Gerhard B. van Huyssteen, Adri Breed and Suléne Pilon

7 'What the hell?!' vs. 'Wat de hel?!': Contrasting the intensifying WHX construction in English and Afrikaans

Abstract: Certain constructions in a language (or even across languages) are conventionally associated with impoliteness; one such construction is the WHX construction, e.g., *what the hell* in English, or *wat de hel* in Afrikaans. It is claimed in this contribution that the latter is not merely a constructional calque from English, but that it is starting to take on a life of its own. Our methodology is in essence contrastive corpus linguistics: we firstly summarise our existing knowledge of the English construction (based on a previous corpus-driven study), before contrasting aspects of the English construction with Afrikaans corpus data. We focus specifically on the characteristics of the different slots in the construction in the two languages, and add an investigation of the constructional relations (attractions or repulsions) between specific WH pronouns (PN.WH) and noun phrases (NP) in the Afrikaans construction. Our investigation finds that: (a) only eight Afrikaans WH pronouns occur in the construction; (b) monosyllabic WH pronouns occur with a much higher frequency than disyllabic ones, but – unlike in English – disyllabic and even polysyllabic items are often found as noun phrases; (c) unlike in English, there is no mentionable difference between the frequency of *wat* 'what' and *hoe* 'how'; (d) similar to English, *hel* 'hell' consistently occurred with the highest frequency in the construction; (e) bare noun phrases occur with the highest frequency in the Afrikaans construction; (f) non-morphemic word-formation strategies, and phonological and graphemic modifications are used in leetspeak to by-pass online moderators; (g) in addition to the semantic domains RELIGION and COSMOGONY activated in the English construction, the Afrikaans construction extends to RELIGION, ANIMALS, NATURAL PHENOMENA, and NEGATIVE EVENTS, but not COSMOGONY; and (h) while the English construction primarily functions as an intensifier, we find that the Afrikaans construction additionally functions as a vehicle for humour, and that this aspect could also be investigated further for English.

Gerhard B. van Huyssteen, North-West University, 11 Hoffman St, 2531 Potchefstroom, South Africa, e-mail: gerhard.vanhuyssteen@nwu.ac.za
Adri Breed, North-West University, 11 Hoffman St, 2531 Potchefstroom, South Africa, e-mail: adri.breed@nwu.ac.za
Suléne Pilon, University of Pretoria, Lynnwood Rd, 0132 Hatfield, Pretoria, e-mail: sulene.pilon@up.ac.za

https://doi.org/10.1515/9783111477084-007

Keywords: Afrikaans, constructional calque, contrastive corpus linguistics, covarying collexeme analysis, taboo language

Warning: This chapter contains examples of language that might be offensive to some users.

Disclaimer: Our classification of racist, homophobic, and other offensive expressions is based on our subjective understanding and analysis of their context and intended meaning, and does not imply endorsement or agreement with these views.

1 Introduction

(Im)politeness is typically associated with emotive psychological states of mind, emotively motivated human behaviour, perceptions and expectations of what is appropriate or not, (dis)agreeable social interactions and relationships, cultural identity, etc. (see for instance Hickey and Stewart 2005: 1; Kádár 2017: 1; Spencer-Oatey 2005: 91). These and other aspects of (im)politeness is most often realised in linguistic behaviour – from gestures and body language to overt expressions in politeness formulae (like greetings), to subtle expressions through sarcasm, humour, and other indirect forms of communication. As such, we view specifically impolite linguistic behaviour as the (semi-/non-)intentional usage of various contextually determined expressive/emotive language acts with some kind of negative (e.g., abusive, rude, dysphemistic, etc.) perlocutionary effect (see, among others, Culpeper 2011: 23; Jay 2008: 268; Jay 2020: 39; see Hoeksema 2019 for a variety of examples).

Hence, language and linguistic analyses more often than not occupy a central position in (im)politeness studies. The predominant approaches in such studies promulgate the idea that impoliteness is not inherent in linguistic constructions, but rather a contextual judgement. However, recent research (see also Van Olmen et al., this volume) started to prove that, in addition to the importance of context, certain constructions in a language are conventionally associated with impoliteness. The argument that follows is that linguistic form and meaning pairings (i.e., constructions) deserve a more prominent place in impoliteness research. For example, Van Olmen, Andersson and Culpeper (2023: 38) conclude that "there do exist constructions in language that are (to a large extent) conventionalised for impoliteness and that impoliteness can be strongly conventionalised across languages". However, to see how widespread such inherently impolite constructions are, more research is of course required.

The aim of this chapter is to contribute to this endeavour by investigating the intensifying WHX construction (e.g., *what the hell. . .?* in English) in Afrikaans (see example (1) below), contrastively with English (example (2) below). We adapted the term "WHX construction" from Hugou (2017), who explained it as follows: "The letter 'W' designates a wh- word (*what, where, when, who*, etc.), or a word which has a similar syntax to wh- words, such as *how*. The letter 'H' refers to a noun phrase such as *the hell*, and the letter "X" stands for the rest of the interrogative clause, i.e. the proposition." In our adaptation, WH represents a WH pronoun (e.g., *what, where, how*, etc. in English), and X the noun plus its pre- and postmodifiers in either a noun phrase (e.g., *the hell*), or prepositional phrase[1] (e.g., *in the hell*). As can be seen in (3), we therefore leave "the rest of the interrogative clause" out of the construction name, since for us WH = WH word, and X = noun plus pre-/postmodifiers in a noun phrase (NP) or prepositional phrase (PP).

(1) ***Wat de fok*** *rook die ou?*
what the fuck smoke the guy
'What the fuck is he smoking?'
(OSMO-2023)

(2) *Hey, this guy doesn't know* ***what the fuck*** *consciousness is either.*
(enTenTen21-2021)

(3) **Intensifying WHX construction**
$[z_{PN.WH}\ [_{NP}\ y_{ART.DEF}\ x_{N}]\ ]$ AND $[z_{PN.WH}\ [_{PP}\ p_{PREP}\ y_{ART.DEF}\ x_{N}]\ ]$
where x = a taboo word (e.g., *fuck*), or euphemism (e.g., *flip* 'freck')

Despite it being indisputably a highly frequent, productive and ubiquitous construction in both Afrikaans and English,[2] research on the Afrikaans WHX construction is non-existent, while Hugou's (2017) corpus study is – to our knowledge – the only comprehensive study of the construction in English. In fact, Hugou (2017: 1)

1 Although prepositional phrases can also be used in the X part of the Afrikaans WHX construction (for example, *wat in die wêreld* 'what in the world'), we narrow our focus in this chapter to look only at noun phrases in the X part (also see footnote 9).

2 Preliminary investigations regarding the same construction in Dutch indicate that it might not be as frequent, productive, or ubiquitous as in Afrikaans and English. For instance, in the Dutch nlTenTen20 corpus (Sketch Engine 2020), comprising 5,890,009,964 words, *wat de hel* ('what the hell') only appears 80 times (relative frequency of 0.0136 per million words; for comparative frequencies in English, see Hugou (2017), and Afrikaans, see Section 6 of this chapter). The Dutch construction has also not been investigated properly, bar some preliminary remarks (Taalprof 2009), and a brief corpus investigation (Van der Wouden 2019).

shows that "no more than spotty attention has been paid to this construction in the existing literature". Recent literature searches using WorldCat[3] and Litmaps[4] also did not reveal any new literature on the construction since Hugou's article. This chapter therefore addresses a gap in the existing literature on this construction as an example of an inherently impolite grammatical construction.

Our research is done contrastively with Hugou's (2017), implying that we focus on those aspects where the Afrikaans construction might differ from its English counterpart. For example, while Hugou is primarily interested in the expression of intensity, using the WHX construction as a case study to explore whether lexical and syntactic variation in a construction correlate with a variation in intensity, our focus is more only on the constructional variation of the construction. We accept his findings on the expression of intensity as a priori for our study (and therefore for Afrikaans as well), which gives us the opportunity to rather focus on the variation in the Afrikaans construction, and to what extent it is similar to, or different from the English construction (i.e., in contrast with the English construction). Such a contrastive approach is pertinent to make the central claim of our research, namely that the Afrikaans construction is a constructional calque that it is starting to take on a life of its own.[5]

We define a constructional calque as a construction schema (i.e., a construction with less specified, schematic slots) from a donor language, which is realised in a recipient language by means of forms (e.g., typically phonological and/or lexical items) that are analogous to the forms in the donor language. Importantly, the donor language's original syntactic pattern, meaning, and function are by and large preserved in the constructional calque. In other words, the construction in the recipient language mimics a construction in the donor language by directly translating each component (slot) in the construction. As Michaelis and Haspelmath (2020: 1121) put it: "the meaning of a composite form in the donor language is rendered by an analogous composite form in the recipient language."

These and other kinds of interlingual "contamination" phenomena are aptly explained from a diasystematic construction grammar (DiaCxG) perspective – a sub-theory in construction grammar (CxG) that explains various emergent multi-/bilingual phenomena (Höder 2018, 2019; Höder, Prentice, and Tingsell 2021). A central notion in DiaCxG is that diasystematic (i.e., language-unspecific) schemas could instantiate specific lower-level idiosyncratic (i.e., language-specific) realisations. From this emerges the view of a multilingual constructicon that contains

3 https://search.worldcat.org/

4 https://app.litmaps.com/

5 See Taalprof (2009) and Van der Wouden (2019) for similar views on the development of the equivalent Dutch construction.

both language-unspecific diasystematic constructions (diacxns), and language-specific idiosyncratic constructions (idiocxns), which can all be active and entrenched on various levels of schematicity. Höder (2019) illustrates this with the case of the Danish idiocxn *fin* [[fi:n] ⇔ [FINE, DELICATE]]$_{DAN}$ and the German idiocxn *fein* [[fa̯in] ⇔ [FINE, DELICATE]]$_{GER}$ that result in the diacxn [[f]. . .[n] ⇔ [FINE, DELICATE]]$_{DIACXN}$. Similarly, constructional calques can be explained as diacxns that emerged due to analogous change based on formal similarities between idiocxns in two (or more) languages. We will return to this perspective in the conclusion of the chapter.
To support our central claim, we want to answer the following specific questions:

1. Which similarities in, and differences between the definite-article part, WH part, and noun phrase part of the English and Afrikaans construction can be identified?
2. What are the constructional relations (attractions or repulsions) between specific WH pronouns, definite articles, and nouns in the Afrikaans construction?

