

This item is the archived peer-reviewed author-version of:

On comparative concepts and descriptive categories, such as they are

Reference:

Van der Auwera Johan, Sahoo Kalyanamalini.- On comparative concepts and descriptive categories, such as they are
Acta linguistica Hafniensia - ISSN 0374-0463 - 47:2(2015), p. 136-173
Full text (Publishers DOI): <http://dx.doi.org/doi:10.1080/03740463.2015.1115636>

On comparative concepts and descriptive categories, *such* as they are

Johan van der Auwera & Kalyanamalini Sahoo

*University of Antwerp
Center for Grammar, Cognition, en Typology
Prinsstraat 13
B-2000 Antwerp
Belgium*

E-mail : johan.vanderauwera@uantwerpen.be
Kalyanamalini.Sahoo2@uantwerpen.be

Corresponding author:

Johan van der Auwera
E-mail : johan.vanderauwera@uantwerpen.be

1. Introduction

Haspelmath (2010a) claims that the concepts or categories that one needs for cross-linguistic analysis, on the one hand, and for language-specific analysis, on the other hand, are ontologically different things.¹ We maintain that this claim is mistaken. We first discuss Haspelmath's claim in general terms. We then substantiate our disagreement with a study of the categorization of the English word *such* and of its 'counterparts' in Dutch and in the Indo-Aryan language Odia. The reason for focusing on *such* is that it has proven most recalcitrant to categorize. We will show that in English its properties make it absolutely unique – that it is a category all by itself – yet similar to familiar categories such as adjective, article and demonstrative. We will then study the constructions that correspond to *such* in Dutch and in Odia and we will show that constructions are no less unique, yet similar to familiar categories in these languages, but also similar to *such*. All of these categories are the same sorts of things. We propose to call *such* and its counterparts 'similatives' and it is a second purpose of this paper to advance the study of similatives.

¹ Glosses follow the Leipzig Glossing Rules and the one gloss which we have used in this article not found in the Leipzig Glossing Rules is EMPH for emphasis marker. A few examples are from the web, but they are innocent and we do not therefore give any source.

2. Haspelmath (2010a)²

There is much in the Haspelmath (2010a) article that we agree with. Thus we consent to the idea that every language should be described in its own terms. He calls this idea 'categorical particularism' and considers it to be an 'old structuralist insight' (2010a: 663), which, in his opinion, a good number of linguists had lost track of but which has now been reinstated in some quarters of the discipline. The language-particular categories needed are called 'descriptive categories'. To engage in comparative linguistics or typology – Haspelmath (2010a) uses the terms interchangeably – he claims that the linguist should devise another set of categories, called 'comparative concepts'. Here we more or less agree. Thus we don't mind that the entities needed for cross-linguistic work are called 'concepts' and the language-specific ones 'categories': the point that Haspelmath (2010a: 666) wants to make is that calling something a concept emphasizes that the concept is constructed by a linguist. We think, however, and we will show, that a category is also made by a linguist. But it is true that the cross-linguistic construct tends to be more abstract and given that cross-linguistic analysis normally involves more data, there is more room for confusion and divergence of opinion. To reflect these differences, one could indeed give the two constructs different labels and for the sake of easing the comparison with Haspelmath's approach, we here also use the different labels.

We agree when Haspelmath (2010a: 664) contends that the task of the specialist of one language should not consist of merely taking a pre-established list or system of

² The 2010a paper is not the only publication in which Haspelmath presents this view (see esp. Haspelmath 2010b, 2011), but it remains the most explicit and direct one. Haspelmath is also not the only linguist to propose a fairly radical split between language-specific and cross-linguistic categories/concepts and aligns himself with e.g. Croft (2001) and Lazard (2005).

comparative concepts and checking to what extent these concepts are found in this one language. Nevertheless, we think that it is perfectly valuable to check whether concepts established on the basis of, say, 25 languages, make sense for a 26th language, and with respect to this 26th language these concepts can indeed be considered 'pre-established'. To that extent we disagree, and we also disagree, most fundamentally, when Haspelmath (2010a: 666) claims that comparative concepts are 'different kinds of entities' and that one should not think that the comparative concepts are only more general, less specific than the descriptive categories.

We claim that comparative concepts and descriptive categories are the same kinds of entities, differing primarily in abstractness. Note that Haspelmath (2010a: 666) does not deny that descriptive categories can be similar across languages nor that they can 'match' a comparative concept. Haspelmath (2010a) does not tell us anything about just how descriptive categories can be similar, precisely because, we presume, Haspelmath holds that one should not really compare these kinds of entities at all, comparison being the task that would involve the ontologically different comparative concepts. How a descriptive category could 'match' a comparative concept is also not explained.

We think that the explanation is not difficult. Consider first a descriptive category in one language and a similar one in two other languages.

Figure 1: here

How do the categories match? Very simply, A and B match because they share features c and d, thus allowing for a possibly interesting comparative concept defined by c and d. Whether it is really interesting depends on whether the similarity makes sense and allows interesting generalizations. To that extent we again agree with Haspelmath

(2010a: 680). A and C also match, for they share features a and b, thus allowing for a possible comparative concept defined by a and b, which will, again, be interesting depending on further descriptive and explanatory work. Finally, B and C match because of properties e and f, defining yet another possibly interesting comparative concept. Figure 2 adds the three possible comparative concepts.

Figure 2: here

If it turns out that both α and β are indeed cross-linguistically interesting and α is much more frequent than β , then one could further call α 'prototypical'. A cross-linguistic category might also be such that its properties are not all or only rarely instantiated in a given language. One could call this category a 'family resemblance' category or, if there are reasons to think that this category is in some interesting sense 'ideal', then one could call it 'canonical' (Brown & Chumakina 2013). Figure 3 represents such a family resemblance concept.

Figure 3: here

There is a huge literature about categories, and their properties, whether necessary and sufficient or only typical and fuzzy, about prototypes, family resemblance (see van der Auwera & Gast 2011 for some references) and recently also about 'canonicalness'. All that matters for us is here is the question whether or not descriptive categories and comparative concepts are the same sorts of entities. Our answer is positive: they are both sets of properties and in one way or another, properties can be shared across languages and in that case they define the comparative concept. Furthermore, the history of our field strongly supports the idea that descriptive categories and comparative concepts are indeed interconnected. In the Western tradition

comparative concepts were for centuries based on the descriptive categories deemed appropriate for Greek, then for Latin and, later still, for Standard Average European and, in particular, English. As mentioned already, there is nothing wrong with trying out a comparative concept supported by 25 languages on the 26th language. On the contrary, it is highly recommended. The 26th language should most certainly be described in its own terms, but that does not mean that one should start from 'categorical' or 'conceptual' scratch each time one sets out to describe a 'new' language. In the present day and age, we have a decent understanding of hundreds of languages, and we see diversity as well as similarity. Of course, applying categories that are justified for many languages on a 'new' language carries a certain risk: it is tempting to just assume that the 26th language would also fit the categories deemed appropriate for the other languages, especially if one has woven a 'theory' around these categories (cp. van der Auwera & Zamorano Aguilar in print for an illustration in the realm of mood and modality).

Note that nothing is said in the above about the nature of the shared features. They could be formal or functional-semantic. It so happens that in all of the examples of the comparative concepts in the Haspelmath paper (e.g. adjective, future tense, question word, reflexive pronoun) meaning or function is most prominent (Haspelmath 2010a: 663). This is no accident. The wider the cross-linguistic coverage of the comparative concept the more prominent its meaning is likely to be, as compared to form. We will show this in what follows too. We start from a descriptive category in one language. We then move to a closely related second language, and we will see how the emerging comparative concept refers to form as much as meaning or function. When we take a strongly diverging third language, however, the comparative concept gets slimmed down more to semantics.

