

A diachronic corpus study of prenominal *zo 'n* 'so a' in Dutch

Pathways and (inter)subjectification\*

Daniël Van Olmen

Lancaster University

Like its English counterpart *such*, Dutch *zo 'n* has identifying and intensifying uses. The established pathway from the former to the latter is found to constitute a proportional rather than a discrete shift here. The strong presence of intensifying uses from the start, as compared to the older Dutch marker *zulk*, is argued to be due to preexisting constructions that are alike formally and convey intensification. *Zo 'n* is also found to have a recognitional and an approximating use. The case is made that the former has evolved out of the identifying use and that the latter is a development which is independent from the other uses functionally but has modeled itself on them formally. Finally, it is argued that the semantic shift from identification to intensification is best captured by the well-known pathway from textual to expressive, although the unidirectionality of this cline is uncertain, and that the change from identification to recognition supports a recent proposal to distinguish immediate and extended intersubjectivity.

## 1. Introduction

The literature on prenominal *zo 'n* 'so a', with the meaning 'such', in Dutch is much more limited than that on its equivalent *zulk* 'such' and their West Germanic cognates (e.g. Demske 2005 on German *solch*; Ghesquière & Van de Velde 2011 on *such* and *zulk*; Hole & Klumpp 2000 on German *so 'n*). There are several reasons why it deserves to be studied in more detail, though.

The multifunctionality that the other markers have been shown to exhibit is also found in *zo 'n*. In (1a), the marker has what Bolinger (1972: 16 on *such*) calls an identifying use. It signals the identity

---

\* An earlier version of this paper, which builds on ideas that were first discussed in Van Olmen & van de Auwera (2014), was presented at the 48th Annual Meeting of the Societas Linguistica Europaea in Leiden in September 2015. Thanks are due to Vittorio Tantucci, Nick Taylor, Johan van der Auwera, the reviewers and the editors of *Functions of Language*.

of an indefinite token to a definite type for which a definite token serves as an example (see Ghesquière & Van de Velde 2011: 772): the speaker covets a coat of the kind instantiated by the specific one that she has just said is beautiful. In (1b), *zo 'n* has what Bolinger (1972: 16) describes as an intensifying use. It involves “the measuring or heightening of the degree of a certain evaluative or scalar notion inherent in the modified element” (Ghesquière & Van de Velde 2011: 782): *zo 'n* can be argued to raise the level of adorableness of the puppy above some presumed standard here.

- (1) a. Wat een chique outfit, erg mooi! De jas vind ik echt prachtig, *zo 'n* jas stond nou net zelf ook op mijn wishlist.

‘What a fancy outfit, very nice! I think the coat is really beautiful, such a coat just happened to be on my wish list too.’

(<http://mounirasmansion.com/herfst-look>, accessed in October 2016)

- b. Sinds drie weken heb ik een puppy. Het is *zo 'n* schatje!

‘I have had a puppy for three weeks. It is such a darling!’

(Van Olmen & van der Auwera 2014: 216)

Bolinger (1972: 61) hypothesizes that this multifunctionality, which is attested in the noun constructions *sort of* and *kind of* (see Brems 2011) and many other markers too, results from “a kind of wholesale migration” from identification to intensification. For markers with the meaning ‘such’, this claim has so far been subjected to little rigorous empirical scrutiny, as Ghesquière & Van de Velde (2011: 767) note. In their own corpus study of the evolution of *such* and *zulk*, (ambiguous) intensifying uses occur from the earliest sources onward. Moreover, overall, they observe a proportional increase of such cases rather than a complete shift to intensification (see Ghesquière & Van de Velde 2011: 787–791). One of the questions that we aim to address in the present article is which scenario of change is supported by an analysis of diachronic corpus data.

As has been pointed out recently, *zo'n* has two additional uses, which it does not share with *such* and *zulk*. Van Olmen & van der Auwera's (2014) examination of a corpus of spontaneous conversations shows that it sometimes functions as an approximator when it precedes a cardinal number.<sup>1</sup> Sentence (2a), where *zo'n* can be paraphrased as 'roughly', is a case in point. They also reveal that it frequently serves to indicate that something is assumed to be known by the addressee. In this use, the speaker employs *zo'n* to evoke in the interlocutors the image of a prototypical instance of what follows or, put differently, an instance that all are familiar with. In (2b), for example, the context contains no definite token through which the indefinite token *colbertjas* 'jacket' can be identified. The noun's lack of any clear evaluative or scalar meaning makes an intensification reading implausible too. Instead, *zo'n* suggests a type of jacket that everyone can picture in their minds.

- (2) a. *Zo'n* twintig kilometer fietsen daar efkens een terrasje doen en dan terugkomen.

'Cycle for roughly twenty kilometers have a quick drink there and then return.'

(Van Olmen & van der Auwera 2014: 219)

- b. Ik was toen zwanger van wie dat weet ik niet meer ... Nou weet ik het ik had *zo'n*

*colbertjas* aan en ik had een dik buik nou weet ik 't.

'I was pregnant then with I do not remember who ... Now I know I was wearing one of those jackets and I had a big belly now I know.'

(Van Olmen & van der Auwera 2014: 219)

---

<sup>1</sup> This data comes from the *Corpus Gesproken Nederlands* (Nederlandse Taalunie 2004) and contains Northern and Southern Dutch. It confirms one of the reviewers' intuition that there are differences between the two varieties. They do not lie in the types of use of *zo'n*, however, but in its overall frequency and the proportions that its various uses account for. For more information, see Van Olmen & van der Auwera (2014: 217–220).

As the translation of (2b) implies, this use is reminiscent of the so-called recognitional function that Himmelmann (2001: 833) distinguishes for demonstratives and articles originating from them: “[It] is characterised by the fact that the intended referent has to be identified via specific, but presumably shared, knowledge. It can always be (and in fact often is) accompanied by a *you know?* or *remember?*-type of tag question.” The same term will be used here to refer to the use of *zo ’n* in (2b). Interestingly, German *so ’n* and Afrikaans *so ’n* also exhibit the recognitional function (see Van Olmen & van der Auwera 2014: 221–223). So it is quite surprising that it has gone largely unnoticed in the literature on West Germanic. It has been described for Norwegian and Swedish, though, from a synchronic perspective (e.g. Ekberg 2011; van der Auwera & Coussé 2016). In our study, we seek to answer the question how the recognitional use, as well as the approximating one, ties in with the identifying and intensifying uses diachronically.

Another reason why *zo ’n* is of interest is its relation to better-studied *zulk*. The latter is already found in Old Dutch. According to Van der Horst (2008: 167, 1102, 1382, 1667), the former emerged only in the 17th century as the co-occurrence of the adverb *zo* ‘so’ and the indefinite article *een* ‘a(n)’, started fusing into one word in the 18th and became frequent in the 19th century. The timing of these events will be analyzed in the present article. Still, it is undeniable that *zo ’n* has encroached on *zulk*. The older marker used to be able to combine with any type of noun. In Present-day Northern Dutch, however, *zo ’n* is employed for singular count nouns as well as for intensifying mass nouns and *zulk* just for identifying mass nouns and for plural nouns, as in (3a) from left to right. In Present-day (non-standard) Southern Dutch, *zo ’n* can even be used for identifying mass nouns and for plural nouns, as in (3b) (see Haeseryn *et al.* 1997: Section 5.6.6).

- (3) a. *zo ’n* jas — *zo ’n* pijn! — *zulk* goud — *zulke* jassen  
           ‘such a coat’ — ‘such pain!’ — ‘such gold’ — ‘such coats’

b. *zo 'n* jas — *zo 'n* pijn! — *zo 'n* goud — *zo 'n* jassen

‘such a coat’ — ‘such pain!’ — ‘such gold’ — ‘such coats’

This paradigmatic relation between the two markers raises the question whether and how the history of the one interacts with that of the other. Our diachronic corpus study of *zo 'n* tries to shed some light on the matter.

A last issue to which an investigation into *zo 'n* can contribute concerns (inter)subjectification. In the research on *such* and *zulk*, there is a debate about what the assumed shift from identifying to intensifying means for this process of semantic change (e.g. Ghesquière & Van de Velde 2011: 791–793). Traugott (2003: 125, 128) posits a unidirectional tendency from non-subjective via subjective to intersubjective meaning. But *such* and *zulk* constitute a counterexample if one agrees with Ghesquière (2009: 317) that creating joint attention is intrinsically addressee-oriented and thus intersubjective. The intensifying use would still be considered subjective: it conveys the speaker’s estimation of the degree of some evaluative or scalar component in the subsequent noun phrase. The identifying one, however, can be said to manage the addressee’s understanding of the noun phrase: it directs the addressee’s attention to a token in the discourse to which the indefinite token is similar. It would be intersubjective, in other words. In this view, *such* and *zulk* involve a shift from intersubjective to subjective. The solution that Ghesquière & Van de Velde (2011: 791) suggest is to return to Traugott’s (1982: 256) earlier hypothesis that semantic change in grammaticalization goes from propositional via textual to expressive. The present article will revisit this debate in order to answer the question where the recognitional use of *zo 'n* fits in. In the discussion, we will examine the usefulness of Tantucci’s (2015) recent proposal to make a distinction between immediate and extended intersubjectivity for *zo 'n*.

The rest of this article is structured as follows. In Section 2, we will introduce the data. Section 3 will focus on the more formal aspects of the development of *zo 'n* and Section 4 on the functional

changes that it has undergone. In Section 5, we will come back to (inter)subjectification and Section 6, finally, will be the conclusion.

## 2. Data

A diachronic corpus study of *zo 'n* requires data that meets a number of criteria. First, as we want to check whether the marker is indeed a 17th-century innovation, the corpus should range from (at least) the 1500s to the present. Second, since the fusion of *zo een* into *zo 'n* and the recognitional use can be argued to have their origins in oral language use, we need historical data that is an approximation of spoken language. Following Culpeper & Kytö (2000) among others, we assume that plays, as well as fictional prose and in particular its dialogue, satisfy this requirement best, despite obvious differences such as fewer hesitations and unfinished clauses. Third, the corpus should be large enough and the authors of the texts varied enough to be able to obtain a sufficient quantity of examples that is as representative of a whole period as possible.

However, to our knowledge, no existing diachronic corpus of Dutch fulfills all these conditions. For that reason, we compiled our own, from the texts in the *Digitale Bibliotheek voor de Nederlandse Letteren* (the Koninklijke Bibliotheek's 2015 digital library of Dutch literature). All available drama and fictional prose from the 16th to the 20th century was downloaded, but translations into Dutch and texts by non-native speakers were excluded from the corpus. Of the remaining plays and works of fictional prose, we included just one text per author, the selection of which was determined mainly by the relative need for more data in one of the two genres and the presumed more colloquial nature of the language. In Table 1, we give an overview of the period sizes of our corpus.