In the next section, we commence with a rather detailed overview of Hugou (2017), in order to firstly summarise those cognitive-pragmatic aspects of the construction that we do not investigate (and hence take as axiomatic for the Afrikaans construction), and secondly to identify the formal characteristics of the English WHX construction (which will be used for comparative purposes throughout the article). Section 3 contains a brief overview of our methodology and our Afrikaans dataset. Using simple frequency counts, research question 1 above is addressed in Sections 4 (the definite-article slot), 5 (the WH pronoun slot), and 6 (the noun phrase slot). By means of a covarying collexeme analysis (Gries and Ellis 2015; Stefanowitsch and Gries 2003, 2005), we investigate in Section 7 whether certain Afrikaans pronouns attract specific nouns (research question 2 above). In Section 8 we summarise and draw conclusions.

2 The English WHX construction

In his article, *The WHX construction (what the hell. . .?) and intensity: A corpus-based study*, Hugou (2017) is primarily interested in the expression of intensity, using the WHX construction as a case study to explore whether lexical and syntactic variation in a construction correlate with a variation in intensity. From a constructionist perspective, it is taken as an axiom that changes in form impose a particular construal, and therefore always have conceptual import (see for example Langacker 1991: 302). Hence, if there is variation in the construction, it should have some effect on the semantics or pragmatics of the construction. He therefore explores how the

WHX construction "is motivated by the meaning of its parts and [how] the way in which they are combined is inherently meaningful" (Hugou 2017: 5), and hypothesises that "there might be a correlation between varying degrees of intensity and the many sub-types of the WHX construction, be they lexical (*what the hell / fuck / heck / holy hell. . .*) or grammatical in nature (*what in the hell / in hell / the hell. . .*)" (Hugou 2017: 12).

He argues that the WHX construction inherits features from, among others, the subject-auxiliary inversion construction in counterfactual conditionals (*Had I known that before!*), and exclamatory-inversion constructions (*Did I know little before!*), as well as WH question constructions (e.g., *What didn't I know before?*). In speech acts, the WHX construction therefore does not express assertiveness, but rather inquiry, i.e., it expresses "an informational need and therefore an unsatisfied emotion from someone, namely a doubt, or any notion, as distinguished from fact" (Hugou 2017: 6). Because the construction "is structurally similar to an interrogative clause, but [. . .] conveys a similar illocutionary force to an exclamatory clause", it is used with a "predominantly emotive function to express different feelings such as anger, surprise or dismay; [it has] a strong subjective quality, in that [it communicates] the speaker's emotional reaction to a situation" (Hugou 2017: 7). Since it is such an emotive construction, the intensity effect "is a mere corollary, something that naturally follows or results from the expression of a strong emotion" (Hugou 2017: 25).

This emotive function makes it especially difficult (if not close to impossible) to pinpoint the exact intensity and range of emotions expressed by examples in text corpora (i.e., without pronunciation, gestures, facial expressions, etc.). He therefore investigates users' perception of intensity in various variants of the WHX construction by means of two questionnaires and follow-up email interviews – a method that falls outside the scope of our current methodology and research. His findings based on these questionnaires include:

- Frequency of use influences the perceived intensity: the more familiar a construction, the higher the intensity attributed to it. Subsequently, the [$z_{PN.WH}$ *the hell*] construction, compared to [$z_{PN.WH}$ *in the hell*] and [$z_{PN.WH}$ *in hell*], is perceived as the most intense due to its high frequency and familiarity.
- The perceived intensity of the WHX construction is closely related to the offensiveness of [x]: taboo words like *fuck* and *hell* enhance the intensity, while euphemisms like *freak*, *heck*, or *blankety blank* diminish it. Somewhat contrary to the first finding, constructions with *fuck* are perceived as more intense than those with *hell*, while *what the hell* is more frequent than *what the fuck*. This is of course unsurprising, since *fuck* has a higher taboo value than *hell*.
- Perceived intensity is also linked to the length of the realisations of the construction: shorter forms of the construction (*what the hell*) are more intense than longer forms (e.g., *wherever the hell*, or *what the thrice cursed hell*). In

fact, longer versions have "an archaic or humorous flavor" (Hugou 2017: 23), similar to other so-called attention-seeking euphemisms (Burridge and Benczes 2019: 189). He points out: "By using a longer variant (e.g. *worldwide hell, flying blue fuck*), the speaker, or rather the hearer is invited to reconceptualize his or her experience, that is view everything from a different perspective" (Hugou 2017: 23). It therefore requires more processing effort, which might distract the hearer from the actual message (Hugou 2017: 24).

– Factors related to the linguistic, socio-cultural, and situational context also play a role in the perceived intensity – as is the case in all other constructions that are heavily influenced by socio-pragmatic factors.

For the description of the constructional variation and the constraints on such variation, he uses corpus data from the *Corpus of Contemporary American English* (COCA). His observations can be summarised in terms of the WH part of the construction; the NP part of the construction; and the construction as a whole. (For ease of reference, these observations are numbered sequentially with small letters.)

2.1 The WH part of the English construction

a) Based on his analyses (Hugou 2017: 13), we can divide English WH (compounded) words according to their frequency rank order into three groups, as in (4).[6]

(4) *what* (48,328)
> *how* (1,532) > *why* (1,445) > *who* (1,367) > *where* (1,264)
> *whatever* (269) > *when* (77) > *whoever* (30) > *wherever* (28) > *however* (4) > *whenever* (1)

b) It is clear that WH compounded words rarely appear in the construction, while *which* and *whichever* never appear in the construction in Hugou's (2017) data.

2.2 The noun phrase part of the English construction

c) The noun phrase part of the construction can either be a bare noun phrase (*what **the hell***), or an NP in a prepositional phrase (*what **in the hell***), which in turn can be internally modified by adjectives (*what the **bloody** hell; what the **flying blue** fuck*).

6 Words are listed from most frequent to least frequent; in parentheses are summations of the frequency counts in Hugou's (2017: 14–16) Table 1 and 2.

In addition, the noun can be remodelled by means of non-morphemic word-formation strategies (*what the eff*; *what the f. . .*; *WTF; what the . . .*), morphemic word-formation strategies,[7] and phonological / graphemic modifications (*what the fudge; what the %!!*; what the fuuuck*).
d) The noun phrase part of the construction shows "great lexical diversity, ranging from high-frequency variants [*what the hell; who the fuck*], to on-the-spur-of-the-moment creations [*why in the sweet world of wonder; what the piss*] which, for the most part, die aborning" [our additions, from his examples] (Hugou 2017: 16).

e) The noun in the noun phrase is mostly monosyllabic (*hell, heck; fuck, freak, fudge; shit; earth, world*), similar to other taboo words (Allan and Burridge 1991: 145–146). Exceptions of course do occur, with *devil* and *God's name* the most noticeable.

f) Words from the Greek or Latin stratum (i.e., "learned words" – (Hugou 2017: 18)) are not likely to occur in the construction, also similar to other taboo constructions.

g) Longer, more ornate variants of the noun phrase (*the worldwide hell; the flying blue fuck*) dilute the intensity of the construction, as was pointed out above. Unsurprisingly, longer noun phrases tend to have an "archaic or humorous flavor" (Hugou 2017: 23).

h) With regards to the semantic fields of nouns, Hugou (2017: 18) concludes that the construction favours words and phrases belonging to RELIGION / COSMOGONY, such as *hell, devil, God's name, world*, and *earth*. In all cases the noun is used as profanity (i.e., "with careless irreverence" – see Allan and Burridge (1991: 37)), and not as blasphemy (i.e., that which "vilifies or ridicules the deity" – see Allan and Burridge (1991: 37)).

i) The two exceptions to the semantic constraints are *fuck* and *shit*, which, as Hugou (2017: 18) correctly points out, have by and large lost the connections to their original semantic fields (SEX and BODILY EFFLUVIA), and have become all-purpose intensifiers.

j) He finds that, as the noun in the noun phrase part of the construction, *hell* and its variants (N=10,127) are used consistently more frequently than *fuck* and its variants (N=1,769). He ascribes this to two potential factors, viz. that:
– His data comes from written genres (novels, magazines, newspapers); [8] and

7 No examples provided by Hugou (2017: 17).
8 A search in The Movie Corpus (containing 200 million words of English data from movie dialogues – see Davies (2019)) supports this suspicion of Hugou (2017: 15). Whereas the ratio between *hell:fuck* is 85:15 in COCA, in The Movie Corpus it is 57:43 (with N for *what the hell*=30,954; *what the fuck*=23,348). Further research is required to confirm the hypothesis that frequencies of the

– there might be a "need to use more appropriate language in certain social situations" (Hugou 2017: 15) (as reflected by the formerly mentioned written genres). Since both *hell* and *fuck* are regarded as general intensifiers in English, it is unsurprising that they occur with high frequency in the construction. Hugou (2017: 18) argues that their high frequency should be explained in terms of their general promiscuity (Taylor 2002: 266) as intensifiers in English.

k) While *hell* and *fuck* have many variants in the data, *devil* has only two: *dickens* (an established euphemism for *devil*), and *deuce*. Hugou (2017: 18) notes that their "dated character may leave little need for creativity in present-day English." He also ascribes *hell*'s survival in the construction to this "dated character" (in addition to *hell*'s aforementioned status a general intensifier)

l) Based on Bybee's (2010) definition and operationalisation of productivity (i.e., a greater variety of items appearing in the free slots of a construction is an indication that the construction is entrenched and conventionalised, and therefore more likely to be productive), he concludes that the construction is "very productive" (Hugou 2017: 16), since it has a rather high type frequency. The large number of hapax legomena of the construction is also an indication of the construction's high hapax-conditioned degree of productivity, as defined and operationalised by Baayen (1993).

m) However, since the construction favours certain semantic areas (RELIGION / COSMOGONY), as well as monosyllabic words, Hugou (2017) concludes that the construction resists full productivity.

n) Note that Hugou (2017) does not investigate variants of the article in the noun phrase, such as *what **da** fuck* (N=2 in COCA).

2.3 The English construction as a whole

o) He identifies nine different syntactic variations of the English construction (see (5) to (13) below, based on Hugou (2017: 13)), of which the equivalents of only (5) and (9) are attested in our Afrikaans data.[9] Of all these sub-constructions, (5) dis-

construction in written and spoken data will differ significantly. Also see Hugou's (2017: 26) suggestions for further research in this regard.

