

The rise of colligations

English *can't stand* and German *nicht ausstehen können*

Olav Hackstein and Ryan Sandell

Ludwig-Maximilians-Universität München

This article examines the lexically parallel English and German constructions *can't stand somebody/something* and *jemanden/etwas nicht ausstehen können* “not tolerate (someone or something)”, from synchronic, diachronic, and quantitative perspectives. Syntactic and semantic restrictions suggest that the usage of *stand* and *ausstehen* in the relevant sense is older than other semantically similar verbs (e.g. English *tolerate*, German *leiden*), while quantitative evidence from corpora shows that the *can't stand* and *nicht ausstehen können* constructions are both colligationally stronger than lexical competitors. Evidence from the history of *stand* indicates that the lexeme *stand* in the Germanic and other Indo-European languages has a long history of being employed in the relevant sense. The restrictions on usage and the colligational strength of the respective English and German constructions are thus argued to result from the antiquity of the construction and functional competition from other lexemes.

Keywords: collostructional analysis, historical linguistics, Indo-European linguistics, Germanic languages, syntactic productivity

1. Introduction

The following study examines the synchrony and diachrony of the idiomatic phrases English *can't stand somebody/something* alongside the synonymous German *jemanden/etwas nicht ausstehen können* “not be able to out-stand someone/something”; we employ the label “*can't stand* construction” for these constructions together. Examples of this construction in English, German, and the parallel construction in Dutch are given in Table 1; some related usages of *stand*, *ausstehen*, and *uitstaan*, and the similar Swedish *utstå* are shown in Table 2.



Table 1. Synchronic constraints on E. *can't stand*, G. *nicht ausstehen können*, Du. *niet uitstaan kunnen*

	CAN NOT VERB _{INF} X _O OBJ	*VERB _{FIN} X _O OBJ (P)
English	<i>She can't stand him.</i>	* <i>She stands him.</i> Certain conditions may license <i>she stands him</i> , e.g. in a negative polarity context with subordinate clause: <i>I don't know how she stands him.</i>
German	<i>Sie kann ihn nicht ausstehen.</i>	* <i>Sie steht ihn aus.</i>
Dutch	<i>Zij kan hem niet uitstaan.</i>	* <i>Zij staat hem uit.</i>

Table 2. Related usages or idioms with the lexical node E. *stand*/ G. *ausstehen*/ Du. *uitstaan* “endure” license non-negated/finite forms

English	<i>(I can't stand him, but) she can stand him.</i> [Contrastive focus construction.] See also <i>to stand out against sth.</i> ≈ <i>to withstand sth.</i> , (OED, s.v. <i>stand</i> , sense 59; phrasal verb <i>stand out</i> , sub 3)
German	<i>Ich habe Todesängste ausgestanden.</i>
Dutch	<i>Ik heb doodsangsten uitgestaan.</i> “I have gone through mortal agonies. I have been scared to death.”
Swed.	<i>Jag får utstå mycket lidande.</i> “I have to endure much pain.”

Characteristic of this construction in English, German, and Dutch is that the verb *stand*, either with the local particle *out*, as in German and Dutch (and perhaps in older English; see further Section 4), or without it, as in Present-day English, transitively governs a direct object, even though *stand* in its core semantics is an intransitive verb of motion. Furthermore, this idiom exhibits three synchronic constraints in German, English, and Dutch (exemplified in Table 1): first, the idiom's node (E. *stand*, G. *ausstehen*, Dutch *uitstaan*) in the sense of “tolerate”, must, under most conditions, be infinitival; second, the infinitival form is tendentially governed by the modal verb *can*; finally, *can* is nearly always negated. More abstractly, the construction in which *stand* serves as a node may be schematically represented (for English) as CAN NOT VERB_{INF} X_O OBJ where *stand* fills the VERB_{INF} slot; x_o stands for other constituents that might intervene between the verb and the (accusative) object. The same semantics of “endure, tolerate” is effectively barred with non-negated, finite *stand*, although non-negated non-finite *stand* may occur under certain further conditions. However, searches in the COCA (Davies, 2008–) and the morphologically tagged corpora of the DeReKo (Kupietz, 2019), on which see further Section 5, revealed no credible instances in English or Ger-

man of finite, non-negated usages of *stand/ausstehen* with the relevant semantics (outside of the specific collocation *Todesängste ausstehen* in German). On the other hand, instances of non-finite, non-negated *stand* in English subordinate clauses headed by *how*, *if*, or *whether* are relatively well-attested, as well as in subordinate clauses standing under a verb such as *think* or *figure* (e.g. *I figure (that) I can stand it*), or a question with *how*. In German, a handful of instances occur in which *können* does not stand under the negative operator *nicht*, but in which negative polarity is introduced through *kein* “no” or *wenig* “few”, e.g. *Umfragen zufolge können immer weniger Briten die Regierungspartei ausstehen* “According to opinion polls, ever fewer Britons are able to tolerate the governing party” (St. Galler Tagblatt, 28.11.2007, p.6).

This paper aims to provide a diachronically oriented account of why *stand* and *ausstehen* are subject to such restrictions, closely bound up with the CAN NOT VERB_{INF} X₀ OBJ construction, whereas other semantically similar verbs, such as E. *tolerate* or G. *leiden* “suffer”, are not. However, the constraints as shown in Table 1 are manifestly secondary. They must have arisen diachronically later, because related usages of the same lexical node E. *stand* / G. *ausstehen* / Du. *uitstaan* “endure” still license non-negated, finite forms of the same verb. Moreover, the related Swedish verb *utstå* “endure, suffer” is not generally subject to the restrictions observable in the West Germanic languages.

The syntactic and semantic restrictions found for E. *stand* in the sense of “endure, tolerate” and its cogeners may constitute evidence for a lexical collocation (i.e. preferred lexical co-occurrence) between *can not* and *stand*, as well as a colligation between *stand* and the transitive CAN NOT VERB_{INF} X₀ OBJ construction mentioned above. On the definition of colligation, see Gries (2009:14), who defines the term as “the co-occurrence of word forms with grammatical phenomena”, such that certain collocations are bound to specific, idiosyncratic grammatical configurations (see also Lehecka, 2015: 5).

This paper will investigate the extent to which the syntactic gaps (evident in both English and German) observed above, as well as morphological gaps (in the case of German *ausstehen*, which virtually exists only as an infinitive or past participle), may be indicative not only of synchronic collocational strength, but also of the relative age of a collexeme. Specifically, in this case, a lexeme that is demonstrably older than semantically similar lexemes takes on more and stronger colligational restrictions.

The organization of the paper is as follows. Section 2 deals with possible synchronic and diachronic factors that cause the emergence of colligations. Colligations may be conditioned synchronically by pragmatic principles or have a diachronic basis. An important diachronic conditioning factor is lexical renewal and competition which can potentially induce older inherited phrases to adopt

formal and semantic specializations. In Section 3, further comparison between *can't stand* with other verbal lexemes that can be employed in meanings similar to that of the *can't stand* construction shows that the *can't stand* construction is subject to a number of syntactic and semantic restrictions that do not obtain for those other verbal lexemes. Comparative Indo-European reconstruction shows that features of the lexeme *stand* can be traced back into Proto-Indo-European, where the usages of “stand” in the Germanic languages are compared to functions of cognate lexemes in other Indo-European languages. The etymologies of functionally competing lexemes, meanwhile, reveal them to be, in almost all cases, younger than the core of *stand/ausstehen*. Section 4 then undertakes a lexicographic examination of *stand* and *ausstehen*, from Old English and Old High German into the present day, alongside etymologically and functionally related verbs, documenting the parallel rise of the same colligations in English and German. Finally, in Section 5, we further substantiate the claim that functional competition from semantically similar lexical items has led to the emergence of the robust collocations *can't stand somebody/something* and *jemanden/etwas nicht ausstehen können*, arguing from the perspective of quantitative corpus linguistics that colligations and restrictions may be indicative of relative linguistic age. We demonstrate that the usage restrictions found with *can't stand/nicht ausstehen können* correspond to higher degrees of quantitatively measurable colligational strength. Appendix 6 finds indications that *can't stand/nicht ausstehen können* also exhibit relatively lower degrees of syntactic productivity.

2. Synchronically and diachronically conditioned colligations

How is the emergence of a colligation – like that between G. *ausstehen* and x_0 OBJECT_{ACC} NICHT VERB_{INF} KÖNNEN – insofar as it involves paradigmatic or syntagmatic constraints, to be explained? One key to answering this question comes from the insight that colligations like E. *can't stand* and G. *nicht ausstehen können* constitute partially automated units of speech production. Bolinger (1976) and proponents of usage-based approaches (e.g. Bybee, 2010) have pointed out the utility and economy of linguistic formulas or open-slot multi-word constructions, which automate the production of communicatively important and hence frequent semantic concepts, thereby also facilitating language acquisition.¹

Crucially, the automatization of so-called ‘chunks’ involves ‘freezing’, i.e. the suspension of a lexeme’s free modifiability at the morphological and combin-

1. As Bybee and Scheibman (1999:577) put it, “[i]n chunking (Haiman 1994), a frequently repeated stretch of speech becomes automated as a processing unit”.

ability at syntactic level (see Hackstein, 2012a: 89 for diachronic exemplification). The observation that linguistic automatization involves linguistic freezing, and that automatization readily applies to frequently repeated stretches of speech, has long been known; this insight is implicit in Meillet's Principle (Meillet, 1937: 31–32: “[r]elics often occur in the most commonly used expressions of a language.”), which points to the propensity of linguistically frequent forms to show linguistic anomalies. Freezing need not, however, depend on frequency. Another preserving effect independent of high token frequency is embedding in multi-word expressions, which contribute to the diachronic petrification of grammatical features. The importance of multi-word expressions or “formulas” in general for natural languages was recognized by Firth (1964); see Bolinger (1976), Wray (2002), and Bozzone (2014) for a recent survey of the literature. Moreover, the petrification of older linguistic features in multi-word formulas has revealed itself to be a universal phenomenon in natural languages (see Bybee's research on chunking and the conservation effect; Bybee, 2006, 2007, 2010). The conserving effect of formulaic frames has also been substantiated in other fields of linguistics, for instance in studies on child language acquisition; for example, Arnon and Clark (2011) demonstrated that children acquire irregular morphology first in formulaic contexts. The same applies to poetic languages (*Kunstsprachen*), which notably preserve traces of older morphology and older syntactic constructions in formulaic expressions; see Bozzone (Forthcoming). For a survey of previous literature, see Appendix 1.