The preceding discussion was very general. We have not used a single example. The next three sections will be very specific and rich in examples. We will zoom in on a word that has proved particularly difficult to categorize, viz. the English word *such*. We then compare *such* to its counterparts in Dutch (section 4) and in Odia (section 5).

3. English *such*

3.1. Introduction

Over the years the categorization of *such* has attracted quite some attention (e.g. Altenberg 1994, de Mönnink 2000, Mackenzie 1997, Wood 2002, Ghesquière & Van de Velde 2011), but no consensus has yet been reached. Table 1 shows some of the categorizations found in the literature.

Table 1 about here

The situation is very much the quandary that Haspelmath (2010a: 668) describes as resulting from linguists attempting to use cross-linguistic concepts to settle language-specific disputes: '[...] linguists inevitably choose the criteria in such a way that they obtain the results that fit their general perspective best. But since perspectives differ, different linguists arrive at different categories [...]'. To get out of this kind of quandary, Haspelmath (2010a: 668) continues, '[...] the solution is to accept that categories are language-particular, and to describe languages in their own terms'. But this advice won't help here: at least the first six categorizations are offered in language-particular accounts by linguists that unquestionably try to describe English in its own right. This makes clear in a forcible way that language-particular categories are no more or no less 'part of the structure of languages' (Haspelmath 2010a: 666) than cross-

linguistic concepts and no more or no less the constructs made by linguists, and that they are interesting or not depending on the wider perspective.³

In what follows we will discuss to what extent *such* would indeed be a member of certain more established categories. The difference with earlier work is double. First, this comparison is much more comprehensive than the earlier ones⁴, for all in all 34 properties will pass the review. Second, we will not force *such* into the categories discussed. The categories chosen for discussion will be the adjective, the article, and the demonstrative. We could have included many more categories, of course, such as the determiner or the pronoun. We have chosen the adjective, for it figures directly in Table 1 and it is the diachronically original category of English *such* (cp. van de Velde 2009: 267-271, Ghesquière & Van de Velde 2011: 780-782). We have taken the article, as it is arguably the purest case of the determiner in English. Non-article determiners have 'real' determiner uses as well as pronominal uses. We have included one such non-article determiner, viz. the demonstrative. The reason for choosing the demonstrative rather than e.g. the possessive is that *such* has a demonstrative element in its meaning⁵, as shown in (1).

³ Another problem is that a language-particular account can be wrong or misguided or that it uses an idiosyncratic terminology, dangers that beset cross-linguistic accounts to no lesser extent.

⁴ We do not claim the description is exhaustive though, primarily because there are some uses which are very rare now and which were probably more common earlier – compare Poutsma (1916: 931-941, 1296-197) on Late Modern English.

⁵ This point has been made in the literature since at least Mackenzie (1997: 99). Compare also the label 'similarity demonstrative' in Umbach & Gust (2014).

- (1) Give me such a hat please.
≈ Give me a hat of that type please.
≈ Give me a that kind of hat please.

We will also see in section 5 that the Odia counterpart to *such* distinguishes between proximal and distal forms, just like the English demonstratives – and the Odia ones for that matter.

The adjective, article and demonstrative that will enter the comparison are obviously the adjective, article and demonstrative of English. They are, in Haspelmath's terms, 'descriptive categories' and the 34 properties are no less 'descriptive'. We come to their relation to comparative concepts in the next section. Note also that we have just called the adjective, the article and the demonstrative 'more established'. That does not mean their definition is unproblematic or that they do not come in subtypes, which will differ with respect to the properties to be investigated. The problem is most acute for adjectives. For example, we take it in what follows that adjectives allow both attributive and predicative uses, illustrated in (2).

- (2) a. Give me a blue book.
b. The book is blue.

This is a property of most adjectives, but not of the adjective *wooden*, for example.

- (3) a. Give me a wooden chair.
b. *The chair is wooden.

Unless noted explicitly, the comparison below involves only the most typical and the great majority of adjectives. This makes the comparison incomplete but it is rich enough to bring out the unique nature of the word *such*. Note also that the properties that pass the review will not be weighed and this too makes for a rather crude comparison.

Relatedly, although we will call some uses marginal, there is also no systematic study of the frequency of this or the other use. We are fully aware of the fact that frequency properties are important (see van der Auwera & Coussé 2014 for a corpus study of the frequency of uses of *such* and of its counterpart in Swedish) and they will to some extent reflect diachrony, but for the very general point that we want to make in this paper, a frequency study is not necessary.

In the examples that follow, we will always use the definite article *the*, and the proximal demonstratives *this* and *these*. Nothing should hinge on this choice.

3.2. *Adnominal uses*

Let's start with the adnominal uses of *such*, which it shares with adjectives, articles and demonstratives.

(4) I like $\left[\begin{array}{c} \text{such} \\ \text{bad} \\ \text{the} \\ \text{these} \end{array} \right]$ books.

This similarity is trivial, but what is less is (4) illustrates a pattern in which *such*, the adjective, the article and the demonstrative combine with a noun without an article. For *such* and for the adjective this is possible only when the noun is count plural, as in (4), or mass, as in (5).

(5) I like $\left[\begin{array}{c} \text{such} \\ \text{bad} \\ \text{the} \\ \text{this} \end{array} \right]$ gold.

The article and the demonstrative, however, also combine with a singular count noun, as shown in (6).

(6) I like $\left[\begin{array}{c} *such \\ *bad \\ \text{the} \\ \text{this} \end{array} \right]$ book.

In the $[x N]_{NP}$ structure, x can be preceded by a plural or count singular quantifier like *all* and the noun can be preceded by an adjective. This property does not discriminate between *such*, adjectives, articles and demonstratives. (7) illustrates this for count plurals and (8) for a mass noun.

(7) I like all $\left[\begin{array}{c} \text{such} \\ \text{bad} \\ \text{the} \\ \text{these} \end{array} \right]$ French books.

(8) Art, theatre, film and all such great stuff

Interestingly, the co-occurrence with the singular count quantifier *each* does discriminate the four word types.

- (9) Each $\left[\begin{array}{l} \text{such} \\ \text{new} \\ \text{*the} \\ \text{*this} \end{array} \right]$ employer shall accept the conditions.

Note that (7) can have two adjectives, thus showing that the adjective is a recursive category. Of course, there are restrictions on the combinability of adjectives, but they need not concern here. Also, when there is more than one adjective, it can be the same adjective, as in (10), in which *bad bad* means 'very bad'.

- (10) I like bad bad books.

such is not recursive, and neither is the article or the demonstrative. However, *such* does have a doubling use, viz. the idiomatic, non-compositional coordination *such and such*, often spelled as *such-and-such* and illustrated in (11), in which it is equivalent to *certain*.

- (11) In estimating the character of a great general, it is not sufficient to consider that he has overcome such and such people, or overrun such and such countries ...

Adjectives, articles and demonstratives do not allow this use. As with the $[x N]_{NP}$, the N in this *such and such* pattern has to be plural count or mass. Since this condition has no

discriminatory value with respect to adjectives, articles and demonstratives, this condition will not count as a separate property. Note that *such and such* also allows a non-idiomatic, compositional reading, which it shares with demonstratives. Articles do not exhibit this use, and with adjectives the result is semantically anomalous. The semantic anomaly is also given a star, though we realize that it can be saved pragmatically.