9 Although sentences such as (6) and (7) have been excluded from our dataset due to our search criteria (see Section 3), they do occur in Afrikaans, but – as far as we know – only in the constructions [$z_{\text{PN.WH}}$ [$_{\text{PP}}$ *op deeske aarde*] . . .*?*] and [$z_{\text{pn.wh}}$ [$_{\text{pp}}$ *op aarde*] . . .*?*] ('WH on (this) earth . . .?'). Since these cases fall outside the scope of this chapter (see footnote 1), it is left for future research.

plays by far the highest frequency in the *Corpus of Contemporary American English,* followed by (6) and (7) with much lesser frequencies (Hugou 2017: 14–15), and the others trailing far behind.

(5) **[$Z_{PN.WH}$ [$_{NP}$ ***the*** X_N] . . .?]** (*Who the hell is this guy?*)

(6) **[$Z_{PN.WH}$ [$_{PP}$ ***in the*** X_N] . . .?]** (*Where in the hell have you been?*)

(7) **[$Z_{PN.WH}$ [$_{PP}$ ***in*** X_N] . . .?]** (*Why in hell can he never be on time?*)

(8) **[***the* **[X_N] . . .?]** (*The hell you doing?*)

(9) **[$Z_{PN.WH}$ [$_{NP}$ ***the*** X_N]?]** (*"What the hell?" Johnny said with surprise.*)

(10) **[$Z_{PN.WH}$ [$_{NP}$ ***the*** X_N q_N] . . .?]** (*What the hell difference does that make?*)

(11) **[$Z_{PN.WH}$ [$_{ADVP}$ ***the*** X_N q_{ADV}] . . .?]** (*How the hell long do I have to keep doing this?*)

(12) **[$Z_{PN.WH}$'s [$_{NP}$ ***the*** X_N q_N] . . .?]** (*What's the fuck are you talking about?*)

(13) **[how'd [$_{NP}$ ***the*** X_N] . . .?]** (*Well, how'd the hell it get there?*)

r) The construction in (9) is a case of aposiopesis, where a sentence is not being completed by the speaker. Hugou (2017: 13) argues that this construction might have developed into a separate construction, especially in cases where it expresses dismissal (*I wasn't supposed to tell you yet, but what the hell!*). It is, however, most often not possible or easy to differentiate between such cases, and it is therefore categorised as a sub-construction of the WHX construction.

s) The construction has a constructional preference for the modal auxiliary *would*, as in *Why the fuck would I want to kill Dominic?* This ties in with the idea that WH questions typically "presents an incomplete proposition" (Hugou 2017: 5), even though the construction as a whole is used to "overtly express [speakers'] stance toward a certain state of affairs: the unsatisfied emotion" (Hugou 2017: 6).

t) While the construction most often appears in direct WH questions (*What the fuck are you doing here?*), it is also found in indirect questions (*I didn't know what the hell I was talking about.*). Nevertheless, the construction most often appears in initial rather than clause-final position, as first noted by Huddleston and Pullum (2002: 916).

u) Since so-called telling questions (Thompson, Fox and Couper-Kuhlen 2015) often exhibit the powerlessness of the speaker, the construction is often found in such questions (*Where the hell did you meet Virgil's wife at? Hell, I've lived next door and I've only caught a few glimpses of her myself.*). The same also applies to rhetorical questions (*How the hell would I know?*), which often "testifies to the attitudinal stance of the speaker" (Hugou 2017: 10). On the other hand, the construction is not associated with so-called reprise questions (*Tracy likes who (*the hell)?*) (Hugou 2017: 11).

v) In nominal relative clauses the construction only appears if it is introduced with a WH-*ever* compound (*I won't sit still for extortion or manipulation or. . . whatever the hell she has in mind.* vs. *What (*the hell) I like is chocolate.*) (Hugou 2017: 9).

w) In short, Hugou (2017: 21) concludes that "the WHX construction functions like expletives".

3 Afrikaans data

Against this existing knowledge of the English construction, we are particularly interested in the following aspects of the Afrikaans equivalent of the construction:

- variation in the definite-article slot of the construction;
- the frequency with which WH words occur in the construction;
- the lexical items that can occur after the definite article to fill the X part of the construction; and
- the relationship or interaction between the various WH words and X.

While Hugou's description is based on COCA data and questionnaires, our description of the WHX construction is based on data from two corpora[10] that are inherently different from his – both in size and nature. He used the *Corpus of Contemporary American English*, which contains more than 1 billion words of text (Davies n.d.), while the two corpora that we use, together contain only about 61 million words (see Table 1). Despite these differences, we will compare our findings about the

10 In the initial phase of our investigation of the Afrikaans WHX construction, we also analysed data from a third corpus, namely the *Language Commission Corpus* (TK; Taalkommissie van die Suid-Afrikaanse Akademie vir Wetenskap en Kuns 2011), which contains only edited data from Afrikaans publications. We found that the construction occurs with a very low relative frequency, and with very little variation in this corpus, most probably due to the intervention of editors of these kinds of texts. We therefore did not include the results from this corpus further in our investigation.

Afrikaans construction with Hugou's (2017) discussion of the English construction throughout, without making any claims that would require comparable corpus data.

Table 1: Corpus size, number of instances (N), and frequency per million (fpm) of the WHX construction.

Corpus	Size	N	fpm
ModCor	9,024,829	382	42
ComCor	51,693,425	2,031	39
Total	**60,718,254**	**2,413**	**40**

It is not clear whether Hugou (2017: 25) used the full *Corpus of Contemporary American English,* or only "written genres (novels, magazines, newspapers)"; we therefore assume that he used the full corpus. Our corpora cover the following:

– The *NWU Corpus of Moderated Comments* (ModCor; Centre for Text Technology (CTexT) 2020) contains commentary posts from a newspaper website, which were deleted by the newspaper's team of moderators because they were deemed to be impolite or inappropriate. This corpus is used for the initial description, since we expect to find good examples of impolite constructions here, given that the data in this corpus was already categorised as impolite/inappropriate by external respondents (the moderators). 382 instances of the construction were found, resulting in 42 instances per million words (see Table 1).
– The *NWU Commentary Corpus* (ComCor; Centre for Text Technology (CTexT) 2023) contains unedited data from the comment section of news websites and blogs. This corpus is therefore in nature very similar to ModCor, but unlike ModCor, the data in ComCor is not exclusively impolite/inappropriate. We therefore expect to find in this corpus innovative uses of the construction, where language users mask impoliteness by not using lexical items with a high taboo value in the noun phrase part of the construction. For this reason, data from ComCor is used to expand and refine the initial description of the construction. We found in total 2,031 instances of the construction, which is proportionally similar to that of ModCor, viz. 39 instances per million words (see Table 1).

To extract data, we used the CQL search string in (14); the strings to the right of the concordance results were considered to be candidates for X – i.e., nouns (*wat de **hel*** 'hell'), nouns with premodifiers (*wat de **flippen hel*** 'what the flipping hell'),

or nouns with postmodifiers (*wat de **hel op aarde*** 'what the **hell on earth**'). As will transpire in the next section, the definite article *de* 'the' occurs much more frequently than any other variations (e.g., *die* 'the') in the definite-article slot of the construction; we therefore excluded variations of the definite article from our dataset, and only use instances with *de* as definite article.[11]

(14) [word="(?i)wat.*|hoe.*|wie|waar.*|hoekom|wanneer"] [word="se"]? [word="(?i)de"]

In our dataset, all instances of grawlix (e.g., *%$&#@*) were regarded as a single type, since it is mostly impossible to determine which specific taboo item was intended by the author. Other subterfuge phenomena, like leetspeak (e.g. *f@k* 'f@ck'), redacting (e.g. *f*k* 'f*ck'), and abbreviation (e.g. *f* < *fok* 'fuck'), were normalised to the intended taboo item. The total number of tokens also includes three instances where the X part was left empty. Once again, these instances are regarded as belonging to a single type, although different taboo items might have been omitted. Spelling errors and typos (e.g. *vok* instead of *fok* 'fuck'), and spelling variants (e.g. *donner*, which is a spelling variant of *donder* 'thunder') were also normalised. However, instances of leetspeak with valid words (e.g. homophones like *vlok* 'flake' instead of *fok* 'fuck'), were left unchanged, since we regard these as examples of creative extensions of the construction. Incorrect spellings of valid words (e.g. **flok* instead of *vlok* 'flake'), were normalised to the conventional spelling.

4 The definite-article part of the WHX construction

As was shown in (14) above, we only included *de* 'the' in the definite-article part of our main corpus searches regarding the WH part, and the X part of the construction. To investigate variation in the definite-article part, we restricted our search to ComCor, and to those two instances of X that occur most in the construction, viz., *hel* and *fok* (see Section 6). The search string is presented in (15).

(15) [word="(?i)wat.*|hoe.*|wie|waar.*|hoekom|wanneer"] [word="se"]? [word=".*"] [word="(?i)hel|fok"]

11 Hugou (2017) also restricted his investigation to only *the* in the definite-article part of the English construction.

After filtering out irrelevant cases (e.g., *mense wat **die** hel se vure laat brand* 'people who stoke the fires of **the** hell'), the frequencies in ComCor are displayed in Table 2. It is quite surprising to note that *de* appears in the construction with a frequency more than ten times higher than *die*. One would have expected *die* to be the most frequent, since it is the prototypical definite article in Afrikaans, and also the highest-frequency word in Afrikaans. The word *de*, on the other hand, is usually only found in archaic fixed expressions from Dutch (e.g., *om **de** dood nie* 'by no means'), surnames (e.g., ***De** Lange*), and loan phrases (e.g., ***de** facto*). This unexpected occurrence can be explained quite easily from a diasystematic construction grammar (DiaCxG) perspective.

Table 2: Frequency of lexical items in the definite-article slot of the WHX construction in ComCor.

Rank	Definite article	N
1	*de*	1,091
2	*die*	89
3	*te*	29
4	*di*	4
5	*da*	1
6	*the*	1
Total		**1,215**

As was pointed out in Section 1, DiaCxG posits that multi-/bilingual speakers have access to both language-specific idiosyncratic constructions (idiocxns) and language-unspecific diasystematic constructions (diacxns), which emerge based on formal similarities between idiocxns in two (or more) languages. To apply these ideas to the Afrikaans WHX construction, it is important to know that most adult Afrikaans speakers can also speak English, albeit often accented, and with various idiosyncrasies (Van Rooy and Wasserman to appear). For example, it is a well-known fact that English [ð, θ] is often substituted for dental(ised) plosives [d̪, t̪] by some Afrikaans speakers.