**Table 1.** Number of words in the diachronic corpus of Dutch

	16th century	17th century	18th century	19th century	20th century
Drama	239,932	858,632	269,677	226,917	232,133
Fictional prose	1,947,614	704,518	678,717	2,279,896	2,328,004
Total	2,187,546	1,563,150	948,394	2,506,813	2,560,137

The unevenness in the corpus is mostly due to the limited amount of readily accessible electronic data and our decision to accept only one text per writer. Any attempt at, for instance, keeping the ratio of drama to fictional prose data constant over time would have led to a much smaller corpus, with an insufficient number of hits for *zo 'n* (see, among others, Smitterberg 2016 on the almost unavoidable compromises that historical corpus linguists need to make). Moreover, the varying difference in size between the two components reflects the changing popularity of the genres: the 17th century was the heyday of playwriting in the Low Countries while 1801 to 2000 was much more the time of the novel. The imbalances will, of course, be taken into account as much as possible in the rest of this article.

From the corpus, we extracted all hits for *zo 'n*, the collocation *zo een* and their spelling variants, using WordSmith 6.0 (see Scott 2012). Non-prenominal uses such as fused *zô 'ntje* ‘a little one like that’ in (4a), which are intriguing but beyond the scope of the present study, were not searched for. Combinations of *zo* with the numeral *een* ‘one’ (pronounced as [e:n], unlike the indefinite article [ən]) like (4b), which cannot normally be fused, and other false positives were filtered out manually. So were cases of *zo en zo een* ‘so and so a’ like (4c).<sup>2</sup> As Ghesquière (2012: 531) argues for *such and such*, this construction calls up some abstract type that is not linked to the context in any specific way but only exists in the mental space set up at the time of speaking. *Soo en soo een wijs* in (4c) could be paraphrased as ‘like this and like that’ and refers to any conceivable fashion of preparing mutton.

---

<sup>2</sup> As suggested by one of the reviewers, who rightly points out that it constitutes a different construction.

Fused *zo 'n* cannot be used in this manner (see van der Auwera & Sahoo 2015: 163).

- (4) a. *Zô'ntje* hew ik ok.

‘I have got such a little one.’

(1984, WNT s.v. *zoo*)

- b. *zo een* als ik er ook draag

‘one of the type that I wear too’

(1955, Han B. Aalberse, *De Liefde van Bob en Daphne*)

- c. ‘s Anderdaeghs komt de Boer met een goedt Schaep by den Priester, die het hem belast, op *soo en so een* wijs toe te bereyden maer hy brengt ‘t in sijn Schaep-stal.  
‘The next day, the farmer arrives with a good sheep at the priest’s, who tells him to prepare it in such and such a way but the farmer takes it to his sheep barn.’

(1659, Jan Pietersz. Meerhuysen, *De Geest van Jan Tamboer*)

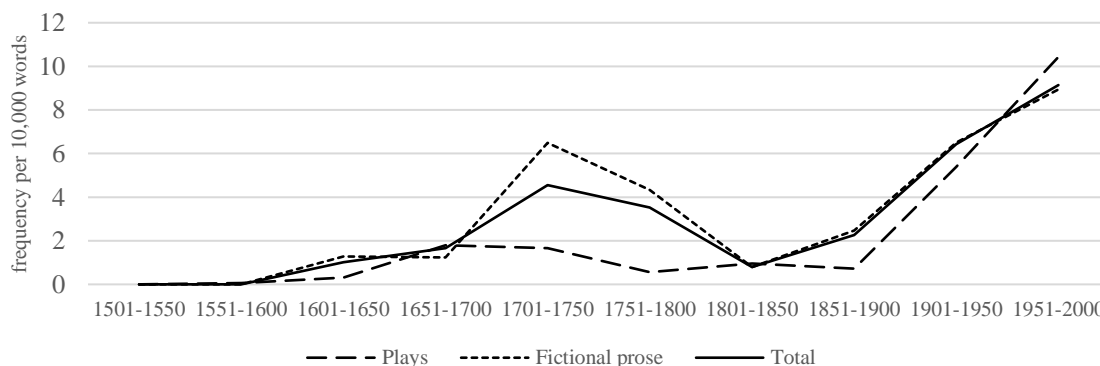
The formal analysis in Section 3 will draw on all the relevant attestations. For the functional analysis in Section 4, however, we can only consider a selection of the examples. Our intention was to obtain, for each century, a random sample of a hundred instances of *zo 'n* from drama and another hundred from fictional prose and a similar sample for *zo een*. But this objective was met for just one of the two variants in just one of the five centuries, i.e. the fused form in the 1900s. In the other centuries, *zo 'n* and/or *zo een* were too infrequent in the drama and/or the fictional prose texts. If it was possible to get an overall sample of two hundred instances of a variant, the limited number of hits in one genre was supplemented with a random sample of hits from the other genre. If not, all attestations were included. Ideally, the way in which our selection of hits was collected would be entirely identical in all centuries. The rationale behind the present approach is two-fold: it provides us with samples that are, on the one hand, sizeable enough to reveal significant trends while still being manageable and,



on the other, inclusive enough of the less common occurrence of one of the forms in one or both genres to bring to light possible differences.

### 3. Formal changes

Let us first look at the rise to prominence of *zo'n*.<sup>3</sup> Figure 1 presents its rates of occurrence per ten thousand words in the drama and prose components and in total for each half-century. For the sake of completeness, the raw numbers of hits for *zo'n* are given in Table 2. The main reason why the periods from Table 1 are split in two here is that the most recent changes are much clearer this way.



**Figure 1.** Relative frequency of *zo'n* from the 16th to the 20th century

**Table 2.** Absolute frequency of *zo'n* from the 16th to the 20th century

1501–	1551–	1601–	1651–	1701–	1751–	1801–	1851–	1901–	1951–
1550	1600	1650	1700	1750	1800	1850	1900	1951	2000

<sup>3</sup> In the rest of this article, *zo'n* will be used as a short-hand for both the original form and the fused form, unless specified otherwise.

Drama	0	1	6	119	24	7	7	11	59	128
Fictional prose	0	1	66	23	141	200	84	301	1000	707
Total	0	2	72	142	165	207	91	312	1059	835

Our corpus indicates that, pace Van der Horst (2008: 1102), the marker emerged not in the 17th but in the 16th century: the second half of the 1500s contains two attestations of *zo een*, of which one is given in (5).

- (5) De mensch leeft niet by tbroot alleen maer by elck woort wt Godts mont reen. Siet, dats der sielen spijsse en leeft in *soo een* wijze.

‘Man does not live by bread alone but by every word from God’s mouth. Look, that is the food of the soul and live in such a way.’

(1562, anonymous, *Het Offer des Heeren*)

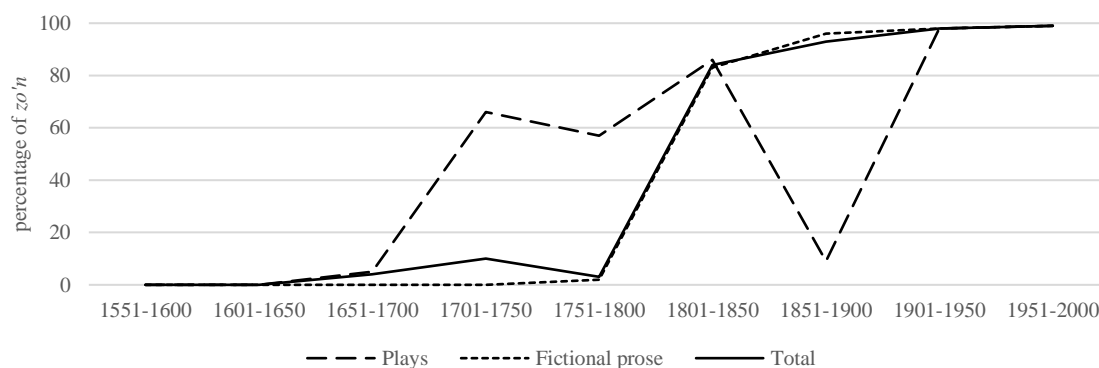
Figure 1 suggests that *zo 'n* then became reasonably common already in the 1700s but took a dip in the 1800s. The data is somewhat skewed, however. In the first/second half of 18th century, the marker has a relative rate of occurrence that is about four/eight times higher in fictional prose than in drama (6.49/4.33 versus 1.66/0.56 instances per ten thousand words,  $G^2 = 51.04/56.40$  and  $p < 0.0001$  for Rayson & Garside’s 2000 log-likelihood test for comparing frequencies across corpora).<sup>4</sup> Moreover,

---

<sup>4</sup> We also see a significant difference between fictional prose and drama in the first half of the 17th century (1.28 versus 0.31 cases per ten thousand words,  $G^2 = 16.06$  and  $p < 0.0001$ ) as well as in the second half of the 19th century (2.47 versus 0.72 cases per ten thousand words,  $G^2 = 24.34$  and  $p < 0.0001$ ). Note that all data in the present study has been analyzed statistically with SPSS (IBM Corp. 2013).

this difference can be put down to just two texts, which make up one third of the 18th-century fictional prose but account for nearly four-fifths of the hits. At any rate, Gries’s (2013: 375–379) correlation test reveals a high positive and significant link between time and relative frequency for the two components separately (in each case,  $\tau = 0.64$  and  $p < 0.01$  for Kendall’s tau) and together ( $\tau = 0.69$  and  $p < 0.01$ ): the higher or, put differently, the more recent the value for time, the higher the rate of occurrence per ten thousand words of the marker. This increase in frequency appears to take off particularly in the 1900s in Figure 1 (see also Ghesquière & Van de Velde 2011: 790 on the period in which *zo ’n* comes to be preferred to *zulk* in combination with adjectives).

An important change in the diachrony of *zo ’n* is the fusion of *zo* and *een* into one word. Figure 2 presents the proportions of the fused form in all of the attestations.



**Figure 2.** Fusion of *zo ’n* from the 16th to the 20th century

The corpus shows that fusion started earlier than claimed by Van der Horst (2008: 1382): the first examples of *zo ’n*, one of which is (6), occur not in the 18th century but in the plays from the second half of the 17th century.

- (6) Goôn is ‘t mooghlijk, dat een Maeght / *Zoo ’n* herte draeght!

‘Gods is it possible that a virgin / carries such a heart!’

