

Marco Angster & Livio Gaeta

Being itself: On German *selbst* in synthetic compounds*

Abstract: German synthetic compounds containing *selbst* have traditionally been treated as reflexive in nature. However, this view cannot cope on the one hand with the basic value of *selbst* as focus particle and not as pronoun and on the other with the non-reflexive value of many compounds whose meaning range from purely intensification of the understood agent of the activity as in *Selbstbau* ‘do-it-yourself-construction’ to anti-causativity like *Selbstentzündung* ‘self-ignition’. In this paper, we will consider a different option, namely to interpret the focus particle *selbst* as a flag for signaling contrastive coreference. This option can be concretely operationalized in constructional terms as a case of constructional idiom.

Keywords: compounds, reflexivity, intensification, constructional idiom

1. Introduction

The German particle *selbst* ‘self’ found in so-called “Reflexive Nominal Compounds” (cf. König 2011) such as *Selbstbestimmung* ‘self-determination’, *Selbstbeherrschung* ‘self-control’, and the like, is normally held to contain a variable-binding pronoun (Fleischer & Barz 2012: 164) and accordingly to “express a reflexive meaning” (König 2011: 113). Typically, these compounds are based on transitive verbs in which the two argument positions referring to an agent and a patient are bound by *selbst*, as shown by examples like *Selbstbeziehung* ‘self-accusation’ and *Selbsthass* ‘self-hate’ where the former profiles a remarkable agent which as “relevant information ... provides the source or agent of the event” and the latter a remarkable patient which is characterized “as a surprising choice for the event in question” (König 2011: 120). Other cases based on intransitive verbs like *Selbstläufer* ‘fast-selling item’ or anti-causative uses of *selbst-* as in *Selbstheilung* ‘self-healing, i.e. something heals by itself’ are deemed as “are very rare and even marginal” (König 2011: 122).

* The paper has been conceived jointly by the authors. However, while §1 and §6–7 have been written jointly by the authors, LG is responsible for §2–3 and MA for §4–5.

However, these uses show that *selbst-* does not necessarily occur with bases derived from transitive verbs, but also – and more interestingly – that it does not imply any co-reference and qualifies for a particular constructional type. Moreover, there are compounds like *Selbstbau* ‘do-it-yourself-construction’ or *Selbstmitnahme* ‘self-service-take-away’ which cannot be interpreted as reflexive in any meaningful way. Finally, *selbst* – which actually differs from the true reflexive pronoun *sich* – partially competes on the one hand with the adjective *eigen* (*Eigen/Selbstlob*) and on the other with the loan prefix *auto-* (*Auto/Selbsthypnose*), which raises the question of the status of these formations, potentially interpretable as compounds or as prefixes.

In this paper, we will first discuss the status of *selbst*-compounds in their structural consistency and their relations to similar constructions containing other morphemes. We will focus mainly on *selbst*-compounds headed by deverbal nouns, either agent or action nouns, although our impression is that our analysis also applies to *selbst*-compounds headed by simplex nouns, by adjectives or participles and even by verbs. However, they will not be directly investigated although they will cursorily be referred to in the paper. In §2 the proper status of *selbst* within German word-formation will be shortly discussed before passing in §3 to discuss recent proposals made in the literature to account for *selbst*-compounds. In §4 we will introduce our empirical investigation discussing our sample extracted from the *deTenTen13* corpus, while in §5 the *selbst*-compounds occurring in the sample will be carefully analyzed. In §6 a treatment in terms of a constructional framework is suggested which crucially relies on the concept of constructional idiom. The final §7 draws a brief conclusion.

2. The proper status of *selbst* within German word-formation

There can be no doubt that the particle *selbst* forms compounds in clear contrast to its English cognate *-self* which never occurs as a free morpheme and in so far might in principle also be qualified as a kind of affix(oid). However, the status of *selbst* is unclear, at least according to the analyses provided in many current handbooks on morphology and word-formation. Therein, *selbst* is normally considered a pronoun on a par with the rest of the pronominal family. In Tab. 1 the result of a quick survey of handbooks on German word-

formation or online platform on German grammar is reported which clearly shows this general orientation:

Tab. 1: The treatment of *selbst* in handbooks on German word-formation

	Oa	F&B	E	H&T	M	L	A	Lo	D
pronoun	✓	✓	✓	✓	✓	✓	✓	✓	
particle									✓
source	Oa = Ortner <i>et al.</i> (1991: 60), F&B = Fleischer & Barz (2012: 165), E = Eichinger (2000: 36), H&T= Harnisch & Trost (2009: 35), M = Motsch (2004: 417), L = Leo-Wortgrammatik, A = Altmann (2011: 35), Lo = Lohde (2006: 76), D = Duden (2022: 838)								

This view, largely reflected in the literature on word-formation, is likely to go back to the traditional approach to *selbst* in terms of a pronoun, as for instance discussed by Becker (1836: 280) where *selbst* is treated as *demonstratives Adjektivpronom* ‘demonstrative adjectival pronoun’ in connection with the corresponding Latin pronoun *ipse*. However, already Becker betrays some dissatisfaction with this definition because in a revised edition of his school’s grammar he employs the more cautious term *Formwort* ‘form-word’: “Man be- greift unter den Demonstrativpronomina auch das Formwort selbst (selber), das insge- mein in der Verbindung mit einem Substantiv oder Substantivpronom gebraucht wird. Dieses Formwort hat meistens den Hauptton, und folgt dem Substantiv oder Substantiv- pronom nach; und es hat dann eine ausschließende Bedeutung: es bezeichnet nämlich, daß nur das genannte Sein und kein anderes gemeint ist” (Becker 1852: 148, original emphasis).¹ In neat contrast to the traditional view, in the most recent edition of Duden (2022: 838) *selbst* is treated as focus particle, as reported in Tab. 1.