- (12) a. He has overcome $\left[\begin{array}{l} *the \text{ and the} \\ *bad \text{ and bad} \\ \text{these and these} \end{array} \right]$ people.
- b.
- c.

We now turn to the combinability of the four word types with constructions with the indefinite singular article *a(n)*. There are two possibilities. Either *x* precedes the indefinite article or the indefinite article precedes *x*. The first constellation singles out *such* and the second the adjective.

- (13) I have never read $\left[\begin{array}{l} \text{such} \\ *bad \\ *the \\ *this \end{array} \right]$ a cheap book.

- (14) I have never read a $\left[\begin{array}{l} *such \\ bad \\ *the \\ *this \end{array} \right]$ cheap book.

In both constructions the N has to be singular count. Note that the fact that the article does not fit the $[a \ x \ N]_{NP}$ or the $[x \ a \ N]_{NP}$ construction also follows from its non-recursive nature. When the indefinite article follows x , *such* also allows the idiomatic $[x \ and \ x \ ADJ]$ pattern, just like with bare nouns, and again adjectives, articles and demonstratives do not allow this pattern. None of the four word types is felicitous with a compositional use.

(15) Today or tomorrow we will go into such and such a town and spend a year there and trade and make a profit.

(16) Give me $\left[\begin{array}{l} *such \ and \ such \\ *French \ and \ French \\ *the \ and \ the \\ *this \ and \ this \end{array} \right]$ a book. [compositional reading]

Next in line are constructions with a definite article, which either follows x or precedes it. The first pattern is uninteresting, for it does not discriminate the four word types (as none can precede *the*), and the second one, unsurprisingly, singles out the adjective.

(17) I have never read the $\left[\begin{array}{l} *such \\ bad \\ *the \\ *this \end{array} \right]$ book

We have nearly completed a review of the properties of the adnominal uses. There are three more points, though. First, there is a set of adjectives that convey a sense of uniqueness, such as *only*, *first*, *last* and superlatives, and they can follow the definite article and precede *such*. In this respect *such* is like adjectives.

(18) The only $\left[\begin{array}{l} \text{such} \\ \text{special} \\ *the \\ *this \end{array} \right]$ place that you should visit is the Grand Canyon.

(19) The first $\left[\begin{array}{l} \text{such} \\ \text{international} \\ *the \\ *these \end{array} \right]$ meetings were held in the twenties.

Second, with respect to meaning the adnominal *such* has two readings, viz. an identifying and intensifying one – a distinction first drawn attention to by Bolinger (1972).

(20) I want *such* a hat, a black one with a feather.
= an indefinite token of a definite type

(21) It was such a nice, sunny day, quite unusual for the season.

= an indefinite but outstanding token of a definite type

No such ambiguity is available for adjectives, determiners or demonstratives. Third, the definiteness is furthermore established anaphorically, cataphorically or exophorically.

Here are examples offered by Mackenzie (1997: 90), culled from a Conan Doyle novel.

For (22) the use is anaphoric: the text preceding the sentence describes a type of presentation. (23) is a cataphoric use. For (24) it is only the non-linguistic context that defines the type. So it is exophoric.

(22) On what occasion would it be most probable that such a presentation would be made?

(23) His mind was prepared for such an end as did eventually overtake him.

(24) I had hardly expected so dolichocephalic a skull or such well-marked supra-orbital development.

The meaning of determiners and demonstratives also involves definiteness and indefiniteness and phoricity. Articles and demonstratives are either definite or indefinite, and when definite, they are also anaphoric, cataphoric or exophoric.

Demonstratives are typically definite, but (25) illustrates an indefinite use.

(25) Yesterday I met this bloke. He wore funny clothes, had a very red face and he spoke Portuguese.

More needs to be said about the cataphoric use. Cataphoric *such* can occur in what could be called a 'correlative' structure. In (23) we see this with *such as* and a finite clause. There are two variants. First, instead of the *as* clause we can have a *that* clause, which expresses result, as in (26), due to Declerck (1991: 314). Second, instead of a clause we can have a noun phrase introduced by *as*, as in (27), due to Quirk et al (1985: 1309).⁶

(26) She is such a good cook that she can prepare any kind of meal.

(27) She still enjoys such books as science fiction, detective stories and historical novels.

None of the correlative cataphoric uses is possible for adjectives, articles or demonstratives.

This concludes the discussion of the adnominal uses of *such*. We have discussed thirteen discriminatory values and we summarize them in Table 2.

Table 2 about here

For three of them *such* is like an adjective, for four different ones it is like a demonstrative, and for three of these four it is also like an article. In short, *such* is really a different kind of word. This conclusion is stronger still, if we look at the other uses of *such*.

⁶ Yet another use, very marginal, has *as* with a *to* infinitive, as in *He was such an important man as to be protected by police night and day*. According to Declerck (1991: 315), this is a variant on *as* followed by a finite clause.

3.3. *Predicative uses*

such allows a predicative (or 'subject complement') use, illustrated in (28), which it shares with adjectives, clearly not with articles and probably not with demonstratives either.

(28) Such is life.

(29) $\left[\begin{array}{l} \text{Bad} \\ * \text{The} \\ * \text{This} \end{array} \right]$ was my experience when I entered the town hall.

The version of (29) with the demonstrative is, of course, a grammatical sentence, but this might be better explained by pointing to the fact that demonstratives have pronominal uses, different from *such* – at least to some extent (see below) – as well as from adjectives and demonstratives.⁷ As to the predicative uses that *such* shares with adjectives, for *such* they seem to be limited to a bare clause-initial position or to a postverbal position in which *such* is followed by either a *that* clause or by a *to* infinitive introduced by *as*.

(30) Life was $\left[\begin{array}{l} * \text{ such.} \\ \text{pleasant.} \end{array} \right]$

⁷ It has also been suggested that in (28) *such* is a pronoun (Quirk et al 1985: 376). Then *such* would pattern with demonstratives.

(31) The noise was $\left[\begin{array}{c} \text{such} \\ * \text{loud} \end{array} \right]$ that I could not understand her.

(32) The noise was $\left[\begin{array}{c} \text{such} \\ * \text{loud} \end{array} \right]$ as to make it impossible to hear anything.

We take it that (33) instantiates the same predicative postverbal use, though with a zero copula and with *such that* being equivalent to *that are such that*.

(33) Find three numbers $\left[\begin{array}{c} \text{such} \\ * \text{special} \\ * \text{the} \\ * \text{these} \end{array} \right]$ that the sum of the first two numbers is equal to ...

We further assume the use in the phrases *such as it is/was* or *such as they are/were*, as in (34) or *if it may be called such* in (35) is also predicative. This is a use shared by adjectives, but the *such* phrase is non-compositionally depreciatory: *such as it is* is equivalent to *for what it is worth*.

(34) The food, $\left[\begin{array}{c} \text{such} \\ \text{rich} \\ * \text{the} \\ * \text{that} \end{array} \right]$ as it is, will be prepared by the nurses.

- (35) The food, if it may be called $\left[\begin{array}{l} \text{such,} \\ \text{rich,} \\ \text{*the,} \\ \text{*that,} \end{array} \right]$ will be prepared by the nurses.

Table 3 summarizes the findings and again we see that *such* is a very peculiar word.