(1668–1669, Johan Beets, *Dichtkunst van Verscheide Stoffen*)

Unsurprisingly, the fused form is found to gain ground on the non-fused one in Figure 2, with a statistically significant correlation between time and the type of form in the fictional prose part and in total ( $\tau = 0.75$  and  $p < 0.05$ ).<sup>5</sup> What may be worth mentioning is that the first meaningful increase of *zo ’n* overall, i.e. between the second half of the 1700s and the first half of the 1800s ( $p < 0.0001$  for Fisher’s exact test), goes from a mere 3.38% to 83.52% of the cases. We cannot readily account for this dramatic shift, though. Possibly, it became acceptable in the 19th century to write down in prose a fused form of the marker that, as the higher proportions in the 18th-century ( $p < 0.0001$  for the difference between the genres in both halves) suggest, was already quite common in speech. The drop in the drama of the second half of the 19th century ( $p < 0.01$  and  $p < 0.0001$  as compared to,

---

<sup>5</sup> In this article, we use Kendall’s tau-b, a non-parametric test that measures the association between two ordinal variables and makes adjustments for ties, to see whether time, for which successive periods are considered an ordinal scale, correlates with a binary variable of the formal or functional history of the marker (e.g. *zo een* versus *zo ’n*), the values for which can be regarded as ordered too (see also Ghesquière & Van de Velde 2011: 770). The coefficient itself ranges from 1 (perfect positive association) to 0 (no association) and -1 (perfect negative association) and SPSS provides a p-value to assess the significance of the correlation. In essence, the test checks if lower/less recent time values prefer the one formal or functional value (e.g. *zo een*) and higher/more recent ones the other (e.g. *zo ’n*) and, as such, it captures what is happening diachronically in a general way. To compare fictional prose to drama or one period to another, we use Fisher’s exact test because our corpus often produces low absolute numbers.

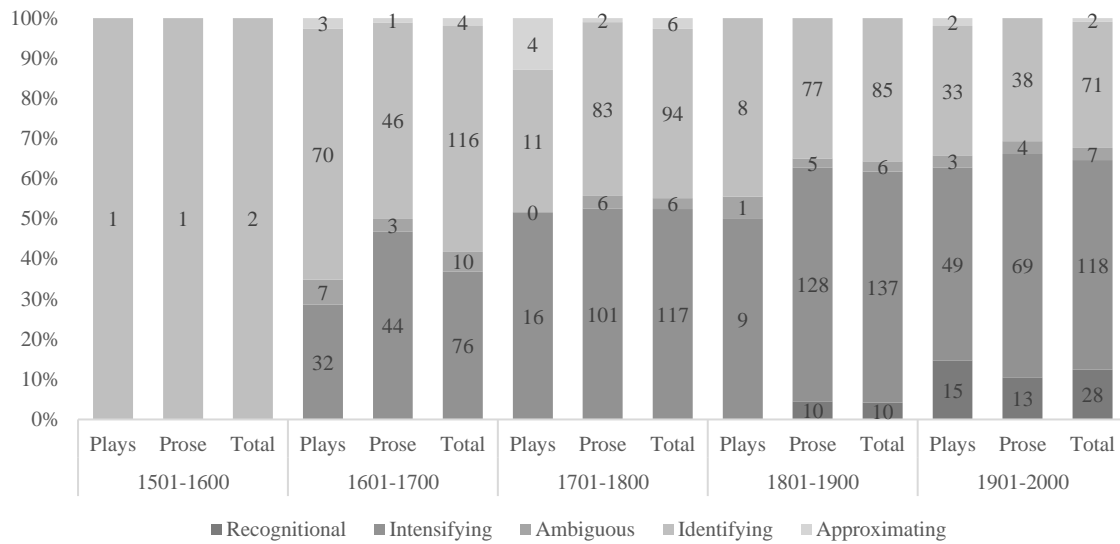
respectively, the first half and fictional prose) is probably just a quirk of the data.<sup>6</sup>

#### 4. Functional changes

An overview of our functional analysis of a sample of hits of *zo 'n* is presented in Figure 3, which provides the raw numbers and the proportions of its different uses. The rightmost bar, for example, is to be read as follows: the 20th-century data contains 28 recognitional cases, 118 intensifying ones, seven instances that are ambiguous between intensification and identification, 71 identifying cases and 2 approximating ones.

---

<sup>6</sup> The division of labor between *zo 'n* and *zulk* in Present-day Dutch (see Section 1) raises the question whether *zo 'n* has undergone an expansion of the types of noun that it can combine with. For reasons of space, we cannot go into detail here. Let the following observations suffice. In our sample, mass nouns occur from the start and no statistically significant differences are found between 1600 and 2000. Furthermore, as all mass nouns in our data are intensified or ambiguous between intensification and identification, nothing can be said about a possible expansion from intensifying to identifying mass noun uses either. In other words, it seems that *zo 'n* has actually been pushing out *zulk* from contexts in which it could be employed from the beginning. There is one exception, though. Only from the 1800s onward (see Van der Horst 2008: 1668) do we attest the occasional example of fused *zo 'n* identifying or intensifying plural nouns. Originating from an adverb and the indefinite article *een*, it was initially restricted to singular nouns. As its formal association with *een* waned through fusion in the 18th and the 19th century (see Figure 2), however, it must have become easier to extend its existing competition with *zulk* in singular count and mass nouns to plural ones.



**Figure 3.** Functions of *zo'n* from the 16th to the 20th century

In Section 4.1, we will focus on *zo'n*'s identifying use and, in Section 4.2, on its intensifying use in relation to identification. Section 4.3 will address the recognitional use and, in Section 4.4, finally, we will deal with the approximating use.

#### 4.1 Identifying use

This function has been characterized as signaling the identity of an indefinite token to a definite type for which another token serves as an example (see Section 1). The main feature used for the classification of the attestations in our sample is the existence of a phoric relation to another referent in the context. Ghesquière & Van de Velde (2011: 774–776) make a distinction between anaphoric and cataphoric cases. In the former, the ‘other referent’ occurs in the preceding discourse. In (7), for instance, the kind of wife that the speaker would not be able to deal with is exemplified by the addressee’s daughter, who has been mentioned earlier and does not know Dutch.

- (7) Indien dat jou dochter geen hollands ken ... zo zou ik se tot myn Vrouw niet begeeren.  
Wat zou ik met *zo een* Vrouw doen?  
'If your daughter does not know Dutch ... I would not want her as my wife. What would I do with such a wife?'  
(1685, Pieter de la Croix, *De Meid Juffrouw*)

In the latter, the 'other referent' occurs in the ensuing discourse. The 'as'-phrase in (8a) and the resultative 'that'-clause in (8b), for instance, identify the type to which the indefinite token belongs. Together with *zo 'n*, they form a noun-phrase-internal, though discontinuous, unit that specifies the head noun (see Ghesquière & Van de Velde 2011: 776).

- (8) a. Of *soo een* man als ick, voldoen kan u begeert.  
'Whether such a man as I, can satisfy your desire.'  
(1678, Aernout van Overbeke, *Geestige Wercken*)
- b. Nou was me vader van *zoo 'n* natuur, dat as ie, met permissie, maar aan 't zweeten  
kommen kon, dan was ie weer klaar.  
'My father was of such a nature that if he, if I may say so, managed to break some  
sweat, he was done again.'  
(1839, Nicholas Beets, *Camera Obscura*)

The distinction between anaphoric and cataphoric uses does not suffice to describe our data, however. Particularly in the drama component, we find the occasional case of exophoric reference. In (9), for instance, no exemplary token for the type evoked by *zo 'n* is present in either the preceding or the following discourse.

(9) N: Ik kwam even langs kantoor omdat ik iets was vergeten.

‘I was just dropping in at the office because I had forgotten something.’

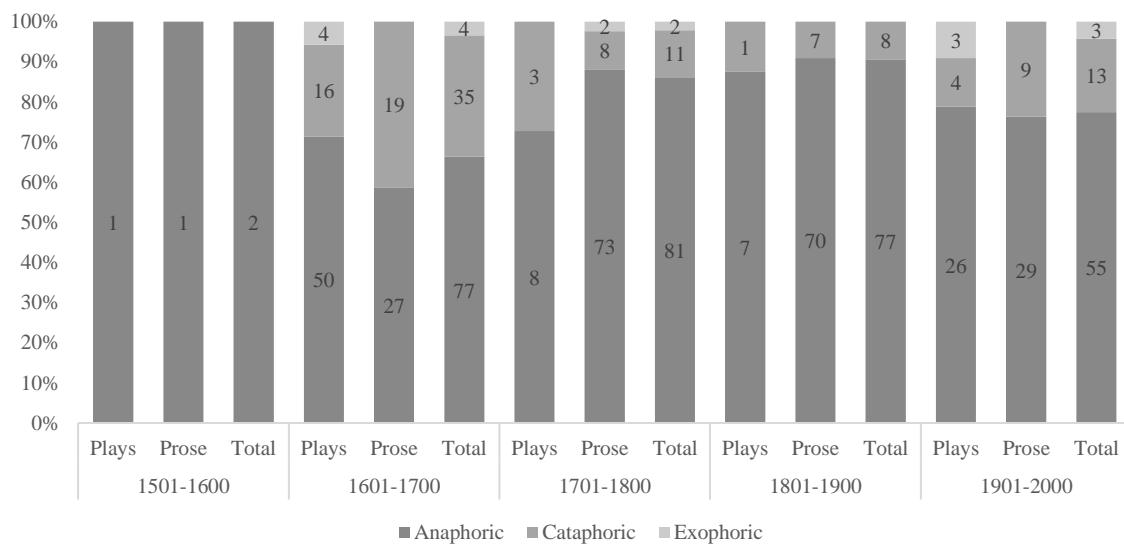
B: In *zo ’n* avondjurk?

‘In such an evening dress?’

(1989, Annie M.G. Schmidt, *En ik dan?*)

The only plausible interpretation here is that *zo ’n* establishes a relation with something extralinguistic, i.e. ‘an evening dress like the one that N is wearing at the moment’ (see Section 4.4 for the possible link between this exophoric use of *zo ’n* and its recognitional use).

Figure 4 gives, for the sake of completeness, the raw numbers and the proportions of the three types of reference of identifying *zo ’n*.



**Figure 4.** Types of reference of identifying *zo ’n* from the 16th to the 20th century

Unlike *zulk*, for which Ghesquière & Van de Velde (2011: 774) observe “a rather dramatic decrease in the originally predominant cataphoric uses”, *zo ’n* has always preferred its anaphoric use, as Figure



4 shows. The proportion of cataphoric cases does appear to decline, i.e. from 30.17% in the 1600s to 11.70% and 9.41% in the 1700s and 1800s respectively but up again to 18.13% in the 1900s. Yet, there is no statistically significant correlation between time and type of reference or difference between the 17th and the 20th century. The relatively marginal role of cataphoric reference in the history of *zo'n* may be connected to Ghesquière & Van de Velde's (2011: 775) observation that, by the 1500s, when *zo'n* emerged, cataphoric *zulk* had already dropped to approximately one seventh of the cases (but see Section 4.2 on *zulk* as an analogical model for *zo'n* in terms of functions). It may also be linked to the following argument by Ghesquière & Van de Velde (2011: 780): "Submodification by means of a complement clause is typical of adjectives, and the loss of the ability of *such/zulk* to take their own (sub)modifiers ties in with the idea that these elements gradually acquire determiner characteristics over time." If discontinuous constructions such as (9) are indeed more a trait of adjectives, it is perhaps not so surprising that the cataphoric use was never very prevalent in *zo'n*. Unlike *zulk*, it comes not from an adjective but from an adverb and a determiner, i.e. the indefinite article *een*.

