

On contrastive reduplication: Adding Dutch to the Germanic typology

ABSTRACT

Reduplication as a productive morphosyntactic process in West Germanic is usually only associated with Afrikaans. The morphosyntax, phonology and semantics of this phenomenon in Afrikaans has been extensively investigated. However, since the start of the new millennium, some papers point out that reduplication is also a productive phenomenon in English, and there are also some hints about its existence in German. While just like in Afrikaans, reduplication can be used for intensification, a new usage seems to have evolved, i.e. the contrastive reduplication to refer to a prototypical item of a set. In this article, we research the occurrence of this new phenomenon in Dutch and compare it with the other West Germanic languages.

1. Introduction

This paper concerns contrastive reduplication in Dutch. Contrastive reduplication (henceforth CR), also called *the double* (Dray 1987) *contrastive focus reduplication* (Ghomeshi et al. 2004; Bross & Fraser 2020), *Identical Constituent Compound* (Hohenhaus 2004; Frankowsky 2022) and *the lexical clone construction* (Horn 1993, 2018)) is a form of reduplication of which the main semantic function is to indicate that the particular lexical item used in the reduplication constitutes a prototype of the set of items to which it can refer, henceforth the ‘prototypical X’ function. One of the classical examples from the seminal paper on CR by Ghomeshi et al. (2004) is that of the ‘salad-salad’, as illustrated below (1).

(1) I’ll make a tune salad, and you make the SALAD-salad. (Ghomeshi et al. 2004: 308)

By using CR, the speaker indicates that the salad should be a prototypical salad, e.g. with lots of greens/lettuce and tomatoes, and not something less prototypical like a tuna salad. The construction is by no means limited to the nominal domain: it can also be found with verbs (2), adjectives (3) and prepositions (4).^{1,2}

(2) Are you LEAVING-leaving? [i.e. are you “really” leaving (for good), or are you just stepping out for a minute]

(3) A: Are you nervous?

B: Yeah, but, you know, not NERVOUS-nervous. [i.e., not “really” nervous]

(4) Lily: You have to get up.

Rick: I am up.

Lily: I mean UP-up.

(Ghomeshi et al. 2004: 312)

Besides this ‘prototypical X’ function, CR has other semantic functions as well, or at least in American English (Ghomeshi et al. 2004; Horn 2018). A second function is that of

¹ However, the construction is limited to lexical items, i.e. a functional preposition like ‘to’ or an auxiliary verb cannot be reduplicated in this way (see Ghomeshi et al. 2004: 313 for examples).

² As can be seen in (2), CR can copy more than just the bare stem (i.e. the *-ing* inflection). This also holds for inflection on nouns (see again Ghomeshi et al. 2004 for more examples and discussion).

intensification. This function is usually seen in CR with relative adjectives. For example, someone who is DRUNK-drunk is very drunk. With absolute adjectives, CR results in an intensification in the meaning range of ‘absolutely’ or ‘completely’. For instance, someone who is NAKED-naked is completely naked (Horn 2018: 246). Thus, as Horn observes, CR with absolute relatives result in the literal meaning of the adjective, which can be seen as a third type of function. This function is not limited to absolute adjectives. An example of CR with a universal quantifier being used with this type of semantic function is given in (5).

- (5) A: So what have you got for me?
 B: Professionally? Well, honestly, nothing.
 A: Nothing?
 B: Well, not NOTHING-nothing. Just nothing much.
 (Horn 2018: 246 – exchange between a singer and her manager on *Nashville*, Feb. 11, 2015)

Finally, there is also the function which Horn (1993, 2018) labels ‘added-value’. This function adds extra meaning to the reduplicated form compared to the non-reduplicated one, i.e. as in (6). Here, the value that is added to the reduplicated form is that of the meaning of living together as lovers.

- (6) A: I hear you guys are, um living together now.
 B. Well, we’re not LIVING TOGETHER- living together. (Horn 1993: 50)

Frankowsky (2022:176) furthermore mentions that CR has an expressive meaning component as well (in the sense of Potts 2007). That is, because reduplication is not a part of the core grammar of German, using it is a more expressive way (i.e. indicating emotion of the speaker or a particular perspective of the speaker towards the content of the utterance) compared to a neutral construction with adjectival or adverbial modification using ‘real/really’. This expressive property has been discussed for CR in Italian by Wierzbicka (1991), denied for German by Mau (2002, i.e. Frankowsky goes against Mau), and considered at least an optional feature of CR in British English by Hohenhaus (2004).

As for the Germanic typology of the phenomenon, most studies concern some variety of English, i.e. Dray (1987), Ghomeshi et al. (2004), Hohenhaus (2004), Horn (1993, 2018), Huang (2009, 2015) among others. To a lesser extent, the phenomenon has been discussed in German (Mau 2002, Finkebeiner 2014, Freywald 2015, Kentner 2017, Bross and Fraser 2020), with the first in-dept corpus study being the one by Frankowsky (2022). As for Dutch, the fact that the phenomenon is emerging in informal Dutch is mentioned in passing by Booij (2019), but as far as we are aware, not concrete study on CR in Dutch exists till date.^{3,4} The

³ We have written a short comparative paper on CR in Dutch and Afrikaans for a Festschrift which will appear in the course of 2022 (Cavirani-Pots & Dirix to appear).