Based on this general observation, we can now propose that in the multi-/bilingual constructicon of Afrikaans speakers, the English idiocxn *the* [[ðə] ⇔ [THE]]$_{ENG}$ and the Afrikaans idiocxn *die* [[di] ⇔ [THE]]$_{AFR}$ results in the diaxcn *de* [[də] ⇔ [THE]], so that *the* is then realised as [də] in the English of many Afrikaans speakers. In a loan construction like the WHX construction, the strong phonological similarities in the immediate textual context of ⟨*the*⟩ also strengthen this diacxn schema, e.g., on the lefthand side [[**w**ɑt] ⇔ [WHAT]]$_{ENG}$ vs. [[**v**ɑt] ⇔ [WHAT]]$_{AFR}$; and on the righthand side [[**h**ɛl] ⇔ [HELL]]$_{ENG}$ vs. [[**ɦ**ɛl] ⇔ [HELL]]$_{AFR}$, or [[f**ʌ**k] ⇔ [FUCK]]$_{ENG}$ vs.

[[fɔk] ⇔ [FUCK]]$_{AFR}$. During the process of calquing, written English ⟨*the*⟩ is then realised in writing as ⟨*de*⟩ instead of the expected ⟨*die*⟩. The fact that ⟨*die*⟩ does occur in the Afrikaans construction (see Section 3), might be a sign of a newer, more recent development, so that we can postulate the following developmental path (where round parentheses indicate non-entrenched items) as hypothesis for future research:

([vɑt di ɦɛl]$_{AFR}$) <
(⟨*wat die hel*⟩) <
⟨*wat de hel*⟩ <
[vɑt də ɦɛl]$_{AFR}$ <
[wɑt də hɛl] <
[wɑt ðɪ hɛl]$_{ENG}$

Therefore, the frequent use of *de* ‘the’ instead of the more prototypical *die* ‘the’, sets the scene for our assertion that the Afrikaans construction is a constructional calque. In the next sections, we investigate this assertion further when discussing the WH part and X part of the Afrikaans construction.

5 The WH part of the WHX construction

5.1 Identification and description (ModCor)

Table 3 presents the raw and relative frequences[12] of WH pronouns used in the WH part of the construction in ModCor, while the relative frequency results are also visualised in Figure 1. Seven out of a potential total of ten WH pronouns (see Kirsten and Breed (2020) for an overview) are used in the construction; (16) to (22) below contain an example of each).[13] As is the case for the English construction (see Section 2, point a, they can be grouped into three frequency categories, viz.:

wat ‘what’ (N=153), *hoe* ‘how’ (N=141)
wie ‘who’ (N=49), *waar* ‘where’ (N=32)
hoekom ‘why’ (N=5), *wanneer* ‘when’ (N=1), *waarom* ‘why’ (N=1).

12 Unless indicated otherwise, relative frequency in this chapter is the frequency per million (fpm) words.

13 For the sake of brevity, we don’t provide glosses for the examples, but rather only translations, since in most cases the Afrikaans words relate almost directly to the English, e.g., *hel* ‘hell’, or *fok* ‘fuck’, and *duiwel* ‘devil’.

Also like in the English construction, the four WH pronouns occurring with the highest frequency are all short (i.e. contain only one syllable) and morphologically simplex (see Section 2, point b). Longer, morphologically complex WH pronouns like *waarom* 'why' occur only once in the construction, and other morphologically complex WH pronouns like *waarop* 'on what', and *waarvoor* 'for what' do not occur at all. Also, none of the so called attributive interrogative pronouns (Kirsten and Breed 2020), like *watter* and *wat se* 'which', occurs in the construction.

Table 3: Frequency of WH pronouns in the WHX construction in ModCor.

Rank	WH pronoun	Translation	N	fpm
1	*wat*	what	153	16.95
2	*hoe*	how	141	15.62
3	*wie*	who	49	5.43
4	*waar*	where	32	3.55
5	*hoekom*	why	5	0.55
6	*wanneer*	when	1	0.11
7	*waarom*	why	1	0.11
Total			**382**	

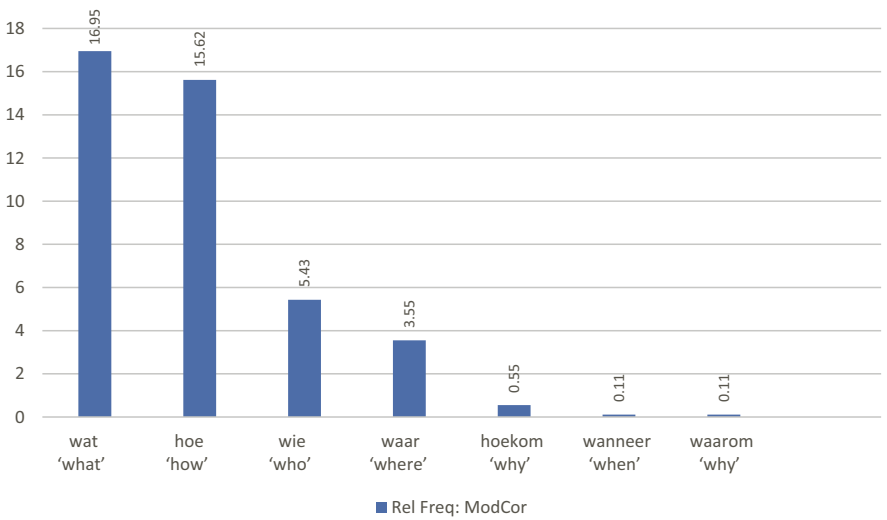

Figure 1: Relative frequency of WH pronouns in WHX in ModCor.

(16) ***Wat*** *de dinges is fout met jou man!!!?????* (ModCor-2018)
'What the thingamajig is wrong with you man!!!?????'

(17) ***Hoe*** *de duiwel kry mens dit reg . . ..* (ModCor-2019)
'How the devil does one manage that . . .'

(18) ***Wie*** *de heng noem jou kind "Success"?* (ModCor-2018)
'Who the heck calls your child "Success"?'

(19) ***Waar*** *de f%%%%% werk dit so* (ModCor-2019)
'Where the f%%%%% does it work like that'

(20) *As dit dan common is* ***hoekom*** *de joos lees jy dit?* (ModCor-2016)
'If it is common then why the devil do you read it?'

(21) ***Wanneer*** *de duiwel praat iemand due waarheid??* (ModCor-2019)
'When the devil does someone speak the truth??'

(22) *. . .* ***waarom*** *de donner is daar al die wette om hulle in blanke besighede en boerderye in te forseer.* (ModCor-2018)
'. . . why the hell [thunder] are there all these laws to force them into white businesses and farms.'

The Afrikaans WH pronouns occurring in the construction also differ from those in the English construction in two ways. Firstly, *wat* 'what' and *hoe* 'how' have very similar frequencies in the Afrikaans construction, while in English *what* has a much higher frequency than *how.* Secondly, *why* has the third highest frequency in the English construction, while the Afrikaans translation equivalents, *hoekom* 'why' and *waarom* 'why', have a much lower frequency in this construction. Since Hugou (2017) pointed out that the construction favours short lexical items, this discrepancy is probably due to the length of the Afrikaans lexical items, making them less likely to occur in the construction. This indicates that, as is the case for the English construction, the Afrikaans construction also prefers shorter lexical items in the WH part.

5.2 Verification and extension (ComCor)

Table 4 presents the raw and relative frequencies of WH pronouns in the WH part of the construction in ComCor. For comparison, the relative frequencies of the WH

pronouns in the construction in ComCor and in ModCor are visualised in Figure 2. From this visualisation it is clear that the distribution of WH pronouns used in the construction is very similar in the two corpora. The relative frequencies of the WH pronouns are also very similar in the two corpora, indicating that the construction occurs with similar frequency in both corpora.

Table 4: Frequency of WH pronouns in the WHX construction in ComCor.

Rank	WH pronoun	Translation	N	fpm
1	*wat*	what	812	15.71
2	*hoe*	how	750	14.51
3	*wie*	who	264	5.11
4	*waar*	where	171	3.31
5	*hoekom*	why	26	0.50
6	*wanneer*	when	4	0.08
7	*waarom*	why	3	0.06
8	*waarvoor*	what for	1	0.02
Total			**2,031**	

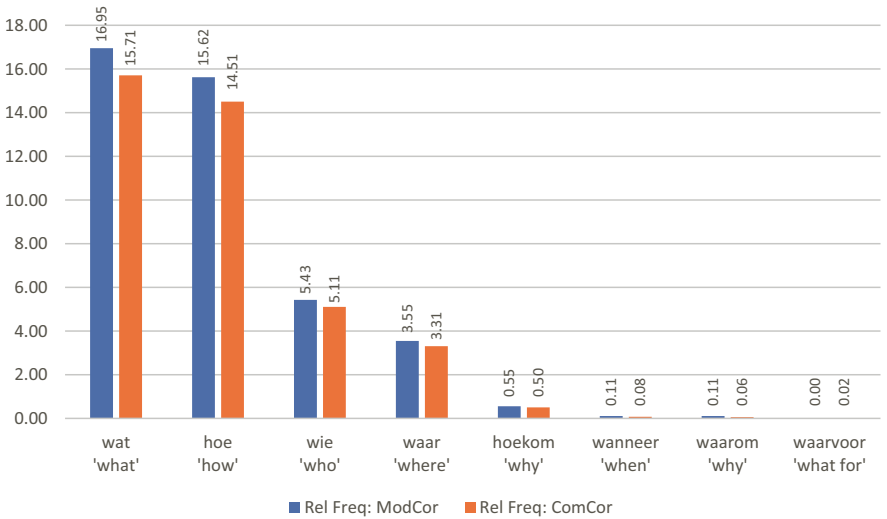

Figure 2: Relative frequencies of WH pronouns in the WHX construction in ModCor and ComCor.

Once again the WH pronouns group together in three frequency categories with *wat* 'what' (N=812) and *hoe* 'how' (N=750) occurring with a much higher frequency than the other WH pronouns. The second frequency grouping again consists of *wie* 'who' (N=264), and *waar* 'where' (N=171), and the disyllabic WH pronouns *hoekom* 'why' (N=26), *wanneer* 'when' (N=4), and *waarom* 'why' (N=3) occur with the lowest frequency. We found one additional WH pronoun, namely *waarvoor* 'what for', occurring once (see (23)) in the WH part of the construction.