#### 4.2 Intensifying use

The main features used for classifying the examples in our sample are the existence of some gradable notion in the noun phrase introduced by *zo'n* and the presence of other indications of subjectivity in the context (see Section 1). This gradable notion can be expressed by an adjective, as in (10a), but also by a noun on its own, as in (10b).

- (10) a. En eer ik wat zeggen kon, daar tastte menheer Van Zuchter na *zoo'n* groote tafelbel; ik weet niet dat ik ooit *zoo'n* tafelbel meer gezien heb.

‘And before I could say anything, Mr Van Zuchter reached for such a big table bell; I

do not think that I have ever seen such a table bell since.’

(1839, Nicholas Beets, *Camera Obscura*)

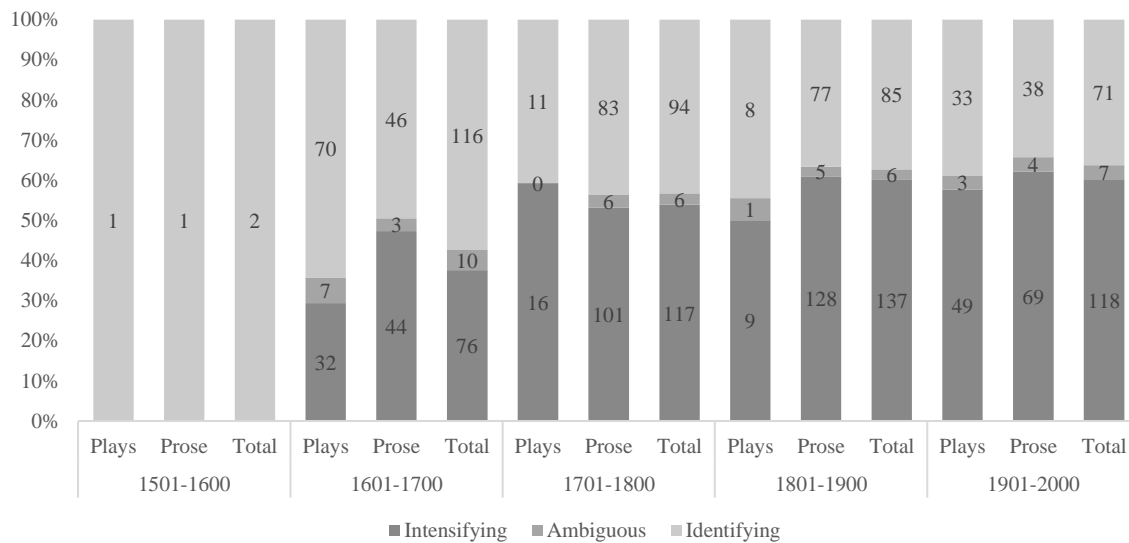
b. Myn Vriendinne Styntje denkt net als ik, en dat doet my *zo een* deugd!

‘My female friend Styntje thinks just like me, and that gives me such pleasure!’

(1782, Betje Wolff & Aagje Deken, *Historie van Mejuffrouw Sara Burgerhart*)

In the former, *zo* ‘n can be argued to boost the size of the bell: ‘a very big table bell’. In the latter, *zo een* affects the degree of enjoyment: ‘so much pleasure’.

As mentioned in Section 1, intensification has been linked to identification diachronically in the literature. To examine this relation in more detail, we repeat Figure 3 here without the recognitional and approximating uses.



**Figure 5.** Identifying and intensifying functions of *zo 'n* from the 16th to the 20th century

As Figure 5 suggests, identification is sometimes ambiguous with intensification. *Zoo een* in (11) is a case in point. It allows an identifying reading: the moderation that the speaker hopes to find in all

soldiers is like the temperance that the aforementioned knight is said to have. It simultaneously evokes an intensifying interpretation, however. As moderation is a gradable notion, *zoo een tempering* can also be understood as ‘this much moderation’. This reading is compatible with the context: the speaker obviously wants soldiers to exhibit this quality to a great extent.

- (11) Een Fransch historischrijver geeft ons de schildery van zoodanigen oorloghsman, in de persoon van een Ridder, die hy aldus beschrijft: Il avoit une valeur jointe à tant de mansuetude, que s’il attiroit par celle-la le respect de tous les hommes, il gaignoit par celle-cy les coeurs de toutes les Dames. Van *zoo een tempering* wensch ik de ziel aller soldaten.

‘A French historian gives us a picture of such a warrior, in the person of a knight, whom he describes as follows: He had virtue combined with such moderation that if he gained the respect of all men with the former, he won the hearts of all women with the latter. Of such moderation I wish the soul of all soldiers to be.’

(1644, Johan de Brune, *Wetsteen der Vernuftten*)

The attestations in our data of the type in (11) are of interest in two regards. First, their infrequency (i.e. only 4.95%, 2.76%, 2.63% and 3.57% of the cases in the 17th to the 20th century respectively) confirms Ghesquière’s (2012: 538) corpus finding for *such*: in contrast to what is claimed for example by Spinillo (2003: 206), *such* is not often vague between identification and intensification in Present-day English. The same seems to hold for its equivalents in different periods and languages. Second, the attestations like (11) resemble the context that Ghesquière & Van de Velde (2011: 787–788) identify as the bridge from the one function to the other for *zulk* and *such*. The intensifying use tends to be associated with noun phrases with adjectives, such as (10a). But this meaning first appears in unmodified noun phrases with gradable head nouns, like (11). In this context, the identifying use

nearly automatically invites the inference of intensification. As Ghesquière & Van de Velde (2011: 788) argue for *such*:

An NP like *such a problem*, for instance, can be reasonably assumed to have originally functioned as a phoric NP referring to a specific kind of problem described or illustrated in the context. The reason for picking up the NP in the ensuing discourse will often be that we are dealing with a kind of problem that is in some way ‘remarkable’, e.g., by means of its importance impact. Over time, this might have led to the inference that the use of *such* entails intensification.

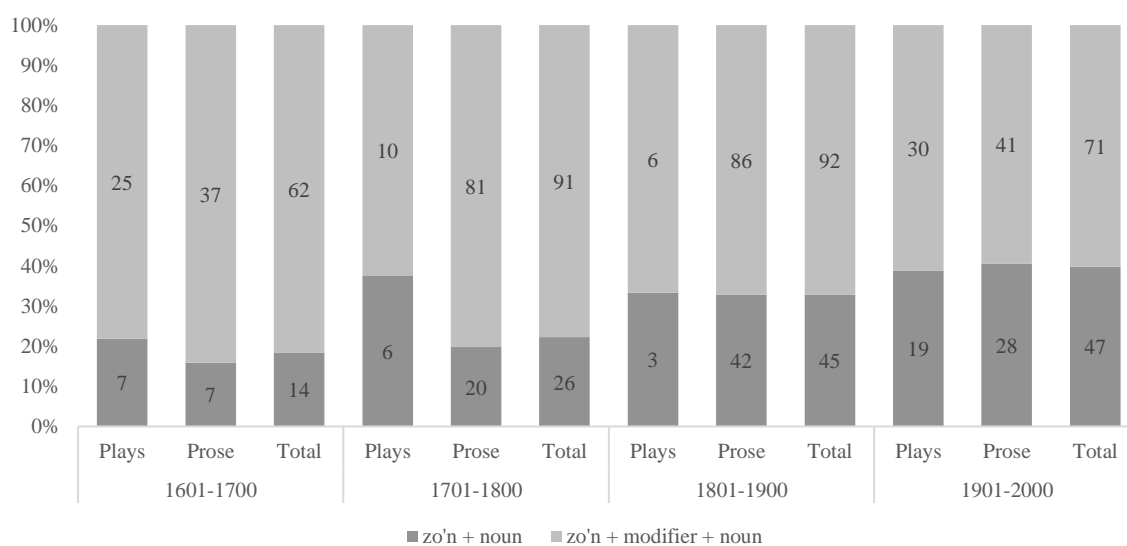
This invited inference can then conventionalize and it goes well with modified noun phrases later on.

As far as the shift from identification to intensification in *zo'n* is concerned, Figure 5 reveals a gradual increase in the proportion of unambiguous intensifying uses versus all uses with an identifying reading: from (0% in the 1500s and) 37.62% in the 1600s to 53.92% in the 1700s and 60.09% and 60.20% in the 1800s and 1900s respectively, with a statistically significant correlation between time and the type of use ( $\tau = 0.67$  and  $p < 0.05$ ). What is essential to note here is that, ignoring the two attestations in the 16th century, we find ambiguous and intensifying cases from the start. Ghesquière & Van de Velde (2011: 788–789) have not dissimilar results for *such* and *zulk*. Hence, they argue that the semantic shift is not a typical, more or less discrete change from a first stage where a marker has only one meaning to a second stage where it has its initial meaning as well as the new one. Their proposal is to regard the shift as a change in proportions instead. They suggest that intensification is an inference which is always possible with an identifying marker introducing a gradable noun, as in (11), and which may semanticize and become the predominant meaning over time.

This scenario is convincing for *such* and *zulk*, which have few or no unambiguously

intensifying uses when they are first attested (see Ghesquière & Van de Velde 2011: 786–787). It is not unproblematic for *zo'n*, though: it appears to be unambiguously intensifying in more than one third of the cases almost from the beginning. One potential explanation has to do with the fact that *zulk* had already undergone proportional change by the time the new marker arrived: did *zo'n* rapidly take over the functional distribution that its existing competitor displayed? When one considers Ghesquière & Van de Velde's (2011: 787) finding that, in their 1501–1700 period, *zulk* was used as a pure intensifier in 38.42% of the cases, one is tempted to accept this hypothesized analogy of the one marker with the other. This proportion closely resembles that of *zo'n* in the 17th century.

A closer examination of the structures in which intensifying *zo'n* occurs makes one wonder, however, whether analogy with *zulk* it is the main/real reason for *zo'n*'s remarkably high percentage of unambiguous intensifying uses from the beginning. Figure 6 presents the absolute numbers and the proportions of intensifying *zo'n* + noun versus intensifying *zo'n* + modifier + noun.