⁴ Though CR in Dutch has not yet been systematically investigated, we think the construction is quite known among speakers, possible under influence of speakers’ awareness of the phenomenon in English. A famous case of CR is the quote in (i) from the TV show *Temptation Island*. This quote was discussed so much that it ultimately got cited in 2019 in an advertising campaign for road safety (<https://www.hln.be/tv/verkeerscampagne-gebruikt-kijken-kijken-uitspraak-van-temptation-haroon-ze-hadden-me-wel-eens-kunnen-bellen~a764c415/>).

- (i) Je hebt kijken en je hebt **kijken- kijken**.
 you have look.INF and you have look.INF look.INF
 ‘There is looking and there is really looking at something.’

main aim of this paper is to add Dutch to the Germanic typology of CR, by presenting and discussing findings from a recent corpus study.

The outline of this paper is as follows. In section 2, we present our methodology and results of the corpus study on Dutch. In section 3, we compare the results of our corpus study with what has been found for English and German. We furthermore discuss CR in Afrikaans, which has not been discussed in any literature on CR in Germanic so far. In section 4, we conclude and present an outlook on future research on CR in Germanic.

2. The Dutch data⁵

In this section we present the Dutch data. In subsection 2.1, we discuss the methodology of the corpus study. In subsection 2.2, we present the findings of the study.

2.1 Methodology

Before presenting the methodology of data collection, it needs to be mentioned that CR is a very recent phenomenon in Dutch, and furthermore a colloquial and spoken phenomenon. This makes it very hard to collect spontaneous data by means of a corpus study for two reasons. First, the most extensive tagged corpus of Dutch, namely the SoNaR corpus (Oostdijk et al. 2013) contains very little spoken data.⁶ Second, it is impossible in this corpus to automatically extract all instances of CR (see Hohenhaus (2004) for discussion of this methodological issue with CR in British English). However, anticipating the creation of a more recent, and larger spoken data corpus, we extracted data from the SoNaR corpus based on a manually made list of reduplications that, according to us, might be able to occur in CR.⁷ This list included a set of frequent nouns, adjectives, adverbs, verbs and prepositions (see Appendix). We looked up all the reduplicated forms of the words on this list with the advanced search option of the corpus, using the ‘lemma’ option for each noun, adjective or verb in order to include inflected forms in the search (e.g. the search ‘jongen jongen’ in which both nouns were indicated as ‘lemma’ can return hits with or without plural inflection on either or both of the nouns). We looked up both the version with a hyphen between the two segments and the version without a hyphen (e.g. both *meisje-meisje* and *meisje meisje*). For both versions, we manually annotated all hits found per reduplicated word for semantic function, unless there were more than 100 hits for that word. In that case, we took a random sample of 100, and annotate only those (i.e. for one lemma maximally 200 hits were annotated: max. 100 with hyphen and max. 100 without hyphen). The semantic functions we annotated for were based on the four functions discussed in Horn (1993), namely: i) intensification, ii) literal meaning, iii) prototypical X, iv) added-value meaning.⁸ In addition

⁵ We have already presented some highlights of this data set in Cavirani-Pots & Dirix (to appear, 2023).

⁶ The biggest spoken data corpus on Dutch, *Corpus Gesproken Nederlands*, is included in the SoNaR corpus.

⁷ The data were extracted between August-October 2021.

⁸ A specific set of hits for the nouns *jongen* ‘boy’, *meisje* ‘girl’, *man* ‘man’ and *vrouw* ‘woman’ that were returned by the search, could not be annotated for any of the annotation categories, but rather instituted cases which can be called ‘vocative’ (often with a pejorative or affective taste to them). These hits are reduplications which are not integrated into the sentence, and function as a sort of discourse marker. An example is given in (ii).

- (ii) Jongen jongen, wat onnozel weer.
Boy boy what silly again
‘Oh boy, how silly.’ (SoNaR, CGN document fv901144)

The same holds for a set the hits of some adverbs (e.g. *snel/gauw* ‘quickly’, *goed* ‘well’). Since these hits do not seem to be true cases of CR, we will not discuss them further in the results section.

to the data from the SoNaR corpus, we added hits found in a recent corpus of Dutch Twitter data.^{9,10}

2.2 Results

The quantitative results of the corpus study are presented per lexical category.¹¹ The semantic function category of ‘added-value meaning’ was not found for any lexical category, and is therefore not given in any of the tables. For each lexical category in turn, we present some examples and discuss interesting findings.

2.2.1 CR with nouns

In Table 1, the quantitative results of CR with nouns are given. As can be seen in the table, the number of hits is low for all cases, with many hapaxes legomena (see Hohenhaus 2004 for similar results for British English). In almost all of the cases, the semantic function is the ‘prototypical X’ function. A few examples are given in (7)-(9).

(7) ...dat ik niet zo'n **meisje-meisje** ben.
 ...That I not such.a girl girl are
 ‘... that I’m not a prototypical/real girl.’ (SoNaR, document ‘Army of darkness’)

(8) [Het] probleem is dat je een **man-man** wilt.
 The problem is that you a man man want
 ‘The problem is that you want a prototypical/real man.’ (SoNaR, document ‘Enough’)

(9) [Ze hebben] een DDR-achtige fixatie op **werk-werk**.
 They have a GDR-like fixation on job job
 ‘They have a GDR-like fixation on a prototypical/real job (i.e. full time, with benefits).’