Nonetheless, the utility of automatization and formularity do not exhaust the conditioning factors that may cause colligations to arise. It is necessary to distinguish between synchronic and diachronic factors, which will be laid out in Sections 2.1 and 2.2.

2.1 Synchronic colligational restrictions

First, colligational restrictions may be described synchronically, the level on which they are part of a speaker's lexical and grammatical knowledge. Consider the constraint on negation observed for E. *can't stand* and G. *nicht ausstehen können* in Table 1. This constraint may be explained by the pragmatic tendency to encode negative emotions indirectly rather than directly. Many languages prefer to encode these speech-acts of disapproving, declining, or rejecting indirectly by negating the positive, thus *I don't approve* or *I can't approve* instead of *I disapprove*, or *I don't like* rather than *I dislike/hate* preferring a syntactically more complex periphrasis. The maxim “Don't address the unpleasant, undesirable directly, and distance yourself from it” holds crosslinguistically and diachronically. For example, senses of antipathy among Homeric Greek heroes or characters in Platonic dialogues are

expressed significantly more often by negated οὐκ ἔτλη [o:k étlɛ:] “I cannot suffer (it)” than by a non-negated psych verb with a negative denotation. In the same vein, E. *can't bear*, G. *kann nicht aushalten* is pragmatically more apposite than *hate/hassen*. In other cases, limiting verbs to negated expressions might add to vernacular expressivity. Sinclair's example *It doesn't budge* (Sinclair, 1998: 16–22), may fall under this rubric, because *It doesn't budge* conveys more emotional involvement than the nonnegated *It is immobile*.

2.2 The diachronic conditioning of colligations

Beyond the synchronic workings of a given colligation, colligations are frequently conditioned diachronically. Two diachronic mechanisms may be observed that can block normally productive morphological and syntactic derivations.

2.2.1 *Lexical renewal and competition induce linguistic specialization via colligation*

The first mechanism involves gradual lexical renewal without replacement and lexical competition. When an innovative lexeme or phrase impinges upon the meaning or function of a semantically competing existing word or phrase, then the two elements may come into competition with one another. As a result, innovative limitations on the selection of grammatical categories across their constituents on a phrasal level may arise for one of the competing elements, with the consequence that certain morphosyntactic derivations may be blocked.

Lexical renewal typically takes effect gradually, often with a phase in which the inherited lexical item and the innovative lexical item coexist. Thus, the renewal of a lexeme within a multi-word phrase may, and perhaps typically does, undergo lexical renewal without fully replacing the inherited variant of the phrase in question containing the older lexeme in all contexts. In such cases, a principle of economy is operative: semantic differentiation is more economical and useful than synonymy and redundancy. Semantic differentiation entails a semantic repartition, which alots a broader sense to the innovative and usually more frequent expression but narrows the meaning of the same phrase's less-frequent forerunner. This phenomenon constitutes a phrasal analogue to Kuryłowicz's 4th “Law” of Analogy (Kuryłowicz, 1945). Furthermore, the semantic specialization of multi-word phrases is part and parcel of the chunking process, in which semantic specialization and morphological freezing represent each other's analogues on different grammatical levels.

This “Kuryłowicz-4” mechanism of lexical renewal, which prompts chunking and grammatical restriction of the inherited phrase, is arguably exemplified by E. *can't stand* and G. *nicht ausstehen können*. These fall into the semantic domain of

psych verbs, which are prone to high degrees of lexical variability and renewal. For instance, alongside E. *can't stand* we find *doesn't like*, *dislikes*, *doesn't suffer*, *can't bear* and alongside G. *nicht ausstehen können*, there is *nicht leiden können*, and *nicht mögen*. The data on synchronic morphosyntactic restrictions imposed upon E. *stand* and G. *ausstehen* treated in Sections 3 and 4 suggest that such restrictions may have arisen due to lexical renewal.

2.2.2 *Diachronically persistent parameters motivate blocking effects and colligations*

The second mechanism giving rise to colligations involves diachronically persistent parameters, which can block productive morphosyntactic derivation. Constructional inheritance can impede otherwise expected morphosyntactic derivations. One may speak in German of *eine abgelegene Hütte* “a remote hut”, while the attributive use of German *weit weg* “far away” is excluded: G. **die weit wege Hütte*. Such a ban on the attributive use of G. (*weit*) *weg* is a case of persistence. G. *weg* is diachronically not a true adjective but arose secondarily from a prepositional phrase meaning “on the way”; see (1).

- (1) Prepositional phrase “on the way”: MHG *ën-wēc* [ɛn'vɛk], by aphoresis > *wēc* [vɛk] > ModG *weg* [vɛk] “away” with prosodic conservation of [ɛ] (contrast ModG noun [vɛ:k]; compare Yid. *avek*, *vek*. See generally *Oxford English Dictionary* (OED; Oxford University Press, 1884–) s.v. *away*, Grimm’s *Deutsches Wörterbuch* (DWB; Grimm, 1854–1961) s.v. *weg*).

Certain inherited idioms like E. *far away* and G. *weit weg* typically occur in spoken vernacular registers, which agrees with the empirical fact that one of the primary vehicles both for linguistic innovation and for the transmission of persistent archaisms is language acquisition. This is in accord with Labov (1989: 96), who pointed out that children “reproduce the historically preserved variable patterns”, acquiring them as sociolinguistic variables (Labov, 2001: 89). An example of a purely orally transmitted, non-standard variable in English is the alternation between [æsk] and [æks] in some American English dialects, which descends from and replicates the Old English variation between *āscian* and *ācsian* (Labov, 1989: 86).

The same phenomenon is demonstrable for multi-word expressions (collocations, instantiations of particular lexemes in a construction). Examples include (2) and (3).

- (2) PWGmc. **an wegan dō* > E. *to do away (with sth)*, G. colloquial (*etwas*) *weg-tun*, Dutch (*iets*) *wegdoen*.

- (3) PIE **uoi-/uaj-* *d^heh₁-* > E. *do woe*, OE *wa dydan* (OED, s.v. *woe*), G. colloquial *wehtun*, OHG *tûot wê* (Notker); further Goth. *wai-dedja* “evil-doer”.

(see Hackstein, 2012b: 16, Pinault, 2015: 170)

The oral transmission of persistent features over long periods of time also applies to the phrases under discussion, E. *can't stand*, and G. *nicht ausstehen können*, which, as emotive expressions, belong primarily to a vernacular register, and are tendentially replaced by E. *can't bear/tolerate* and G. *nicht leiden/ertragen können* in more formal registers. It can thus be argued that the transitive use of E. *can't stand*, and G. *nicht ausstehen können* is due to inheritance from PIE (see Appendix 2); the otherwise intransitive use of E. *stand* and G. *stehen* is blocked by inheritance.

3. Colligations, restrictions, and the linguistic age of E. *stand* and G. *ausstehen*

The comparison of E. *can't stand* and G. *nicht ausstehen können* with verbs and phrases of parallel formation across the Germanic languages shows that the lexically competing verbs and phrases differ in their linguistic age. The following trend emerges: diachronically older lexemes occur primarily in the context of strong colligations and exhibit more restrictions on their licit co-occurrence, while comparatively younger lexemes exhibit fewer colligations and restrictions, and yet younger lexemes may lack colligations and grammatical restrictions altogether.

The construction E. *can't stand* and G. *nicht ausstehen können* is the oldest; it can, in essence, be traced back to PGmc. (see Tables 1 and 2 above) and partially projected even further into Proto-Indo-European. The second oldest lexical competitor of G. *ausstehen* considered here is *nicht leiden*, which cannot have been used to mean “endure” before the end of the 1st Millennium CE. Finally, the youngest competitor is G. (*nicht*) *mögen* in the sense “(dis)like”, which arose in the relevant sense in the New High German era.

In examining a potential correlation between linguistic age and colligational strength, the data in Table 3 shows that the presumably oldest lexeme, *ausstehen*, is highly restricted (five restrictions) and essentially confined to just a single colligation, occurring almost exclusively in collocation with *nicht* and *können*; the second oldest, *leiden*, shows three colligations, and some restrictions on its use (Table 4), while the most recent, *nicht mögen*, exhibits neither any clear colligations nor any apparent grammatical restrictions (Table 5). These facts suggest that the fixity of colligations to some few or just one collocation and the

increase in restrictions can be indicative of linguistic age. Although a stable colligation need not be of great antiquity, the establishment of a restricted colligations may be indicative of linguistic archaism relative to lexemes or phrasal units that unrestrictedly enter into productively built phrases. In sum, the stronger the colligation and greater the number of grammatical constraints, the older the construction may be.