On the other hand, also in certain analyses going beyond word-formation, *selbst* is held to form a unitary paradigm in combination with the personal pronoun: “das Reflexivpro- nomen weist nun nicht länger ein defektives, sondern ein vollständiges Paradigma auf, das z.B. im Bereich des Pers.Pronomens 3.Sg.Mask. lautet: *er (selbst), seiner (selbst), sich (selbst), sich (selbst)*”² (Leys 1973: 153). Leys supports this assumption by referring to two empirical

¹ [Under the demonstrative pronouns one also understands the form-word *selbst (selber)*, which is generally used in connection with a substantive or substantive pronoun. This form-word usually carries the main stress, and follows the substantive or substantive pronoun; and it then has an exclusive meaning: namely, it indicates that only the named being and no other is meant, our translation].

² [Thus, the reflexive pronoun does not display a defective, but a complete paradigm, which sounds for in- stance in the domain of the 3rd sg. masc. personal pronoun: *er (selbst), seiner (selbst), sich (selbst), sich (selbst)*, our translation].

considerations. First, the reflexive value extends also beyond the personal pronouns, for instance to the indefinite pronoun *man*, as in the following example:

- (1) *das heißt, führt man sich nach außen hin hinreichend mit bürgerlich konventionell auf, kann man, innerlich entschlossen, man selbst ... sein.*

‘that is, if one behaves sufficiently conventionally bourgeois to the outside world, one can, with inner determination, be oneself’.

The second fact is its obligatory occurrence in certain contexts (see also Duden 2022: 741), as for instance in: *er ließ mich mich *(selbst) sein* ‘he let me be myself’. In correspondence of this view, *selbst* is attributed by König (2011: 113) the main function of binding a variable connected with a verbal argument:

More often than not, the first part of these derived nouns [*scil. selbst-* in *Selbstlob* ‘self-praise’, etc.] provides information about the noun’s argument structure: Typically, the two argument positions inherited from the underlying transitive verb must be bound by the same (generalized) quantifier or, putting it somewhat loosely, they must be filled by the same argument or by metonymically related arguments. In other words, the relevant nominals express a reflexive meaning.

Clearly, binding variables is the typical function usually attributed to pronouns as grammatical morphemes deprived of any intrinsic referential value. The problem with this analysis as generalized quantifier is twofold. First, as already observed above, *selbst* cannot not be treated as a reflexive pronoun, but rather as an intensifying or focus particle, and in fact Duden (2022: 838) correspondingly attributes *selbst* to the group of focus particles (see for a detailed discussion also Siemund 2007: 718–720). This is clearly shown in examples like (2), where no pronominal, let alone reflexive, interpretation is possible:³

³ Lohde (2006: 76) treats *selbst* as an “undeclinierbares Demonstrativpronomen” [undeclinable demonstrative pronoun] characterized by a “hohe Kompositionsaktivität” [high compounding activity]. To be sure, *selbst* can be used for contrastive focus, even in the absence of any explicit referential expression, as for instance in *Selbst gebackener Kuchen schmeckt am besten* ‘Homemade cake tastes best’. In Metrich & Faucher (2009: 776–782) *selbst* is attributed to two different word classes, respectively the pronouns and the focus particles. The assignment to the pronominal class is justified on the basis of its reference to a given (possibly implicit) constituent with an emphasis on its exclusive interpretation, as already observed by Becker above. This difference is overtly manifested in a complementary syntactic behavior: while in its pronominal value *selbst* is postposed to the modified constituent, in its focus value it is generally preposed. Although this difference is not completely reliable as shown by Duden’s (2022: 438) example: *Selbst Anna / Anna selbst wusste nichts*. ‘Even Anna knew nothing’, it holds generally true, but does not necessarily speak in favor of a different treatment in terms of word classes. In both cases, we can equally speak of focus particles although with a different value, respectively a scalar and a contrastive or exclusive one.

(2) *Helga hat die Haare ihrer Großmutter selbst gewaschen.*

‘Helga herself washed her grandma’s hair.’

In this regard, Eisenberg (2020: 190) concludes: “Obwohl reflexiviert wird und man sogar von emphatischen Reflexiva oder Intensifikatoren spricht, gehört *er selbst* nicht mit *sich* in dasselbe Paradigma”.⁴ The paradigmatic independence of the reflexive pronoun and of the particle *selbst* is well known to König (2011: 113) who in fact speaks of means of “derivation” employed to express reflexivity in compounds:

The first component of these nominals can be derived (both diachronically and synchronically) from intensifiers (Engl. (X-)self; Ger. *selbst*, *eigen*; Russ. *sam*; Mand. *zǐjǐ*, etc.) and the compounds can be paraphrased in terms of at least one of the uses normally distinguished for intensifiers.

The problem with this approach is that a synchronic and a diachronic view need not coincide or, even worse, they can even stand in contrast. In this sense, the term derivation refers to completely different things in the two perspectives. Thus, it might appear justified in diachronic terms to assume that a certain paradigmatic relation holds within a word pair while synchronically the opposite is true. For instance, the Latin verb *secāre* ‘to cut’ provides the derivational base for the instrument noun *seca* ‘tool for cutting, saw’ attested in Medieval Latin. In Italian, however, the noun *sega* ‘saw’ is likely to be the basis for the verbal conversion *segare* ‘to saw’. Thus, diachronically *sega* is a nominalization from *segare*, while synchronically the opposite is true.⁵ In this light, the term “derivation” used indiscriminately in a synchronic and a diachronic perspective ingenerates a certain confusion.⁶ We rather prefer to keep the two perspectives distinct. On the one hand, there is the issue of derivation intended in word-formation terms as the way how synchronically a certain complex word is formed or has to be analyzed. On the other hand, there is the diachronic issue of the lexical basin wherefrom certain morphemes are recruited in the course of time and give rise to synchronic processes of word-formation. While both perspectives

⁴ [Although it’s reflexivized and one even speaks of emphatic reflexives or intensifiers, *er selbst* does not belong in the same paradigm, our translation].