Lemma	Total # hits	# intensification	# literal meaning	# prototypical X
<i>Meisje</i> ‘girl’	19	0	0	19
<i>Vrouw</i> ‘woman’	1	0	0	1
<i>Man</i> ‘man’	1	0	0	1
<i>Werk</i> ‘work’	3	1	0	3
<i>Drank</i> ‘drink’	1	1	0	0
<i>Rap</i> ‘rap’	1	0	0	1

Table 1: quantitative results for CR with nouns

CR with the lexical item *meisje* ‘girl’ is the only CR which seems to be lexicalized¹², or at least more commonly used. Interestingly, we also find hits in the plural, and among these, both possible inflectional patterns are found (see section 3 for comparison with the inflection possibilities of CR in English and German). In the first pattern, the inflection only occurs on

⁹ This corpus was created by our colleague Tim Van de Cruys (KU Leuven), which we kindly thank for giving us the opportunity to use it.

¹⁰ Due to time constraints, we only searched for the list of nouns and adjectives. We hope to be able to do the same for the other lexical categories as well in the near future.

¹¹ If no hits were found for a given lexical item that was searched for (see the list in the Appendix), this item is not listed in the given Table.

¹² It is the only reduplication found in the online Algemeen Nederlands Woordenboek of the Institute of the Dutch Language (<https://anw.ivdnt.org/article/meisjesmeisje>), which lexicalizes it as ‘meisje(s)meisje’ with ‘jongensmeisje’ (boy girl) as antonym.

the second segment of the reduplication, whereas in the second pattern the inflection occurs on both segments. Examples are given in respectively (10) and (11).

(10) **Meisje-meisjes** en bimbo's.

Girl girl.PL and bimbo's

'Prototypical/real girls and bimbo's (slightly trashy but beautiful women who are not super intelligent).'

(SoNaR untitled document)

(11) Ik hou van **meisjes meisjes**.

I love of girls.PL girls.PL

'I love prototypical/real girls.'

(Twitter corpus)

Another interesting hit with the lemma *meisje* 'girl' is the one given in (12). In this sentence, it appears as if the CR is used rather as an attributive adjective than a noun. Of course, more similar data is needed to see if this can be done more commonly, before making claims about this being a grammatical option in Dutch. However, we do think it is worth investigating this possibility in the future (see section 4).

(12) [Een] prachtig **meisje meisje** jurkje.

A beautiful girl girl dress

'A beautiful dress for prototypical/real girls.'

(Twitter corpus)

2.2.2 CR with adjectives

The quantitative results of CR with adjectives are presented in Table 2. Here we can see that with the exception of *geel* 'yellow', all other lemmas have more than one hit. Among the semantic functions, intensification is the most frequent, after that the 'prototypical X' function, and finally also some cases in which we see the 'literal meaning' function. In line with what Horn (2018) observed for English, also in Dutch, this meaning arises with absolute adjectives (*dood* 'dead', *wakker* 'awake' and *zot* 'crazy'). For each semantic function, an example is given in respectively (13)-(15).

(13) **Gek gek** word ik ervan.

Crazy crazy become I it.of

'It makes me go really crazy.'

(SoNaR, untitled document)

(14) Hersendood mss, maar nie **dood-dood**.

Brain.dead maybe (abbr.) but not dead dead

'Maybe brain dead, but not literally dead.'

(SoNaR, untitled document)

(15) Is ie nu **wit- wit** of toch een beetje grijs?

Is it now white white or rather a bit gray

'Is it really white or rather a bit gray?'

(Twitter corpus)

Lemma	Total hits	#	% intensification	% literal meaning	% prototypical X
<i>Geel</i> 'yellow'	1		0	0	1
<i>Blauw</i> 'blue' ¹³	6		5	0	1

¹³ Reduplication of *blauw* 'blue' is part of a Dutch idiom *iets blauw-blauw laten* lit. 'let something blue blue', which means to let something go and not talk about it anymore. Many such hits were returned, but as these are

<i>Rood</i> ‘rood’	3	0	0	3
<i>Wit</i> ‘white’	3	0	0	3
<i>Zwart</i> ‘black’	2	0	0	2
<i>Blij</i> ‘happy’	4	4	0	0
<i>Boos</i> ‘angry’	4	3	0	1
<i>Kwaad</i> ‘angry’	2	2	0	0
<i>Mooi</i> ‘beautiful’	13	11	0	2
<i>Lelijk</i> ‘ugly’	2	2	0	0
<i>Ziek</i> ‘ill’	6	3	0	3
<i>Dood</i> ‘dead’	4	0	4	0
<i>Wakker</i> ‘awake’	4	0	4	0
<i>Vies</i> ‘disgusting’	3	2	0	1
<i>Lief</i> ‘sweet’	5	3	0	2
<i>Gek</i> ‘crazy’	3	3	0	0
<i>Zot</i> ‘crazy’	4	2	2	0

Table 2: quantitative results for CR with adjectives

Among the hits, we only found hits without inflection markers, which means that at this point we cannot say whether different inflectional patterns can or cannot occur. However, it is definitely something that would be worth investigating in the future (see section 4). One possibility is that they can only be used in a predicative, which is uninflected.

2.2.3 CR with adverbs

In Table 3, the results of CR with adverbs are displayed. As can be seen, with three different lexical items referring to ‘quickly’, reduplication is common, and exclusively used as intensification. An observation that can be made from the data is that this adverb reduplication is often used to signal irritation, disapproval or secrecy by the speaker/writer. A few examples are given in (16)-(18).