Table 3. Colligations and Restrictions for *G. ausstehen*

(PIE >) <i>G. nicht ausstehen können</i> : 1 strict colligation involving 5 restrictions	
√a. Negated present tense auxiliary construction	Sie kann ihn nicht ausstehen. “She cannot stand him.”
😞 b. *Nonnegated present tense auxiliary construction	*Sie kann ihn ganz gut ausstehen.* **“She can stand him quite well.”
😞 c. *Negated present tense non-auxiliary construction	*Sie steht ihn nicht aus. **“She does not stand him.”
😞 d. *Nonnegated present tense non-auxiliary construction	*Sie steht ihn aus. **“She stands him.”
😞 e. *Nonnegated past tense non-auxiliary construction	*Sie stand ihn aus. **“She stood him.”
😞 f. *Nonnegated perfect passive non-auxiliary construction	*Er ist wohl ausgestanden. **“He is stood well.”

* A reviewer holds that German non-negated *ganz gut ausstehen* may be well-formed, though it is the intuition of the native speaker Author 1 that it is not. The phrase “ganz gut ausstehen” does not occur whatsoever in either of the two tagged subcorpora of the DeReKo (Tagged-C and Tagged-C2) employed in Section 5, nor in the considerably larger subcorpus W – Archiv der geschriebenen Sprache, whereas “ganz gut leiden” occurs in those three subcorpora together 73×. Similarly, Google searches (on July 15, 2020) reveal only 582 tokens of “ganz gut ausstehen” versus more than 107000 (!) tokens of “ganz gut leiden”. It seems therefore altogether possible that “(ganz) gut ausstehen” may be sufficiently rare as to be ungrammatical or nearly so for some (perhaps most) native speakers.

While the sort of restrictions found on *ausstehen* as seen above may give some indication as to the relative age of competing linguistic structures, which may be further substantiated through quantitative analysis (see Section 5), the absolute age of a lexeme and importantly, the age of its semantics, can be gauged only by philological facts (chronology of attestation) and the application of comparative reconstruction (this section and Appendix 2). Along these lines, the history and etymology of E. *stand* “endure” and its lexical competitors shows that (i)

Table 4. Colligations and Restrictions for G. *leiden*(OHG >) G. *nicht leiden können*: 3 weak colligations, 3 restrictions

✓a. Negated present tense auxiliary construction	Sie kann ihn nicht leiden. "She cannot suffer him."
✓b. Nonnegated present tense auxiliary construction	Sie kann ihn ganz gut leiden. "She can suffer him quite well."
😞 c. *Negated present tense non-auxiliary construction	*Sie leidet ihn nicht. *"She does not suffer him."
😞 d. * Nonnegated present tense non-auxiliary construction	*Sie leidet ihn. *"She suffers him."
😞 e. * Nonnegated past tense non-auxiliary construction	*Sie litt ihn. *"She suffered him."
✓f. Nonnegated perfect passive non-auxiliary construction	Er ist wohlgeleitten. "He is suffered well = he is well-liked."

Table 5. Colligations and Restrictions for G. *mögen*G. (*nicht*) *mögen*: no colligations, no restrictions

✓a. Negated present tense auxiliary construction	Sie soll ihn (angeblich) nicht mögen. "She supposedly does not like him."
✓b. Nonnegated present tense auxiliary construction	Sie soll ihn (angeblich) mögen. "She supposedly likes him."
✓c. *Negated present tense non-auxiliary construction	Sie mag ihn nicht. "She does not like him."
✓d. * Nonnegated present tense non-auxiliary construction	Sie mag ihn. "She likes him."
✓e. * Nonnegated past tense non-auxiliary construction	Sie mochte ihn. "She liked him."
✓f. Nonnegated perfect passive non-auxiliary construction	Er ist wohl (von den meisten) gemocht worden. "He has probably been liked (by most people)."

E. *stand* "endure" (ii) *like*, and (iii) *endure*, *tolerate* synchronically reflect lexical competitors along a trajectory of descending time depth, and the same holds for

(i) G. *ausstehen*, (ii) *leiden*, (iii) *mögen*. More extensive details on the following are available in Appendix 3.

E. *stand (out)* and G. *ausstehen* “*to stand up against, endure sth.” are represented in at least three Indo-European branches (Germanic, Slavic, Indo-Iranian; cf. Vedic *úd sthā-* “stand up [against an obstacle]”); Slavic (see Russian *vosstat’* “to rise against, withstand”); and Proto-Germanic **us-standa-* continued in Goth. *us-standan* “rise, be resurrected (with endurance)”, and in OHG *ir-standan* “rise, be resurrected”, OHG *irstān* “rise; endure”, OE *astandan* “stand up; abide; withstand”. NHG *ausstehen*, Du. *uitstaan*, E. *stand out* “endure”, Sw. *utstå* represent an inherited particle verb construction Gmc. **standa- ūt*, **sta(j)i/a- ūt* which, due to its vernacular register, remained unattested in the OHG and OE literary transmission (Sections 4.2 and 6 below). Furthermore, within West Germanic, the status of the preverb as inseparable and unstressed hints at the compound’s linguistic age. Differently, OE *astandan* and ME *ostonden* lose their unstressed preverb (*a-/o-*), merging formally with *stand*, while remaining semantically and collocationally distinct. The anomalous transitivity of E. *stand* “tolerate, endure”, G. *ausstehen* continues a persistent parameter which is inherited from the verb’s stem formation, see Appendix 2 sub b) and c). Both verbs are barred from their intransitive use, which otherwise is standard for G. *stehen* and E. *stand (upright)* (a blocking effect by persistence, as per Section 2.2.2 above). In contrast to E. *stand (out)* and G. *ausstehen*, their lexical competitor G. *erleiden*, *leiden* “to experience, incur sth.”, E. *not like* (as stative psych verb) is confined to Germanic and therefore cannot be reconstructed beyond Germanic. Rather G. *leiden* with the sense “tolerate” reflects a semantic innovation (cf. Goth. *ga-leiþan* “come, go”) not older than the 8th century CE. Finally, the third and linguistically youngest layer is made up of E. *endure*, *tolerate*, G. *mögen*. E. *endure* and *tolerate* are loanwords, first attested in English in the 14th and 16th centuries, respectively. G. *mögen* in the sense “like” is recent, having arisen in the Modern German period.

4. Colligations in English *can’t stand* and German *nicht ausstehen können*

Thus far, we have mustered evidence that E. *can’t stand* and G. *nicht ausstehen können* exhibit more colligational effects than do other verbs that permit similar readings in a construction with a negated modal verb *can/können*. These phenomena, as laid out above in Section 3 for G. *ausstehen*, may be argued to correlate with linguistic age; that *can’t stand* and *nicht ausstehen können* contain a lexical core that is older than many competing lexemes that may occur in the construction is born out by the reconstructable prehistory and etymology of those lexemes

(see Section 3). With these facts in mind, we now turn to the philological investigation of the respective English and German lexemes and constructions, beginning from the Old English and Old High German periods. The objective is to document the functional development of *stand* and *ausstehen*, focusing on emerging colligational restrictions in connection with modal verbs and negation. By and large, the OED and DWB contain sufficient data for the documentation of the relevant colligational restrictions; these are supplemented with further data and lexicographical assessments for English from Bosworth and Toller (2021; hereafter “Bosworth-Toller”), the *Dictionary of Old English* (DOE; Cameron et al., 2018), the *Dictionary of Old English Web Corpus* (DOEC; dePaolo et al., 2009), and *Middle English Dictionary* (MED; Lewis, 1952–2001), and, for German, the *Referenzkorpus Altdeutsch* (RA; Donnhäuser et al., 2015), the *Referenzkorpus Mittelhochdeutsch* (RM; Wegera & Klein, 2016), and the *Mittelhochdeutsches Wörterbuch* (MWB; Hirzel, 2006–). Further details, textual examples, and statistics are provided in Appendix 4.

4.1 English: (can't) stand someone/something (out)²

The original core semantics as a **verb of position**, meaning “take, have or keep an upright position” (see OED s.v. *stand* 11a) is attested from the Old English period. Example (4) shows that combination with negation and a modal verb, *magan* “be able, can” was entirely possible.

- (4) Ne mæg hūs nāht lange standan on þām heān munte...
 (Boethius, *De consolatione philosophiae* 12.26.30; DOEC)
 “A house cannot stand for long on that miserable hill...”

A semantic expansion to a **verb denoting physical persistence or endurance** is already found already in Old English as well, as illustrated in Example (5).

- (5) ... gif satanas (*sic*) winð ongen hine sylfne he bið tōdæled & he standan ne
 mæg ac hæfð ende (Gospel of Mark 3:26; DOEC)
 “... if Satan (*sic* SG, not PL) starts up against himself, he is divided, and he cannot stand, but rather has an end.”

