$). When *easily* co-occurs, for instance, with *would* ($G^2 = 1,853$, neg) or *will* ($G^2 = 1,653$, neg), they match the ability sense of *easily* with *would* and *will*'s meanings of hypotheticality or prediction, respectively:

(14) (a) Truly, this sandwich **would easily** handle the needs of two people. (NEWS, 2006)
       (b) Whatever you lift **will easily** slide off the blade and onto the plate. (MAG, 2010)

That said, these examples are ambiguous with respect to the scope of modification: *easily* could also qualify *handle* or *slide off*. Yet ability meanings are compatible with the main senses of *will* (prediction) and *would* (hypotheticality), which is why neither *would easily* nor *will easily* are particularly odd. Contrary to *would well*, their statistical repulsion is only apparent: it is due to the inability of CCA to distinguish between subsenses if, like

here, subsenses are not explicitly coded (see below). Note that apparent repulsion largley concerns bigrams of high token frequency (*would also*: 5,136); true repulsion is rare here (*will well*: 17; see table 5). In a sense, high raw frequency 'overrides' mathematical dissociation such that 'apparent repulsion' bigrams are also perceived as unit-like.

The phenomenon of apparent repulsion leads us to a shortcoming of the collostructional method in the current context, because CA glosses over polyfunctionality. This is particularly obvious with modal auxiliaries, which enter the analysis as one type for each. Put simply, polysemy is poorly represented in CA (but see Gilquin 2013 for using CA in the context of constructional polysemy). If a modal auxiliary has an infrequent subsense, then any collocation that is restricted to that subsense is mathematically disadvantaged. A hypothetical workaround would involve the manual annotation of each MOD + ADV observation for subsenses. This is clearly unfeasible for a data-hungry method such as CA. Given the logic of CA, an analysis of manually annotated data would *in*crease both the individual association scores and the number of MOD + ADV types at the idiom-like end of the continuum. From this angle, the current approach underestimates the extent of unit-like MOD + ADV. Yet CA fares better than raw frequency and potentially better than collocation-based analyses (Evert 2004). A collocation-style analysis based on transitional probabilities did not bring out sensible pattern in the context of modal idiomaticity.

A limitation of a different kind is the exclusion of negated uses (*can't possibly*), which were not queried to begin with (rather than subsuming them under their positive forms). Following from the behavior of contracted forms, if negated modal auxiliaries were included, they should probably be treated as separate types: {will} would enter the analysis as five types (*will, 'll, will not, won't, 'll not*). This would reshuffle the ranks, since, for instance, *even* is particularly frequent in negation (*won't even, can't even*). However, separating them could lead to the identification of further interesting sub-patterns.

The methodological limitations are reminders that collostructional analyses uncover tendencies and/or latent patterns, not fixed or deterministic results. Yet the results are robust in the sense that the general idea – i.e. assigning a crucial role to probabilistic information – remains unaffected. Idiomaticity clines can, therefore, be seen as part of speaker knowledge, both for modal constructions and beyond (Stefanowitsch & Gries 2003; Wulff 2008; Hilpert 2016).

6 Concluding remarks

Previous research on MOD + ADV collocation made an implicit categorical distinction between idioms and non-idioms. The current analysis illustrated the added value in going beyond this distinction. While the notions idioms and harmony are useful for individual aspects of modality, they cover neither the full range nor the complexity of (post-)modal adverbial modification. Hence, they are less suitable to address the question why *'d rather* or *could possibly* sounds more natural than *can rather* or *should possibly*. This motivated the collective analysis of the relevant

collocational phenomena from a distributional angle as a function of (quantitative) cohesion.

Cohesion was operationalized in terms of statistical association as a systematic measure of idiomaticity. The stronger the cohesion, the higher the unit-status and the higher the degree of idiomaticity. Greater cohesion, not necessarily higher frequency, is more conducive to idiosyncrasy, which is why modal idioms are at the extreme end of the continuum. The same argument holds for contraction compared to their full forms, that is, higher statistical cohesion leads to greater idiom-like behavior (rather than frequency). At the other end, many forms of dissociation signal modification beyond MOD + ADV. In other words, cohesion is predictive in *both* directions: greater association correlates with unit-status and idiomaticity, while greater dissociation correlates with forward or sentential modification. The collostructional method is better suited to distinguish between idiomatic and less idiomatic or unidiomatic sequences than raw frequency.