(16) Moest dat nu zo **snel- snel** gebeuren?
 Must.PAST that now so quickly quickly happen
 ‘Was it really necessary to do that so quickly?’ (SoNaR, document ‘Bravo Bruno’)

(17) Dali’s; die ze **gauw- gauw** wegmoffelen achter de klerkast.
 Dali’s, which they quickly quickly put.away behind the wardrobe.
 ‘Dali’s, which they very quickly put away behind the wardrobe.’
 (SoNaR, document ‘kunsthandel’)

(18) En dees doenek **rap rap** int geniep...
 And this.one do.I quickly quickly in.the secret
 ‘And I do this one quickly and in secret...’ (SoNaR, untitled document)

Lemma	Total # hits	# intensification	# literal meaning	# prototypical X
<i>Snel</i> ‘quickly’	120	120	0	0
<i>Gauw</i> ‘quickly’	40	40	0	0
<i>Rap</i> ‘quickly’	59	59	0	0

not cases of CR, they were ignored. Actually, it’s not really a reduplication, but short for “iets wat blauw is, blauw laten”.

<i>Traag</i> ‘slowly’	1	1	0	0
-----------------------	---	---	---	---

Table 3: quantitative results for CR with adverbs

Even though as native speakers we do not find it hard to imagine to use adverbial CR with the ‘prototypical X’ or ‘literal meaning’ function, e.g. respectively (19) and (20), no such hits were found in the corpus. The fact that we did not find such cases might be due to the methodological issues discussed in section 2; we will come back to this in section 4.

(19) Ze praat niet **raar-raar** van ‘huh heeft ze een spraakgebrek’, maar met een raar accent.

She talks not strange strange of huh has she a speech disorder but with a strange accent
‘She doesn’t talk in a strange-strange way, like as if she has a speech disorder, but with a strange accent.’

(20) Hij rent best snel. Niet **snel- snel**, maar voor zijn leeftijd toch niet verkeerd.

He runs quite quickly. Not quickly quickly but for his age yet not bad
‘He runs quite fast. Not literally fast, but for his age it’s not bad.’

2.2.4 CR with verbs

In Table 4, the quantitative results of CR with verbs are given. As can be seen, there are a couple of hapaxes legomena, but also a number of lemmas who occur in a CR more than once. The most frequent semantic function is intensification, but there are a few cases of the ‘literal meaning’ function (with the lemma *praten*) and of the ‘prototypical X’ function as well. Of each semantic function, an example is given in (21)-(23) respectively. Note that the ‘intensification’ function with verbs results in a durative aspectual flavor.

(21) ‘n boek beginnen en blijven **lezen lezen** tot het uit is ja.

a book start and continue read read until it out is yes

‘To start a book and keep on reading until it’s finished, yes.’

(SoNaR, CGN document
fv400171)

(22) **praat-praat** met haar, waarmee ik bedoel dat je een tweegesprek voert

Talk talk with her with.which I mean that you a conversation.in.two executes

‘Literally talk with her, with which I mean a (proper) conversation in two.’

(SoNaR, untitled document)

(23) ja maar meer **kijken kijken**. Kijkend lezen zal ik zeggen.

Yes but more watch watch watching read shall I say

‘Yes but more real/prototypical watching. Reading by watching I would say.’

(SoNaR, document fv400108)

Lemma	Total # hits	# intensification	# literal meaning	# prototypical X
<i>Werken</i> ‘work’	6	5	0	1
<i>Slapen</i> ‘sleep’	1	0	0	1
<i>Kijken</i> ‘watch’	5	3	0	2
<i>Lezen</i> ‘read’	1	1	0	0
<i>Lopen</i> ‘lopen’	2	1	0	1
<i>Rennen</i> ‘run’	2	2	0	0
<i>Praten</i> ‘talk’	3	1	2	0

<i>Huilen</i> ‘cry’	2	2	0	0
---------------------	---	---	---	---

Table 4: quantitative results for CR with verbs

Interestingly, the example given in (22) is an inflected rather than an infinitival form. This is again something that should be investigated with more adequate data in the future (see section 4).

2.2.5 CR with prepositions

Finally, the quantitative results of CR with prepositions is given in Table 5. Only two hits were found; given in (23)-(24). Both are cases of the ‘literal meaning’ function.

(23) Ondanks geen zin, laptop nu **uit- uit** (niet alleen standby-uit).

Even.though no desire laptop now off off not just standby off

‘Even though I don’t feel like (to stop), laptop literally switched off now (not just on standby)’ (SoNaR, untitled document)

(24) Kom eens naar achter. **Achter- achter?** Ja, **achter- achter**.

Come once to the.back the.back the.back? Yes, the.back the.back

‘Can you come to the back. To the literal back? Yes, to the literal back.’