**Figure 6.** Modification of intensifying *zo'n* from the 17th to the 20th century

Intensifying *zulk* exhibits an increase of modified instances from 0% in Early Middle Dutch to 92%

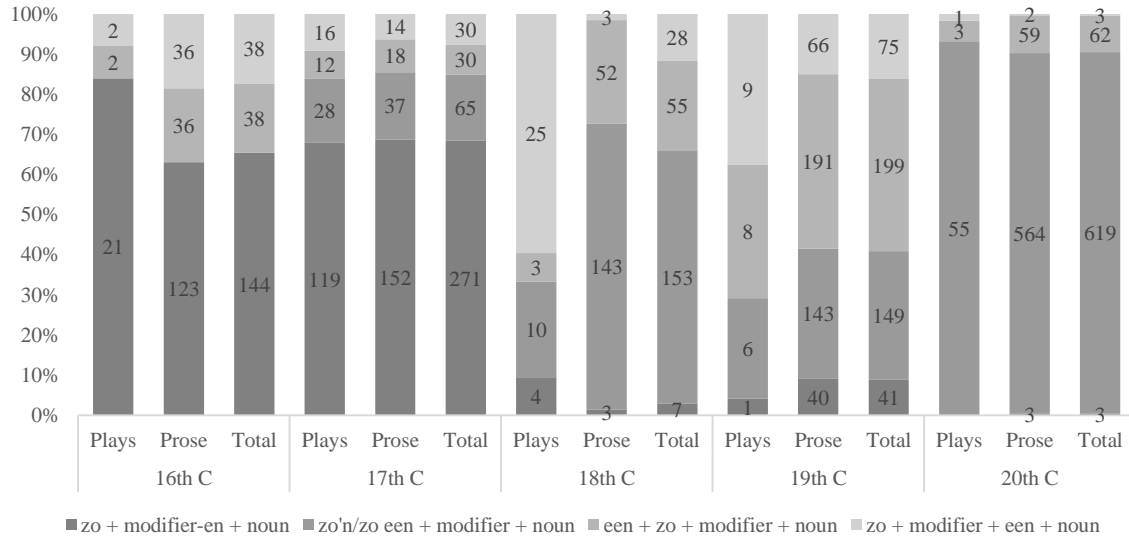
in Present-day Dutch and has a proportion of 45.59% in 1501–1700 (see Ghesquière & Van de Velde 2011: 785). Intensifying *zo 'n*, by contrast, shows a gradual decrease from 81.58% and 77.78% in the 1600s and 1700s respectively to 67.15% and 60.17% in the 1800s and 1900s respectively. The difference between the 17th and the 20th century is statistically significant ( $p < 0.01$ ). The two markers appear to prefer different structures in Early Modern Dutch, in other words. This fact indicates that the more recent one's large number of intensifying uses early on cannot be entirely attributed to analogy with the older one.

The answer to the question where *zo 'n*'s many modified cases in the 1600s come from lies in the other contexts in which *zo* combines with a modifier and a noun. Van der Horst & Van de Velde (2003: 237–238) mention the three constructions in (12a) to (12c), in addition to (12d).

- (12) a. *zo groot een man*  
       ‘so tall a man’  
       b. *een zo grote man*  
       ‘a so tall man’  
       c. *zo grooten man*  
       ‘so tall-*en* man’  
       d. *zo 'n grote man*  
       ‘such a tall man’

The first two correspond to their (unidiomatic) English translations. The *en*-ending of the adjective in the third one has traditionally been regarded as an enclitic form of the indefinite article *een*. Van der Horst & Van de Velde (2003: 245–247), however, argue that it is an old, lingering case ending which was eventually reanalyzed as the indefinite article, giving rise to the construction in (12a). Figure 7 gives the raw numbers and the proportions in the whole corpus (not just the sample) of the

four variants.



**Figure 7.** Constructions of *zo*, modifier and noun from the 16th to the 20th century

Two remarks are in order, though. First, we acknowledge that, though semantically equivalent, the constructions are probably all slightly different pragmatically. A study of the factors that have an effect on the choice of construction would be interesting but is beyond the scope of the present article. Second, Van der Horst & Van de Velde (2003: 239) point out that it is hard, if not impossible, to determine whether *-en* is just a masculine accusative ending (and should not be taken into account) or the enclitic in (12c). The reasons are that, with ongoing deflexion, Late Middle Dutch and Early Modern Dutch case marking is quite erratic at times and that, from Late Middle Dutch onward, accusative forms are sometimes used for typically nominative functions. Figure 7 therefore includes all formal matches for the construction in (12c) and thus overestimates its actual occurrence. We believe, however, that it still gives us a fairly good idea of the general trends.

The constructions in (12a) to (12c) predate (12d) and provide a template with a modifier, of

which *zo 'n* makes immediate use in the 1600s.<sup>7</sup> The preference for the construction with *zo 'n* then increases over time. *Zo* + modifier-*en* and *zo* + modifier + *een* seem to become obsolete in the 1700s and 1900s respectively while *een* + *zo* + modifier survives into the 20th century but makes up only 9.02% of the cases in this period (see also Van der Horst & Van de Velde 2003: 245). The data suggests, in other words, that the semantic shift to intensification in *zo 'n* is not just a matter of proportional change, unlike that in *such* and *zulk*, or of analogy with *zulk*. The marker's relatively high percentage of intensifying uses from the start is directly connected with its numerous early instances with a modifier. Moreover, it appears to have this structure under the influence of a number

---

<sup>7</sup> As (ia) shows, this template includes mass nouns, which may explain the early occurrence of examples such as (ib) and the limited expansion of the types of noun that *zo 'n* can combine with (see footnote 6).

- (i) a. Zoo quam *zoo* groot *een* vreesse in de ghuesghezinde ... zoo datmen niewers ... en hoorde eenen psalm resoneren.

‘In this way such great dread came upon the Beggars ... that nowhere did one ... hear a psalm resonate.’

(1566-1568, Marcus van Vaernewyck, *Van die Beroerlicke Tijden in die Nederlanden en Voornamelick in Ghendt*)

- b. Het meeste gout, en zilver, wort, door een regel van Staat, tot geld geslagen, op dat *zoo een* aangenaam metaal ... by niemand ergens, in een cabinet, zou beschimmelen.

‘Most gold and silver, is, in accordance with government regulation, coined, so that such pleasant metal ... would not gather dust, in anyone's cabinet, anywhere.’

(1644, Johan de Brune, *Wetsteen der Vernuftten*)



of older constructions that are formally and functionally similar and that it gradually replaces.

### 4.3 Recognitional use

This function has been characterized as indicating that the referent of the noun phrase starting with *zo 'n* is presumed to be known by the addressee in that it is a typical instance of some type (see Section 1). The features used for the classification of the hits in our sample include: the presence of phrases like 'you know'; the absence of a token through which the indefinite token can be identified; and, as Ekberg (2010: 60) argues for Swedish *sån* 'such', the not uncommon possibility of substituting *zo 'n* for the indefinite article with only a small change in meaning.<sup>8</sup>

The cases of *zo 'n* in (13) can serve as examples here. The noun *poftertjeskraam* 'mini pancake booth' in (13a), for instance, does not have any straightforward evaluative or scalar semantics. An intensification reading is therefore very unlikely. Moreover, the context does not contain any explicit referent through which identity to some type can be conveyed. An identification interpretation is thus ruled out as too. The second mention of the mini pancake booth in (13a), for example, is not cataphorically linked to *zo 'n*. Rather, it refers back to the instance that is brought into the story by *zo 'n*. It is also in this sense of introducing something new into the discourse that recognitional *zo 'n* can quite easily be replaced by *een*.

---

<sup>8</sup> As one of the reviewers rightly points out, this last feature is quite vague and does not apply to all recognitional uses. Moreover, as another reviewer argues, it holds for *zo 'n* in its cataphoric use too, as in (8). *Zo 'n man als ik* 'such a man as me' can easily be replaced by *een man als ik* 'a man like me'. Unlike recognitional *zo 'n*, the cataphoric use does obviously involve the presence of a token by means of which the indefinite token can be identified. The absence of such a token also helps in classifying recognitional uses in which *zo 'n* cannot simply be replaced by *een*.

- (13) a. Laastleden varkemart hebben ze nog een jong gezel, een die hier, zei ik maar zeggen, vreemd was, opgepakt. Der komt ommers altijd op varkemart hier *zo 'n* poffertjeskraam? Nou, hij mocht zoo bij die poffertjeskraam staan te kijken...  
 ‘Last pigs market they arrested a young guy, one that was, let’s say, unknown around here. There is always this mini pancakes booth at the pigs market? Well, he was hanging around that mini pancakes booth...’  
 (1839, Nicholas Beets, *Camera Obscura*)
- b. Niet helemaal onze soort maar een schat van een kind hoor. Ach *zo 'n* danseresje weet je wel. Heel aardig figuurtje.  
 ‘Not entirely our kind of people but a delightful girl for sure. Well one of those ballerinas you know. Very nice figure.’  
 (1989, Judith Herzberg, *Kras*)

The difference with the indefinite article is, of course, that *zo 'n* presents the mini pancake booth as being of a type with which the interlocutor is familiar. It affects *danseresje* in (13b) in the same way. In this example, the appeal to shared knowledge is emphasized by *weet je wel* ‘you know’.

Figure 3 tentatively suggests that the recognitional use is a recent development of *zo 'n*. It is first attested in 1839 and makes up 4.20% of the 19th-century cases and 12.39% of the 20th-century ones.<sup>9</sup> These percentages are much lower than the proportion of recognitional uses of *zo 'n* in Present-

---

<sup>9</sup> Admittedly, the 19th- and 20th-century numbers are quite low, which makes one of the reviewers wonder if the absence of recognitional uses from earlier periods is not just an accident. We recognize that this scenario is a possibility. To test it, one would need to consult a very sizeable diachronic corpus but this task, though potentially interesting, is beyond the scope of the present article. We also

day Dutch spontaneous conversations, i.e. 48.50% (see Van Olmen & van der Auwera 2014: 219). One could argue that speech represents a further stage of evolution in which this use becomes more frequent. In all probability, though, its relatively small numbers here are mainly due to the fact that it involves some kind of contact between speaker and addressee and is therefore more likely to occur in spoken than in written language. For that reason, a comparison of the present data with, say, a collection of 21st-century fictional prose and drama might reveal another small increase in the proportion of recognitional uses but that number is unlikely to be close to the percentage of instances found in the spontaneous conversations.

Interestingly, the recognitional use is never ambiguous with any of the other functions of *zo 'n* in our data. Given that semantic change of the kind relevant here often, though not always, involves bridging contexts (see Section 4.2; Heine 2002: 84; Traugott & Dasher 2002: 34–41), this fact raises the question if it is related to identification and/or intensification (see Section 4.4 on approximation), and if so, how. Considering that recognition appears to follow the (proportional) change from identification to intensification, one might be inclined to see it as a further evolution of the intensifying use. There is, however, no obvious semantic connection between the two uses.