Note at this juncture that statistical repulsion was interpreted in a different way than in most CA applications for traditional 'closed' slot–filler constructions. In the latter, a repelled lemma is usually an 'odd' use of that lemma *in* a pattern (or construction) under investigation. In the case of MOD + ADV, repulsion identifies an increased probability that the lemma has relevance for something *outside* the pattern. This follows from an application of CA that does not take a pre-defined construction or 'node' (in the CxG sense) as its starting point, but rather a linear sequence.

Two points deserve a brief comment. First, the results are in line with work in constructionist frameworks, which handle scalar categories and gradience rather well (Langacker 1987, 2000; Goldberg 1995, 2006; Wulff 2008). Usage-based approaches assume that speakers make use of statistical information: strongly associated or highly frequent items are stored, or entrenched, and therefore more quickly activated (Langacker 1987). In a usage-based perspective, idiomaticity is a function of distributional properties (which may take a number of forms) that speakers derive from their linguistic environment. There is a growing body of research which suggests that approaching this knowledge by corpus-based means is psychologically plausible (Gries, Hampe & Schönefeld 2005; Wulff 2008, 2009; Ellis & Simpson-Vlach 2009; see overview in Stefanowitsch & Flach 2016).

Second, the study adds to recent work on the integration of modal auxiliaries into a Construction Grammar model (see the papers in a special issue of *Constructions and Frames*; see Cappelle & Depraetere 2016a). This is not a trivial task since modal auxiliaries fail to meet many classic criteria for constructionhood (Hilpert 2016). To be sure, there is little disagreement that modal idioms are constructions, i.e. learnt form–meaning parings with unpredictable formal and/or semantic properties (Goldberg 1995). Yet the constructional status is less clear for modal auxiliaries, and the ubiquity of gradience in idiosyncrasy and compositionality of MOD + ADV bigrams adds to this problem. As a way out, Hilpert (2016) suggests a combinatorial perspective, which views collocational relationships between modals and infinitives as part of constructional meaning. Such an approach shifts

the focus away from constructions as static schemas, and highlights more dynamic, connective links between constructions in the network. The idiomaticity perspective on MOD + ADV can be seen as a prime empirical case study of the underlying idea that constructional knowledge is knowledge of connections: links will be stronger for more attracted items and will be most extreme for modal idioms. Conversely, the greater the repulsion, especially for true repulsion, the weaker the links between them (or the stronger the link of an adverb with another unit of meaning). In other words, looking at modal collocation in this way works around the problem that we would otherwise have of assuming arbitrary categorical thresholds between (modal) idioms and non-idioms.

Author's address:

Institut de langue et littérature anglaises
Université de Neuchâtel
Espace Tilo-Frey 1
2000 Neuchâtel
Switzerland
susanne.flach@unine.ch

References

Aijmer, Karin. 2013. Analyzing modal adverbs as modal particles and discourse markers. In Liesbeth Degand, Bert Cornillie & Paola Pietrandrea (eds.), *Discourse markers and modal particles. Categorization and description*, 89–106. Amsterdam: John Benjamins.

Auwera, Johan van der & Astrid De Wit. 2010. The English comparative modals: A pilot study. In Bert Cappelle & Naoaki Wada (eds.), *Distinctions in English grammar: Offered to Renaat Declerck*, 127–47. Tokyo: Kaitakusha.

Auwera, Johan van der, Dirk Noël & An Van linden. 2013. *Had better*, *'d better* and *better*: Diachronic and transatlantic variation. In Juana I. Marin-Arrese, Marta Carretero, Jorge Arús Hita & Johan van der Auwera (eds.), *English modality: Core, periphery and evidentiality*, 119–53. Berlin: De Gruyter Mouton.

Cappelle, Bert & Ilse Depraetere. 2016a. Modal meaning in Construction Grammar. *Constructions and Frames* 8(1), 1–6.

Cappelle, Bert & Ilse Depraetere. 2016b. Response to Hilpert. *Constructions and Frames* 8(1), 86–97.

Celle, Agnès. 2009. Hearsay adverbs and modality. In Raphael Salkie, Pierre Busuttil & Johan van der Auwera (eds.), *Modality in English: Theory and description*, 269–93. Berlin: Mouton de Gruyter.