(SoNaR, ‘Barbershop 2 back in business’)

Lemma	Total # hits	# intensification	# literal meaning	# prototypical X
<i>Uit</i> ‘off’	1	0	1	0
<i>Achter</i> ‘back’	1	0	1	0

Table 5: quantitative results for CR with prepositions

2.2.6 Summary

In this section, we have presented the results of the corpus study on Dutch CR. We have seen that CR occurs with all lexical categories. The intensification, ‘literal meaning’ and ‘prototypical X’ function were found, but not the ‘added-value meaning’ as reported by Horn (2018) for English. As for the CR’s with nouns, we saw that *meisje-meisje* (girl girl) occurred several times. Furthermore, both inflectional patterns for the plural were found (*meisje-meisjes* and *meisjes meisjes*). Most other noun CR’s were hapaxes legomena. In the CR with adjectives, we saw that the most frequent semantic function was intensification, however, both other functions were found as well. CR with absolute adjectives resulted in a ‘literal meaning’ function, whereas those with gradable adjectives had a ‘prototypical X’ interpretation. All CR with adverbs had the intensification function. CR with verbs mostly had the intensification function, but there were a few cases of the other two functions as well. Finally, we only found two instances of CR with prepositions, both cases with the ‘literal meaning’ function.

3. A Comparison to other West Germanic languages

In this section, we compare the findings of our own corpus study with previous studies on CR in West Germanic languages, namely American English, British English and German. Furthermore, we also discuss CR in Afrikaans.

3.1. American and British English CR

CR has been investigated for American English by among others Dray (1987), Ghomeshi et al. (2014) and Horn (1993, 2018). All studies indicate that it is a colloquial phenomenon

which occurs mostly in spoken and informal language. The lexical categories that can undergo CR are nouns, adjectives, verbs, adverbs and lexical prepositions (Ghomeshi et al. 2004). These findings are in line with what we have found for Dutch.

Ghomeshi et al. (2004) discuss that what makes English CR so interesting is the fact that it is not restricted by prosodic rules, unlike other cases of reduplication that have been investigated by the Prosodic Morphology research program (McCarthy and Prince 1986 et seq.). That is, whereas it has been shown that phonological reduplication is restricted to the copying of prosodic constituents such as a heavy syllable, a foot, or a prosodic word, this restriction clearly does not hold for CR in English. This can be shown on the one hand by examples in which the copied segment is smaller than the prosodic length of the primary segment, but bigger than a syllable or a foot, as in (25). On the other hand, there are examples in which something larger than a prosodic word is copied (26)-(28).

(25) ... and here are the GLOVE-gloves.

(26) ...you mean thought-about-it considered it or just CONSIDERED-IT-considered-it.

(27) I don't LIKE-HIM-like-him.

(28) OUT-OF-HER-MIND-out-of-her-mind. (Ghomeshi et al. 2004: 321)

Note that with respect to idioms, it holds that only the entire idiom can be reduplicated, not just a part of it:

(29) *OUT-out of her mind. (Ghomeshi et al. 2004: 321)

A similar finding to (25) was found in the Dutch data with the lemma '*meisje*', i.e. *meisje-meisjes* (cf. example (10)). With respect to the findings in (26)-(28), we cannot confirm whether this is possible in Dutch as well, as we have not searched for such strings in the corpus as of yet. Furthermore, given the limitations of the corpus (see section 2), it is unlikely we would find such cases. However, as native speakers we do not find it hard to imagine them being possible. Some examples are given in (30)-(31).

(30) A: Z'n haar zit weer door de war.

His hair sits again through the mess

'His hair is messy again.'

B: Ja, maar niet DOOR-DE-WAR-door-de-war, het valt best mee.

Yes, but not through the mess through the mess, it falls quite with

'Yes, but not completely messy, it's actually not that bad.'

(31) Ik vind dit echt OVER-DE-TOP-over-de-top.

I find this really over the top over the top

'I really find this way too much.'

For Dutch, it also seems very odd or even ungrammatical to reduplicate only part of the fixed expression:

(32) ??Z'n haar zit DOOR- door- de- war.

His hair sits through through the mess

Intended: ‘His hair is messy again.’

(33) ??Ik vind dit echt OVER-over-de-top.

I find this really over over the top.

Intended: ‘it really find this way too much.’

This is again something that should be investigated in depth in future work on CR in Dutch.

Compared to our Dutch findings, it seems that CR in American English is a more productive construction than it is in Dutch. This is not very surprising, as it might in fact be the case that the emergence of the construction in Dutch is influenced by the high degree of contact with (informal) English, especially via social media and streaming platforms such as Netflix. It might also be the case that CR in American English is even more productive by now, as the most in-depth study on the phenomenon, Ghomeshi et al. (2004), is almost two decades old now. This is probably also the case for British English, which was investigated with a corpus study by Hohenhaus (2004). Even though he did find quite some cases of CR, the number of cases with respect to the size of the corpus he used, the British National Corpus, is very low. However, as he mentions, this should be interpreted as CR being underrepresented in the corpus, rather than the phenomenon being marginal or even negligible. That is, the composition of the corpus is 90% written text vs. 10% spoken language. In other words, Hohenhaus (2004) faced the same methodological issue as our study. A second reason he mentions for the low number of CR in the corpus is the fact that CR often requires specific contextual conditions, namely that of a contrast with something else (e.g. contrasting a green/prototypical salad with a tuna salad). As he states, ‘such specific contextual conditions cannot be expected to be particularly frequent in corpora in which even the relevant register, freely spoken language, especially dialogue, is underrepresented’ (Hohenhaus 2004: 311). As for Dutch, then, also for English it would be interesting to look into CR using different methods/data sources (see section 4).