Table 3 about here

3.4. *Pronominal uses*

Now we come to pronominal uses. In the case of *such*, the use is very limited. Bare *such* only (or at least primarily) occurs as a subject and it can, compositionally, combine with *and such*. For demonstratives there are no limitations, and for adjectives and articles pronominal uses are impossible (examples based on Quirk et al. 1985: 376).

- (36) but $\left[\begin{array}{l} \text{such} \\ \text{*bad} \\ \text{*the} \\ \text{these} \end{array} \right]$ have yet to be reported.

- (37) but $\left[\begin{array}{l} \text{such and such} \\ \text{*bad and bad} \\ \text{*the and the} \\ \text{these and these} \end{array} \right]$ have yet to be reported.

- (38) I very clearly saw
- | |
|--------|
| *such. |
| *bad. |
| *the. |
| this. |

Then there are two further pronominal uses. The first is in the phrase *and such* (also occurring as *and suchlike*), in which *such* means 'such things/persons' or 'the like'. This use is unavailable for articles and adjectives, but also for demonstratives.

- (39) There you can buy bags, wallets, and
- | |
|---------|
| such. |
| *bad. |
| *the. |
| *these. |

The second pronominal use is in the non-compositional phrase *such and such*, a use unavailable to adjectives, articles and demonstratives.

- (40) Yoga is great, I know; such and such told me about it but I can't touch my feet.

Table 4 about here

3.5. *Minor uses*

such also has a nominal use, illustrated in (41), and it occurs in *none such*, in which *such* is neither a nominal nor a pronominal.

(41) ... but $\left[\begin{array}{l} \text{many such} \\ * \text{many good} \\ * \text{many the} \\ * \text{many these} \end{array} \right]$ have yet to be reported

(42) ... but $\left[\begin{array}{l} \text{none such} \\ * \text{none good} \\ * \text{none thee} \\ * \text{none these} \end{array} \right]$ have/has⁸ yet to be reported

Another use is what could be called 'depictive' (Himmelman & Schultze-Berndt (eds.) 2005). It is typical for adjectives, as in (43), and while *such* does not allow it, *as such* does.

(43) I like it $\left[\begin{array}{l} * \text{such.} \\ \text{raw.} \\ \text{the} \\ \text{this.} \end{array} \right]$

(44) I like it $\left[\begin{array}{l} \text{as such.} \\ * \text{as raw.} \\ * \text{as the.} \\ * \text{as this.} \end{array} \right]$

⁸ Quirk et al (1985: 376) allow *none such* with both singular and plural agreement.

With *as such* we are considering a phrase. The combination of *such* and *as* also exists in the opposite order, as we have already seen in (32) and (34). In the postnominal use in (45) *such as* is equivalent to *like*. This use is not available for adjectives, determiners or demonstratives

- (45) People

such
*bad
*the
*these

 as John and Bill should be avoided.

Such does not have any adverbial use. Neither do articles have them or adjectives⁹, but the singular demonstratives do have them.

- (46) I had never seen him

*such
*bad
*the
this

 crazy.

Table 5 about here

3.6. Morphological properties

The final properties concern morphology. The question whether the four word types are sensitive to gradation (comparative and superlative), number and spatial deixis (distance

⁹ One could disagree here, given the existence of the correlative *the* (*the sooner the better*) and the fact that some adjectives can do adverbial work (*The fast runner ran fast*).

with respect to speaker gives) gives us the answers in Table 6. (47) to (49) illustrate these.

Table 6 about here

(47) This car is

*sucher.
faster.
*the-er.
*thiser.

(48)

Such
French
The
This/these

 book/books is/are expensive.

(49) Give me

such a
a green
the
this/that

 book.

With a 34rd and final property we state that *such* has an interrogative-relative counterpart in the word *which*.

(50) Which books do you like?

(51) The books which I like are *Burmese Days* and *1984*.

This link is seldom stated in synchronic descriptions, reflecting that this link is synchronically very weak. If we were to investigate the word *which* in terms of all the properties discussed so far, we would notice that *such* and *which* are indeed very different. But there is nevertheless a connection. It is semantic as well as formal, most clearly between *such* and interrogative *which*. Not unlike *such* making reference to two sets of things, viz. the indefinite token(s) of a definite type, interrogative *which* involves a set of definite entities, in (51) the books that one is talking about, and there is also set of the books that the hearer likes, the membership of which is unknown (and in that sense indefinite). However weak the link, adjectives, articles and demonstratives do not have interrogative-relative counterparts.

3.7. Conclusion

We have now come to an end of the comparison. 34 features are found to be discriminatory. The properties and similarities are summarized in Table 7. For more than half of the parameters (19, to be exact), *such* is different from adjectives, articles as well as demonstratives. *such* is best as a demonstrative, with altogether 9 similarities, less good as an article and least good as adjective. Still, *such* can be taken as a subcategory of the article, the demonstrative or the adjective, but then it would be a special subcategory, with a majority of unique properties. So one can just as well consider *such* to be a category of its own, with, to be sure, an overlap with the other categories, but a minimal one.

Table 7 about here

One can retort to this conclusion in more than one way. First, one might claim that there is not really one *such* word, but several ones. For those uses in which *such* is like adjective, it is an adjective, and for those uses in which it is like a demonstrative, it is a demonstrative. This is a defensible approach, but at a deeper level, the original problem reappears. If there is indeed a *such* adjective as well as a *such* demonstrative etc., what kind of word – what category – is it that allows *such* as such to display this multiple category membership? We are not dealing with an accidental homonymy, but with polysemy: all the uses involve a sense of similarity.

Second, given that *such* is a hyperindividual item, one could argue that there is no need to categorize it. It would then just be an individual item resisting a neat categorization. It would be an item that falls in between categories. This is also a defensible approach. And, indeed, if *such* were a quirk of English only, then why should one go beyond the listing of properties given so far and say that the 34 properties define a category. And, very much in the spirit of Haspelmath's, just like languages deserve to be studied in their own right and not forced into the moulds of other languages, so words deserve to be studied in their own right and not be forced into the adjective, demonstrative or article categories. In this way this study of *such* is ideal grist to the mill of Haspelmath's 'categorial particularism'. However, in case other languages also have *such* words, with similar properties, the situation will be different. In the next section, we will explore this issue with respect to Dutch.

4. Dutch *zulk* and *zo'n*

4.1. Introduction

Dutch has a *such* word, too, and in fact even two words or, better, constructions, viz. the

etymological counterpart *zulk* and also *zo'n* 'so a'. We will review these constructions in terms of the properties deemed relevant for English (and we will add one property). Of course, we now compare *zulk* and *zo'n* to the adjectives, articles and demonstratives of Dutch. So this will again be a language-particular 'descriptive' analysis. We assume, however – and we are not aware that anyone could take issues with this – that the adjectives, articles, and demonstratives of Dutch are at least very similar to the ones of English, so that we can legitimately compare *zulk* and *zo'n* to *such*. To illustrate this point and to anticipate on the discussion: in English *such a book* is grammatical and *such the book* is not, and likewise, in Dutch *zulk een boek* is grammatical and *zulk het boek* is not. Neither English *the* nor *a* will have exactly the same properties as Dutch *het* and *een*, but this does not stand in the way of observing a strong similarity between English *such* and Dutch *zulk* and to phrase it in terms of 'sharing a property': both *such* and *zulk* can be followed by indefinite articles in their languages, but not by a definite articles. So while the language-particular analysis that follows does refer to the language-particular category of e.g. the Dutch definite article, what makes this article different from the English one is irrelevant, and the similarity between *such* and *zulk* can thus also be stated in terms of what the two articles have in common, to wit, in terms of the 'comparative concept', be it that is found both in English as well as in Dutch, and that it is thus ontologically no different from the latter's 'descriptive categories'. This double layering, we claim, holds good for all the properties to be discussed. Note also that just like for *such*, the analysis of *zulk* en *zo'n* will necessarily be crude, largely because we will treat major and minor uses, frequent and infrequent uses, in the same way. For example, we know that *zulk een boek* illustrates an infrequent and archaic use, but it still hasn't left the language.