---

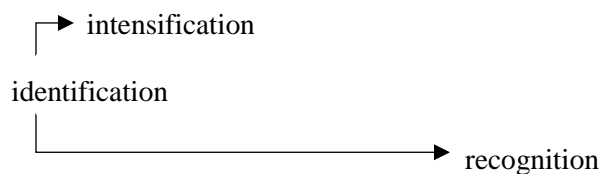
acknowledge that, as the reviewer writes, “more recent prose and plays [could just] have become better at integrating what is typically considered to be spoken language”. A few comments are in order, however. First, as discussed in Section 2, historical plays and fictional dialogue have generally been regarded as approximating spoken language. The fact that not a single case of recognitional *zo 'n* is found in almost five million words of pre-1800 drama and fiction is perhaps telling after all. Second, Dutch would not be the only language in which the recognitional use is a recent development. For Norwegian and Swedish, it has been argued to be an innovation of adolescents in multilingual settings (see Ekberg 2010, 2011). Third, and finally, even if the timing of the emergence is slightly different, the argument that the recognitional use must have developed out of the intensifying one still holds.

Recognitional *zo'n* does, for instance, not depend on the existence of some gradable notion in the noun phrase, as (13a) shows. It does have something in common with identification: it too signals the identity of an indefinite token to a type. The difference is that this type is not instantiated by another token in the context but instead belongs to the knowledge assumed to be shared between the interlocutors. To be clear, the addressee does not actually need to be familiar with it. What the speaker essentially does with recognitional *zo'n* is to *present* the indefinite token as belonging to a known type. This may sometimes require the addressee to then imagine a plausible match for the intended point of comparison. In short, it seems reasonable that the recognitional use is a further development of the identifying use or, as Ekberg (2011: 222) claims for *sân*, that “the deictic/comparative meaning inherent in the primary comparative use [i.e. exophoric identification] has been extended to the mental domain”.

Because of the lack of ambiguous cases, it remains to be seen how the shift from identification to recognition takes place, however. In our view, a crucial role is played by: pragmatic markers like *je weet wel* in (13b) and *hè* ‘right?’ in other instances; interrogative intonation, as the question mark in (13a) implies; and perhaps even forms of non-verbal communication such as tilting one’s head sideways or turning one’s eyes upward to steer the addressee’s attention to something that is not really there (though more research on this topic is obviously necessary). What these linguistic and extra-linguistic signs share is that they engage the addressee in the search for an entity to which the indefinite token is similar. When there is no token in the context that can act as an example of some type, the speaker can only (over)use identifying *zo'n* if the point of comparison can somehow be retrieved from the one source that is left to the interlocutors, i.e. their experience and imagination. A phrase such as ‘you know’ functions as an explicit pointer to the place where it is supposed to be found: it induces the addressee to draw on this knowledge to fill the gap created by the lack of another referent and to make sense of an otherwise incongruous identifying marker. In a sense, this phrase can be said to stand in for the ‘other referent’: identification can be paraphrased as ‘a referent like

something in the context’, recognition as ‘a referent like something that you know’. The other (extra-)linguistic signs mentioned above fulfill a comparable purpose. *Hè* and interrogative intonation, for instance, can be argued to present the absence of a point of comparison for *zo’n* as a question to which the addressee is presumed to be able to give an answer. The forms of non-verbal communication, lastly, appear to point to some referent with which the addressee is expected to be familiar because of its imaginary status and which they have to fill in themselves.

As an interim conclusion, Figure 8 sums up the diachronic semantic relations between the uses of *zo’n* discussed in Sections 4.1 to 4.3.



**Figure 8.** Diachronic semantic relations between the uses of *zo’n*, preliminary version

Both the intensifying and the recognitional use have been argued to arise from the identifying use. The difference in distance between identification and intensification, on the one hand, and identification and recognition, on the other, aims to reflect the fact that the intensifying uses emerged almost immediately (see Figure 5) and the recognitional ones only arose in the 1800s in our data (see Figure 3).

#### 4.4 Approximating use

This function is relatively easy to classify: it involves a cardinal number and can be paraphrased as ‘roughly’. As Figure 3 shows, *zo’n* is first found to serve as an approximator in the 1600s and continues to be used in this way, though infrequently, to the present day. Importantly, the data does

not contain any cases in which approximation is ambiguous with identification, intensification or recognition. Given that semantic change typically involves bridging contexts, this fact may be taken to suggest that this use is not directly related to the other functions of *zo 'n*. Supporting evidence for the claim comes from the marker's cognates in German and Afrikaans. *Son* and *so 'n* share all its uses apart from the approximating one (see Van Olmen & van der Auwera 2014: 222).<sup>10</sup> The question thus arises as to how exactly the approximating use came about.

Part of the answer probably lies in the hedging use of the restrictive modal adverb *zo*. According to the WNT (s.v. *zoo*), the Dictionary of the Dutch Language, it has been used on its own to attenuate what is expressed in the following word or phrase, to make it more vague or to indicate that it should not be taken as very precise from at least the 1500s onward. When it co-occurs with cardinal numbers, as in the 16th-century example in (14a) and the present-day one in (14b), the resulting meaning is approximation.

(14) a. ... daer drij *zoo* vier schepen laghen.

‘... where roughly three or four ships were docked.’

(1566–1568, Marcus van Vaernewyck, *Van die Beroerlicke Tijden in die Nederlanden en Voornamelick in Ghendt*)

b. Er zit een clubje van *zo* veertig, vijftig fans.

‘A club of roughly forty, fifty people is there.’

(Van Olmen & van der Auwera 2014: 224)

---

<sup>10</sup> Further support comes from the fact that, in our corpus, *zo 'n* is attested with plural nouns as an approximator a century before this combination became (marginally) acceptable in its other uses (see footnote 6).

This approximating adverb is found to collocate with the indefinite article in the 17th century only when the cardinal numbers follow rather than precede the noun and the noun is singular as a result, as in (15a). In this example, the numerals *drie* ‘three’ and *vier* ‘four’ are conjoined by *of* ‘or’ but they can also be combined asyndetically, as (15b) aims to show. The conjunction *of* is obligatory in the case of just one number, as in (15c). Note that, as (15d) and (15e) make clear, these constructions already express approximation without *zo* and that the adverb therefore seems to function as an additional hedge.

- (15) a. *soo een* mutsje brandewijn drie of vier  
       ‘roughly three or four brandies’  
       (1669, Isaak Vos, *Klucht van de Mof*)
- b. ‘so’ + ‘a’ + noun + cardinal number (+ ‘or’) + cardinal number
- c. ‘so’ + ‘a’ + noun + ‘or’ + cardinal number
- d. De gewone kerkenraadt ... vergadert gewoonlyk een week drie of vier voor Paaschen.  
       ‘The ordinary church council ... usually has a meeting roughly three of four weeks  
       before Easter.’ [literally ‘a week three or four’]  
       (1727, David van Hoogstraten & Matthaëus Brouërius van Nidek, *Groot Algemeen  
       Historisch, Geographisch, Genealogisch en Oordeelkundig Woorden-boek*)
- e. een week of vier  
       ‘roughly four weeks’ [literally ‘a week or four’]

The structures in (15b) and (15c) disappear from the data after 1700 (though the latter is still acceptable nowadays). The more recent attestations all look like (16) in that *zo een* — and later on *zo ’n* — is followed first by a numeral and then a plural noun.

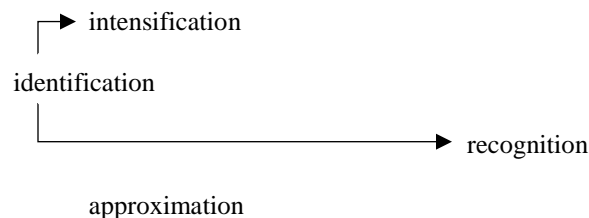
(16) Daar heb ik, als ik het alles nareeken, *zo een* twee dagen en drie uren in de magt des Satans geweest.

‘I was under Satan’s control, if I count everything again, for roughly two days and three hours.’

(1782, Betje Wolff & Aagje Deken, *Historie van Mejuffrouw Sara Burgerhart*)

What appears to have happened here is the following. An initial approximating construction with the indefinite article (‘a day or two’) was, at some point, further hedged by *zo* (‘so a day or two’). The combination of this modal adverb and *een* was then reanalyzed as a unit, most likely by analogy with the emerging identifying and intensifying marker *zo ’n* (see Figures 1 and 2). Like modifiers with this marker, the numerals moved to a prenominal position and the nouns were pluralized accordingly (‘so a two days’).

Figure 9 summarizes the diachronic semantic relations between all uses of *zo ’n* examined in Section 4. The unconnected status of approximation is meant to capture the fact that, semantically, this function is not directly linked to any of the other functions. Rather, it results from the reanalysis, as a unit, of the indefinite article in an approximating construction and a hedging adverb. The horizontal position of approximation reflects the fact that the first attestations of approximating *zo een* coincide with the first intensifying uses (see Figure 3).



**Figure 9.** Diachronic semantic relations between the uses of *zo ’n*, final version



In Section 4, we have shown that there is a proportional shift from identification to intensification in *zo'n*. As mentioned in Section 1, this change has been argued to constitute a counterexample to Traugott's (2003) hypothesis of unidirectional change from non-subjective via subjective to intersubjective meaning (see Ghesquière & Van de Velde 2011: 791–793). She regards subjective meaning as the encoded expression of the speaker's/writer's "belief state or attitude toward what is being said and how it is being said" (Traugott 2003: 125). Intersubjective meaning is understood as the encoded expression of the speaker's/writer's "attention to the 'self' of the addressee/reader in both an epistemic sense (paying attention to their presumed attitudes to the content of what is said) and in a more social sense (paying attention to their 'face' or 'image needs' associated with social stance and identity)" (Traugott 2003: 128).

In its intensifying function, *zo'n* expresses the speaker's appraisal of the extent of some gradable notion inherent in the noun phrase and is thus subjective in the traditional sense. In fact, Traugott (2006: 343–344) herself gives such uses of Bolinger's (1972) identifiers as an example of so-called subjective degree and focus modifiers, which "involve the speaker's assessment and evaluation of intensity, position on a scale, ordering of alternatives, etc.". In its original identifying use, however, *zo'n* has an intersubjective meaning if one subscribes to Ghesquière's (2009: 317) position that establishing joint attention is inherently addressee-oriented and hence intersubjective. Her view is based on Diessel's (2006: 476) characterization of the function of joint attention in demonstratives as coordinating both the speaker's and the addressee's "attentional focus in the speech situation" and as playing "an important role in the organization of discourse". Identification can be described in those terms as well: it manages the addressee's interpretation by steering their attention to a token in the context that helps them identify the relevant type. Under this analysis, a subjective use has developed out of an intersubjective one, which goes against Traugott's (2003) hypothesis.

According to Ghesquière & Van de Velde (2011: 791), Traugott's (1982) earlier view of unidirectional semantic change, i.e. from propositional via textual to expressive, offers a possible way out. The textual category, on the one hand, "as understood then, included various connectives such as *and* and *therefore*, as well as anaphoric and cataphoric pronouns, topicalizers, relativizers, complementizers, etc." (Traugott 2010: 31). Identification may be argued to fall into this category too, as it relates two elements in the discourse to each other (the indefinite token and the token serving as an example of the type). The expressive category, on the other hand, encompassed what Traugott (2010: 31) essentially describes as non-textual (inter)subjective meanings (see also Traugott & Dasher 2002: 94). Intensification fits this description. Thus, from the perspective of Traugott's (1982) cline, the identifying use of *zo'n* is an 'acceptable' textual ancestor of its expressive intensifying use.