3.2 German CR

The emergence of CR in German has been mentioned and/or discussed in some recent literature (Mau 2002, Finkebeiner 2014, Freywald 2015, Kentner 2017, Bross and Fraser 2020). The first in-depth corpus study is Frankowsky (2022). He investigated CR in two corpora (deTenTen13 and DECOW16, see Frankowsky 200:163 for description and references) which comprised 20 billion tokens each. Even though these corpora consist of written text only, all of the texts were collected from the web. This means that they contain a lot of non-edited material, which is often interaction-oriented writing (blog posts, chats and fora). His corpus search focused on CR with nouns only. In the two corpora, he found 1858 cases of CR. There were 458 different lexical items, of which 262 were hapaxes legomena. The most productive lexical items occurring in CR were nouns, namely *Chef* ‘boss’, *Film* ‘movie’, and *Mädchen* ‘girl’, which he assumes to have been lexicalized in German. Since the study did not investigate CR with other lexical categories besides nouns, it is impossible to compare this formal restriction on CR with Dutch and English CR. However, he does discuss inflection patterns. His findings show that if CR in German are inflected, they almost always inflect as a whole word (in line with Finkbeiner 2014: 186). Inflection is rare, though, since 70% of all CR found in his study consist of bare stems, e.g. in singular for nouns, without inflection for adjectives, and as infinitive for verbs. This is in line with the Dutch findings; we found only very few cases of CR with inflection. As for the semantic functions of CR in German, Frankowsky reports the ‘prototypical X’ function as well. He furthermore

mentions that in many cases, the exact meaning of the CR is explained directly after. Two examples are given in (34) and (35).¹⁴

(34) *Es gibt ja **Musiker-Musiker**, also die studierten, die ihre ganzes Leben lang Unterricht genommen haben, und dann gibt es Herz-Musiker, zu denen wir her gehören.*

‘There are **musicians musicians**, that is, those who have studied all their lives, and then there are musicians at heart [lit. heart musicians] that we rather belong to.’

(35) *Eigentlich handelt es sich hier um eine runde Schachtel voll mit Salat. Also **Salat-Salat**, kein Salat mit so Sachen. Sondern einfach nur Kopfsalat ein bis zwei Kirschtomaten, Karotte und Mais und Balsamico Dressing.*

‘This is actually a round box full of lettuce. So, **salad salad**, no salad with much stuff in it. Just lettuce, one or two cherry tomatoes, carrots, and corn and balsamic dressing.’

(Frankowsky 2022: 174)

As his study only focused on CR with nouns, no mention is made of CR with larger constituents, such as the English cases discussed in Gohmeshi et al. (2004) ((26)-(28) above).

Taken together, even though Frankowsky (2022) only investigated CR with nouns, the results look quite similar to that of Dutch CR. That is, there are a few CR’s with nouns that seem lexicalized (e.g. Dutch *meisje-meisje* and German *Chef-Chef*), most CR’s occur uninflected and have the ‘prototypical X’ function. It would be interesting to see a future investigation of CR in German in which other lexical categories are included, as well as larger constituents such as verbs with complements, fixed expressions and idioms.

3.3 Afrikaans CR

In contrast to the other West Germanic languages, reduplication is a productive means of word formation in Afrikaans. It has a range of semantic, morphosyntactic and pragmatic functions, a few examples of which are given in (36)-(38) (Botha 1988, Conradie 2007, Den Besten et al. 2012, Van Huyssteen 2000, 2004, Van Huyssteen & Wessing 2007, among many others).

(36) Die klomp kinders is toe **blêr-blêr** agter hom aan.

The group of children is then bleat bleat after him on

‘The group of children then followed him bleating.’

(Conradie 2003: 15)

(37) *Flinkdink*, die vinnige spelletjie vir **slim- slim** mense.

Flinkdink, the fast game.DIM for clever clever people

‘Flinkdink, the fast game for very clever people.’

(Conradie 2003: 20)

(38) Sy **voel- voel** met haar voet hoe warm die water is.

she feel feel with her foot how warm the water is

‘She tentatively puts her foot into the water to feel how warm it is.’ (Combrink 1978: 78)

(39) Ons het **huisie- huisie** gespeel en dit was fantasties

we have house.DIM house.DIM play.PP and it be.PRT fantastic

‘We played house and it was fantastic.’

(Korpusportaal, untitled document)

¹⁴ The examples are presented without glosses in Frankowsky (2022), and as the glosses are not necessary for the purposes of these examples here, we have not added them either.

As can be seen from these examples, reduplication can be used to create an adverbial form out of a verb which can replace a present participle (36), and it can be used to semantically intensify (37) or attenuate (38) the non-reduplicated form (see the references above for a full discussion of all the functions of reduplication in Afrikaans). It is also used in creating new words, like the name of a game (39).

As far as we are aware, no mention has been made about the possibility of CR in Afrikaans. We have investigated this with a corpus search, using the corpus *Korpusportaal* (ViVA 2018, see Cavirani-Pots and Dirix (to appear, 2023) for the details of the methodology). We found thousands of examples pertaining to about 500 different lemmas over all words classes, but the only hit we found that seems a clear case of CR is a noun reduplication, as given in (39).

- (39) Ons doen van Eagles- blues, **blues- blues** en selfs Briels- blues...
we do of Eagles blues blues blues and even Briels blues
'We'll be playing a bit of Eagles blues, prototypical blues, and even Briels blues...'
(Korpusportaal, untitled document)

As can be seen, this case of CR is one with an English loan word. This might mean that it was borrowed from English, either from its South African variant or directly from American English. It might thus be interesting to investigate CR in South African English in the future, in particular as there is also considerable language contact with Afrikaans. Even though the same methodological issues to investigate CR hold for Afrikaans as for the other West Germanic languages, it is clear that CR is not a productive phenomenon in Afrikaans. Whether the phenomenon will be used more in the future, possibly due to language contact with English, remains to be investigated in the future.