Of the 244 instances of the form *standan* (usually infinitive) in the DOE, 28 instances are attested as part of a negated modal construction (like *standan ne mæg* in (4)), predominantly with *magan* “can, be able, may”, but also *motan* “must; can” and *sculan* “must, should”. The constructional components that ulti-

2. All examples in this section stem from the OED (accessed November 4, 2019), unless otherwise indicated, and are cited following the information given in the OED.

mately grow into colligational restrictions are thus already present at an early date. However, clear transitive uses of OE *standan*, or any instances with the sense “endure, tolerate”, are difficult to come by, perhaps entirely absent, in the Old English corpus.

Meanwhile, prefixed OE *a-standan* (where *a-* < PGmc. **uz-* > OHG *ir-*, as in OHG *irstantan*, Goth. *us-* in *usstandan*) is attested with the meaning “stand firm, resist destruction; survive, endure, persevere”; as in Example (6). This verb is not yet, however, clearly susceptible to the reading “tolerate” (see DOE s.v. *astandan*). Likewise, any clear usages of a phrasal verb **standan ūt* “stand out = tolerate” are unknown in the Old English corpus.

- (6) seo studu ... þe sēo molde on hongode, sēo gesund & ungehrinen from þæm
fýre **astod** & āwunade
(Bede, *Ecclesiastical History* 8.180.29; DOE s.v. *astandan*)
“that post... on which the dust hung in the linen cloth remained safe and
untouched by the fire”

Middle English *stōnden*, as reflected in the MED, largely continues the usage of OE *standan*, but the sense “make a stand in battle, stand and fight, resist and enemy in battle; withstand” is much more prominently represented, constituting some of the earliest syntactically transitive instances of a form of *stand*. Instances of the infinitive under a negated modal verb or with other negative polarity item are well-represented; examples include (7) and (8).

- (7) His strok **may no man stand** (c. 1330, *Sir Tristrem*, 2335; MED)
“No man can (with)stand a blow from him.”
(8) There **myghte none** stande hym a **stroke**.
(1470–85 Malory *Morte d'Arthur* x. lxxiv. 543)
“There could no one (with)stand a blow from him”

This transitive usage of *stōnden* is entirely parallel to a prominent usage of *astōnden* “stand one’s ground, offer resistance; withstand, resist (an enemy, a blow, etc.)”, as per the MED. Given that the *a-* prefix of English (< PGmc. **uz-*) was often apherized (see OED s.v. *a-*, prefix), it is possible that some senses of *stand*, especially relating to physical endurance, derive from Old English *astandan*, which has fallen together with *stand* from the Middle English period (the latest citations of *astand* in the OED are c. 1400, in the MED c. 1500). From the Middle English period, *astand* appears transitively in senses approaching “tolerate” (see also MED s.v. *astōnden*); an example is shown in (9).

- (9) Theih bien londes and ledes, **ne may** hem non **astonde**.
(c. 1330 *Political Songs of England*, 338; MED)

“They are lands and peoples (which) no one can tolerate.”

Transitive usages of *stand* in the EEBO corpus (1470–1700; Davies, 2017) between 1600 and 1700 predominantly attest to resistance or endurance of an explicitly physical kind, e.g. *stand a blow*, *stand a fight* (see Appendix 4 for a fuller list). Such instances are possible with or without negation, and with or without a modal verb; see Example (10) below.

- (10) *volp*: o god, sir! i were a wise man, **would stand the fury of a distracted cuck-old**
(1616, *The workes of Beniamin Ionson*; EEBO)

From c. 1600, transparent usages of *stand* as a psych verb, including with pronominal objects (equivalent to the Present-day English usage shown in Tables 1 and 2) are first attested, and a semantically equivalent expression *stand out* OBJECT first appears.³ Interestingly, *stand out* OBJECT (~ 120× in the EEBO) tends to fill the object slot with noun phrases belonging to the very same semantic sphere – an adverse thing or event – as the objects with *stand* OBJECT above; see Examples (11) and (12). Corpus searches for instances of *stand out* (ARTICLE) NOUN/PRONOUN and *stand* (ARTICLE) NOUN/PRONOUN *out* in the COCA (Davies, 2008–), COHA (Davies, 2010–), BNC (Davies, 2004), and Hansard (Davies, 2015), strongly suggest that *stand out* ADVERSE EVENT/THING was seemingly moribund by the 19th century (at least in American English and the British Parliament): it occurs but only 4× or 5× in COHA (4× 19th century, perhaps only one 20th century instance) and 1× (1888) in the Hansard corpus. The latest unambiguous attestations of *stand out* (ARTICLE) NOUN/PRONOUN known to us are found in 1949 (COHA), where *stand out the night* occurs twice (thanks to an anonymous reviewer for bringing this attestation to our attention).

- (11) i **can not stand** nosing of Candlestickes, or euphuing of similes, alla sauoica.
(1593, *Pierces supererogation*; EEBO)
- (12) if noble lisbon **could not stand it out**, where is that city so resolv'd, and strong....
(1655, *The Luisad, or Portugals historical poem*; EEBO)

The data from Early Modern English indicate that the reading of *stand* as a psych verb, indicating **mental/emotional endurance**, had emerged by at least this period, and the usage of *stand* (*out*) in the relevant sense became increasingly restricted to the context provided by the CAN NOT VERB_{INF} X_O OBJ construction in the 18th and 19th centuries, as evidenced in Examples (13) and (14).

3. Note that this usage must be distinct from *stand out* in the sense “be noticeable; be better”, e.g. *the bright lettering stands out well from/against a dark background*; *her work stands out from the rest as easily the best*, which remains in use in Present-day English.

(13) It is often said, such an **one cannot stand** the Mention of such a Circumstance.
(1710 R. Steele *Tatler* No. 225. 2)

(14) Soapy **couldn't stand out** against the big ranchmen when they got together and meant business. He had to pull his freight.
(1936, W. Raine *Crooked Trails and Straight*; COHA)

Finally, note that Present-day (British) English preserves *stand out* “endure” in the idiom *stand out against something*, which can be paraphrased as “continue to resist”, as in (15).⁴

(15) I have had **to stand out with** my editor once or twice on that point.
(1887 G. R. Sims *Mary Jane's Mem.* 296)

4.2 German: From *(ir)stan(tan)* to *nicht ausstehen können*⁵

As in English, the cognate OHG verbs *stān* and *stantan* are well-attested as basic **verbs of position**, and in contexts that admit of readings relating to **physical endurance**, seen in (16).

(16) Fóra sinen óugon **stent** alle ménnisgon, úbile joh guáte.
(Otfrid, *Evangelienbuch* 5.20; RA)
“Before his eyes stand all people, evil and good.”

As mentioned under 3, the particle verb *ausstehen* (MHG *uʒstān*) is first directly attested in the Middle High German period, but only with senses other than “endure, tolerate”, as the example in (17) shows.

(17) dô muost mein herr etlich tag dô still ligen, das die pferd ein wênig **ausstunden**.
(c. 1470 *Tetzel Rozmit*.160; MWB)
“you have to remain for a few days, my lord, so that the horses might rest (*ausstunden*) a little”

Nevertheless, the OHG prefixed verb *irstantan*, whose primary sense is “stand up, arise”, is susceptible to the reading “endure” in some instances, although employed exclusively intransitively, as, for instance, in (18).

(18) Soso uúas Ionas in thes uuales uuámbu thrí taga inti thriio naht... Thie
Nineiscun mán **arstantent** in tuome mit thesemo cunne inti furniderent iz...
(Tatian 57; RA)

4. Compare further the fixed English idiom *stand the test of time*, as well as Modern Icelandic *standast próf* “to pass an exam”.

5. Citations in this section stem from the DWB (accessed through <http://woerterbuchnetz.de/DWB> (4 November 2019), unless otherwise indicated.

“So Jonah was in the belly of the whale for three days and three nights... The man from Nineveh endured in the trial with this creature and humiliated it...”

Similarly, MHG *erstân* knows at least one clear instance in which this verb is clearly to be read in the sense “endure, tolerate”, shown in (19).

- (19) ir habt gestriten einen strît, / daz mich immer wunder hât, / wie ez iwer lîp
erstât (Garel von dem blüenden Tal 8325–7; Walz, 1892)
 “you have fought such a battle, that it amazes me, how your bodies stand it”

From the late 16th century, *ausstehen* comes to be attested in abundance (the earliest occurrence in DT: 1583), at which point its primary sense is “endure, tolerate”; an early such usage is given in (20). It may be inferred that *ausstehen* perhaps belonged primarily to the vernacular register of the language, being clearly documented once its primary sense became that of a psych verb, and sources became more abundant. *ausstehen* has conceivably replaced a possible function of older OHG *irstantan* / MHG *erstân* in the sense of “endure”.

- (20) das er von dem verbotenen Bawm gessen, so hilfft es jn doch nicht sondern
 mus die gedrawte straff **ausstehen**
 (1583, L. Polycarp, *Ein Christliche Leichpredigt* 23; Berlin-
 Brandenburgischen Akademie der Wissenschaften, 2020)
 “that he ate from the forbidden tree did him no good; rather, his wife had to
 endure punishment”

Older German does preserve traces of the inherited use of *ausstehen* as intransitive “step out”, as seen under (21) below, and has developed the transitive phrasal verb *etwas ausstehen* “to out-stay, endure something or somebody”, as in (22). Like English *to stand (out)*, *ausstehen* has adopted three restrictions: to the negated form (Restriction 1, seen in (23)), use with a superordinate modal verb (Restriction 2, seen in (24)), and the restriction of the modal verb to *can* (Restriction 3, illustrated in (25)).