⁵ The same can be repeated for the case of *selbst* vs. *das Selbst* ‘the self’, see Angster (to appear: 197).

⁶ In this respect, an anonymous reviewer wonders “how far is the first constituent of *selbst*-compounds in any synchronic or diachronic sense derived from *selbst*? The first constituent IS *selbst*; there is no derivation process at all.”

are perfectly legitimate, in this paper we will insist on the former, casting the latter aside for further research.⁷

The second general problem with the view of *selbst* as a reflexive pronoun is that German displays a proper reflexive pronoun, which cannot occur in compounds with reflexive meaning: **Sich-Lob*, **Sich-Zerstörung*, etc. One might be tempted to interpret this restriction as positive evidence in support of the suppletive function of *selbst* with regard to *sich*.⁸ However, pronouns – with the remarkable exception of *selbst* – cannot generally be employed in compounds, unless they are employed in the nominative, their citation form, without any possible function of variable-binder (cf. Fleischer & Barz 2012: 164):

- (3) *Ichkult* ‘self-cult’, *Ichform* ‘first person’, *Ichgefühl* ‘sense of self’, *Ich-Erzähler* ‘first-person narrator’, *Ichsucht* ‘egoism’, *Wirgefühl* ‘We-feeling’, *Wirbewusstsein* ‘we-consciousness’, *Wirgruppe* ‘we-group’, *Dubeziehung* ‘you-relation’

This is actually an old observation, going back at least to Edmondson & Plank (1978: 394):

Among other things such accounts [*scil.* the emphatic reflexive account] still fail to explain why it is not the pronoun that shows up in complex words but invariably the intensive or emphatic element alone (**himself-educated*). This becomes particularly evident in German since here intensives differ in form from ordinary reflexives.

They observe in particular the ungrammaticality of **Sichbildnis* ‘REFL-image, **Michselbstkritik* ‘myself-criticism’, **sichselbstgenügsam* ‘REFL-self-undemanding’.⁹ Notice that the lack of functional activation of the pronouns which only appear in the citation form is also hinted at by Fleischer & Barz (2012: 165) who seek in this way to justify their scarce occurrence in compounds:

Die Gründe für die geringe Kompositionsaktivität der Pronomina werden in ihren ‘situationsbestimmten Funktionswerten’ gesehen, womit sie in dieser Hinsicht den ‘situationsvariablen grammatischen Kategorien Tempus und Modus’ entsprechen.¹⁰

⁷ See Chapter 6 in Angster (to appear) for a survey of diachronic and contact-related issues concerning contrastive coreference formations within a cross-linguistic perspective.

⁸ While in other cases an allomorphic variant enters a compound like *Sonderkennzeichen* ‘special character’ with regard to the adjective *besondere* ‘special’ (cf. Elsen 2014: 62), *selbst* cannot be treated as allomorphic with regard to *sich*.

⁹ To be sure, compounds containing the reflexive pronoun reinforced by *selbst* are actually found, as for instance *Sichselbstgleichheit* ‘REFL-self-equality’. They are likely to be interpreted as containing a full-fledged phrase in modifier position (see §4 below).

¹⁰ [The reasons for the low compositional activity of the pronouns are seen in their “situation-determined functional values”, which in this respect correspond to the “situation-variable grammatical categories of tense and mood”, our translation].

The crucial point with pronouns within compounds, however, is that they are normally employed in a substantially non-anaphoric way, which goes well beyond the possible neutralization of, for instance, the tense or mood values of nominalized verbs. To cope with this, Edmondson & Plank (1978: 394) suggest that “intensification proper rather than reflexivization is the concept that is crucially relevant here”. We will discuss this crucial suggestion in §3 below when we will propose to treat these derivatives as contrastive coreference compounds.

On the other hand, *selbst* competes with other bona fide word-formation (“derivational” in our meaning) means like the prefix *auto-* and the adjective *eigen* ‘own’ which apparently display the same reflexive function although they are clearly not immediately connected to any pronominal function:¹¹

- (4) a. *Eigenlob/Selbstlob* ‘self-praise’, *Eigensucht/Selbstsucht* ‘egoism’
 b. *Autobiographie/Selbstbiographie* ‘autobiography’, *Autohypnose/Selbsthypnose* ‘autohypnosis’

Clearly, in these cases no anaphorical relation can be assumed. Again, König (2011: 114) is well aware of this, but instead of distinguishing the grammatical (i.e. pronominal) and the lexical perspective he rather prefers to assume that “[i]n addition to varying degrees of lexicalization, there may be several types and layers of such reflexive compounds in a language.” In other words, the label “reflexive compounds”, which calls into play the reflexive function typically assigned to pronouns, is stretched so much to cover cases which are essentially different from reflexivity and are rather encompassed by intensification to which we turn in the next section.

3. *Selbst*-compounds as contrastive coreference compounds

Before discussing our approach, we will briefly introduce König’s account which is useful to distinguish different types of *selbst*-compounds. We already mentioned above that König

¹¹ For reasons of space, we have to refer to Angster (to appear) for a further discussion of both lexical types. Very briefly, while *auto-* is strongly restricted to the non-native stratum of the lexicon, *eigen-* appears to display a wider spectrum of productivity which also covers non-argumental nouns like *Eigentor* ‘own goal’ and rather refers to an attributive-possessive interpretation.