(23) *Nou as alles dit deur die staat betaal word,* ***waarvoor*** *de F@K is sy salaris dan????* (ComCor-2019)
'Now if everything is paid for by the government, what the f@ck is his salary for then????'

6 The x part of the WHX construction

6.1 Identification and description (ModCor)

Table 5 contains the lexical items that are used in the X part of the WHX construction in ModCor.[14] The 382 instances consist of only 30 types, which include a type marked [GRAWLIX] (e.g. *hoe de* ***#$@&&*** 'how the #$@&&' – 28 instances), and a type consisting of [EMPTY] slots (3 instances). Examples with lexical items occurring with a frequency N>10 are presented in (24) to (31).

Table 5: Frequency of lexical items in the X part of the WHX construction in ModCor (N>10).

Rank	x	Translation	N
1	*hel*	hell	182
2	*donder*	thunder; interjection	29
3	[GRAWLIX]		28
4	*duiwel*	devil	25
5	*moer*	coffee sediment; beat up; interjection	24
6	*fok*	fuck	22
7	*joos*	devil	14
8	*ongeluk*	accident	12

14 Given that the NS are in principle an open list, across corpora, we do not provide relative frequencies for NS. Where comparisons are made, we report on the rank order in the respective corpora, based on raw frequencies.

Table 5: continued

Rank	x	Translation	N
9	*heng*	heck	7
10	*dinges*	thingamajig	5
11	*bliksem*	lightning; interjection	4
12	[EMPTY]		4
13	*vlok*	flake < fuck	4
14	*donkie*	donkey	3
15	*hoender*	chicken	2
16	*drommel*	idiot	2
17	*flip*	flip < fuck	2
18	*blikskottel*	tin plate < *bliksem* (see above)	1
19	*blou dinges*	blue thingamajig	1
20	*dêm*	damn	1
21	*vonk*	spark < fuck	1
22	*hek*	gate; heck < hell	1
23	*hemel*	heaven	1
24	*hoenders*	chickens	1
25	*otter*	otter	1
26	*poes*	pussy (person); vulva; interjection	1
27	*vak*	subject < fuck	1
28	*vloek*	curse (potentially < fuck)	1
29	*water*	water	1
30	*wetter*	yard scoundrel	1
Total			**382**

(24) *. . . want hoe de **hel** gaan jy aan na so iets?* (ModCor-2017)
'. . . because how the hell do you go on after something like that?'

(25) *Hoe de **donder** kan dit veilig wees vir kinders . . .* (ModCor-2019)
'How the heck [thunder] can it be safe for children'

(26) *ek kyk na sy foto en wonder- hoe de *****kry jy so iets reg?* (ModCor-2017)
'I look at his picture and wonder – how the ***** do you manage something like that?'

(27) *Met alle respek gesê, hoe de **duiwel** is jy 16 in graad 7.* (ModCor-2019)
'With all due respect, how the devil are you 16 in grade 7.'

(28) *Hoe de **moer** redeneer jy?* (ModCor-2018)
'How the hell [beat up] do you reason?'

(29) *hoe de **fok** neem die sap iemand in diens wat op parool is . . .* (ModCor-2018)
'How the fuck does the SAP employ someone who is on parole . . .'

(30) *. . . hoe de **joos** kry die regstelsel dit reg?????* (ModCor-2019)
'. . . how the devil does the legal system manage that????'

(31) *Hoe de **ongeluk** verduidelik jy dít dán?* (ModCor-2017)
'How the heck [accident] do you explain it then?'

Like in English (compare Section 2, point c), the X part of the Afrikaans WHX construction mostly contains only a bare noun with only one example of an NP with an adjective as modifier, viz. *blou dinges* 'blue thingamajig'. As is the case for the English construction, several of the high frequency Afrikaans nouns are also remodelled by non-morphemic word-formation strategies,[15] e.g., *f.. < fok* 'fuck'; *m. . . < moer* 'beat up'; and *d. . .r < donder* 'thunder'. We also see instances of phonological or graphemic modification of the noun in the X part, e.g., *vlok* 'flake' < *fok* 'fuck'; *vonk* 'spark' < *fok* 'fuck'; *vak* 'subject'< *fok* 'fuck'; etc.

Regarding the lexical diversity of the Ns, the Afrikaans construction is also like the English one: lexical items range from high-frequency items (e.g., *hel* 'hell'), to once-off creations (e.g., *water* 'water'; *vloek* 'curse'). We also do not see any Ns in the ModCor list that originate from the Greek or Latin stratum (see Section 2, point f). A number of nouns on the ModCor list are also, similar to English, used as all-purpose intensifiers and have largely lost the connections to their original semantic fields. These include *fok* 'fuck'; *heng* 'heck'; *bliksem* 'thunder'; *flip* 'flip'; *dêmn* 'damn'; and *poes* 'pussy' (see Section 2, points d and i).

Regarding the use of euphemisms, several interesting examples can be identified in the Afrikaans ModCor dataset. Like English, the word *duiwel* 'devil' has a euphemised form, namely *joos* 'devil' (N=14). A number of other nouns also have euphemistic variants, namely *hek* 'heck' for *hel* 'hell'; *vak* 'subject', *vlok* 'flake', *vloek* 'curse', *vonk* 'spark', and *flip* 'flip' for *fok* 'fuck'; and *donkie* 'donkey' and *drommel* 'rascal' for *donder* 'thunder' (also compare Section 2, points c and k).

Unlike the English construction, we also do not see any longer, ornate examples of noun phrases that might add an "archaic or humorous flavor" (Hugou 2017: 23)

15 While several of these remodelled examples occur in our data, only the normalised types are presented in Table 5.

to the Afrikaans construction. It is possible that these types of noun phrases are absent in the ModCor results because the humorous nature of these words might not have been regarded as impolite by the moderators, and were therefore not removed from the comments (see Section 2, point g).

According to Hugou (2017), most of the nouns in the English construction are monosyllabic, with some exceptions (such as *devil* and *God's name* as the most noticeable). The results for the Afrikaans construction differ in this regard. Only about half of the nouns on this list (i.e., 14 out of the 30 noun types) are monosyllabic (such as *hel* 'hell', *moer* 'beat up' and *joos* 'devil'), while 12 are disyllabic (such as *duiwel* 'devil', *bliksem* 'lightning' and *hoender* 'chicken'), and three are even polysyllabic (namely *ongeluk* 'accident', *blikskottel* 'tin plate' and *blou dinges* 'blue thing') (see again Section 2, point e).

With regards to semantic fields, Hugou (2017) determined that the English construction favoured words belonging to the category RELIGION (such as *hell, devil,* and *God's name),* and COSMOGONY (such as *earth* and *world)* (see Section 2, point h). The set of Afrikaans nouns in ModCor also include words belonging to the category of RELIGION, such as *hel* 'hell', *hemel* 'hemel', *duiwel* 'devil', and *jissis* 'Jesus'. However, no clear examples of nouns related to COSMOGONY have been attested. We also observe three additional semantic categories, namely ANIMALS (e.g., *otter* 'otter' and *donkie* 'donkey'), NATURAL PHENOMENA (e.g., *donder* 'thunder' and *bliksem* 'lightning'), and NEGATIVE EVENTS (e.g., *moer* 'beat up' and *ongeluk* 'accident'). Note that, as is the case in English, many of the words in the latter two categories are often used as interjections, as illustrated for *bliksem* 'lightning' in example (32), which implicates that the original semantic fields (NATURAL PHENOMENA and NEGATIVE EVENTS) are perhaps not part of the nouns' construal anymore – even more so when used in the construction.

(32) *O **bliksem**, waar val jy uit die bus uit?*
'O damn [lightning], where are you from?'

In his English dataset, Hugou (2017: 15) found that *hell* and its variants are used consistently more frequently than *fuck* and its variants (see Section 2, point j. He ascribes this to two potential factors: his data comes from written genres (novels, magazines, newspapers), and that there might be a "need to use more appropriate language in certain social situations." We see a similar phenomenon in the Afrikaans ModCor dataset, where *hel* 'hell' (N=182) is by far the most frequently used noun in the construction. However, even though *fok* 'fuck' is the most frequent Afrikaans swearword in most contexts (Van der Merwe 2022), *fok* 'fuck' is only the fourth most frequent noun on the ModCor list, and the frequency difference between *hel* 'hell' and the second most frequent noun, namely *donder* 'thunder'

(N=29), is substantial. We therefore see similar results to those of Hugou (2017), but the frequency difference between *hel* in Afrikaans and the rest of the nouns on the list is notable. We must keep in mind that our Afrikaans dataset and Hugou's dataset differ, as the ModCor dataset comprises unedited language, while Hugou's dataset has edited language. However, the reason for the large difference in frequency between *hel* 'hell' and the rest of the Afrikaans nouns on the list may be that contributors to the ModCor dataset were aware that their comments might be removed due to impolite noun choices, and might therefore have opted for the more appropriate, less offensive *hel* 'hell', rather than the inappropriate, offensive *fok* 'fuck'.

Lastly, based on Hugou's (2017) interpretation of Bybee's (2010) definition and operationalisation of productivity (see Section 2, point l), we can also conclude that the WHX construction in Afrikaans is productive. About half of the nouns on the ModCor list are hapax legomena, which serves as an indication of the construction's high hapax-conditioned degree of productivity (see again Baayen 1993). Nonetheless, even though the Afrikaans nouns shows greater variation in terms of semantic categories, the construction in Afrikaans, similar to English, resists full productivity – compare Section 2, points m and l.

6.2 Verification and extension (ComCor)

To verify and extend our list of possible noun phrases, we again compare it with data from ComCor. Table 6 presents the frequencies of noun phrases found in the X part of the WHX construction in ComCor. In total, 2,031 tokens (of 72 different types) were found, with more than 40 new lexical items that can be used in the X part of the WHX construction.