4.2. *zulk*

For the adnominal uses, *zulk* is exactly *such*, except that it lacks a counterpart to the idiomatic *such and such* construction and that it lack a clausal correlative introduced by the counterpart of *as*. Thus (52) only has the compositional meaning and (53) is ungrammatical.

(52) *Ik wil zulke en zulke boeken.*

I want such and such books

'I want such and such books.'

(53) **Je moet zaken onderwijzen als later nuttig zullen zijn.*

you must things teach as later useful will be

'You must teach things as will be useful later.'

Thus *zulk* fits the counterparts to the [*all* x ADJ N] construction, when N is plural count or mass, and the [*each* x ADJ N] construction, and *zulk* is not recursive.

(54) *Ik wil al zulke goede boeken.*

I want all such good books

'I want all such good good books.'

(55) *Elk zulk ongeval is er één te veel.*

each such accident is there one too much

'Each such accident is one too many.'

(56) **Ik wil zulke zulke boeken.*

I want such such books

'*I want such such books.'

zulk also fits the counterparts to the [x a ADJ N] and the [*the* ADJ-UNI x N] construction, but not the [*a/the/this* x ADJ N] construction.

(57) *Ik hou van zulk een boek.*

I hold of such a book

'I like such a book.'

(58) **Ik hou van een zulk boek.*

I hold of a such book

'*I like a such book.'

(59) *De eerste zulke toestellen zouden in de*

the first such machines should in the

Zuid-Italiaanse stad Sala Consilina zijn opgedoken.

South-Italian city Sala Consilina be appeared

'The first such machines would have appeared in the South-Italian city of Sala Consilina.'

zulk has an identifying and intensifying use and the notions of definiteness and phoricity are relevant in the same way as they are for *such*.

(60) *Ik wil zulke boeken, die hier en niet die daar!*

I want such books, those here and not those there

'I want those books, the ones over here, not the ones over there.'

(61) *Het waren echt zulke mooie boeken, ongelooflijk mooi!*

it were really such beautiful books incredibly

beautiful

'They were really such beautiful books, incredibly beautiful.'

And, finally, Dutch *zulk* allows correlative result clauses introduced by the counterparts of *that* and correlative NP introduced by the counterpart of *as* (examples due to Ghesquière & Van de Velde 2011: 774).

(62) *Het waren zulke goede musici dat ze nooit zonder werk zaten*

it were such good musicians that they never without work sat

'They were such good musicians they were never without work.'

(63) *... met zulke uiteenlopende gebieden als complexe analyse, topologie*

with such diverse fields as complex analysis, topology

en ...

and ...

'... with such diverse fields as complex analysis, topology and'

Table 8 surveys the similarities for the adnominal use.

Table 8 about here

We see that *zulk* is similar to *such* in 11 out of the 13 relevant properties. Thus in the adnominal domain, *zulk* and *such* are very similar. For all the other uses, however, the differences loom large. Different from *such*, *zulk* has no predicative, nominal or depictive uses and there is nothing corresponding to *none such* and *such as* 'like'.

(64) **Zulk is het leven.*

such is the life

'Such is life.'

(65) **Het lawaai was zulk dat ik haar niet kon verstaan.*

the noise was such that I her not could understand

'The noise was such that I could not understand her.'

(66) *Het lawaai was zulk als het onmogelijk te maken*

the noise was such as it impossible to make

iets te horen.

something to hear

'The noise was such as to make it impossible to hear something.'

(67) **Zoek drie getallen zulk dat de som van de eerste ...*

search three numbers such that the sum of the first

'Find three numbers such that the sum of the first ...'

(68) **Dit is mijn familie, zulk als ze is.*

this is my family such as she is

'This is my family, such as it is.'

(69) *Vele zulke moeten beschreven worden
many such must described become
'Many such must be described'

(70) *Geen zulke moeten beschreven worden.
none such must described become
'None such must be described.'

(71) *Ik hou er zulk van.
I hold there such of
'I like it such.'

(72) *Ik hou er als zulk van.
I hold there as such of
'I like it as such.'

(73) *Mensen zulk als John en Bill moeten vermeden
people such as John and Bill must avoided
worden.
become
'People such as Johan and Bill must be avoided.'

zulk does have a pronominal use, however, and it even has a special neuter singular form *zulks*. There are no restrictions to a subject position and the non-compositional uses of *and such* and of *such and such* are absent.

(74) *Ik heb zulke nooit gezien.*
I have such never seen
'I have never seen such things/people a thing.'

(75) *Ik heb zulks nooit gezien,*

I have such never seen

'I have never seen such a thing.'

(76) **Je kan er boeken kopen, kranten en zulke.*

you can there books buy newspapers and such

'You can buy books there, newspapers and such.'

(77) **Zulke en zulke heeft me verteld dat ...*

such and such has me told that

'Such and such has told me that'

Like *such*, *zulk* lacks an adverbial use; Dutch uses *zo* 'so'.

(78) *Hij was niet *zulk overtuigend.*

he was not such convincing

'He was not so convincing.'

Table 9 surveys the non-adnominal uses of *such* and *zulk*.

Table 9 about here

With respect to morphology, *zulk* is insensitive to gradation and deixis, but *zulk* is sensitive to number, and as we have already illustrated, in its pronominal use illustrated in (75), *such* is sensitive to gender, too.

(79) **Dit boek is zulker.*

this book is sucher

*'This book is sucher.'

(80) *Zulk goud/zulke hoeden vind je nergens.*

such gold such hats find you nowhere

'You won't find such good/hats anywhere.'

(81) *Geef me zulke hoeden.*

give me such hats

'Give me such hats.'

[no indication about the distance of the hats relative to the speaker]

Like for *such* there is an interrogative-relative counterpart, viz. *welk*.

(82) *Welke boeken verkies je?*

which boeken prefer you

'Which books do you prefer?'

(83) *De boeken welke ik verkies zijn Burmese Days en 1984.*

the books which I prefer are *Burmese Days* and *1984*.

'The books which I prefer are *Burmese Days and 1984*.'

Table 10 about here

35 properties have now passed the review, the 34 properties that were discriminatory for *such* and the gender sensitivity property which distinguishes *zulk* from *such*. For 17 of them *zulk* and *such* have identical values. Remember that

language-internally *such* was closest to demonstratives but the similarity involves only 9 (out of 34) values. Now we hit upon a more important similarity, for it concerns 17 (out of 35) values. Thus *such* is much more similar to a word in another language than to any word in its own language. Moreover, in its own language the other word, i.e. *zulk*, would also be very individual, though not as hyperindividual as *such* is.¹⁰ This similarity is beginning to be worthy of a label and as *such* and *zulk* express similarity, we propose to call them 'similatives'. Of course, *such* and *zulk* still differ. At the most general level, each of the properties has a language-specific dimension: when reference was made to an adjective, for example, it was the English adjective for English and the Dutch one for Dutch. But since we assume that English and Dutch adjectives are very similar, we can abstract from the difference and state the relevant property in terms of the 'Anglo-Dutch adjective' or, for simplicity's sake, in terms of the adjective *tout court*. At a more specific level, *zulk* has no counterparts to the phrasal uses of *such as* and *such that*, and it does not have the idiomatic *such and such* uses. It also lacks predicative and depictive uses and to that extent *zulk* is less adjectival than *such* (cp. Ghesquière & Van de Velde 2011: 780). With respect to the 13 features pertaining to the adnominal use (features 1 to 13), the similarity is very high, however: for those 13 features there is an identity for 11 of them, the only two features that disturb the picture, viz. the idiomatic use of *such and such* and the correlative use with a finite clause starting with *as*. At least for English *such* and Dutch *zulk*, the adnominal uses are the core features of the similative category.