3.4 Summary

In this section, we have compared the findings of our own corpus study on CR in Dutch with other studies into CR in American and British English and German. We have furthermore discussed CR in Afrikaans, which seems not yet to have emerged in that language. Even though the different studies are from different periods in time (ranging from 2004 until present), all face similar methodological issues, and all had a slightly different focus or set of data, we can tentatively say that among the West Germanic languages, CR is most productive in American English and the least so in Afrikaans. In between are British English, and then Dutch and German. This is represented in the following cline:

- (40) American English > British English > Dutch/German > Afrikaans

Given that the (semi-)productivity of the phenomenon is very recent, we have to wait for the future in order to see whether this cline remains as it is, or whether CR becomes more productive in Dutch, German and/or Afrikaans as well. Generally, it also looks like reduplication in general and CR in particular occurs mostly with shorter words of Germanic stock, although there are some exceptions.

4. Conclusion and outlook

The aim of this paper was to add Dutch to the Germanic typology of CR. We have shown that the phenomenon exists in this language as well, and is possible with nouns, adjectives, adverbs, verbs and lexical prepositions. As was already mentioned in section 2, our study faced methodological issues, for which we see two directions for future research to be resolved. First, it would be interesting to work with very recent data such as Twitter data and

other social media data on a large scale, and write Python scripts to automatically extract all cases of reduplication. Second, it would be worthwhile to combine corpus data with data from elicitation experiments. This second type of data would allow us to investigate the acceptability of different inflection patterns, CR with idioms, the attributive use of adjectives and noun CR's, and target all four semantic functions for all lexical categories. It would furthermore allow us to investigate whether there is an effect of age of the speaker on the acceptability judgments, as well as their daily use of English. These extralinguistic factors potentially have a corroborating effect on accepting CR with larger segments, or accepting more lexical items to occur in CR's. Finally, in future work we hope to investigate the expressive dimension (i.e. the perspective of the speaker on the utterance) that has been mentioned as an additional function of CR.

Besides the empirical dimension, a future direction we hope to take is to contribute to the theoretical analysis of CR. The phenomenon has been analyzed by some as compounding (Hohenhaus 2004, Finkebeiner 2014, Freywald 2015) and by others as syntactic reduplication (Ghomeshi et al. 2004, Travis 2001, 2003, Bross and Fraser 2020). However, both types of analyses still face technical issues (see Ghomeshi et al. 2004 and Bross and Fraser 2020 for discussion), and furthermore neither type of analysis is able to explain how the semantic functions come about, nor why CR often carries expressive meaning. We hope to address these topics in future work, but for now we hope to have shown that CR is an emergent phenomenon in Dutch, which has many interesting sides worth investigating.

References

- Botha, R. P. 1988. Form and meaning in word formation. A study of Afrikaans Reduplication. Cambridge: Cambridge University Press.
- Booij, G. 2019. *The Morphology of Dutch*. Oxford: Oxford University Press.
- Bross, F. and Fraser, K. 2020. Contrastive focus reduplication and the modification puzzle. *Glossa: a Journal of General Linguistics* 5(1):47, 1-18.
- Cavirani-Pots, C. and Dirix, P. To appear (2023). Reduplication or reduplicación-reduplication. A contrastive study of reduplication processes in Afrikaans and Dutch. In: Bekker, I. and T. Biberauer (Eds.), *Ex Africa semper aliquid novi: Linguistic shorts in honour of Andries Coetzee for his 50th birthday, Stellenbosch Papers in Linguistics (special issue)*.
- Conradie, J. 2003. The iconicity of Afrikaans Reduplication. In: W.G. Müller, and O., Fischer (Eds.) *From Sign to Signing*. Amsterdam: Benjamins, 203-224.
- Combrink, J. G. H. 1978. Afrikaans: its origin and development. In: W. Lanham and K. P. Prinsloo (Eds.), *Language and communication studies in South Africa*. Oxford and Cape Town: Oxford University press, 69-95.
- Den Besten, H., Luijckx, C. and Roberge, P.T. 2012. Reduplication in Afrikaans. In: T. van der Wouden (Ed.), *Roots of Afrikaans. Selected writings of Hans den Besten*. Amsterdam: Benjamins, 195-219.
- Dray, N. 1987. Doubles and modifiers in English. Unpublished M.A. thesis, University of Chicago.