Table 6: Frequency of X in the WHX construction in ComCor.

	x	Translation	N
1	*hel*	hell	801
2	*fok*	fuck	381
3	*duiwel*	devil	151
4	*joos*	devil	93
5	[GRAWLIX]		87
6	*moer*	coffee sediment; beat up; interjection	86
7	*donder*	thunder; interjection	75
8	*ongeluk*	accident	60
9	*dinges*	thingamajig	39
10	*poes*	pussy (person); vulva; interjection	39
11	*heng*	heck	37

Table 6 (continued)

	x	Translation	N
12	*hoenders*	chickens	26
13	*drommel*	idiot	20
14	*swernoot*	scoundrel	16
15	[EMPTY]		8
16	*donkie*	donkey	8
17	*hoender*	chicken	8
18	*bliksem*	lightning; interjection	7
19	*hek*	gate; heck < hell	6
20	*vlok*	flake < fuck	6
21	*wetter*	yard scoundrel	6
22	*flip*	flip < fuck	5
23	*vy*	fig < fuck	4
24	*otter*	otter	3
25	*dêm*	damn	2
26	*flippen hel*	flipping hell	2
27	*heck*	heck < hell	2
28	*hemel*	heaven	2
29	*hoppende fok*	hopping fuck	2
30	*jissis*	Jesus	2
31	*Moses*	Moses	2
32	*vet*	fat < fuck	2
33	*vloek*	curse (potentially < fuck)	2
34	*vrek*	die (like an animal) < fuck	2
35	*zir*	zir	2
36	*actual hel*	actual hell	1
37	*bliksemstraal*	lightning bolt	1
38	*blikskottel*	tin plate < *bliksem* (see above)	1
39	*clue*	clue	1
40	*dammit*	damn it	1
41	*deksels*	lids < *donder* (see above)	1
42	*diekens*	Dickens < devil	1
43	*dooie dinges*	dead thingamajig	1
44	*falala*	falala (nonsensical, perhaps < fuck)	1
45	*fokken hel*	fucking hell	1
46	*fokietie fok fok*	fuckity fuck fuck	1
47	*frieken hel*	freaking hell	1
48	*grote griet*	great Griet < *grote God* ‘great God’	1
49	*heilige fok*	holy fuck	1
50	*hel op aarde*	hell on earth	1
51	*herrie*	Harry < hell	1
52	*hoeders*	keepers (potentially < *hoenders*)	1

Table 6 (continued)

	x	Translation	N
53	*iets*	something	1
54	*kees*	monkey	1
55	*kont*	cunt	1
56	*kriewel*	squirm	1
57	*liewe goeie heilige fok*	dear good holy fuck	1
58	*moeder*	mother < *moer* (see above)	1
59	*oertel*	asshole	1
60	*os*	ox	1
61	*pienk*	pink < pussy	1
62	*poefies*	shits	1
63	*pok*	small pox < fuck	1
64	*te hel*	too hell	1
65	*vak*	subject < fuck	1
66	*vakbond*	trade union < fuck	1
67	*vonk*	spark < fuck	1
68	*vrommel*	crumple	1
69	*vrug*	fruit (potentially < fuck)	1
70	*water*	water < *wetter* (see above)	1
71	*wharra*	Wharra (nonsensical)	1
72	*Yaris*	Yaris	1
Total			**2,031**

The results from the ComCor *verify* that Afrikaans nouns are modified phonologically and graphemically (such as *vlok* 'flake' and *fonk* 'spark' for *fok* 'fuck'), as well as being remodelled by non-morphemic word-formation strategies (such as *d@nner* for *donder* 'thunder', and *m..r* for *moer* 'beat up'). We also do not see any words on the ComCor list that originate from the Greek or Latin stratum. Furthermore, in the ModCor dataset, we noted that, unlike English, the Afrikaans construction does not have such a strong preference for monosyllabic nouns, as several of the Afrikaans nouns in the ModCor dataset were disyllabic or even polysyllabic. This observation is even more relevant to the ComCor dataset, wherein 29 of the nouns (phrases) occurring in the x part of the construction are disyllabic and 15 are polysyllabic. In other words, most of the nouns in the list are actually non-monosyllabic words, and the preference for these types of nouns is apparently not a feature of the Afrikaans construction. Lastly, based on the number of hapax legomena in the ComCor list (37 types out of the total 72 types), we can confirm that, in terms of the operationalised definition of productivity mentioned in 2m above, this construction could be considered productive in Afrikaans.

Additionally, we can *expand* our initial description of the construction. While only one case of a noun modified by an adjective was found in ModCor (namely *blou dinges* 'blue thing'), numerous examples can be found in ComCor. Compare for instance *flippen hel* 'flipping hell'; *hoppende fok* 'hopping fuck'; *actual hel* 'actual hell'; *dooie dinges* 'dead thingamajig'; *fokken hel* 'fucking hell'; *grote griet* 'Great Griet'; *heilige fok* 'holy fuck'; and *liewe goeie heilige fok* 'dear good holy fuck'. Furthermore, in addition to the same high-frequency nouns (such as *hel* 'hell'), a number of additional, on-the-spur-of-the-moment creations have been attested, such as *yaris* 'Yaris car'; *falala* 'falala' (nonsensical); *vakbond* 'trade union'; and *kriewel* 'squirm'.

In addition to the nouns belonging to the category RELIGION that we have already identified in ModCor, we can now add several proper names and/or (euphemistic) modifications of such religious proper names, e.g., *Moses*; *jissis* 'Jesus'; *grote Griet* 'great Griet' < 'good God'; *diekens* 'Dickens' < 'devil'; *Joos/Josie* < 'devil'; and *Yaris* < (probably) 'Jesus'. We also see numerous intensifiers of specifically *hel*, such as *fokken hel* 'fucking hel'; *frieken hel* 'freaking hell'; *flippen hel* 'flipping hell'. While we did not find any noun phrases belonging to the category of COSMOGONY in ModCor, we found one example in ComCor, namely *hel op aarde* 'hell on earth'. Two new nouns belonging to the category of ANIMALS could also be added, namely *os* 'ox' and *kees* 'baboon', and one new noun belonging to the category of NATURAL PHENOMENA, namely *bliksemstraal* 'thunderstrike'. Additionally, we see one new noun in the category of NEGATIVE EVENTS, namely *vrek* 'die like an animal'. On the other hand, it could also possibly simply be a modification of *fok* 'fuck', since we find numerous examples of leetspeak with valid words – compare for *fok* 'fuck' the following potential candidates: *vakbond* 'trade union'; *vet* 'fat'; *vrommel* 'crumple'; *vrug* 'fruit'; *vy* 'fig' (all starting with [f]); and *pok* 'pox' (rhyming with *fok*). For hel 'hell' we found the modifications *heck*, *heng*, and *herrie*; and for *duiwel* 'devil', among others, *deuvel.* While one might be tempted to add a new semantic category EXCREMENT for the noun *poefies* 'shit' (diminutive, plural), we consider it to rather fit under the larger umbrella of euphemistic words being used in written conversations in the public domain.

While we did not see any nouns in the ModCor list that could be regarded as adding an "archaic or humorous flavor" (Hugou 2017: 23) to the WHX construction, we do find in the ComCor dataset several noun phrases that might be construed as humorous, for example *hoppende fok* 'hopping fuck', *dooie dinges* 'dead thingamajig', *foketie fok fok* 'fuckity fuck fuck', *grote griet* 'Great Griet', and *liewe goeie heilige fok* 'dear good holy fuck'.

7 Constructional relations between the WH part and X part

We now turn to our second research question, namely, to investigate the attraction between specific WHs and specific Xs in the Afrikaans WHX construction, using covarying collexeme analysis. Whereas the results in the previous sections could be compared easily to Hugou's (2017) results for English, the results below cannot, since he did not investigate this aspect of the English construction.

Covarying collexeme analysis is a technique developed by Stefanowitsch and Gries (2003), and Gries and Stefanowitsch (2004a, 2004b) as part of the larger set of techniques called collostructional analysis. Collostructional analysis uses a statistical method to measure the strength of association within or between constructions (also see Gilquin 2010: 195), while covarying collexeme analysis is specifically used to measure the association strength between variables in a construction. Stefanowitsch and Gries (2005: 9) explain that this method is suitable for investigating constructions with "two (or more) slots which may be associated with sets of items whose semantic properties we want to investigate with respect to each other".

To investigate the attraction between the WHs and the Xs in the Afrikaans WHX construction, we made use of a combined dataset (ComCor and ModCor). We used this dataset to analyse all individual instances of the 8 WH pronouns in the combined 2,414 occurrences of the construction, assessing all possible collocations with all 74 types of X items in this combined dataset. The analyses have been performed using the statistical software R, using the Coll.analysis 4.0 script (Gries 2022).

Table 7 presents the results from the covarying analysis of the combined ComCor and ModCor data. Only significant attractions (i.e., collocations with a likeliness ratio LLR>3.84) are included in the table. We also did not include items where the noun phrase appears only once in the construction, and thus also only once in the combination, because it will obviously appear statistically significant.

Table 7: Results of covarying analysis of the combine ComCor and ModCor data.

WH_PN	Translation	RudeNP	Translation2	Freq	FREQOFSLOT1	FREQOFSLOT2	RELATION	LLR
hoe	how'	*moer*	beat up'	64	891	110	Attr.	21.411
hoe	how'	*duiwel*	devil'	94	891	176	Attr.	21.334
waar	where'	*fok*	fuck'	58	203	403	Attr.	19.623
wie	who'	*fok*	fuck'	78	313	403	Attr.	15.917
wat	what'	*hel*	hell'	437	965	983	Attr.	13.737
hoe	how'	*hoenders*	chickens'	19	891	27	Attr.	12.564
hoe	how'	*ongeluk*	accident'	41	891	72	Attr.	12.230
hoekom	why'	*fok*	fuck'	13	31	403	Attr.	11.131
waar	where'	*flippen_hel*	flipping hell'	2	203	2	Attr.	9.920
wie	who'	*blikskottel*	tin plate'	2	313	2	Attr.	8.181
hoe	how'	*hoender*	chicken'	8	891	10	Attr.	7.809
waarom	why'	*donder*	thunder'	2	4	104	Attr.	7.242
hoe	how'	*dinges*	thingamajig'	24	891	44	Attr.	5.732
waar	where'	*poes*	pussy'	8	203	40	Attr.	5.317
waar	where'	*grote_griet*	great Griet'	1	203	1	Attr.	4.955
waar	where'	*os*	ox'	1	203	1	Attr.	4.955
waar	where'	*vrommel*	crumple'	1	203	1	Attr.	4.955
wanneer	when'	*fok*	fuck'	3	5	403	Attr.	4.753
wie	who'	*dêm*	damn'	2	313	3	Attr.	4.638
wie	who'	*hemel*	heaven'	2	313	3	Attr.	4.638
wie	who'	*hel*	hell'	145	313	983	Attr.	4.605
wanneer	when'	*duiwel*	devil'	2	5	176	Attr.	4.213
wie	who'	*dooie_dinges*	dead thingamajig'	1	313	1	Attr.	4.088
wie	who'	*kont*	cunt'	1	313	1	Attr.	4.088
hoe	how'	*heck*	heck'	2	891	2	Attr.	3.988
hoe	how'	*Moses*	Moss'	2	891	2	Attr.	3.988

hoe	how'	*water*	water'	2	891	2	Attr.	3.988
hoe	how'	*zir*	zir'	2	891	2	Attr.	3.988
hoe	how'	*poes*	pussy'	9	891	40	Rep.	−3.913
wie	who'	*ongeluk*	accident'	4	313	72	Rep.	−4.445
wanneer	when'	*hel*	hell'	0	5	983	Rep.	−5.239
wat	what'	*swernoot*	scoundrel'	2	965	16	Rep.	−5.942
wie	who'	*duiwel*	devil'	13	313	176	Rep.	−6.015
wie	who'	*drommel*	idiot'	0	313	22	Rep.	−6.143
wie	who'	*moer*	beat up'	6	313	110	Rep.	−7.097
wat	what'	*hoenders*	chickens'	4	965	27	Rep.	−8.252
hoe	how'	*vlok*	flake'	0	891	10	Rep.	−9.241
wat	what'	*moer*	beat up'	29	965	110	Rep.	−9.392
wat	what'	*duiwel*	devil'	51	965	176	Rep.	−9.967
waar	where'	*joos*	devil'	1	203	106	Rep.	−12.436
hoe	how'	*hel*	hell'	321	891	983	Rep.	−13.067
hoekom	why'	*hel*	hell'	3	31	983	Rep.	−15.136
hoe	how'	*fok*	fuck'	89	891	403	Rep.	−48.764