(2011: 122) considers *selbst*-compounds to be mainly based on transitive verbs, while those derived from intransitive bases are deemed as marginal: we will come back to this issue below. On the basis of transitive verbs, König – relying on a terminology which refers to the typological literature on intensifiers and focus particles¹² – generally distinguishes between two main types (see more generally on intensification König & Siemund 1999). The first one is called “exclusive adverbial type” insofar as it replicates the value shown by the focus particle *selbst* used adverbially with regard to the VP containing the reflexive marker *sich* as in (5) where its syntactic independence is emphasized by the possibility of interposing the phasal adverb *normalerweise* and the negation *nicht*:

- (5) a. *Dieses Dorf [verwaltet sich (normalerweise)] selbst.*
 ‘This village has self-administration.’
- b. *Dieses Dorf verwaltet sich nicht selbst.*
 ‘This village does not have self-administration.’

The focus particle *selbst* elicits an exclusive interpretation of the subject agent insofar as it excludes other possible agents, including the speaker. This stands in neat contrast to the inclusive interpretation obtained in the following sentence in which the other possible agents are included:

- (6) *Die Landesregierung kann dem Dorf nicht vorwerfen, dass es sich nicht verwaltet, denn sie tut es selbst nicht.*
 ‘The provincial government can’t blame the village for not having self-administration because it doesn’t have it itself.’

In correspondence of the exclusive interpretation in (5), in the expression containing the *selbst*-compound *die Selbstverwaltung des Dorfes* ‘the self-administration of the village’ only an exclusive adverbial interpretation is possible in which it is suggested that *selbst* flags “not only the patient — as one would expect — but also the agent” (König 2011: 119), i.e. that the activity is carried out without help, interference, delegation, and the like.

¹² König and collaborators largely employ the term intensifier and focus particle as synonyms. However, since especially in word-formation intensifiers and intensification is connected with an array of expressions including augmentatives, pejoratives, and the like, we rather prefer to speak of focus particle for *selbst* in order to avoid any possible confusion.

The second type is called adnominal because the focus particle is strictly connected to the reflexive marker (7a) and cannot be separated by a negation (7b):

- (7) a. *Karl kritisiert (oft) [sich selbst].*
 ‘Charles often criticizes HIMSELF.’
- b. *Karl kritisiert nicht sich selbst.*
 ‘Charles is not criticizing HIMSELF.’

In contrast to the adverbial type, in *Selbstkritik* ‘self-criticism’ it is suggested by *selbst* that against any expectation “the agent in activities of criticizing ... also chooses himself as a victim, patient or target” (König 2011: 119).¹³

Accordingly, the two different types profile respectively remarkable agents (8a) and remarkable patients (8b) in correspondence of the related syntactic configurations in which *selbst* serves as focus particle modifying either the verb (‘adverbial intensifier’ in König’s 2011 terms, cf. (5) above) or the reflexive pronoun (‘adnominal intensifier’ in König’s 2011 terms, cf. (7) above):

- (8) a. *Selbstverwaltung* ‘self-administration’, *Selbstzensur* ‘self-censorship’, *Selbstbeziehung* ‘self-accusation’, *Selbstbestimmung* ‘self-determination’, *Selbsthilfe* ‘self-help’, *Selbstachtung* ‘self-respect’, *Selbstentmündigung* ‘self-incapacitation’, *Selbstentleibung* ‘suicide’, *Selbstverpflichtung* ‘voluntary acceptance of an obligation’, *Selbstaufgabe* ‘self-abandonment’, *Selbstbedienung* ‘self-service’
- b. *Selbstkritik* ‘self-criticism’, *Selbstgefälligkeit* ‘complacency’, *Selbstzufriedenheit* ‘self-satisfaction’, *Selbstironie* ‘self-irony, self-mockery’, *Selbstinszenierung* ‘self-fashioning, self-styling’, *Selbstüberschätzung*, ‘overestimate of one’s potential’, *Selbstgespräch* ‘talking to oneself’, *Selbstvertrauen* ‘self-confidence’, *Selbstbefriedigung* ‘masturbation’, *Selbsthass* ‘self-hate’, *Selbstbild* ‘self-image’, *Selbstreflektionen* ‘self-reflections’, *Selbstverständnis* ‘self-image’

¹³ However, as pointed out by an anonymous reviewer, *sich selbst kritisieren* can also be used with the exclusive interpretation in which the negation separates the reflexive and the focus particle: *Karl kritisiert sich nicht selbst, sondern wird von anderen kritisiert* ‘Karl is not criticized (lit. criticizes himself) by himself but by others’. Even the positive sentence can be interpreted in this exclusive manner when *selbst* is appropriately stressed: *Karl kritisiert sich SELBST* ‘Karl criticizes himself by himself’.

While this account copes well with the general function of *selbst* of flagging the unexpected co-reference of the understood typical participants of an activity or an event, respectively a remarkable agent or patient, the neat distinction as maintained by König does not seem to stand an empirical scrutiny. First, far from being marginal, intransitive verbs appear to be largely compatible with *selbst*-compounds as shown by the following examples from the Internet:

- (9) a. *Mir sind kürzlich wieder Selbstlacher über den Weg gelaufen. Also die Spezies, welche eine vermeintlich humorvolle Bemerkung macht und dann selbst lachend den Kopf (leicht in den Nacken gedrückt) rechts - links schwenkend den tosenden Beifall des entgeisterten Publikums erheischen möchte.*

‘Recently, self-laughers crossed my path again. Namely, the species that makes a supposedly humorous remark and then swings its head (slightly pushed back) to the left and right, laughing and wanting to garner the thunderous applause of the aghast audience’.

- b. *Kohl war ein berüchtigter „Selbsttelefonierer“, der sogar als Kanzler Kreisvorsitzende persönlich anwählen konnte.*

‘Kohl was a notorious ‘self-telephone operator’ who, even as chancellor, was able to dial district chairpersons personally.