Most importantly, we now also have a comparative concept, at least one that captures the similarities between English *such* and Dutch *zulk*. This comparative

¹⁰ We would see this if we tabulated the similarities of *zulk* to Dutch adjectives, articles and demonstratives in the same way as Table 7 does for English *such*.

concept is defined with the 17 properties that *such* and *zulk* have in common. In Figure 4 we present the *such* and *zulk* categories and the matching cross-linguistic concept in the format presented in Figure 1.

Figure 4: here

We cannot see any harm in calling the comparative concept 'similative' too. *such* is the English similative, *zulk* a Dutch one and their similarity, defined with 17 features, defines a cross-linguistic similative, but, of course, so far only an Anglo-Dutch one, and only with respect to *such* and *zulk*.¹¹

4.3. *zo'n*

Let us briefly turn to the second similative of Dutch, viz. *zo'n*. *Zo'n* is originally a phrasal unit. *'n* is short for the indefinite article *een* 'a', the phrase stood for 'so a' and was thus limited to singular count nouns. *Zo'n* is in the process of taking over the functions of *zulk* and it has progressed furthest in the Southern, Belgian variant of Dutch. In this variant *zo'n* has lost its original singular count character.

- (84) *zo'n boeken*
so.a books
'such books'

¹¹ If the same term is used for both descriptive categories and comparative concepts, some linguists, including Haspelmath, give the first ones an initial capital. We don't follow this practice in this paper: calling something a Dutch or English similative is clear enough.

- (85) *zo'n goud*
 so.a gold
 'such gold'

In this respect (property 1) *zo'n* is thus different from *zulk*. And there are a few more differences. Different from *zulk*, *zo'n* does not combine with a NP headed by an indefinite article (property 6), no doubt because it is obstructed by the etymological indefinite article 'n that is part of *zo'n*. The origin probably also obstructs the combination with phrases like *the first* (property 8). Thus, though *zo'n* has bleached number, it has not bleached indefiniteness.

- (86) **zo'n een grote hond*
 so.a a big dog
 'such a big dog'
- (87) **de eerste zo'n bijeenkomsten*
 the first so.a meetings
 'the first such meetings'

Then *zo'n* is invariable and thus insensitive for number and gender (features 32 and 35), so there is no counterpart to the *zulk-zulke-zulks* variation and there is also no interrogative-relative counterpart (feature 34). Nevertheless, in terms of the number of specifically simulative properties of English *such*, *zo'n* shares 14 of them, which is, interestingly, higher than the number of features that *zo'n* shares with *zulk*. Meanwhile the comparative category of the Anglo-Dutch simulative, which holds good for *such*, *zulk* and *zo'n*, has lost five features. But the three of them still share 12 features. This

shows that postulating a 'similative' comparative concept is a sensible decision. If Dutch and English just had *zulk*, *such* and *zo'n* as strange words defying any sensible categorization, one could more easily treat them as unique irregular lexical units not worthy of a category label.

Figure 5 about here

For a full language-specific account of *zo'n* more needs to be said. Thus *zo'n* developed something close to a so-called 'recognitional' use (Himmelmann 2001: 833), paraphrasable by 'a typical' or 'one of those' and when it precedes a numeral it can mean 'more or less' (Van Olmen & van der Auwera 2014).

(88) *Zij heeft zo'n appartement aan zee, weet je wel.*
she has so.a apartment at sea know you well
'She has one of those apartments at the coast, you know.'

(89) *Er waren zo'n twee honderd betogers.*
there were so.a two hundred demonstrators
'There were about two hundred demonstrators.'

In both respects *zo'n* is different from both *such* and *zulk*, but facts like these do not influence the content of the comparative concept nor do they bear on the basic conclusion: comparative concepts and descriptive categories are the same sorts of things. They are made of exactly the same material, viz. features, and the decision to posit them depends on the goals of the linguist. Thus a linguist of English may indeed want to recognize the unique language-specific nature of *such* with a label like

'similative' or (s)he may prefer to say that it is special kind of demonstrative ... or article or adjective. In the same vein, the comparative linguist may or may not want to recognize that despite the very special language-specific properties of *such*, *zulk* and *zo'n*, they have a good deal in common and use a label for this. No harm is done either by using the same label as the one used for language-specific needs. On the contrary, by being a semantic label, there is every chance that the label covers what it is essential to this concept and categories.

For a better understanding of the value of the 'similative' concepts, more languages should, of course, be looked at, for one can retort to the preceding argumentation that we have really only shown that 'similatives' can be of value in the context of contrastive linguistics – the linguistics of typically just two languages, and usually closely related ones. This objection is easy to retort, for the simple reason that a comparative concept, even if it only holds for just two languages, is still a comparative concept. Nevertheless, in the next section, we briefly discuss one more language, and one that is only distantly related.

5. Odia

Odia, also known as Oriya, is an Indo-Aryan language spoken in the eastern part of India¹². Of particular interest for our purposes is that Odia has no obvious definite or indefinite articles, although indefiniteness and definiteness can be marked in other ways. We have seen in features 1 to 13 that both English and Dutch similatives are very sensitive to the presence or absence of articles, so it will be interesting to see what will

¹² We adopt the following transcription conventions: the symbol T and D stand for the retroflex voiceless and voiced stops respectively, N stands for the retroflex nasal, R for the retroflex flap, vowel symbols repeated twice indicate long vowels, and aspiration is indicated by adding an extra *h*.

remain of the comparative concept based so far on two article prominent languages.

For indefiniteness, the closest Odia construction with an indefinite article use is the word *goTie* (or its short form *goTe*). It derives from a morphological indefiniteness marker *-e*, attached to a classifier *goTaa* 'a unified whole, complete entity' (Sahoo 1996: 50). It is used as an indefinite article but also as a numeral 'one'.

- (90) *mun goTe ghara kiNili*
I one house bought
'I bought a/one house.'

Then there are two morphological strategies. The first one is the above-mentioned suffix *-e*. Sahoo (1999: 103) derives it from a numeral *eka* 'one'.¹³ It has to combine with a classifier, and the latter can either precede or follow the noun.

- (91) *bastaa-e chaauLa*
sack-INDF rice
'a sack of rice'

- (92) *jotaa haL-e*
shoe pair-INDF
'a pair of shoes'

¹³ *eka*, different from *goTie/goTe*, is used in mathematics and occasionally in formal speech. In everyday language *goTie/goTe* is the numeral 'one'.

In (93) *-e* combines with a classifier which is itself followed by a general, default singular classifier *-Taa*, which functions as a suffix.

- (93) *ghaasa gochhaa-Taa-e*
grass bundle-Taa-INDF
'a bundle of grass'

The second morphological indefiniteness marker *-Te* derives from *-Taa-e*.