In our view, Ghesquière & Van de Velde's (2011) proposal is attractive in at least two respects. First, it is based on a substantial improvement on Traugott's (2003) theory of (inter)subjectification. She characterizes subjective and intersubjective meaning in such a way that "there cannot be intersubjectification without some degree of subjectification" (Traugott 2003: 125). The fact that directionality is built into her definitions means, as Narrog (2011) among others points out, that the alleged unidirectionality of the process is simply not testable (see also Ghesquière *et al.* 2012: 140–143). The notion to which Ghesquière & Van de Velde (2011) appeal to describe identification, by contrast, has been defined autonomously. With reference to Diessel (2006) and in line with work by Breban (2010) on the grammaticalized uses of English adjectives of comparison and Carlier & De Mulder (2010) on Latin determiners, Ghesquière (2009) considers meanings concerned with guiding the addressee's interpretation — as expressed by, *inter alia*, focus, backgrounding and joint attention markers — as textually intersubjective. Second, linguists who do not wish to call such meanings intersubjective still have to acknowledge their textual nature in the sense of Traugott (1982: 248). The phoricity of identifying *zo'n*, for instance, contributes to the coherence of the discourse. Put differently, even if one rejects the semantic change from identification to intensification as a

counterexample to the cline from subjective to intersubjective,<sup>11</sup> one must admit that the cline from textual to expressive captures the shift under discussion very well. *Zo'n* can thus be said to provide evidence for the earlier hypothesis about semantic change.

Some reservations are in order, however. The fact that the shift from identification to intensification supports Traugott's (1982) cline should not entice us to instantly reinstate it as the unidirectional path that meaning in general is expected to take. Narrog (2012: 35–40), for instance, shows that so-called concessive *may* in (17a) follows epistemic *may* in (17b). The former can be regarded as textual since it creates logical connections in complex clauses and in discourse. The latter can be considered expressive in that it is non-textually subjective: it conveys the speaker's evaluation of the likelihood of the proposition.

- (17) a. Incredible though it *may* seem, she owes me money.  
b. He *may* be coming later.

To be clear, Narrog (2012) does not use the term 'expressive'. He makes a distinction between textual, subjective and intersubjective meanings, which are oriented toward the text, the speaker and the

---

<sup>11</sup> In fact, one anonymous reviewer takes this position, stressing that Traugott's (2003) cline from non-subjective via subjective to intersubjective is a hypothesis about semasiological change and, more precisely, about *adding* (inter)subjective meaning to the original meaning of constructions. In this sense, the shift from identification to intensification is, strictly speaking, not a counterexample: the intensifying use indeed adds subjective meaning to the original identifying use of *zo'n*. However, it does constitute a counterexample if the notions of subjective and intersubjective meaning are defined independently from clines of semantic change and if one regards meanings oriented toward guiding the addressee's interpretation as intersubjective.

addressee respectively. We are inclined to agree with his overall assessment that “[i]t may be premature to assume a single directionality (e.g. subjective > intersubjective > textual; subjective > textual > intersubjective; etc.) to which all items in all domains would obey” (Narrog 2012: 46–47).

In addition, to our knowledge, insufficient attention has been paid to the potential of change of constructions that are intersubjective by nature. Haspelmath (2004: 25), for example, argues that “some elements such as demonstratives ... are never created by grammaticalization from full lexical items and have simply always been demonstratives”. If he is correct, they can be said to start out not at the usual non-subjective or propositional end of the aforementioned clines of unidirectional semantic change but at the intersubjective or textual stage.<sup>12</sup> In the same vein, markers such as *zulk* and *zo'n* come from a word that appears to be inherently deictic, i.e. the Proto-Germanic demonstrative *\*sa*, and thus intersubjective in the sense of Ghesquière (2009). What is more, such intrinsically intersubjective items have been found to be able to undergo (some degree of) subjectification. Even Traugott (2010: 59) admits that *say* in (18), for instance, “derives from an intersubjective, imperative use, but is subjectified over time”: it conveys wonder to what is said afterward.

(18) ‘*Say*,’ Mr Cooper said, ‘you’ve got quite a baby here.’

(Van Olmen 2013: 255)

Under Ghesquière & Van de Velde’s (2011) analysis, the shift from identification to intensification — though slightly different, of course — would be another example of a development from inherently intersubjective to subjective. We leave it for further research to determine how common this type of

---

<sup>12</sup> This point again assumes, pace one of the reviewers, that the concepts used in those clines can be characterized independently (see also footnote 11).

evolution is and if subjectification is perhaps better thought of as a process of semantic change that can apply to any ‘basic’ meaning, regardless of its intrinsic non-subjective, subjective or intersubjective character.

In addition to confirming the proportional change of *zo ’n* from identifying to intensifying uses, Section 4 has made the case that its recognitional use is a further development of its identifying one. The question now is how recognition fits in with (inter)subjectification. As this use presents the referent of the noun phrase as being of a type known to the addressee, it is fairly obvious that, like identification, recognition is an intersubjective notion. It is not clear, though, whether the change, for example, instantiates any established pathway between types of intersubjective meaning.

Besides the textual type, Ghesquière *et al.* (2012) make a distinction between responsive and attitudinal intersubjectivity. Responsive intersubjectivity “is conveyed by linguistic items eliciting not just a cognitive operation on behalf of the hearer but a certain (speech) act or behaviour promoting discourse continuity or discourse cooperation” (Ghesquière *et al.* 2012: 135). It covers, *inter alia*, question tags and turn-taking devices. Recognitional *zo ’n* is not one of these linguistic items: it may cause the addressee to think of something to which the indefinite token is similar but it does not require any real action from them. This is evidenced by the fact that, in both (13a) and (13b), repeated as (19) here, the speaker just keeps on talking after using *zo ’n*.

- (19) a. Laastleden varkemarkt hebben ze nog een jong gezel, een die hier, zei ik maar zeggen, vreemd was, opgepakt. Der komt ommers altijd op varkemarkt hier *zoo ’n* poffertjeskraam? Nou, hij mocht zoo bij die poffertjeskraam staan te kijken...  
‘Last pigs market they arrested a young guy, one that was, let’s say, unknown around here. There is always this mini pancakes booth at the pigs market? Well, he was hanging around that mini pancakes booth...’  
(1839, Nicholas Beets, *Camera Obscura*)

- b. Niet helemaal onze soort maar een schat van een kind hoor. Ach *zo 'n* danseresje weet je wel. Heel aardig figuurtje.

‘Not entirely our kind of people but a delightful girl for sure. Well one of those ballerinas you know. Very nice figure.’

(1989, Judith Herzberg, *Kras*)

Attitudinal intersubjectivity stands for “meanings which code the speaker’s image of (his/her relation to) the hearer, and attention to the face needs and social self of the hearer” (Ghesquière *et al.* 2012: 132). It concerns honorifics and hedges, among other things.<sup>13</sup> The recognitional use does meet this description to some extent: one could argue that *zo 'n* suggests that the addressee is co-responsible for the identification of the type and that, as such, it “is interactive and implies a certain degree of closeness between the participants” and hence “functions as a positive politeness marker” (Ekberg 2011: 222 on *sån*). The possibility of a shift from textual to attitudinal intersubjectivity is not discussed by Ghesquière *et al.* (2012: 140–143). It would go well with Traugott’s (1982) hypothesis, however, if the expressive stage — just like Halliday & Hasan’s (1976) interpersonal component of the linguistic system, on which it is based — is taken to include intersubjective meanings (see also Traugott 2010: 31). Yet, recognitional *zo 'n*’s primary function continues to be the management of the addressee’s understanding, by drawing their attention to an indefinite token as being of a known type.

---

<sup>13</sup> As one of the reviewers rightly points out, approximating *zo 'n* is intersubjective too. It essentially functions as a hedge (see Section 4.4) and would therefore almost certainly be classified as a marker of attitudinal intersubjectivity in Ghesquière *et al.*’s (2012) typology. The reason why approximation is not discussed any further here is that Section 5 focuses on changes in (inter)subjective meaning. As this use is not directly related to the other uses of *zo 'n* in our view, there is simply no straightforward semantic shift in the marker to or from approximation that needs to be considered.

The kind of intersubjectivity in Ghesquière *et al.*'s (2012: 145) typology that it matches the best is probably still the textual one, which is “specifically oriented toward steering the hearer’s interpretation (e.g. focus and backgrounding devices and elements creating a joint focus of attention)”. Unlike identifying *zo ’n*, though, it does not really relate two elements in the discourse to each other and is therefore less clearly textual.

A distinction that may be relevant for the shift from identifying to recognitional uses is a recent one made by Tantucci (2013, 2015) between immediate and extended intersubjectivity. The immediate type is concerned with “the here-and-now of the conversation” and “is generally limited to the two actual personas — SP/W and AD/R [i.e. speaker/writer and addressee/reader respectively] — engaging in the conversation and their respective inferred feelings and attitudes” (Tantucci 2015: 5). It closely corresponds to intersubjectivity as it is generally understood in the literature (e.g. Traugott’s 2003: 128 stress on the role of the interlocutors in her definition of intersubjective meaning as the expression of the speaker’s consideration for the addressee’s self). In the other type, “the notion of ‘you’ [i.e. the addressee] turns into the notion of ‘you and someone else’” (Tantucci 2015: 5). In other words, extended intersubjectivity invokes some third party, typically a generic one or society as a whole, “indirectly supporting the conversation, personally or impersonally determining what it is to be approved or denied emotionally or epistemically between the interlocutors” (Tantucci 2015: 6). Evidential verb + *guo* in Mandarin Chinese can serve as an example. It derives from an experiential perfect construction and can be paraphrased as ‘it is known’. Rather than calling upon some direct or indirect evidence, the speaker using verb + *guo* can be argued to appeal “to a form of shared knowledge that can be confirmed by whatever third party in society” or, put differently, to ground their statement evidentially “in what a rather small or large group of people is assumed to know” (Tantucci 2013: 224). An important claim about the two types of intersubjectivity is that they are on a unidirectional cline of semantic change: the immediate type is presumed to precede the extended one (see Tantucci 2015: 15).