What is interesting in these examples is the fact that the speakers who coined *Selbstlacher* and *Selbsttelefonierer* described the meaning of their coinage by using either the same focus particle *selbst* or a corresponding expression like *persönlich*. Second, König emphasizes that in the adnominal type human referents and especially patients are involved as a rule; moreover, the latter appear in concomitance with verbs denoting activities or states typically directed away from the agent. The occurrence of unexpected patients co-referencing with the agent is claimed to motivate the alleged non-existence of compounds like **Selbstrasur* ‘self-shaving’, **Selbstvorbereitung* ‘self-preparation’, **Selbtscham* ‘self-shame’ in contrast to the occurring *Selbstgespräch* ‘talking to oneself’, *Selbstmord* ‘suicide’ and *Selbstkontrolle* ‘self-control’.¹⁴ In other words, the adnominal compounds are held to express remarkable

¹⁴ As pointed out by an anonymous reviewer and discussed below, these compounds are in fact attested. In the Internet we found examples like: *der erste mechanische Apparat zur Selbstrasur* ‘the first mechanical device for self-shaving’, *Selbstvorbereitung auf die Prüfung in Biologie* ‘self-preparation for the biology exam’, *vererbte Tendenzen zu Angst und Selbtscham* ‘inherited tendencies towards fear and self-shame’, etc.

reflexivity and are therefore particularly suited for verbs denoting “[b]ody care and grooming ... normally performed on oneself rather than on others” as well as for “deverbal nominalizations such as ‘washing’, ‘shaving’, ‘dressing’, ‘preparation’ or ‘defense’ ... normally interpreted in a reflexive sense” (König 2011: 122).

There are exceptions, though, which cry for an explanation. A first group of exceptions come from cases where an adnominal *selbst*-compound is attested in spite of the occurrence of unremarkable patients such as *Selbstenthaarung* ‘self-depilation’, *Selbstrasierer* ‘self-shaver’¹⁵, *Selbstwascher* ‘self-washer’, already attested in the Twenties of the last century as a name for washer, and the established *Selbstverteidigung* ‘self-defense’, which is also mentioned by König. These cases can be easily multiplied, and we will discuss below other examples coming from our investigation. What is worse is that it is not clear in which sense these activities are generally directed towards oneself: one can easily figure out that many such activities are normally directed towards other persons or things (for instance washing, or defending) and even profile traditional professions like barber, beautician, and the like. In fact, *Selbstrasierer* mentioned above appears in a text in which it is normal to have barbers for shaving, but of course this is not the only possible value shown by the compound. Thus, it is not easy to tell whether an activity displays an other-directed character.

A second group of exceptions consists of not uncommon adnominal *selbst*-compounds based on verbs displaying an anti-causative alternation like *Selbstentzündung* ‘self-ignition’, *Heuselbstentzündung* ‘hay-self ignition’, *Datenselbstzerstörung* ‘data self-destruction’, in which no human patient is involved and the entity involved also appears as a modifier. Note that in these compounds the entity involved behaves like a true subject insofar as it can also appear as the unique realized argument, as also maintained by König (2011: 115): “[i]n keeping with the reflexive meaning of the compounds under analysis, however, only one argument can be realized, viz. the argument corresponding to the subject of the underlying verb”. Accordingly, we observe *die Selbstentzündung von Heu*, *die Selbstzerstörung der Daten*, etc. What these cases have in common with the previous examples is the effect

¹⁵ Example from the Internet: *dass er sich selbst dann und nur dann rasieren könnte, wenn er kein ‚Selbstrasierer‘ wäre* ‘that he himself then and only then could shave if he were not a ‘self-shaver’’. Again, note the occurrence of the reflexive marker *sich* reinforced by *selbst*.

crucially connected with the unexpected occurrence of a patient which is depicted as activating the process independently of external forces. Notice that it is basically our world knowledge that forces a different interpretation of compounds like *Blutdruckselbstmessung* ‘blood pressure self-measurement’, *Zyklusselbstbeobachtung* ‘(menstrual) cycle self-observation’, where the adverbial interpretation is due to the availability of a reading by which a person (doctor or nurse) – and not an automatic machine operated by a patient – normally carries out the operation. On the other hand, in such adverbial interpretation it is possible to realize as an argument the object of the verb, as shown by the following examples from the Internet:

- (10) a. *Nur die Selbstbeobachtung des Zyklus gibt Aufschluss.*
 ‘Only the self-observation of the cycle gives information’.
- b. *Für die Selbstmessung des Blutdrucks stehen Ihnen verschiedene Geräte zur Auswahl.*
 ‘There are various devices to choose from for measuring your blood pressure yourself’

This substantially restricts König’s claim about the realization of the subject argument because in these examples both the subject and the object are overtly expressed: the subject agent is referred to by *selbst* in its adverbial exclusive interpretation and the object patient is expressed by the object genitive. At the same time these examples show the contrastive value of *selbst* in compounds as hinting at an unexpected referent. Finally, it has to be stressed that the same *selbst*-compound can be interpreted as adnominal or as adverbial, depending on the context. To make one clear example based on one of the body-care verbs which are claimed to be incompatible, namely *waschen* ‘to wash’, we find the following examples in the Internet, where *Selbstwaschung* is clearly adnominal in (11a) and adverbial in (11b):

- (11) a. *Das liest sich nicht wie eine politische Analyse, sondern wie die rituelle Selbstwaschung eines ehemaligen Autonomen und Geisteswissenschafts-Studenten, der völlig verzweifelt seine eigene Vergangenheit ausradieren möchte.*
 ‘This reads not like a political analysis, but like the ritual self-washing of a former autonomist and liberal arts student who is desperate to erase his own past’.

- b. *Das während der Bestrahlung austretende Serum wurde als Selbstwaschung der Wunde betrachtet.*

‘The serum that escaped during irradiation was considered to be self-washing of the wound’.

For these reasons, instead of speaking of reflexive nominal compounds we prefer to use the label of contrastive coreference compounds (cf. Angster to appear), in which any reference to reflexivity is avoided while the focus is put on the context-related contrastive effect of *selbst*. Instead, the term contrastive coreference points to the general function of *selbst* of referring to an unexpected (and therefore contrastive) agent/subject or patient/object involved as participant in the activity or the event depicted by the deverbal head.