- (21) So schied er ab von Babylon, in die insel Sagena kam, da stund man aus in
 gottes nam. (16th c., H. Sachs I, 171d; see also DWB 1.985, s.v. *ausstehen* sub 1)
 “So he departed from Babylon, and came to the island of Sagena, where one
 appeared (= “stood out”) in god’s name”
- (22) nach allen prüfungen, die ihr **ausgestanden** habt
 (18/19th c., Goethe 14, 196; see also DWB 1.985, s.v. *ausstehen* sub 4)
 “... after all of the trials that you have endured”

Restriction 1: negative polarity item, colligation with negation, giving *steht nicht aus*:

- (23) er **steht** den schmerz **nicht** aus, er überwältigt ihn. (18th c., Zachariä 1, 179)
 “he cannot tolerate the pain, it overwhelms him”

Restriction 2: restriction to infinitival use after modal verbs *kann, darf ... nicht ausstehen*:

- (24) das kein unglück so grosz ist, es sei geistlich oder leiblich, das ich **nicht künde ausstehen** und überwinden. (Luther 6, 345a)
 “... that no misfortune, whether spiritual or physical, is so great that I cannot endure it and overcome it”

Restriction 3: further restriction of modal verb to *kann, kann nicht ausstehen*:

- (25) sie konnte ihren herrn vater **nicht** eher **ausstehn**, bis er u. s. w. (18/19th c., Klinger 1, 168)
 “she could not stand her father, until he etc. ...”

4.3 Summary of attested diachronic developments

The available evidence points towards the diachronic development of restrictions on the English (phrasal) verb *stand (out)* “bear, tolerate” and German *ausstehen* along the following path, binding themselves to transitive negated modal verb constructions:

- i. **Semantic narrowing:** *stand (out)* and *ausstehen* are originally verbs of position, used to denote physical endurance, but which have undergone a semantic narrowing to become psych verbs denoting a negative psychological attitude towards a person, thing, or experience. *stand (out)* and *ausstehen* have likely replaced older forms OE *astandan* and OHG *irstantan*, containing a prefix etymologically related to E. *out*, G. *aus*.
- ii. **Negative polarity:** Both verbs adopt a constraint of occurring only in the negated form; the non-negated use persists only in fixed idioms like *Ängste (Todesängste) ausstehen*.
- iii. **Negated modal verb and nonfinite constraint:** both *stand (out)* “bear, tolerate” and *ausstehen* are increasingly restricted to infinitives under the negated modal *can/können*.

5. Quantitative analysis of colligations, cohesion, and linguistic age

Considering the deeper historical background of *stand (out)* and *ausstehen* in the foregoing, further evidence pointing towards a strong colligational relationship between the CAN NOT VERB_{INF} X₀ OBJ construction and the verbs *stand/ausstehen*

in Present-day English and German can be mustered. In this section, synchronic evidence that speaks in favor of clear and well-established colligations will be systematically assessed; this evidence of colligational strength is in turn argued to be indicative of the colligation's relatively greater antiquity.

5.1 Collostructional analysis and constructional productivity

Potentially, cohesion between the lexemes *stand* or *ausstehen* and the relevant English and German constructions could become evident prosodically through the univerbation or phonetic reduction of elements in the phrase, independent of any regular synchronic phonological processes or broader sound changes in progress. Since, however, no transparent phonetic or phonological traces of chunking are available for the phrases *can't stand* and *nicht ausstehen können*, evidence pointing towards a strong colligational relationship between the relevant lexemes and the construction must be assessed from a purely quantitative perspective, with the help of large-scale corpora. The methods applied here fall under the label of COLLOSTRUCTIONAL ANALYSIS (Gries & Stefanowisch, 2003; Evert, 2004: Chapter 3).

To compare the colligational behavior of the phrases under discussion with synonyms or semantically similar lexemes is useful to this end. For instance, the collostructional strength of English *can't stand sb./sth.* can be compared with that of *can't tolerate someone/something*. For German, the collostructional strength of *jdn./etw. nicht ausstehen können* can be compared with that of *jdn./etw. nicht leiden können*, etc. These are specific instantiations of the more general, abstract construction characterized by the presence of a transitive verb in its infinitival form standing under the modal verb CAN and negation, which was introduced in Section 1: CAN NOT VERB_{INF} X_O OBJECT_{ACC}.

The synonymous usages of this construction that concern us here (*can't stand someone/something*, *can't bear someone/something*, etc.) can in turn be compared with a sample of other non-synonymous verbal collexemes (e.g. *make*, *kill* and their German counterparts) in the construction, to provide a fuller picture of the extent to which verbs such as *stand* and *bear* cohere with this construction. Section 5.2 presents the data, quantitative collostructional analysis, and assessment for for English (5.2.1) and German (5.2.2), respectively. Details on the procedures of data collection are available in Appendix 5. Likewise provided in Appendix 6 is an examination of the PRODUCTIVITY of specific instantiations of the *can't stand* construction (e.g. CAN NOT STAND_{INF} X_O OBJECT, CAN NOT TOLERATE_{INF} X_O OBJECT, etc.), following the methodology for the corpus-based measurement of morphological productivity developed in Baayen (1989, 1992) and further applied to larger syntactic units in Zeldes (2012) and Bozzone (Forthcoming). That investi-

gation is intended to test the hypothesis that constructions with quantitatively lower relative productivity may be suspected of being diachronically older.

5.2 Colligational strength in the CAN NOT VERB_{INF} X_O OBJECT_{ACC} construction: Data collection and quantitative analysis

The question to be quantitatively investigated here is: how strongly does a lexeme like *stand* or *tolerate* prefer to occur in the context of the CAN NOT VERB_{INF} X_O OBJ construction? Intuitively put: colligational strength is greater the more that the tokens of a lexeme occur in a given construction, beyond what the rates of occurrence of the lexeme and the construction would independently predict.

How best to determine the degree of collocational (in the case of word sequences) or colligational (in the case of partially abstract syntactic constructions) strength remains a debated issue. Conservatively, one might conclude that no single statistic decisively represents collostructional/colligational strength in absolute terms. For this reason, several different statistics will be reported and interpreted. Examination of the correlation between various association measures in Levshina (2015: 235–239) indicates that there may be three major categories: measures that indicate the dependence of a collexeme on a given construction (e.g. RELIANCE of Schmid, 2000, or ΔP with collexeme cue of Ellis, 2006), measures that indicate how strongly a collexeme is drawn to a construction (e.g. ATTRACTION of Schmid, 2000, or ΔP with constructional cue of Ellis, 2006), and some measures that indicate general cohesion between a collexeme and a construction (e.g. Pearson's Chi-Squared Test or the chi-squared log-likelihood ratio test). In this connection, the general conclusion of Wiechmann (2008) is worth reporting: his MINIMUM SENSITIVITY POINT ESTIMATE (the smaller of RELIANCE and ATTRACTION) showed the best correlation to a regression model fit to a psycholinguistic experiment concerned with identifying fixed processing units (i.e. idioms or constructions; see Kennison, 2001, for these results). See further Evert (2004: especially Chapter 4) on the mathematical properties of assorted association measures.

To present a full portrait of collostructional strength, we report in this section all association measures mentioned above. For calculating these statistics, four values are needed: the frequency of the lexeme in the construction, the frequency of the lexeme overall, the frequency of the construction overall, and the frequency of constructions containing the POS (part of speech) of the lexeme (i.e. the frequency of all verbal constructions). These quantities reach into the thousands or hundreds of thousands of occurrences in the corpora consulted, and therefore certainly contain some false positives and false negatives.

5.2.1 English

Quantitative data on the token frequency of the relevant construction, with specific verbal collexemes, the token frequency of those verbal collexemes, and the token frequency of verbal constructions in general in English was harvested from the Corpus of Contemporary American English (COCA; Davies, 2008–).⁶ Table 6 reports the total token frequency of the specific verbal collexemes in the transitive CAN NOT VERB_{INF} construction and the corresponding token frequency of the verbal lemma in the COCA.⁷ 2×2 contingency tables for each verbal collexeme were then constructed, to allow for the calculation of association measures. The contingency table for *stand* (Table 7) is given as an example.⁸

Table 6. Frequency of English Verbal Lemmata and Occurrence in the CAN NOT VERB_{INF} X_O OBJECT_{ACC} Construction

Verbal collexeme	CAN NOT VERB _{INF} X _O OBJECT _{ACC} Freq.	Verbal lemma Freq.
<i>Stand</i>	2266	197108
<i>Tolerate</i>	340	7866
<i>Bear</i>	818	29943
<i>Do</i>	7341	3676671
<i>Make</i>	5093	1219896
<i>Kill</i>	337	134186
<i>Love</i>	147	159950

On the basis of the respective contingency tables, the following summary statistics as measure of colligational strength were calculated for each verbal collexeme: RELIANCE and ATTRACTION (Schmid, 2000), ΔP with both collexemic and constructional cue (Ellis, 2006), MINIMUM SENSITIVITY POINT ESTIMATE (Wiechmann, 2008), Pearson's chi-squared statistic, and the chi-squared log-likelihood ratio (*G*-Test). All statistics were calculated using *R* version 3.5.2 (*R*

6. All queries were first conducted on May 1, 2019, then later updated and corrected on October 17 and October 21, 2019.

7. In the case of *stand* and *bear*, manual checking of the data allowed us to exclude tokens that did not correspond to the semantic reading “cannot endure, cannot tolerate”, e.g. the collocations *bear witness* and *bear children* in the case of *bear*.