The question now is whether and how this distinction applies to *zo'n*. In its identifying use, *zo'n* could be analyzed as immediately intersubjective. The speaker employs it to steer the addressee's understanding of some noun phrase, by calling upon their awareness of the on-going discourse — in the case of anaphoric or cataphoric reference — or of the situation — in the case of exophoric reference, where the addressee is directed to find the referent that serves as an example of the type that the indefinite token belongs to. Identification is thus primarily concerned with the context at the time of speaking or, in other words, with the here-and-now of the conversation. In its recognitional function, by contrast, *zo'n* could be regarded as a case of extended intersubjectivity. When the speaker of (19a), for instance, presents the mini pancake booth as of a familiar type, they rely on the knowledge that the addressee is assumed to share with them. Yet, the knowledge that is evoked between the interlocutors is not specific to them (see also Tantucci 2016: 161 on the intersection between extended intersubjectivity and the common ground). The speaker is talking not about a kind of mini pancake booth with which only the current speech participants are familiar but about one which it is usually found at a market and, which, in the speaker's view, everyone is supposed to know. Roughly the same holds for (19b): the speaker is suggesting that the girl has the characteristic properties of a ballerina, of which their interlocutor, like anyone else, is taken to be aware. Recognitional *zo'n* can thus be argued to attribute to the addressee the same 'state of mind' that a third party would have in the situation, which explains the sense of typicality that it calls up. It may therefore be analyzed as an instance of extended intersubjectivity (Tantucci p.c.). To sum up, the change from identification to recognition seems to support the two types of intersubjectivity that Tantucci (2015) distinguishes and the directionality that he postulates for them. Further research will need to test how robust his hypotheses are for semantic change in general, though.

## 6. Conclusion



In this article, *zo'n* has been shown to emerge and fuse into one word earlier than previously claimed in the literature. Regarding identification and intensification, our study confirms previous findings that the change from the former to the latter is of a proportional rather than a discrete nature and that ambiguity between the two meanings is a relatively rare phenomenon. *Zo'n* has been found to be used in its intensifying function more frequently than *zulk* in their respective first stages, though. This difference has been attributed to influence from pre-existing constructions, in which *zo* combined with a modifier and a noun to convey intensification. The under-described recognitional use of *zo'n* has been shown to be a relatively recent innovation and has been argued to have evolved out of its identifying use. As to the marker's fourth and last function, the approximating one, we have made the case that it is a development independent from the other uses and linked to prior constructions expressing approximation. As far as the relation between all these developments and the process of (inter)subjectification is concerned, finally, we have endorsed the earlier view that the traditional cline of semantic change from (propositional via) textual to expressive best reflects the shift from identifying to intensifying uses. But we have also pointed out that the cline is not unproblematic as a hypothesis of unidirectional change. Expressive items may become textual too, for one thing. In addition, the shift from identifying to recognitional uses has been argued to support a recent proposal in the literature to distinguish immediate intersubjectivity from its further development of extended intersubjectivity.

## References

- Auwers, Johan van der & Evie Coussé. 2016. *Such* and *sådan* — the same but different. *Nordic Journal of English Studies* 15(3). 15–32.
- Auwers, Johan van der & Kalyanamalini Sahoo. 2015. On comparative concepts and descriptive

- categories, *such as they are*. *Acta Linguistica Hafniensia* 47(2). 136–173.
- Bolinger, Dwight. 1972. *Degree words*. The Hague: Mouton.
- Breban, Tine. 2010. *English adjectives of comparison: Lexical and grammaticalized uses*. Berlin: De Gruyter.
- Brems, Lieselotte. 2011. *Layering of size and type noun constructions in English*. Berlin: De Gruyter.
- Carrier, Anne & Walter De Mulder. 2010. The emergence of the definite article: *ille* in competition with *ipse* in Late Latin. In Kristin Davidse, Lieven Vandelanotte & Hubert Cuyckens (eds.), 241–275.
- Culpeper, Jonathan & Merja Kytö. 2000. Data in historical pragmatics: spoken interaction (re)cast as writing. *Journal of Historical Pragmatics* 1(2). 175–199.
- Davidse, Kristin, Lieven Vandelanotte & Hubert Cuyckens (eds). 2010. *Subjectification, intersubjectification and grammaticalization*, Berlin: De Gruyter,
- Demske, Ulrike. 2005. Weshalb Demonstrativpronomina nicht immer Determinantien sind [Why demonstrative pronouns are not always determiners]. In Franz Josef D’Avis (ed.), *Deutsche Syntax: Empirie und Theorie [German syntax: Empiricism and theory]*, , 53–80. Gothenburg: University of Gothenburg, 53–80.
- Diessel, Holger. 2006. Demonstratives, joint attention and the emergence of grammar. *Cognitive Linguistics* 17(4). 463–489.
- Ekberg, Lena. 2010. Extended uses of *sån* (‘such’) among adolescents in multilingual Malmö. In Roger Källström & Inger Lindberg (eds.), *Young urban Swedish: Variation and change in multilingual settings*, 49–65. Gothenburg: University of Gothenburg.
- Ekberg, Lena. 2011. Joint attention and cooperation in the Swedish of adolescents in multilingual settings. In Friederike Kern & Margret Selting (eds.), *Ethnic styles of speaking in European metropolitan areas*, 217–237. Amsterdam: Benjamins.
- Ghesquière, Lobke. 2009. From determining to emphasizing meanings: the adjectives of specificity.

*Folia Linguistica* 43(2). 311–343.

Ghesquière, Lobke. 2012. Identifying and intensifying uses of prenominal *such*: A data-based approach. *International Journal of Corpus Linguistics* 17(4). 516–545.

Ghesquière, Lobke, Lieselotte Brems & Freek Van de Velde. 2012. Intersubjectivity and intersubjectification: typology and operationalization. *English Text Construction* 5(1). 128–152.

Ghesquière, Lobke & Freek Van de Velde. 2011. A corpus-based account of the development of English *such* and Dutch *zulk*: identification, intensification and (inter)subjectification. *Cognitive Linguistics* 22(4). 765–797.

Gries, Stefan Th. 2013. Elementary statistical testing with R. In Manfred G. Krug & Julia Schlüter (eds.), *Research methods in language variation and change*, 361–381. Cambridge: CUP.

Haeseryn, Walter, Kirsten Romijn, Guido Geerts, Jaap de Rooij & Maarten C. van den Toorn. 1997. *Algemene Nederlandse spraakkunst [Comprehensive Dutch grammar]*. Groningen: Nijhoff.

Halliday, M.A.K. & Ruqaiya Hasan. 1976. *Cohesion in English*. London: Routledge.

Haspelmath, Martin. 2004. On directionality in language change with particular reference to grammaticalization. In Olga Fischer, Muriel Norde & Harry Perridon (eds.), *Up and down the cline: The nature of grammaticalization*, 17–44. Amsterdam: Benjamins.

Heine, Bernd. 2002. On the role of context in grammaticalization. In Ilse Wisner & Gabriele Diewald (eds.), *New reflections on grammaticalization*, 83–101. Amsterdam: Benjamins.

Himmelmann, Nikolaus P. 2001. Articles. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals: Volume 1*, 831–841. Berlin: De Gruyter.

Hole, Daniel & Gerson Klumpp. 2000. Definite type and indefinite token: the article *son* in colloquial German. *Linguistische Berichte* 182. 231–244.

Horst, Joop M. van der. 2008. *Geschiedenis van de Nederlandse syntaxis [History of Dutch syntax]*.

- Leuven: Leuven University Press.
- Horst, Joop M. van der & Freek Van de Velde. 2003. Zo vreemd een groep [So strange a group]. *Taal en Tongval* 15/16. 237–250.
- IBM Corp. 2013. *IBM SPSS statistics for Windows: Version 22.0*. Armonk: IBM Corp.
- Koninklijke Bibliotheek. 2015. *Digitale bibliotheek voor de Nederlandse letteren [Digital library of Dutch literature]*. The Hague. <<http://www.dbnl.org>> (accessed in July 2015)
- Narrog, Heiko. 2011. The limits of (inter)subjectification. Paper presented in December at the University of Leuven.
- Narrog, Heiko. 2012. Beyond intersubjectification: textual uses of modality and mood in subordinate clauses as part of speech-act orientation. *English Text Construction* 5(1). 29–52.
- Nederlandse Taalunie. *Corpus gesproken Nederlands 1.0 [Corpus of spoken Dutch 1.0]*. The Hague.
- Rayson, Paul E. & Roger Garside. 2000. Comparing corpora using frequency profiling. *Proceedings of the Association for Computational Linguistics Workshop on Comparing Corpora at their 38th Annual Meeting*, 1–6.
- Scott, Mike. 2012. *WordSmith tools version 6*. Stroud: Lexical Analysis Software.
- Smitterberg, Erik. 2016. Extracting data from historical material. In Merja Kytö & Päivi Pahta (eds.), *The Cambridge handbook of English historical linguistics*, 181–199. Cambridge: CUP.
- Spinillo, Miriàngela. 2003. On *such*. *English Language and Linguistics* 7(2). 195–210.
- Tantucci, Vittorio. 2013. Interpersonal evidentiality: the Mandarin V-*过* *guo* construction and other evidential systems beyond the ‘source of information’. *Journal of Pragmatics* 57. 210–230.
- Tantucci, Vittorio. 2015. From immediate to extended intersubjectification: a gradient approach to intersubjective awareness and semasiological change. *Language and Cognition*. <<http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=10088734&fileId=S1866980815000265>> (accessed in March 2016)

- Tantucci, Vittorio. 2016. Textual factualization: the phenomenology of assertive reformulation and presupposition during a speech event. *Journal of Pragmatics* 101. 155–171.
- Traugott, Elizabeth C. 1982. From propositional to textual and expressive meanings: some semantic-pragmatic aspects of grammaticalization. In Winfred P. Lehmann & Yakov Malkiel (eds.), *Perspectives on historical linguistics*, 245–271. Amsterdam: Benjamins.
- Traugott, Elizabeth C. 2003. From subjectification to intersubjectification. In Raymond Hickey (ed.), *Motives for language change*, 124–139. Cambridge: CUP.
- Traugott, Elizabeth C. 2006. The semantic development of scalar focus modifiers. In Ans van Kemenade & Bettelou Los (eds.), *Handbook on the history of English*, 335–359. Oxford: Blackwell.
- Traugott, Elizabeth C. 2010. (Inter)subjectivity and (inter)subjectification: a reassessment. In Kristin Davidse, Lieven Vandelanotte & Hubert Cuyckens (eds.), 29–71.
- Traugott, Elizabeth C. & Richard B. Dasher. 2002. *Regularity in semantic change*. Cambridge: CUP.
- Van Olmen, Daniël. 2013. The imperative of *say* as a pragmatic marker in English and Dutch. *Journal of Germanic Linguistics* 25(3). 247–287.
- Van Olmen, Daniël & Johan van der Auwera 2014. Over *zo'n* and *zo meer* [On *zo'n* and so on]. In 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* [*Pattern and argument: A double festschrift on the occasion of the retirements of William Van Belle and Joop van der Horst*], 215–228. Leuven: Leuven University Press.
- WNT. 2010[1864-1998]. *Woordenboek der Nederlandsche taal* [*Dictionary of the Dutch language*]. Leiden. <<http://gtb.inl.nl/?owner=WNT>> (accessed in July 2015)