As a general frame to approach compounding, we will conveniently rely on Scalise & Bisetto’s (2009) classification in which three basic types of compounds are assumed:

Tab. 2: Scalise & Bisetto’s (2009) classification of compounds

Compounds					
Subordinate		Attributive		Coordinate	
Endocentric	Exocentric	Endocentric	Exocentric	Endocentric	Exocentric
<i>Nachttisch</i> ‘night table’	<i>Taugenichts</i> ‘good-for-nothing’	<i>Rotwein</i> ‘red wine’	<i>Rothaut</i> ‘redskin’	<i>Hosenrock</i> ‘pant skirt’	[<i>Mutter-Kind</i>]- <i>Kur</i> ‘mother-child-cure’

Except for exocentric subordinate compounds going back to frozen sentences like *Tauge-nichts* ‘lit. be.good.for-nothing’, in German compounds exocentricity is generally based on the referential properties of the compounds which need not coincide with the right-most constituent providing the morphological head. Accordingly, *Rothaut* ‘redskin’ is exocentric although its morphological head is *Haut* which provides the general morphological properties of the compound, namely [+ feminine, plural: ‘-e’]. Note that exocentric coordinate compounds are normally contained as modifiers in endocentric subordinate compounds like [*Mutter-Kind*]-*Kur* ‘mother-child-cure’, [*Hals-Nasen-Ohren*]-*Arzt* ‘neck-nose-ear-doctor’, etc. (cf. Fleischer & Barz 2012: 70). In the next section, we will discuss the constructional typology of *selbst*-compounds as instances of the more general category of contrastive coreference compounds as they are found in a large text corpus. They are normally, although – as we will see in §5 below – not exclusively, endocentric subordinate

compounds. In particular, since they contain a deverbal head, they are labeled synthetic or verbal-nexus subordinate compounds by Scalise & Bisetto (2009). The synthetic compounds contrast with non-synthetic or ground compounds because their head-modifier relation reflect or correspond to an argument relation while this is not true for ground compounds where the modifier provides the ground against which the head is interpreted, as in the pair *Taxifahrer* ‘taxi driver’ vs. *Sonntagsfahrer* ‘Sunday driver’ (see Baroni *et al.* 2007, Gaeta & Zeldes 2012). In the next sections we will focus on synthetic compounds because they are normally held to unveil argument relations which are crucial for our understanding of the role of the particle *selbst* as a marker of contrastive coreference.

4. Extracting and filtering the corpus data

In this section we will briefly sketch the procedure adopted to gather examples of synthetic compounds including *selbst*-compounds as their right constituent from the corpus *deTenTen13* which is a 16.5 billion-token web-based German corpus of the *TenTen Corpus Family* (cf. Jakubiček *et al.* 2013) available on the platform *SketchEngine*: <https://www.sketch-engine.eu/>. An example of this type of formation is reported in (12a).

- (12) a. *Steuer-selbst-bemessung*
 tax-self-calculation
 ‘self-calculation of the taxes’
- b. [[X]_α [*selbst* [Y]_β]

In (12b) a hierarchical structure representing a general schema for synthetic *selbst*-compounds (henceforth SSC) is reported which constitutes the object of our investigation. The schema in (12b) is a rough simplification of SSCs identifying the flat structure queried through the corpus in search of SSCs. In this flat structure <X*selbst*Y> the unspecified strings <X> and <Y> correspond to any case-insensitive character. Based on this flat sequence of characters, the corpus *deTenTen13* was queried using the CQL query reported in Tab. 3 along with the resulting number of tokens. We report in the same table also the number of tokens obtained from a similar query intended to yield potential *selbst*-compounds.

Tab. 3: Potential *selbst*-compounds occurring in the *deTenTen13* corpus

target	query	tokens
potential synthetic <i>selbst</i> -compounds	[word="(?)?.+selbst.+"]	157 852
<i>selbst</i> -compounds	[word="(?)selbst.+"]	7 228 869

The results of the two queries do not overlap, i.e. no string yielded by the first query can also be the output of the second one and vice versa. The figures in Tab. 3 show that the set potentially containing synthetic *selbst*-compounds is essentially much more restricted than the other containing *selbst* as initial string. This is not unexpected given that synthetic *selbst*-compounds imply *selbst*-compounds insofar as a complex word derived by another (complex) word is normally less frequent than its base.

To explore the results of the query targeting synthetic *selbst*-compounds (henceforth “the SSC-query”) we extracted an exploratory sample containing 1000 examples (0.63% of the results) and built a frequency list of the types yielded by the query. This list of types does not necessarily provide different lexemes: it can simply provide different strings that a simple tokenizer would set apart, henceforth referred to as “string-types”. For instance, we collected 17 different string-types containing the prefix *un-* which mostly are different morphological or graphematic forms of the adjective *selbstständig* ‘independent’, of its deadjectival noun *Selbstständigkeit* ‘independence’ and of the adjective *selbstbewusst* ‘self-confident’. Accordingly, the 17 different string-types can be lumped together providing the three lexical types: *unselbstständig*, *Unselbstständigkeit* and *unselbstbewusst*, independently of any variation.

The results provided by the exploratory sample are quite noisy and a substantial amount of the types must be filtered. In Tab. 4 we summarize the typology of undesired results which is largely based on the analysis of the type of <X> in the flat structure <XselbstY>.