The results clearly indicate that some WH pronouns and specific nouns attract each other within the WHX construction. For example, there is a very strong attraction between *hoe* 'how' and *moer* 'beat up' (LLR=21.411), stronger than any other WH pronoun with any other noun in this construction. It is also clear that we cannot assume that all highly frequent noun phrases (such as *hel* 'hell', *fok 'fuck'*, and *duiwel* 'devil') show a strong collocational preference for all pronouns that occur in the construction. For example, one might assume that *hel* 'hell' (N=983) would show a strong attraction to all or most WH pronouns in this construction, but there is in fact only an attraction relationship between *wat* 'what' (LLR=13.737) and *hel* 'hel', and *wie* 'who' and *hel* 'hel' (LLR=4.605). While combinations with other WH pronouns do occur and are used frequently, the attraction is never statistically significant. In fact, there is even a strong collocational repulsion between *hel* 'hell' and *hoe* 'how' (LLR=-13.067) and *wanneer* 'when' (LLR=-5.239). The strong repulsion between *hel* and *hoe* is despite the high frequency of this collocation (N=321). From this result, we see again that the covarying collexeme analysis is not suitable to identify prototypical and entrenched constructions.

Similarly, we observe that there is a strong attraction between *fok* 'fuck' and *waar* 'where' (LLR=19.623) and *wie* 'who' (LLR=15.917), and a weak (but still significant) attraction with *wanneer* 'when' (LLR=4.753). However, *fok* 'fuck' shows an extremely strong repulsion with *hoe* 'how' (LLR=-48.764, the strongest collocational value in our results). *Duiwel* 'devil' and *hoe* 'how' (LLR=21.334), on the other hand, are strongly attracted, while *duiwel* 'devil' and *wie* 'who' (LLR=-6.015) and *wat* 'what' (LLR=-9.967) are in a repulsion relationship. *Joos* 'devil', the fourth most frequent noun in our ComCor dataset, does not show any significant attractions, except for a repulsion with *waar* 'where' (LLR=-12.436).

Another interesting observation is that there are strong attractions between *hoe* 'how' and *moer* 'beat up' (LLR=21.411), *hoenders* 'chickens' (LLR=12.564), and *hoender* 'chicken' (LLR=7.809). This could be due the assonance effect created by combining these words in the construction, contributing to a humorous effect. The fact that the *wat* 'what' and the nouns *moer* 'beat up' and *hoender*s 'chicken' show a repulsive relation (LLR=-9.392 and LLR=-8.252 respectively) strengthens our assertion.

As noted, the aim of the covarying collexeme analysis is primarily to enable the identification of semantic patterns in the way items in the two slots collocate with each other. Although we were able to identify strong attractions and repulsions for various collocations, we do not see any clear semantic patterns emerging from our results. However, this is not entirely surprising, given the nature of the two items that can fill these slots. The WH pronouns are function words that carry little semantic weight (serving mainly as an interrogative to request specific information), while the X items function to convey taboo value or humour. The con-

struction's function is ultimately to attract attention rather than to convey semantic content, and it may thus be better considered a pragmatic construction.

8 Conclusion

Given the fact that the Afrikaans WHX construction had not been previously described in the literature, we set out to provide such a description to get supporting evidence for our assertion that the Afrikaans construction is a constructional calque, based on the English WHX construction (Hugou 2017), but it is starting to take on a life of its own. By conducting a corpus linguistic investigation of the construction in two Afrikaans corpora, we focused specifically on the characteristics of the various open slots in the construction, and the constructional relations (attractions or repulsions) between specific WH pronouns and noun phrases. Throughout, we contrasted our findings with those of Hugou (2017) for English. Our investigation finds the following:

a. Only eight Afrikaans WH pronouns occur in the construction, viz. *wat* 'what', *hoe* 'how', *wie* 'who', *waar* 'where', *hoekom* 'why', *wanneer* 'when', *waarom* 'why' and *waarvoor* 'what for'.
b. In both corpora the four monosyllabic WH pronouns occur with a much higher frequency than the disyllabic ones, with *wat* 'what', followed by *hoe* 'how' consistently occurring with the highest frequency. The fact that the Afrikaans construction clearly favours monosyllabic WH pronouns and the fact that *wat* 'what' occurs in this part most frequently, correlates well with what Hugou (2017) reported for the English construction. However, unlike in English, disyllabic and even polysyllabic items are often found as noun phrases. This indicates that Afrikaans speakers are elaborating on the construction, proving that the construction is taking on a life of its own.
c. One major difference between the Afrikaans and English construction is the frequency difference between *wat* 'what' and *hoe* 'how': while Hugou (2017) reports a very large difference between the frequency of *what* and *how*, in Afrikaans there is no mentionable difference between the frequency of the two WH pronouns. Nonetheless, distributional matters aside, this illustrates the existence of a diacxn schema clearly: the English idiocxn *what* [[wɒt] ⇔ [WHAT]]$_{ENG}$ and the Afrikaans idiocxn *wat* [[vɑt] ⇔ [THE]]$_{AFR}$ results in the diaxcn [[w/v . . . t] ⇔ [WHAT]] (and so forth for the other pronouns).
d. We identified 76 unique noun phrases that can occur in the X part of the Afrikaans construction. In addition, grawlix was also used in this part often, and in

some instances the x part was left empty. Identical to English, *hel* 'hell' consistently occurred in this part of the construction with the highest frequency. Five other lexical items make up the 6 most frequently occurring nouns in both our corpora, viz. *donder* 'thunder'; *duiwel* 'devil'; *fok* 'fuck'; *joos* 'devil'; and *moer* 'beat up'. Apart from high-frequency items, in-the-spur-of-the-moment nouns (e.g. *vakbond* 'trade union' and *Yaris*) also occur, as is the case in the English construction. The high number of hapax legomena that occur in the x part of the Afrikaans construction indicates that, just like the English construction, this construction could be considered productive in Afrikaans (specifically in unedited, informal, written contexts).

e. While bare noun phrases occur with the highest frequency in the Afrikaans construction, nouns can also be modified, e.g. *hoppende fok* 'hopping fuck'.

f. Noun phrases in the Afrikaans construction can also contain lexical items that are remodelled by non-morphemic word-formation strategies, e.g. *.f.* < *fok* 'fuck', and *MMMMMMMMMMMM* < *moer* 'beat up'. Instances of phonological and graphemic modification were also attested quite frequently, e.g., *vlok* 'flake' < *fok* 'fuck', and *pok* 'pox' < *fok* 'fuck'. This is most probably the result of the context of our data, where users might resort to leetspeak to by-pass moderators removing their comments.

g. Hugou (2017) identified only two semantic categories, viz. RELIGION and COSMOGONY, for the x part of the English construction. Our analysis of the Afrikaans counterpart reveals a broader range of categories, including RELIGION, ANIMALS, NATURAL PHENOMENA, and NEGATIVE EVENTS, but not COSMOGONY (bar one disputable example). This testifies to the elaboration of the construction schema in Afrikaans.

h. Lastly, while Hugou (2017) posits that the English construction primarily functions as an intensifier, our findings suggest an additional function for the Afrikaans construction, namely as a vehicle for humour. Examples such as *hoppende fok* 'hopping fuck', *fokietie fok fok* 'fuckity fuck fuck', *zir*, *falala*, *Yaris*, *grote griet* 'great Griet', *hoenders* 'chickens', *otter* 'otter', *oertel* 'asshole', and *vakbond* 'trade union' indicate a strong tendency towards comedic expression. This highlights an important function of the Afrikaans construction – a feature that was not explicitly identified by Hugou (2017), given his specific focus on intensification. This therefore does not mean that this is a unique feature of the Afrikaans construction, because humorous examples of the English construction abound, e.g. *what the flying blue fuck* (an example taken from Hugou 2017: 16).

The large number of telling similarities between the Afrikaans and English construction confirm our thesis that the Afrikaans construction is a constructional calque, based on the English WHX construction. However, while the construction might be borrowed, it is already developing into new directions, taking on a life of its own. This is specifically noticeable in the occurrence of disyllabic and even polysyllabic noun phrases in the Afrikaans construction, a broader range of semantic categories of the noun phrases, and the emergence of humour as an important function of the construction. However, we should note that we suspect that one might also see similar effects in comparable English genres, and in the WHX construction in other languages – something that should be examined in future research.