8. A reviewer observes that the value obtained for the token frequency verbal constructions (100535787) in the COCA may not be entirely accurate, since the query used (see Appendix 5) likely results in an overcount. Since this same value was used in constructing the contingency tables for all collexemes, the general tendency of the statistics, and the qualitative conclusions derived therefrom, should not be adversely affected.

Table 7. Contingency Table for *stand* and the CAN NOT VERB_{INF} X_O OBJECT_{ACC} Construction

	CAN NOT Cx	¬ CAN NOT Cx	TOTAL
<i>Stand</i>	2266	194842	197108
¬ <i>Stand</i>	158908	100179771	100338679
TOTAL	161174	100374613	100535787

Core Team, 2018); Pearson’s chi-squared statistic and the likelihood ratio were obtained using the function `assocstats()` from the package *vcd* (Meyer et al., 2017); the other measures of colligational strength were calculated using functions written by the second author. All statistics are reported rounded to five decimal places, reported in Table 8.

Table 8. Measures of Colligational Strength for English Verbs in the CAN NOT VERB_{INF} X_O OBJECT_{ACC} Construction

Verbal collexeme	ATTRACTION	RELIANCE	ΔP collexeme	ΔP construction	MINIMUM SENSITIVITY	Pearson’s χ^2	χ^2 likelihood ratio
<i>stand</i>	0.01406	0.0115	0.00991	0.01212	0.0115	12076.5	5071.4
<i>tolerate</i>	0.00211	0.04322	0.04162	0.0020	0.00211	8514	1599.9
<i>bear</i>	0.00508	0.02732	0.02572	0.00479	0.00508	12374.7	3122.7
<i>do</i>	0.04555	0.002	0.00041	0.00899	0.002	369.17	343.28
<i>make</i>	0.0316	0.00417	0.0026	0.0195	0.00417	5102.9	3545.1
<i>kill</i>	0.00209	0.00251	0.00091	0.00076	0.00209	69.256	58.993
<i>love</i>	0.00091	0.00092	−0.00069	−0.00068	0.00091	46.844	55.416

The surveyed collexemes may be divided into four groups on the basis of measures of colligational strength: *stand* alone, *tolerate* and *bear*, *do* and *make*, and *kill* and *love*. The last group, *kill* and *love*, is only very weakly associated with the CAN NOT VERB_{INF} X_O OBJECT construction, and *love* is in fact repelled from it (see ΔP). For all collexemes surveyed here other than *love*, more tokens were observed in the construction than would be independently expected. *kill* and *love* thus exhibit patterns of usage largely independent of the construction. *Make* and *do*, meanwhile, exhibit the highest degrees of attraction to the construction, but the comparatively lower degrees of reliance. Conversely, *bear* and *tolerate* show the highest degrees of reliance, but comparatively lower degrees of attraction. Finally, in the case of *stand*, this collexeme is neither as reliant on the CAN NOT VERB_{INF} X_O OBJECT construction as its synonyms *bear* and *tolerate*, nor as strongly attracted

to the construction as the very high frequency lexemes *do* and *make*. Precisely because *stand* shows neither a very small value for ATTRACTION nor a small value for RELIANCE, it has the highest MINIMUM SENSITIVITY value of all collexemes surveyed (exceeding 0.01).

These patterns largely replay themselves in the χ^2 -statistics, where *stand*, *tolerate*, *bear* exhibit the largest values Pearson's χ^2 -test, while *do* and *make* have comparatively larger likelihood ratio values, and *kill* and *love* the smallest values. *Bear*, in fact, shows the largest χ^2 -value overall (12374.7), but which is not considerably larger in the range of values seen here than the value for *stand* (12076.5). The χ^2 likelihood-ratio test points in a similar direction, where *kill* and *love*, followed by *do* return the smallest values; *make*, however, obtains a larger value than that of *tolerate* and the value for *stand* is largest (5071.4).

Overall, these results lend themselves most readily to the interpretation that *stand* is most strongly associated with the CAN NOT VERB_{INF} X_O OBJECT construction among all of the collexemes surveyed. In particular, *stand* appears to outrank its potential synonyms in this construction, *tolerate* and *bear*, in terms of ATTRACTION and the degree to which the construction cues the collexeme *stand*. On the other hand, because *tolerate* and *bear* are considerably less frequent than *stand* (see Table 7) above, their usage is more dependent upon the construction under investigation. With respect to the control collexemes *do*, *kill*, and *love*, *stand* and *bear*, and to a lesser extent, also *tolerate* and *make*, are clearly more strongly associated with CAN NOT VERB_{INF} X_O OBJECT_{ACC}.

From these quantitative results, we can draw the conclusion that *stand* in the sense typically assumed by *can't stand sb./sth.* has a stronger connection to the CAN NOT VERB_{INF} X_O OBJECT than most verbs of English. This comparatively strong colligational association is then to be attributed to the restrictions placed on *stand* in the relevant sense discussed in Sections 1 and 3. If a phrasal verb **stand out* directly cognate to German *ausstehen* and Dutch *uitstaan* were in use in contemporary English, the English results for **stand out* might then more closely resemble the quantitative picture seen for German *ausstehen* in 5.2.2 below, with greater dependence upon the CAN NOT VERB_{INF} X_O OBJECT_{ACC} construction – the reason that *stand* does not exhibit greater dependency on this construction is partly a consequence of the generally high frequency of this lexeme, which occurs in a wide range of distinct usages.

5.2.2 German

To acquire frequency data on the colligation under investigation for German, queries were applied to morphologically tagged subcorpora TAGGED-C and TAGGED-C2 of the Deutsches Referenzkorpus (DeReKo; Kupietz, 2019) through

the COSMAS-II interface.⁹ In contradistinction to English, the fact that, in German, different constituent orders occur in main clauses as opposed to (most types of) subordinate clauses, different query terms were needed to accommodate and find these different possible word orders. In the following, we represent the relevant construction for German as x_o OBJECT_{ACC} NICHT VERB_{INF} KÖNNEN, since we take the constituent order in subordinate clauses to reflect the underlying order of constituents prior to the V-to-T-to-C raising (V2; see Holmberg, 2015) of tensed verbs and verbal auxiliaries in main clauses. At least one further option is syntactically possible in main clauses: the topicalization or focusing of the verbal infinitive in initial position of a main clause, giving rise to structures such as *Ausstehen konnte er sie überhaupt nicht* “stand her, he absolutely could not”. This structure in an abstract form proved difficult to query properly in COSMAS-II, and searches with concrete verbal collexemes revealed it to be very uncommon. Therefore, instances of x_o OBJECT_{ACC} NICHT VERB_{INF} KÖNNEN with topicalization/focus of the infinitive are excluded from the results below.

Table 9 reports the total token frequency of the specific verbal collexemes in the transitive x_o OBJECT_{ACC} NICHT VERB_{INF} KÖNNEN construction and the corresponding token frequency of the verbal lemma in the TAGGED-C and TAGGED-C2 subcorpora of the DeReKo.

Table 9. Frequency of German Verbal Lemmata and Occurrence in the x_o OBJECT_{ACC} NICHT VERB_{INF} KÖNNEN Construction

Verbal collexeme	x_o OBJECT _{ACC} NICHT VERB _{INF} KÖNNEN Freq.	Verbal lemma Freq.
<i>ausstehen</i> “stand (out)”	718	5938
<i>aushalten</i> “tolerate”	150	13775
<i>leiden</i> “suffer”	672	88363
<i>ertragen</i> “bear, endure”	397	16231
<i>tun</i> “do”	809	674154
<i>machen</i> “make”	2728	2651446
<i>töten</i> “kill”	41	98558
<i>lieben</i> “love”	95	75608

2x2 contingency tables for each respective verbal collexeme were then constructed. The contingency table for *ausstehen* is given as an example (Table 10). Just as with the English data above, summary statistics as measures of colligational

9. All queries were first conducted on May 8, 2019, then updated and corrected on October 20, 2019.

strength were then calculated, using the same array of statistics explained. Results are reported in Table 11.

Table 10. Contingency Table for *ausstehen* and the x_o OBJECT_{ACC} NICHT VERB_{INF} KÖNNEN Construction

	NICHT VERB _{INF} KÖNNEN Cx	¬ NICHT VERB _{INF} KÖNNEN Cx	TOTAL
<i>ausstehen</i>	718	5220	5938
¬ <i>ausstehen</i>	360524	260791754	261152278
TOTAL	361242	260796974	261158216

Table 11. Measures of Colligational Strength for German Verbs in the x_o OBJECT_{ACC} NICHT VERB_{INF} KÖNNEN Construction

Verbal collexeme	ATTRACTION	RELIANCE	ΔP collexeme	ΔP construction	MINIMUM SENSITIVITY	Pearson's χ^2	χ^2 likelihood ratio
<i>ausstehen</i>	0.00199	0.12092	0.11954	0.00197	0.00199	61423.1	5090.3
<i>aushalten</i>	0.00042	0.01089	0.00951	0.00036	0.00042	901.20	358.41
<i>leiden</i>	0.00186	0.0076	0.00622	0.00152	0.00186	2477.1	1195.4
<i>ertragen</i>	0.0011	0.02446	0.02308	0.00104	0.0011	6257.6	1540.9
<i>tun</i>	0.00224	0.001	-0.00018	-0.00034	0.001	16.424	17.198
<i>machen</i>	0.00755	0.00103	-0.00036	-0.0026	0.00103	243.50	267.18
<i>töten</i>	0.00011	0.00042	-0.00097	-0.00026	0.00011	66.776	92.252
<i>lieben</i>	0.00026	0.00126	-0.00013	0.00005	0.00026	0.87962	0.90774

In contrast to the English data, the German data may first be divided into two groups on the basis of the ΔP statistics: one the one side stand *ausstehen*, *aushalten*, *leiden*, and *ertragen*, which appear to be genuinely attracted to and reliant upon the x_o OBJECT_{ACC} NICHT VERB_{INF} KÖNNEN construction; on the other stand *tun*, *machen*, *töten*, and *lieben*, which are repelled from and occur independently of the construction under investigation. In the case of *tun*, *machen*, and *töten*, the statistically significant χ^2 -values result from observed frequencies that are markedly lower than their corresponding expected values (e.g. for *machen*, Observed = 2728, Expected = 3668). *tun*, *machen*, *töten*, and *lieben* thus do not build colligations with the NICHT VERB_{INF} KÖNNEN construction.