Tab. 4: Typology of results extracted from the *deTenTen13* corpus

Tag	Type frequency	Token frequency	Properties
URL	24	24	noise: URLs containing <i>selbst</i>
NEG	40	236	negation: <i>un-</i> , <i>nicht-</i>
PFX	33	226	prefixes: <i>ver-</i> , <i>ent-</i>
SICH	23	23	(prepositional) phrases containing <i>sich selbst</i> and modifying Y
/	374	491	results including synthetic <i>selbst</i> compounds
total	494	1000	

The tag URL identifies the strings constituting a URL in which by accident the string <selbst> is included and constitute a type of noise not infrequent in web-based corpora such as the *deTenTen13*. The following tags identify results that are undesired as far as the scope of this paper is concerned, but that not necessarily are uninteresting from a linguistic point of view. NEG identifies the strings in which a negation precedes a *selbst*-compound. In the exploratory sample we obtained 23 types and 174 tokens prefixed with the negative prefix *un-* (e.g. *unselbst(st)ändig* ‘unindependent’, *unselbstbewusst* ‘unconfident’) and 17 types and 62 occurrences of *selbst*-compounds modified by the syntactic negation *nicht* (e.g. *nichtselbst(st)ändig* ‘not independent’, *nicht-selbstblockierend* ‘not self-blocking’). Together they constitute 8.1% of the types and 23.6% of the tokens of the exploratory sample.

The tag PFX identifies strings in which a verbal prefix precedes the string *selbst*, which in this case is part of a verb formed by means of so-called prefix conversion (see Elsen 2014: 217) or parasynthesis (see Iacobini 2020 for this term). The latter identifies derivatives in which a lexical base is converted into a verb in the absence of any suffix, but with concomitant occurrence of a prefix.¹⁶ In the exploratory sample only two prefixes occur: *ver-* and *ent-*. The prefix *ver-* is found in a series of word forms the verb *verselbstständigen* ‘to become

¹⁶ The alternative view assuming that in derivatives like *verselbstständigen* and *entselben* the prefix constitutes the head and has to be made responsible of the derivational process, as suggested for instance by Elsen (2014: 217), has to be rejected for two main reasons. Besides the consistent productivity of the pattern, these prefixes are also found as modifiers of verbal bases like in *kaufen* ‘to buy’ / *verkaufen* ‘to sell’, *laden* ‘to load’ / *entladen* ‘to unload’, etc. where the appearance of the prefix does not alter the lexical properties of the base, for instance the inflectional class. The same can be repeated for so-called verb particles, which display the additional property of separability: *Ufer* ‘shore’ / *ausuferen* ‘to overflow’, *nehmen* ‘to take’ / *ausnehmen* ‘to exempt’. Thus, we should assume that verb prefixes and particles behave in different ways with regard to the crucial property of headedness in dependence of the lexical bases they are combined with, which is generally not the case in German where prefixes do not display head properties. For instance, prefixes like *ur-* and *un-* combine both with nouns and adjectives but don’t display the property of headedness.

independent' or in compounds in which the action noun *Verselbstständigung* 'gaining of independence' is present as a modifier – e.g. *Verselbstständigungsprozess* 'process of gaining of independence'. The action noun *Entselbstung* from the verb *entselbsten* 'to deprive oneself of the self' is the only other type and occurrence to which we attributed this tag. Overall, the results identified with the tag PFX constitute 6.7% of types and 22.6% of the occurrences of the exploratory sample.

The last tag identifying noise in the exploratory sample refers to string-types containing the reflexive pronoun *sich* and is probably also the most interesting, though outside the scope of the present contribution. The results tagged with this label are compounds in which a phrase constitutes the modifier. The phrase includes *selbst* in its use as focus particle added to the reflexive pronoun *sich* such as in *Sich-selbst-ausprobieren* '(the) trying out of oneself'. The heavy reflexive *sich selbst* can also be included into a prepositional phrase, as in *Über-sich-selbst-lustigmachen* '(the) making fun of oneself'. At any rate, only 4.7% of types and 2.3% of occurrences correspond to these phrase-modified compounds.

The remaining examples (75.7% of the types and 49.1% of the occurrences) constitute the object of our analysis and the best candidates for being instances of synthetic *selbst*-compounds. To enhance the significance of the query results we extracted a new 1000 examples-sample (henceforth "the working sample") after filtering the SSC query excluding the groups of undesired results summarized in Tab. 4 above. After this filtering we obtain 80 114 occurrences from which we extract another random sample of 1000 occurrences corresponding to 1.25% of the filtered results. At the end, this final sample is almost twice more representative compared to the exploratory 1000 example-sample.

The frequency list obtained from the working sample yields 703 string-types out of 1000 tokens (type/token ration 0.7), showing a higher variety compared to the exploratory sample (0.5). This is not surprising considering the figures in Tab. 4, where the groups of results NEG and PFX, filtered out of the working sample, were also the ones with the lowest type/token ratio (0.17 and 0.15 respectively).

The working sample needed some further filtering, mainly to exclude dozens of typos, after which we got 655 string-types corresponding to 947 tokens. In this additional filtering process, we also lumped together string-types having the same lexical elements in <X> and <Y> (i.e. before and after <*selbst*> in the flat structure), but diverging because of the

presence of hyphens (especially as far as X is concerned), or because they display different word-forms (especially in <Y>), as in *Alkoholselbsttest* ‘alcohol-self-test’ vs. *Alkohol-Selbsttests* ‘alcohol-self-test (NOM.PL / GEN.SG)’ in which <X> is *Alkohol* and <Y> is *Test*.

Concerning the part of the compounds corresponding to <Y> in the flat structure, we also counted as a single lexical type all <Y>s in which a lexeme is the base of the *selbst*-compound, and additional, more external layers of compounding are found with the same base:

- (13) *Internet-Selbstlernzentrum* ‘internet self-learning center’
Internet-Selbstlernmodule ‘internet self-learning modules’
Online-Selbstlernmodule ‘online self-learning modules’
Online-Selbstlernen ‘online self-learning’
Online-Selbstlernphasen ‘phases of online self-learning’

All examples in (13) are labeled as containing the same string <lern*> followed by further lexical material, with the exception of *Online-Selbstlernen* which is labeled as <lernen>. In this way we reduce the variety of labels at this stage of the analysis so that we can more easily survey the phenomenon. This choice is also based on the assumption that the right-most lexemes in <Y> (e.g. *Phase* or *Zentrum*) are structurally more external than the lexemes in <X>. The opposite is also possible: e.g. *Modul* in *Online-Selbstlernmodul* ‘online self-learning module’ is likelier to be more deeply embedded than *online* or *Internet*. We will discuss the different possible patterns of embedding of the <X>s and <Y>s in the next section.