Lastly, as native speakers of Afrikaans, we know that the WHX construction is sometimes univerbated, e.g., *watdefok* < *wat de fok*. In addition, such univerbations could also be phonologically and/or orthographically deformed, e.g., *warrefok* < *waddefok* < *watdefok* < *wat de fok*, where *waddefok* is a progressively assimilated form of *watdefok*, and *warrefok* is a rhotacized form of *waddefok* (see Wissing 2020). Although such changes (univerbations and deformations) fall outside the scope of our current research (and therefore our original searches in the corpora), we did an exploratory investigation of changes to *wat de hel* and *wat de fok* only (i.e., the most frequent WH pronoun, with the two most frequent nouns). The results in Table 8 clearly show a strong tendency to change those constructions with *fok* in the X part of the construction – a somewhat surprising result, given that *hel* occurs much more frequently in the WHX construction (see Table 6). We surmise, on the one hand, that since *fok* is generally considered to be much more taboo than *hell*, users might be trying to euphemise the construction with *fok*, or otherwise simply to trick profanity checkers (i.e., univerbation as a kind of leetspeak). It could, on the other hand, also be an indication of lexicalisation in the making: the phrasal *wat de fok* is slowly becoming a lexical item *waddefok*, which might in the future develop a life of its own. Further research into other such forms, e.g., *hoe-de-fok* 'how-the-fuck', could potentially shed light on the development of new taboo words.

Table 8: Frequency of contracted forms in the WHX construction in ComCor.

Rank	Contracted form with *wat* and *fok*	N	Contracted form with *wat* and *hel*	N
1	*waddefok*	82	*warrehel*	6
2	*watte fok*	8	*waddehel*	5
3	*watdefok*	7	*wate hel*	1
4	*warrefok*	5	*wattehel*	1
5	*warre fok*	4		
6	*wadde fok*	3		

Table 8 (continued)

Rank	Contracted form with *wat* and *fok*	N	Contracted form with *wat* and *hel*	N
7	*watefok*	2		
8	*wattefok*	2		
9	*waddiefok*	1		
10	*watde fok*	1		
11	*wate fok*	1		
	Total	116	**Total**	13

Ethical matters

Ethical clearance for the research project was obtained through the Language Matters Ethics Committee of the North-West University (ethics number: NWU-00632-19-A7).

Declaration of interests

In the table below, the roles of the authors are defined in terms of the Contributor Roles Taxonomy (CRediT; https://credit.niso.org/).

Role	Van Huyssteen	Breed	Pilon
Conceptualisation	X	X	X
Data curation	X	X	X
Formal analysis	X	X	X
Funding acquisition	–	X	–
Investigation	X	X	X
Methodology	X	X	X
Project administration	X	–	–
Resources	X	X	X
Software	–	–	–
Supervision	–	–	–
Validation	X	X	X
Visualization	–	–	–
Writing (original draft)	X (§ 1–3, 8)	X (§ 4–7)	X (§ 4–7)
Writing (review and editing)	X	X	X

This research is partially funded by the Suid-Afrikaans Akademie vir Wetenskap en Kuns, and partially made possible through barter agreements with BlueTek Computers, Afrikaans.com, and WatKykJy.co.za. The *Woordeboek van die Afrikaanse Taal* (WAT), *Handwoordeboek van die Afrikaanse Taal* (HAT) and Centre for Text Technology (CTexT) of the North-West University are hereby also acknowledged for generously supplying the project with material from their respective databases.

Van Huyssteen is a director of the not-for-profit company Viridevert NPC (CIPC registration number: 2016/411799/08), who owns and manages the website vloek.co.za. This website was developed specifically for this project, and this conflict of interest has been approved by NWU.

References

Allan, Keith & Kate Burridge. 1991. *Euphemism and dysphemism: Language used as shield and weapon.* New York & Oxford: Oxford University Press.

Baayen, Harald. 1993. On frequency, transparency and productivity. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology 1992*, 181–208. Dordrecht: Springer Netherlands.

Burridge, Kathryn & Reka Agnes Benczes. 2019. Taboo as a driver of language change. In Keith Allan (ed.), *The Oxford handbook of taboo words and language*, 180–199. Oxford: Oxford University Press.

Bybee, Joan. 2010. *Language, usage and cognition.* Cambridge: Cambridge University Press.

Centre for Text Technology (CTexT). 2020. NWU Korpus Gemodereerde Kommentaar 1.0 [NWU Corpus of Moderated Comments 1.0]. In North-West University (ed.) 1.0 edn. Potchefstroom: Queried on Sketch Engine: app.sketchengine.eu.

Centre for Text Technology (CTexT). 2023. NWU Kommentaarkorpus 2.2 [NWU Commentary Corpus 2.2]. In North-West University (ed.) 2.2 edn. Potchefstroom: Queried on Sketch Engine: app.sketchengine.eu.

Culpeper, Jonathan. 2011. *Impoliteness: Using language to cause offence*. Cambridge: Cambridge University Press.

Davies, Mark. 2019. The Movie Corpus. https://www.english-corpora.org/movies/.

Davies, Mark. n.d. Corpus of Contemporary American English (COCA). https://www.english-corpora.org/coca/.

Gilquin, Gaëtanelle. 2010. *Corpus, cognition and causative constructions.* Amsterdam: Benjamins.

Gries, Stefan Th. 2022. Coll.analysis 4.0. A script for R to compute perform collostructional analyses. https://www.stgries.info/teaching/groningen/index.html.

Gries, Stefan Th. & Nick C. Ellis. 2015. Statistical measures for usage-nased linguistics. *Language Learning* 65 (S1). 228–255.

Gries, Stefan Th. & Anatol Stefanowitsch. 2004a. Co-varying collexemes in the into-causative. In Michel Achard & Suzanne Kemmer (eds.), *Language, culture, and mind*, 225–236. Stanford: CSLI.

Gries, Stefan Th. & Anatol Stefanowitsch. 2004b. Extending collostructional analysis: A corpus-based perspective on 'alternations'. *International Journal of Corpus Linguistics* 9(1). 97–129.

Hickey, Leo & Miranda Stewart. 2005. Introduction. In Leo Hickey & Miranda Stewart (eds.), *Politeness in Europe*, 1–12. Clevedon: Multilingual Matters.

Höder, Steffen. 2018. Grammar is community-specific: Background and basic concepts of Diasystematic Construction Grammar. In Hans C. Boas & Steffen Höder (eds.), *Constructions in contact: Constructional perspectives on contact phenomena in Germanic languages*, 37–70. Amsterdam: Benjamins.

Höder, Steffen. 2019. Phonological schematicity in multilingual constructions: A diasystematic perspective on lexical form. *Word Structure* 12(3). 334–352.

Höder, Steffen, Julia Prentice & Sofia Tingsell. 2021. Additional language acquisition as emerging multilingualism: A Construction Grammar approach. In Hans C. Boas & Steffen Höder (eds.), *Constructions in contact: Language change, multilingual practices, and additional language acquisition*, 309–338. Amsterdam: Benjamins.

Hoeksema, Jack. 2019. Taboo terms and their grammar. In Keith Allan (ed.), *The Oxford handbook of taboo words and language*, 160–179. Oxford: Oxford University Press.

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

Hugou, Vincent. 2017. The WHX construction (what the hell. . .?) and intensity. A corpus-based study. *Lexis* (10).

Jay, Timothy B. 2020. Ten issues facing taboo word scholars. In Nico Nassenstein & Anne Storch (eds.), *Swearing and cursing*, 37–52. Berlin: De Gruyter Mouton.

Jay, Timothy B. & Kristin Janschewitz. 2008. The pragmatics of swearing. *Journal of Politeness Research* 4 (2). 267–288.

Kádár, Dániel Z. 2017. *Politeness, impoliteness and ritual: Maintaining the moral order in interpersonal interaction*. Cambridge: Cambridge University Press.

Kirsten, Johanita & Adri Breed. 2020. Interrogative pronouns. *Taalportaal*. Taalportaal Consortium: https://taalportaal.org/taalportaal/topic/pid/topic-20190516105546177. Last Modified May 2020. Accessed 20 December 2024.

Langacker, Ronald W. 1991. *Concept, image, and symbol: The cognitive basis of grammar.* Berlin: Mouton de Gruyter.

Michaelis, Susanne Maria & Martin Haspelmath. 2020. Grammaticalization in creole languages: Accelerated functionalization and semantic imitation. In Walter Bisang & Andrej Malchukov (eds.), *Volume 2 Grammaticalization scenarios from Africa, the Americas, and the Pacific*, 1109–1128. Berlin: De Gruyter Mouton.

Sketch Engine. 2020. Dutch Web corpus 2020 (nlTenTen20). https://www.sketchengine.eu/nltenten-dutch-corpus/.

Spencer-Oatey, Helen. 2005. (Im)politeness, face and perceptions of rapport: Unpackaging their bases and interrelationships. *Journal of Politeness Research* 1(1). 95–119.

Stefanowitsch, Anatol & Stefan Th. Gries. 2003. Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8(2). 209–243.

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

Taalkommissie van die Suid-Afrikaanse Akademie vir Wetenskap en Kuns. 2011. Taalkommissiekorpus 1.1. Potchefstroom: North-West University.

Taalprof. 2009. Wat voor den duivel is dat nou weer? [What the devil is that again?]. *Taalprof Plein.* https://taalprof.blogspot.com/2009/07/wat-voor-den-duivel-is-dat-nou-weer.html.

Taylor, John R. 2002. *Cognitive grammar*. Oxford: Oxford University Press.

Thompson, Sandra A., Barbara A. Fox & Elizabeth Couper-Kuhlen. 2015. *Grammar in everyday talk: Building responsive actions.* Cambridge: Cambridge University Press.

Van der Merwe, Mart-Mari. 2022. *Prototipiese Afrikaanse taboewoorde [Prototypical Afrikaans taboo words]*. Pretoria: University of Pretoria.

Van der Wouden, Ton. 2019. Hoe de fuk. *Neerlandistiek: Online Tijdschrift voor Taal- en Letterkunde.* https://neerlandistiek.nl/2019/11/hoe-de-fuk/.

Van Olmen, Daniel, Marta Andersson & Jonathan Culpeper. 2023. Inherent linguistic impoliteness: The case of insultive YOU+NP in Dutch, English and Polish. *Journal of Pragmatics* 215. 22–40.

Van Rooy, Bertus. 2021. A Diasystematic Construction Grammar analysis of language change in the Afrikaans and English finite verb complement clause construction. In Hans Boas & Steffen Höder (eds.), *Constructions in contact 2: Language change, multilingual practices, and additional language acquisition*, 109–137. Amsterdam: Benjamins.

Van Rooy, Bertus & Ronel Wasserman. to appear. English of Afrikaans speakers. In Raymond Hickey (ed.), *New Cambridge history of the English language*. Cambridge: Cambridge University Press.

Wissing, Daan. 2020. Consonant related processes. *Taalportaal*. https://taalportaal.org/taalportaal/topic/pid/topic-20170830162004808. Last Modified May 2020. Accessed 20 December 2024.