Conversely, the first four verbal collexemes to appear to build part of colligations with this construction, with varying degrees of strength. The two χ^2 -statistics and MINIMUM SENSITIVITY all point to the same overall ranking of colligational strength: *ausstehen* (strongest), *ertragen*, *leiden*, *aushalten* (weakest). In terms of Schmid (2000)'s ATTRACTION and RELIANCE, remarkable is the comparatively

high degree of RELIANCE exhibited by *ausstehen* (about $5\times$ greater than its nearest competitor, *ertragen*), and the relatively strong degree of ATTRACTION exhibited by *leiden*. In the former case, high RELIANCE but middling ATTRACTION points to a collxeme that is highly dependent upon the NICHT VERB_{INF} KÖNNEN construction; in diachronic terms, it appears that *ausstehen* has become frozen in this colligation – the verbal lexeme cannot be employed in German with great freedom (see examples in Sections 1 and 3). *leiden*, meanwhile, represents a high-frequency verb that occurs in the context of the the NICHT VERB_{INF} KÖNNEN construction significantly more often than would be expected, hence its relatively greater values for ATTRACTION and ΔP Construction. *ertragen*, then, appears to fall into a place similar to *ausstehen*, although it has not yet become as dependent on NICHT VERB_{INF} KÖNNEN as the latter verb. In summary, the connection between *ausstehen* and the NICHT VERB_{INF} KÖNNEN construction is remarkably strong: *ausstehen* exceeds the value for every other collxeme on every measure of colligational strength employed.

The overall conclusion may again be drawn that x_o OBJECT_{ACC} NICHT *ausstehen* KÖNNEN represents a strong colligation in comparison to other transitive verbs generally in German and potentially synonymous verbs as well. In contrast with English *stand*, however, *ausstehen*'s locus of occurrence is much more dependent upon the relevant construction, whereas English *stand*, while attracted to CAN NOT VERB_{INF} x_o OBJECT_{ACC}, is not nearly so dependent on it. As mentioned at the end of Section 5.2.1, the divergent behavior of x_o OBJECT_{ACC} NICHT *ausstehen* KÖNNEN versus CAN NOT *stand* x_o OBJECT_{ACC} must be due to the fact that English *stand* in the relevant sense is morphologically identical to other forms of *stand*, in a wide range of senses. If a phrasal verb **stand out* were still in use in contemporary English, we might predict for it to similarly be very dependent upon the CAN NOT VERB_{INF} x_o OBJECT construction.

6. Conclusion

The present case study on E. *stand* “endure, tolerate” and G. *ausstehen* (as well as Du. *uitstaan* and Swed. *utstå*), as most clearly transmitted in the *can't stand* construction, showcases the lifecycle of a verb and its construction over several millennia. Departing from an original verb of position in PIE, the verb in question comes to denote physical/cognitive endurance, finally becoming a psych verb of (negated) mental/emotional endurance in West Germanic, usually embedded in a specific constructional context. The high degree of colligational strength between E. *stand* in the relevant sense and G. *ausstehen*, respectively, and the

CAN NOT VERB_{INF} X_O OBJECT_{ACC} construction may be attributed to lexical competition with other innovative verbs of the same or similar meaning, following the “Kuryłowicz-4” mechanism discussed in 2.2.1. Indeed, *stand* “endure, tolerate” and *ausstehen* are effectively limited to occurring in the CAN NOT VERB_{INF} X_O OBJECT_{ACC} construction. As shown in Section 3 and Appendix 2, these Germanic lexemes directly continue the PIE verbal root lexeme **steh₂* “stand (upright) firmly”, which built a causative stem PGmc. **standa-* active “to cause sb/sth to stand upright”, middle “to cause oneself to stand upright (against sb/sth)”, and which metaphorically meant “to be able to sustain, to bear, endure”. Comparative evidence from Indo-Iranian and Slavic, and across the diachrony of the Germanic languages, suggests that the verb optionally permitted the local and directional preverbs Gmc. **ūt* “on high”, and **ut-s* “upward”, as orally transmitted, vernacular variables (see Section 2.2.2, Section 4 on E. *stand (out)*, and G. *ausstehen* as vernacular lexemes, and Appendix 2). PIE **ut-s* “upward” prefixed to **standa-* produced PWGmc. **urstanda-*, the ancestor of OE *astandan* and OHG *irstantan*, which are attested with the sense “endure”; the prefixed verb **urstanda-* likely competed with a vernacular particle verb construction **standa- üt*, **sta(j)i/a- üt* which underlies EModE *stand out* and G. *ausstehen*.

Evidence from the OE and OHG corpora indicates that the continuants of PGmc. **standa-* could readily enter into constructions with negation and modal verbs. In West Germanic, the verb with particle **ūt* adopted three colligational restrictions, namely, (i) negative polarity and negated use, (ii) use with modal verbs, involving negation transfer to the modal verb, (iii) restriction of the negated modal verb to E. *can not*, G. *nicht können*, Du. *niet kunnen*. The rise of these three restrictions can be attributed to lexical competition with other etymologically younger lexemes of parallel function, thereby leading to the virtual restriction of *stand* “endure, tolerate” and *ausstehen* to usages in the CAN NOT VERB_{INF} X_O OBJECT_{ACC} construction. The contrastive collostructional analysis in Section 5 and the analysis of constructional productivity in Appendix 6 ultimately indicate that there might exist a relationship between the age of a lexeme and the extent to which it becomes reliant upon a particular construction. Naturally, not all lexemes with a long prehistory are destined to become fossilized within a given constructional frame – whether such constructional reliance with concomitant restrictions arises depends precisely upon the emergence of lexical competition. Expanding upon the observation of Kuryłowicz (1945) concerning competing morphological forms, it is then the oldest lexeme among a set of competitors that is expected to be pushed into a specialized function. Whether such a tendency holds more generally requires the investigation of the diachrony of more lexemes and colligations across a wider typological array of languages. In the case at hand, though, the

marriage of comparative reconstruction and quantitative corpus linguistics indicates that *stand* in the sense “endure, tolerate” has likely undergone precisely such a development.

The following abbreviations are employed in this article

E.	English
G.	German
Du.	Dutch
OE	Old English
ME	Middle English
(E)ModE	(Early) Modern English
OHG	Old High German
MHG	Middle High German
ModG.	Modern German
Goth.	Gothic
PIE	Proto-Indo-European
(P)(W)Gmc.	(Proto)-(West-)Germanic
(Hom.) Gk.	(Homeric) Greek
Lat.	Latin
(Ved.) Skt.	(Vedic) Sanskrit
Yid.	Yiddish

References

- Arnon, I., & Clark, E. V. (2011). *Why Brush Your Teeth* is better than *Teeth* – Children’s word production is facilitated in familiar sentence-frames. *Language Learning and Development*, 7(2), 107–129. <https://doi.org/10.1080/15475441.2010.505489>
- Baayen, R. H. (1989). *A Corpus-Based Approach to Morphological Productivity: Statistical Analysis and Psycholinguistic Interpretation*. PhD Dissertation, Vrije Universiteit Amsterdam. Amsterdam: Centrum voor Wiskunde en Informatica.
- Baayen, R. H. (1992). Quantitative aspects of morphological productivity. In G. E. Booij & J. van Marle (Eds.), *Yearbook of Morphology 1991*. Kluwer. https://doi.org/10.1007/978-94-011-2516-1_8
- Berlin-Brandenburgischen Akademie der Wissenschaften. (2020). *Deutsches Textarchiv: Grundlage für ein Referenzkorpus der neuhochdeutschen Sprache* [German Text Archive: Foundation for a Reference Corpus of the New High German Language]. <http://www.deutschestextarchiv.de/>
- Bolinger, D. (1976). Meaning and Memory. *Forum Linguisticum* 1(1), 1–14.
- Bosworth, J., & Toller, T. N. (1921). *An Anglo-Saxon Dictionary, Based on the Manuscript Collections of the Late Joseph Bosworth*. Oxford University Press. <http://www.bosworthtoller.com/>