Coates, Jennifer. 1983. *The semantics of the modal auxiliaries*. London: Croom Helm.

Davies, Mark. 2008. *The Corpus of Contemporary American English: 450 million words, 1990-present* (mid-2015 offline version). corpora.byu.edu/coca

Declerck, Renaat. 2009. 'Not-yet-factual at time t': A neglected modal concept. In Raphael Salkie, Pierre Busuttil & Johan van der Auwera (eds.), *Modality in English: Theory and description*, 31–54. Berlin: Mouton de Gruyter.

Denison, David & Alison Cort. 2010. *Better* as a verb. In Kristin Davidse, Lieven Vandelanotte & Hubert Cuyckens (eds.), *Subjectification, intersubjectification and grammaticalization*, 349–84. Berlin: De Gruyter Mouton.

Depraetere, Ilse & Susan Reed. 2006. Mood and modality in English. In Bas Aarts & April McMahon (eds.), *The handbook of English linguistics*, 269–90. Malden, MA: Blackwell.

Ellis, Nick C. & Rita Simpson-Vlach. 2009. Formulaic language in native speakers: Triangulating psycholinguistics, corpus linguistics, and education. *Corpus Linguistics and Linguistic Theory* 5 (1), 61–78.

Evert, Stefan. 2004. The statistics of word cooccurrences. Word pairs and collocations. PhD dissertation, Universität Stuttgart. www.stefan-evert.de/PUB/Evert2004phd.pdf

Flach, Susanne. 2017. *Collostructions: An R implementation for the family of collostructional methods* (version 0.1.0). https://sfla.ch/collostructions/

Geurts, Bart & Janneke Huitink. 2006. Modal concord. In Paul J. E. Dekker & Hedde H. Zeijlstra (eds.), *Concord phenomena and the syntax semantics interface*, 15–20. Malaga: ESSLLI.

Gilquin, Gaëtanelle. 2013. Making sense of collostructional analysis: On the interplay between verb senses and constructions. *Constructions and Frames* 5(2), 119–42.

Goldberg, Adele E. 1995. *Constructions: A Construction Grammar approach to argument structure*. Chicago: University of Chicago Press.

Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

Greenbaum, Sidney. 1970. *Verb-intensifier collocations in English: An experimental approach*. The Hague: Mouton.

Greenbaum, Sidney. 1974. Some verb-intensifier collocations in American and British English. *American Speech* 49(1–2), 79–89.

Gries, Stefan Th. 2010. Useful statistics for corpus linguistics. In Aquilino Sánchez & Moisés Almela (eds.), *A mosaic of corpus linguistics: Selected approaches*, 269–91. Frankfurt: Peter Lang.

Gries, Stefan Th., Beate Hampe & Doris Schönefeld. 2005. Converging evidence: Bringing together experimental and corpus data on the association of verbs and constructions. *Cognitive Linguistics* 16(4), 635–76.

Gries, Stefan Th. & Anatol Stefanowitsch. 2004. Covarying collexemes in the *into*-causative. In Michel Achard & Suzanne Kemmer (eds.), *Language, culture, and mind*, 225–36. Stanford, CA: CSLI.

Grosz, Patrick. 2010. Grading modality: A new approach to modal concord and its relatives. In Martin Prinzhorn & Sarah Zobel (eds.), *Proceedings of Sinn und Bedeutung 14*, 185–201. Vienna. www.univie.ac.at/sub14/proc/grosz.pdf

Halliday, M. A. K. 1970. Functional diversity in language as seen from a consideration of modality and mood in English. *Foundations of Language* 6(3), 322–61.

Hilpert, Martin. 2011. Dynamic visualizations of language change: Motion charts on the basis of bivariate and multivariate data from diachronic corpora. *International Journal of Corpus Linguistics* 16(4), 435–61.

Hilpert, Martin. 2014. *Construction Grammar and its application to English*. Edinburgh: Edinburgh University Press.

Hilpert, Martin. 2016. Change in modal meanings: Another look at the shifting collocates of *may*. *Constructions and Frames* 8(1), 66–85.

Hoye, Leo. 1997. *Adverbs and modality in English*. London: Longman.

Huddleston, Rodney D. & Geoffrey K. Pullum *et al.* 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.