- Finkebeiner, R. 2014. Identical constituent compounds in German. *Word Structure* 7, 182-213.
- Freywald, U. 2015. Total Reduplication as a productive process in German. *Studies in Language* 39, 905-945.
- Frankowsky, M. 2022. Extravagant expressions denoting quite normal entities: Identical constituent compounds in German. In: Eitelmann, M. and D. Haumann (Eds.), *Extravagant Morphology. Studies in rule-bending, pattern-extending and theory-challenging morphology*, 155-179. Amsterdam: John Benjamins Publishing.
- Ghomeshi, J., Jackendoff, R., Rosen, N. and Russell, K. 2004. Contrastive focus reduplication in English (the salad-salad paper). *Natural Language and Linguistic Theory* 22, 307-357.
- Hohenhaus, P. 2004. Identical constituent compounding: A corpus-based study. *Folia Linguistica* 38, 297-331.
- Horn, L. R. 1993. Economy and redundancy in a dualistic model of natural language. In: Shore, S. and M. Vilkuna (Eds.), *Sky 1993. 1993 yearbook of the Linguistic Association of Finland*, 33-72. Helsinki: Mouton de Gruyter.
- Horn, L. R. 2018. The lexical clone: pragmatics, prototypes, productivity. In: Finkebeiner, R. and U. Freywald (Eds.), *Exact repetition in grammar and discourse*, 233-264. Berlin/Boston: Mouton de Gruyter.
- Huang, Y. 2009. Neo-Gricean pragmatics and the lexicon. *International Review of Pragmatics* 1, 118-153.
- Huang, Y. 2015. Lexical cloning in English: A neo-Gricean lexical pragmatic analysis. *Journal of Pragmatics* 86, 80-85.
- Kentner, G. 2017. On the emergence of reduplication in German morphophonology. *Zeitschrift für Sprachwissenschaft* 36, 233-277.
- Mau, T. 2002. *Form und Funktion sprachlicher Wiederholungen*. Frankfurt a.M., Berlin, Bern, Bruxelles, New York, Oxford, Wien: Peter Lang (= Forum Linguisticum Vol. 38).
- McCarthy, J. and Prince, A. 1986. 'Prosodic Morphology' ms., University of Massachusetts, Amherst and Brandeis University.
- Oostdijk, N., Reynaert, M., Hoste, V. and Schuurman, I. 2013. The Construction of a 500-Million-Word Reference Corpus of Contemporary Written Dutch. In: P. Spyns & J. Odiijk (eds.), *Essential Speech and Language Technology for Dutch. Results by the Stevin programme*. Springer, 2013, 219-247.
- Potts, C. 2007. The expressive dimension. *Theoretical Linguistics* 33, 165-198.
- Travis, L. 2001. The syntax of reduplication. In: *Proceedings of Northeast Linguistics Society (NELS) 31*, GLSA, 454-469.

Travis, L. 2003. Reduplicatoin feeds syntax. In: Somesfalean, S. (Ed.), *Proceedings of Canadian Linguistics Association (CLA) 2003*, Université du Québec à Montréal. 236-247.

Van Huyssteen, G. B. 2000. *Die reduplikasiekonstruksie in Afrikaans: enkele aspekte van 'n kognitiewe gebruiksgebaseerde beskrywingsmodel vir Afrikaans*. Unpublished PhD dissertation. Potchefstroom: Potchefstroomse Universiteit vir Christelike Hoër Onderwys.

Van Huyssteen, G. B. 2004. Motivating the composition of Afrikaans Reduplications. In: Radden G and K.-U. Panther (Eds). *Studies in Linguistic Motivation*. Berlin: Mouton de Gruyter, 269–292.

Van Huyssteen, G. B. and Wesseling, D. 2007. Datagebaseerde aspekte van Afrikaanse reduplikasies. *Southern Africal Linguistics and Applied Language Studies*, 24:3, 419-439.

ViVA, 2018. *Virtuele Instituut vir Afrikaans*. <http://viva-afrikaans.org>.

Wierzbicka, A. 1991. *Cross-Cultural Pragmatics*. Berlin/New York: Mouton de Gruyter.

Appendix: list of searched lexical items

Nouns

- *jongen* (boy)
- *meisje* (girl)
- *man* (man)
- *vrouw* (woman)
- *kind* (child)
- *werk* (job (Netherlandic Dutch))
- *job* (job Flemish Dutch))
- *salade* (salad)
- *drank* (drink)
- *vlees* (meat)

Adjectives

- *rood* (red)
- *geel* (yellow)
- *groen* (green)
- *blauw* (blue)
- *zwart* (black)
- *wit* (white)
- *blij* (happy)
- *boos* (angry)
- *kwaad* (angry)
- *mooi* (beautiful)
- *lelijk* (ugly)
- *ziek* (ill)
- *dood* (dead)
- *levend* (alive)
- *wakker* (awake)
- *lekker* (tasty)
- *vies* (disgusting)
- *lief* (sweet)

- *streng* (strict)
- *dom* (stupid)
- *stom* (stupid)
- *slim* (smart)
- *gek* (crazy (Netherlandic Dutch))
- *zot* (crazy (Flemish Dutch))

Adverbs

- *snel* (quickly)
- *gauw* (quickly)
- *traag* (slowly)
- *langzaam* (slowly)
- *rap* (quickly (Flemish Dutch))
- *lief* (kindly)
- *streng* (strictly)
- *slim* (smartly)
- *goed* (well)
- *slecht* (badly)

Verbs

- *werken* (to work)
- *slapen* (to sleep)
- *kijken* (to watch)
- *horen* (to hear)
- *lezen* (to read)
- *lopen* (to walk)
- *zitten* (to sit)
- *rennen* (to run)
- *praten* (to talk)
- *trainen* (to train)
- *huilen* (to cry (Netherlandic Dutch))
- *wenen* (to cry (Flemish Dutch))
- *lachen* (to laugh)

Prepositions

- *aan* (at)
- *op* (on)
- *onder* (below)
- *boven* (above)
- *in* (in(to))
- *uit* (out (of))
- *achter* (behind)
- *voor* (in front of)
- *door* (through)
- *af* (off)