- Bozzone, C. (2014). *Constructions: A New Approach to Formularity, Discourse, and Syntax in Homer* [Doctoral dissertation, University of California, Los Angeles]. eScholarship University of California. <https://escholarship.org/uc/item/6kgoq4cx>
- Bozzone, C. (Forthcoming). Homeric constructions, their productivity, and the development of Epic Greek. In L. van Beek (Ed.), *Proceedings of the Conference "Language Change in Epic Greek and other Oral Traditions"*. Brill.
- Bybee, J. (2006). Form usage to grammar: The mind's response to repetition. *Language* 82(4), 711–733. <https://doi.org/10.1353/lan.2006.0186>
- Bybee, J. (2007). *Frequency of Use and the Organization of Language*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195301571.001.0001>
- Bybee, J. (2010). *Language, Usage, and Cognition*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511750526>
- Bybee, J., & Scheibman, J. (1999). The effect of usage on constituency: The reduction of *don't* in English. *Linguistics*, 37(4), 575–596. <https://doi.org/10.1515/ling.37.4.575>
- Davies, M. (2004). *British National Corpus* (from Oxford University Press). Retrieved July 28, 2020 from <https://www.english-corpora.org/bnc>
- Davies, M. (2008–). *The Corpus of Contemporary American English (COCA): 600 million words, 1990-present*. Retrieved October 21, 2019 from <https://www.english-corpora.org/coca/>
- Davies, M. (2010–). *The Corpus of Historical American English (COHA): 400 million words, 1810–2009*. Retrieved July 28, 2020 from <https://www.english-corpora.org/coha>
- Davies, M. (2015). *Handsard Corpus*. Part of the SAMUELS project. Retrieved July 28, 2020 from <https://www.english-corpora.org/hansard>
- Davies, M. (2017). *Early English Books Online*. Part of the SAMUELS project. Retrieved July 28, 2020 from <https://www.english-corpora.org/eebo>
- Cameron, A., Amos, A. C., dePaolo, A., Liuzza, R., & Momma, H. (2018). *The Dictionary of Old English: A–I*. Dictionary of Old English Project. <https://doe.utoronto.ca/>
- dePaolo, A., Wilkin, J. P., & Xiang, X. (2009). *Dictionary of Old English Web Corpus*. <https://tapor.library.utoronto.ca/doecorpus/>
- Donnhauser, K., Gippert, J., & Lühr, R. (2015). *Referenzkorpus Altdeutsch*. <https://www.deutschdiachrondigital.de/>
- Ellis, N. (2006). Language acquisition as rational contingency learning. *Applied Linguistics* 27(1), 1–24. <https://doi.org/10.1093/applin/ami038>
- Evert, S. (2004). *The Statistics of Word Cooccurrences: Word Pairs and Collocations* [Doctoral dissertation, Universität Stuttgart]. <http://elib.uni-stuttgart.de/opus/volltexte/2005/2371/>
- Firth, R. (1964). *Papers in Linguistics 1934–1951*. Oxford University Press.
- Gries, S. Th. (2009). *Quantitative Corpus Linguistics with R: A Practical Introduction*. Routledge. <https://doi.org/10.4324/9780203880920>
- Gries, S. Th., & Stefanowitsch, A. (2003). Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics*, 8(2), 209–43. <https://doi.org/10.1075/ijcl.8.2.03ste>
- Grimm, J. (1854–1961). *Deutsches Wörterbuch von Jacob und Wilhelm Grimm* [German Dictionary by Jacob and Wilhelm Grimm]. Hirzel. <http://www.woerterbuchnetz.de/DWB>


- Hackstein, O. (2012a). When words coalesce: Chunking and morphophonemic extension. In H. Craig Melchert (Ed.), *The Indo-European Verb. Proceedings of the Conference of the Society for Indo-European Studies, Los Angeles 13–15 September 2010* (pp. 87–104). Reichert.
- Hackstein, O. (2012b). Das Gedächtnis der Sprache. Sprachwandel und Gegenwart – welche Bedeutung besitzt die sprachliche Vergangenheit für die Gegenwart? [The Memory of language. Language change and the present day – what significance does the linguistic past have for the present?] *Akademie Aktuell. Zeitschrift der Bayerischen Akademie der Wissenschaften* 01/2012, 12–17. <https://www.academia.edu/14253166/>
- Haiman, J. (1994). Ritualization and the development of language. In W. Pagliuca (Ed.), *Perspectives on Grammaticalization*. (pp. 3–28). John Benjamins. <https://doi.org/10.1075/cilt.109.07hai>
- Hirzel. (2006–). *Mittelhochdeutsches Wörterbuch*. <http://www.mhdwb-online.de/>
- Holmberg, A. (2015). Verb second. In T. Kiss & A. Alexiadou (Eds.), *Syntax: Theory and Analysis*. de Gruyter.
- Kennison, S. (2001). Limitations on the use of verb information during sentence comprehension. *Psychonomic Bulletin and Review*, 8, 132–138. <https://doi.org/10.3758/BF03196149>
- Kupietz, M. (Ed.). (2019). *Das Deutsche Referenzkorpus – DeReKo*. Retrieved October 21, 2019 from <https://cosmas2.ids-mannheim.de/cosmas2-web/>
- Kuryłowicz, J. (1945). Le nature des procès dits analogiques [The nature of so-called analogical processes]. *Acta Linguistica*, 5, 15–37. <https://doi.org/10.1080/03740463.1945.10410880>
- Labov, W. (1989). The child as linguistic historian. *Language Variation and Change*, 1, 85–97. <https://doi.org/10.1017/S0954394500000120>
- Labov, W. (2001). *Principles of Linguistic Change, Volume 2: Social Factors*. Blackwell.
- Lehecka, T. (2015). Collocation and colligation. In J.-O. Östman & J. Verschueren (Eds.), *Handbook of Pragmatics Online* (pp. 1–20). John Benjamins. <https://benjamins.com/online/hop/>. <https://doi.org/10.1075/hop.19.col2>
- Levshina, N. (2015). *How to do Linguistics with R*. John Benjamins. <https://doi.org/10.1075/z.195>
- Lewis, R. E. (Ed.) (1952–2001). *Middle English Dictionary*. University of Michigan Press, 1952–2001. Online edition in *Middle English Compendium*. Frances McSparran, et al. (Eds.). University of Michigan Library, 2000–2018. Available online at <http://quod.lib.umich.edu/m/middle-english-dictionary/>. <https://doi.org/10.3998/mpub.6334>
- Meillet, A. (1937). *Introduction à l'étude comparative des langues indo-européennes* [Introduction to the Comparative Study of the Indo-European Languages] (8th ed.). Hachette.
- Meyer, D., Zeileis, A., Hornik, K., Gerber, F., & Friendly, M. (2017). *vcd: Visualizing Categorical Data* (Version 1.4-5) [Computer software]. <https://cran.r-project.org/web/packages/vcd/index.html>
- Oxford University Press. (1884–). *The Oxford English Dictionary Online*. <https://www.oed.com>
- Pinault, G.-J. (2015). The formation of Buddhist languages, as exemplified by the Tocharian evidence. In M. Malzahn, M. Peyrot, H. Fellner, & T.-S. Illés (Eds.), *Tocharian Texts in Context, International Conference on Tocharian Manuscripts and Silk Road Culture, Vienna, June 25–29th, 2013* (pp. 159–185). Hemen.
- R Core Team. (2018). *R: A language and environment for statistical computing* (Version 3.5.2) [Computer software]. R Foundation for Statistical Computing. <https://www.R-project.org/>

- Sinclair, J. (1998). The lexical item. In E. Weigand (Ed.), *Contrastive Lexical Semantics* (pp. 1–24). John Benjamins. <https://doi.org/10.1075/cilt.171.02sin>
- Schmid, H.-J. (2000). *English Abstract Nouns as Conceptual Shells: From Corpus to Cognition*. de Gruyter. <https://doi.org/10.1515/9783110808704>
- Walz, M. (1892). *Garel von dem blühenden Tal: Ein höfischer Roman aus dem Artussagenkreis von dem Pleier* [Garel of the Flowering Valley: A Courtly Novel from the Arthurian Cycle by Der Pleier]. Wagner'sche Universitäts-Buchhandlung.
- Wegera, K.-P., & Klein, T. (2016). *Referenzkorpus Mittelhochdeutsch*. <https://www.linguistics.ruhr-uni-bochum.de/rem/>
- Wiechmann, D. (2008). On the computation of collostruction strength: Testing measures of association as expressions of lexical bias. *Corpus Linguistics and Linguistics Theory* 4(2), 253–90. <https://doi.org/10.1515/CLLT.2008.011>
- Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511519772>
- Zeldes, A. (2012). *Productivity in Argument Selection: From Morphology to Syntax*. de Gruyter. <https://doi.org/10.1515/9783110303919>


Appendices

The appendices are available from <https://doi.org/10.1075/ijcl.20022.hac.additional>.

Address for correspondence

Olav Hackstein
 Lehrstuhl für Historische und Indogermanische Sprachwissenschaft
 Ludwig-Maximilians-Universität München
 Geschwister-Scholl-Platz 1, Postfach 130
 D-80539, München
 Germany
olav.hackstein@lrz.uni-muenchen.de
 <https://orcid.org/0000-0003-0884-5502>

Co-author information

Ryan Sandell
 Lehrstuhl für Historische und Indogermanische Sprachwissenschaft
 Ludwig-Maximilians-Universität München
ryan.sandell@gmail.com
 <https://orcid.org/0000-0003-1579-9209>

Publication history

Date received: 25 February 2020

Date accepted: 13 April 2022

Published online: 18 July 2022