Jacobson, Sven. 1975. *Factors influencing the placement of English adverbs in relation to auxiliaries*. Stockholm: Almqvist & Wiksell International.

Jacobsson, Bengt. 1980. On the syntax and semantics of the modal auxiliary *had better*. *Studia Neophilologica* 52, 47–53.

Langacker, Ronald W. 1987. *Foundations of cognitive grammar*, vol. I: *Theoretical prerequisites*. Stanford, CA: Stanford University Press.

Langacker, Ronald W. 2000. A dynamic usage-based model. In Michael Barlow & Suzanne Kemmer (eds.), *Usage-based models of language*, 24–63. Stanford, CA: CSLI.

Lyons, John. 1977. *Semantics*. Cambridge: Cambridge University Press.

Mitchell, Keith. 2003. *Had better* and *might as well*: On the margins of modality? In Roberta Facchinetti, Frank Palmer & Manfred Krug (eds.), *Modality in Contemporary English*, 129–50. Berlin: Mouton de Gruyter.

Nuyts, Jan. 2001. *Epistemic modality, language, and conceptualization: A cognitive-pragmatic perspective*. Amsterdam: John Benjamins.

Palmer, F. R. 1990. *Modality and the English modals*, 2nd edn. London: Longman.

Palmer, F. R. 1995. Negation and the modals of possibility and necessity. In Joan L. Bybee & Suzanne Fleischman (eds.), *Modality in grammar and discourse*, 453–71. Amsterdam: John Benjamins.

Perkins, Michael R. 1983. *Modal expressions in English*. Norwood: Ablex.

Simon-Vandenbergen, Anne-Marie. 2008. *Almost certainly* and *most definitely*: Degree modifiers and epistemic stance. *Journal of Pragmatics* 40(9), 1521–42.

Simon-Vandenbergen, Anne-Marie & Karin Aijmer. 2007. *The semantic field of modal certainty: A corpus-based study of English adverbs*. Berlin: Mouton de Gruyter.

Stefanowitsch, Anatol. 2009. Bedeutung und Gebrauch in der Konstruktionsgrammatik. Wie kompositionell sind modale Infinitive im Deutschen? *Zeitschrift für Germanistische Linguistik* 37(3), 565–92.

Stefanowitsch, Anatol & Susanne Flach. 2016. The corpus-based perspective on entrenchment. In Hans-Jörg Schmid (ed.), *Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge*, 101–27. Berlin: De Gruyter.

Stefanowitsch, Anatol & Stefan Th. Gries. 2003. Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8(2), 209–43.

Stefanowitsch, Anatol & Stefan Th. Gries. 2005. Covarying collexemes. *Corpus Linguistics and Linguistic Theory* 1(1), 1–43.

Traugott, Elizabeth Closs. 2016. Do semantic modal maps have a role in a constructionalization approach to modals? *Constructions and Frames* 8(1), 97–124.

Van linden, An. 2010a. Extraposition constructions in the deontic domain: State-of-affairs (SoA)-related versus speaker-related uses. *Text & Talk* 30(6), 723–48.

Van linden, An. 2010b. The clausal complementation of *good* in extraposition constructions: The emergence of partially filled constructions. In Ursula Lenker, Judith Huber & Robert Mailhammer (eds.), *English historical linguistics 2008: Selected papers from the fifteenth International Conference on English Historical Linguistics (ICEHL 15), Munich, 24-30 August 2008*, 95–120. Amsterdam: John Benjamins.

Van linden, An. 2012. *Modal adjectives: English deontic and evaluative constructions in synchrony and diachrony*. Berlin: De Gruyter Mouton.

Wulff, Stefanie. 2008. *Rethinking idiomaticity: A usage-based approach*. London: Continuum.

Wulff, Stefanie. 2009. Converging evidence from corpus and experimental data to capture idiomaticity. *Corpus Linguistics and Linguistic Theory* 5(1), 131–59.

Xiao, Tangjin. 2009. 'We can probably go there': English modal satellite adverbs and modality supplementing in discourse. *Linguistics and the Human Sciences* 5(3), 251–79.

Zeijlstra, Hedde H. 2007. Modal concord. In Tova Friedman & Masayuki Gibson (eds.), *SALT XVII*, 317–32. Ithaca, NY: Cornell University.