- (94) *pilaa-Te*
child-INDF
'a child'

Definiteness can be marked by various affixes, viz. *-ka*, *-Taa*, *-taka*, *-guDaa*, etc. and through the change of word-order. First, the definite affix *-ka*, which also derives from the numeral *eka* 'one', is used in the post-nominal post-classifier position.

- (95) *chaauLa* *bastaa-ka*
rice sack-DEF
'the sack of rice'

Second, the *-Taa* suffix, which we have just met as a default classifier component of the indefinite *-Te*, has a definite meaning when it is attached to a singular noun or to a classifier in postnominal position.

(96) *pilaa-Taa*

child-DEF

‘the child’

(97) *ghaasa gochhaa-Taa*

grass bundle-DEF

‘the bundle of grass’

For mass nouns without classifier, the suffix *-taka* is used.

(98) *paaNi-taka*

water-DEF

‘the water.’

For plural count nouns there are two postnominal definiteness markers, *-guDaa* and *-maane* (the second only for humans and gods).

(99) *pilaa-guDaa/maane kuaaDe gale*

child-PL.DEF where went

‘Where did the children go?’

When the noun is accompanied by a numeral, *-Taa* can also convey definiteness, but only when it attaches to the numeral in postnominal position.

(100) *bahi dui-Taa*
book two-DEF
'the two books'

The numeral can also precede the noun and combine with *-Taa*, but then the meaning is indefinite.

(101) *dui-Taa bahi*
two-INDF book
'two books'

In looking for a counterpart of *such*, by translating sentences illustrating the adnominal uses of *such*, there are two candidates, viz. *emiti* and *semiti*. They both translate adnominal *such/zulk/zo 'n*, but with the first form the standard for the similarity is close to the speaker and it means 'like this', whereas the meaning is 'like that' for the second form. In other words, Odia similatives are sensitive to place deixis, like demonstratives in English and Dutch – and no less so the demonstratives in Odia. *emiti* is proximal, and *semiti* is distal. (102) illustrates an identifying use of *emiti* and (103) an intensifying one.

(102) *mora emiti goTe tinikoNiaa kaLaa Topi darakaara*
my such.PROX one triangular black hat necessary
'I want such a triangular black hat.'

(103) *se emiti bhala roseiaa*
 he such.PROX good cook
 ‘He is such a good cook!’

In what follows we check whether the potential Odia similatives *emiti* and *semiti* fit the comparative concept arrived at after the study of *such*, *zulk* and *zo'n*. In the preceding paragraph we have already discussed 2 of the 12 properties. The first is property 9, illustrating the availability of both intensifying and identifying uses. The second property is 33: different from the English and Dutch similatives the Odia similatives are sensitive to place deixis. Of the remaining 10 properties, 5 survive, viz. properties 10, 12, 27, 30 and 31. Thus *emiti* and *semiti* are definite and indefinite in the same way as *such*, *zulk* and *zo'n*, and they can be anaphoric, cataphoric and exophoric (property 10).

(104) *John sabubeLe office-ku Derire aase: aau mote emiti abhyaasa*
 John always office-to late come and me such.PROX habit
moTe bhala laage-ni
 at.all good feel-NEG

‘John always comes to the office late, and I don’t like such habits.’

(105) *emiti chhaatra ebe kaain!*
 such.PROX student nowadays where !
ekalavya nija buDhaa aanguThi kaaTi taanka
 ekalavya own thumb having cut his
guru DroNanku gurudakshinaa deithile
 teacher Drona.to donation had. given

'Nowadays we don't find such students. Ekalavya had cut his own thumb finger and had donated it to his teacher Drona.'

(106) *mora emiti phone nuhan, semiti phone-Te darakaara*
mine such.PROX phone be.not such.DIST phone -INDF
required

'Not this type of phone, I want one like that.'

Like *such*, *zulk* and *zo'n*, *emiti* and *semiti* can be followed by a *that* clause expressing a result (property 12).

(107) *se emiti bhala roseiaa je se sabu prakaara khaadya*
he such.PROX good cook that he all sorts food
raandipaariba
can.cook

'He is such a good cook that he can prepare all kinds of food.'

There are no depictive or adverbial uses for *emiti* and *semiti* and the forms do not accept gradation (properties 27, 30, and 31).

The remaining properties do not hold for *emiti* and *semiti*. Thus the Odia similatives do not follow the word for *each*, but instead precede it (property 2).

(108) **pratyeka emiti bhala kheLaaLi-ku swarna padaka miLiba*
 each such.PROX good player-ACC gold medal gotten
 ‘Each such good player will be given a gold medal.’

(109) *emiti pratyeka bhala kheLaaLi-ku swarna padaka miLiba*
 such.PROX each good player-ACC gold medal gotten
 ‘Each such good player will be given a gold medal.’

Different from *such*, *zulk* and *zo'n* Odia *emiti* and *semiti* can be repeated and even combined, with *emiti semiti* meaning 'trivial' (property 3).

(110) *se beNiTaa-ku emiti emiti duithara bulaai, taa upare*
 she braid-ACC such.PROX such.PROX twice turn, on that
bobby pin lagaai-delaa
 bobby pin fix-gave
 ‘She turned the braid like this twice, and fixed the bobby pin on that.’

(111) *eiTaa emiti semiti abhijoga nuhen, hatyaa abhijoga,*
 this such.PROX such.DIST accusation be.NEG, murder accusation,
taahaa puNi double murder
 that too double murder
 ‘This is not a trivial accusation, this is murder accusation, and that’s double murder too.’

Odia has nothing that corresponds to *such N as N* (property 13). Bare *emiti* and *semiti* do not have pronominal uses either (property 21) - for a pronominal uses they take the -*aa* suffix.¹⁴

(112) *mora emiti-aa-Te darakaara*
 my such.PROX-NOM-INDF necessary
 ‘I want one like this.’

For property 7, the impossibility of *such* to occur in between *a/the/this* and (ADJ) N, the verdict is mixed. It does apply to *emiti* and *semiti*, with respect to *the*, for the simple reason that there is no *the* word in Odia, and, as (113) shows, it also applies to *this*.

(113) **mun ei emiti kharaapa bahiTaa kebe bi paDhinathili*
 I this such.PROX bad book-Taa never EMPH had.not.read
 *‘I had never read this/that such bad book.’

However, property 7 does not apply to *goTie*, if at least we can take it as an indefinite article.

(114) *mora goTie emiti mercedes car darakaara*
 mine one such.PROX Mercedes car required
 ‘I want such a Mercedes car.’

¹⁴ It is interesting – though not relevant for our purpose – that the forms with *-aa* can be used adnominally too, but only in the identifying use. The adnominal *-aa* uses also feel 'less polished' and rural.

The *emiti goTie* order is also acceptable, but in this order we have the numeral reading – compare (115) with (116) and (117), sentences with the numeral ‘two’, and with (118), a sentence with the formal ‘one’ numeral *eka*.

(115) *mora emiti goTie mercedes car darakaara*

mine such.PROX oneMercedes car required

‘I want one such Mercedes car.’

(116) *mora emiti duiTaa mercedes car darakaara*

mine such two Mercedes car required

‘I want two such Mercedes cars.’

(117) **mora duiTaa emiti mercedes car darakaara*

mine two such Mercedes car required

‘I want two such Mercedes cars.’

(118) *emiti eka bagichaare mun baandhichi mora ghara*

such.PROX one garden.in I have.set.up my house

‘I have set up my house in such a garden’

(from the song *emiti eka bagichaa* ‘such a garden’ by Akshya Mohanty)

Property 7 thus only applies partially and to simplify matters, we will give it a negative evaluation.

Table 11 summarizes the values for *emiti* and *semiti*. We keep the wording of the earlier tables and the features thus explicitly refer to English *such*, but they also apply to Dutch *zulk* and *zo'n*.

Table 11 about here

The above is not, of course, a full analysis of *emiti* and *semiti*. Most importantly, it does not tell us to what extent *emiti* and *semiti* are, language-internally, similar to adjectives and demonstratives, and, interestingly, it does not show that *emiti* and *semiti*, like *such* and *zulk*, have interrogative and relative counterparts – two different ones even. What we targeted here is whether there is still a comparative concept, and there clearly is one. Figure 6 visualizes it.

Figure 6 about here

After three languages and five simulative categories the simulative concept has become more and more a matter of semantics. There is, for example, no reference any more to a definite or an indefinite article. The simulative now is something adnominal that expresses non-gradable similarity that can be identifying and intensifying and that can express a result. How much of this survives the testing of yet a fourth language is not our concern here. The point is that the English Dutch Odia comparative simulative is no different in kind from the language-particular ones.

6. Conclusion

To sum up, we have seen that the categorization of English *such* is a tricky matter. It is not obvious that there is one good categorization and that the others are bad: it at least partially depends on the goals of the linguist. And this is one respect in which language-particular categorization is not different from comparative conceptualization. We have also argued that it makes sense to recognize the uniqueness of *such* with a label of its own, viz. 'simulative', and that part of the attractiveness lies in the fact that Dutch also

has a unique construction and even two, with similar properties, also then 'similatives'. We have shown how the three constructions are similar and how they define a 'similative' comparative concept, definable in terms of feature sharing, with features proper to cross-linguistic analysis as well as language-particular analysis. Truly typological comparison should obviously not be limited to two related languages and we brought in Odia to sketch the perspective of further comparison and we have seen how the new comparative concept drops at least most of its formal properties, as expected. It is important to stress that we do not claim that all languages will have similatives nor that similatives will have to be (significantly) different from adjectives or demonstratives, nor even that the identification of a comparative concept will always be a matter of isolating the properties that are common to descriptive categories along the simple model represented in Figure 2 – rather than the more complex model represented in Figure 3. Finally, we hope to have contributed to a better understanding of such words as English *such*.

References

- Altenberg, Bengt. 1994 'On the functions of *such* in spoken and written English', Nelleke Oostdijk & Pieter de Haan (eds.), *Corpus based research into language*. Amsterdam: Rodopi, 224-240.
- Biber, Douglas, Stig Johanson, Geoffrey Leech, Susan Conrad & Edward Finnegan 1999. *Longman grammar of spoken and written English*. London: Longman.
- Bolinger, Dwight. 1972. *Degree words*. The Hague: Mouton.
- Brown, Dunstan & Marina Chumakina. 2013. 'What there might and what there is: an introduction to Canonical Typology', Dunstan Brown, Marina Chumakina & Greville G. Corbett (eds.) *Canonical morphology and syntax*. Oxford: Oxford University Press, 1-19.
- Collins COBUILD English grammar*. (1990-1992). London: Collins.
- Croft, William. 2001. *Radical construction grammar*. Oxford: Oxford University Press.

- Declerck, Renaat. 1991. *A comprehensive descriptive grammar of English*. Tokyo: Kaitahusha.
- de Mönnink, Inge. 2000. *On the move: The mobility of constituents in the English noun phrase: a multi-method approach*. Amsterdam: Rodopi
- Ghesquière, Lobke & Freek Van de Velde. 2011. "A corpus-based account of the development of English *such* and Dutch *zulk*: Identification, intensification and (inter)subjectivity", *Cognitive Linguistics* 22: 765-797.
- Greenbaum, Sidney. 1996. *English grammar*. Oxford: Oxford University Press.
- Haspelmath, Martin. 2007. 'Pre-established categories don't exist-consequences for language description and typology', *Linguistic typology* 11: 119-132
- Haspelmath, Martin. 2010a. 'Comparative concepts and descriptive categories in cross-linguistic studies', *Language* 86: 663-687.
- Haspelmath, Martin. 2010b. 'Framework-free grammatical theory', Bernd Heine, Bernd and Heiko Narrog (eds.) *The Oxford handbook of grammatical analysis*. Oxford: Oxford University Press, 341-365.
- Haspelmath, Martin. 2011. 'The indeterminacy of word segmentation and the nature of morphology and syntax'. *Folia Linguistica* 45: 31-8
- Himmelmann, Nikolaus P. 2001. 'Articles', Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals. Volume 1*. Berlin: De Gruyter, 831-841.
- Himmelmann, Niklaus & Eva Schultze-Berndt. eds. 2005. *Secondary predication and adverbial modification: The typology of depictives*. Oxford: Oxford University Press.
- Lazard, Gilbert. 2005. 'What are we typologists doing"?", Zygmunt Frayzyngier, Adam Hodges and David S. Rood (eds.), *Linguistic diversity and language theories*. Amsterdam: Benjamins. 1-23.
- Mackenzie, J. Lachlan. 1997. 'Grammar, discourse and knowledge: the use of *such*', Jan Aarts, Inge de Mönnick and Herman Wekker (eds.), *Studies in English language and teaching: In honour of Flor Aarts*. Amsterdam: Rodopi, 85-105.
- Poutsma, H. 1916. *A grammar of Late Modern English for the use of continental, especially Dutch, students. Part II. The parts of speech. Section 1, B. Pronouns and numerals*. Groningen: Noordhoff
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.

- Sahoo, Kalyanamalini. 1996. *The DP-analysis of Oriya and English Noun Phrases*. M. Phil. dissertation. Hyderabad: Central Institute of English and Foreign Languages.
- Sahoo, Kalyanamalini. 1999. 'The syntax of definiteness in Oriya', In Rajendra Singh (eds.) *The yearbook of South Asian languages and linguistics*. London: Sage Publications, 101-117.
- Umbach, Carla & Helmar Gust. 2014. 'Similarity demonstratives', *Lingua* 149: 74-93.
- Van de Velde, Freek. 2009. *De nominale constituent: structuur en geschiedenis*. Leuven: Leuven University Press.
- van der Auwera, Johan & Volker Gast. 2011. 'Categories', Jae Jung Song (ed.) *The Oxford handbook of linguistic typology*. Oxford: Oxford University Press, 165-189.
- van der Auwera, Johan & Evie Coussé. 2014. '*such* and *sådan* – the same but different', Symposium on Parallel Corpora, Lund.
- van der Auwera, Johan & Alfonso Zamorano Aguilar (in print) 'The history of modality and mood', Jan Nuyts & Johan van der Auwera (eds.) *The Oxford handbook of modality and mood*. Oxford: Oxford University Press.
- Van Olmen, Daniel & Johan van der Auwera. 2014. 'Over *zo'n* en *zo meer*', Freek Van de Velde, Hans Smessaert, Frank Van Eynde & Sara Verbrugge (eds.) *Patroon en argument. Een dubbelfeestbundel bij het emeritaat van William Van Belle en Joop van der Horst*. Leuven: Leuven University Press, 215-228.
- Wood, Johanna L. 2002. 'Much about *such*', *Studia Linguistica* 56: 91-117.