Based on this procedure, we identified 510 lexical types obtained by the combination of 451 different <X>s and 147 different <Y>s (or the first lexeme in complex <Y>s). The most frequent <X>s found in series of lexical types are listed in Tab. 5 along with their string-type frequency:

Tab. 5: Most frequent lexical types of <X> in the sample

Type of <X>	Lexical type frequency	String-type frequency
<i>Online-</i>	9	13
<i>Schüler-</i>	5	5
<i>Blutzucker-</i>	4	7
<i>Frauen-</i>	4	10

Type of <X>	Lexical type frequency	String-type frequency
<i>Künstler-</i>	4	4
<i>Massen-</i>	4	7
<i>Patienten-</i>	4	8
<i>Teil-</i>	4	7
<i>Internet-</i>	3	4
<i>Fähigkeits-</i>	2	4
<i>Familien-</i>	2	4
<i>Migrant*-</i>	2	9
<i>Hi-Fi-</i>	1	4
<i>Krebs-</i>	1	8
<i>schein-</i>	1	18
<i>Schlaganfall-</i>	1	5
<i>Solo-</i>	1	9
<i>Sucht-</i>	1	7
<i>Vollzeit-</i>	1	4

In Tab. 6 we report the list of the most frequent <Y>s accompanied by their string-type frequency.

Tab. 5: Most frequent lexical types of <Y> in the sample

<Y>	Lexical type frequency	String-type frequency
<i>-hilfe*</i>	97	146
<i>-ständig*</i>	28	65
<i>-verwalt*</i>	22	26
<i>-mord*</i>	21	25
<i>-test</i>	18	20
<i>-bewusst*</i>	15	16
<i>-zünder*</i>	15	17
<i>-bau*</i>	13	16
<i>-behalt*</i>	10	11
<i>-kontrolle*</i>	9	12
<i>-bedienung*</i>	8	8
<i>-verständnis</i>	8	8
<i>-verteidigung*</i>	8	10
<i>-organisation</i>	7	13
<i>-versorger*</i>	7	9
<i>-versuch</i>	7	7
<i>-darstellung</i>	6	6
<i>-verständlich*</i>	6	6
<i>-bild</i>	5	5
<i>-management</i>	5	5

<Y>	Lexical type frequency	String-type frequency
-bestimmung*	4	5
-erfahrung*	4	5
-konzept	4	6
-auskunft	2	5

It can be observed that the number of different <X>s is much bigger than the number of <Y>s. Very few elements found on the left side of <selbst> occur in more than one compound – e.g. *Online-* ‘online’, *Schüler-* ‘pupils’, *Blutzucker-* ‘blood sugar’, *Frauen-* ‘women’s’, *Künstler-* ‘artist’s’, *Massen-* ‘mass’, *Patienten-* ‘patients’, *Teil-* ‘partially’ are the only elements occurring in more than three lexical types. In contrast, the elements on the right side of <selbst> display a higher productivity in terms of lexical types, especially *-hilfe** ‘help’, which occurs in 97 different lexical types. It must be noted that if we had also included the most external elements on the right of *-hilfe* the number of lexical types would have been further increased. For instance, the combination of *Krebs-* ‘cancer’ and *-hilfe** ‘help’ counts as a single lexical type in the tables above, but it actually corresponds to three different compounds: *Krebsselbsthilfe* ‘cancer self-help’, *Krebsselbsthilfegruppe* ‘cancer self-help group’, *Krebsselbsthilfeorganisation* ‘cancer self-help organisation’. Our choice of simplifying the labeling for the right side of SSCs has the advantage to highlight how many modifiers combine with each *selbst*-compound, showing that compounds with *-hilfe** account for 21.5% and the first eight elements in Tab. 6 for 50% of the list of the elements in <X>.

5. Describing the data

In the previous section we have shown the method to extract the 451 lexical types that constitute our candidates for the analysis of SSCs from the *deTenTen13* corpus. In this section we will conduct a qualitative survey of the data and we will provide examples of the types of compounds yielded by the flat structure <XselbstY>. It is worth stressing that – by turning from the computation of strings in a flat structure to the analysis of different structural configurations – we are going to model the compounds in hierarchical structures of

the kind shown in (12b) above. This implies that not always what we have found as <X> or <Y> in the flat structure corresponds to [X]_α or [Y]_β in the hierarchical structure.¹⁷

Among the 451 lexical types, i.e. compounds, we have nouns (14a), but also adjectives (14b), present (14c) and past participles (14d).¹⁸

- (14) a. *Maschinen-selbst-diagnose* ‘engine self-diagnosis’; *Speisen-selbst-findung* ‘food self-discovery’; *Perioden-selbst-kosten* ‘prime costs of (a) period’
- b. *IT-selbst-ständig* ‘IT-self-employed’ *kriegsfähig-selbst-ständig* ‘independent and capable of conducting a war’, *schein-selbst-ständig* ‘apparently independent’; *ewig-selbst-verständlich* ‘eternally obvious’; *aggressiv-selbst-bewusst* ‘aggressive-self-confident’, *sprach-selbst-bewusst* ‘language-wise self-confident’, *super-selbst-bewusst* ‘super self-confident’
- c. **8cm-selbst-haftende** *Flex-Schablonen* ‘8cm self-adhesive flex-patterns’; **dauerhaft-selbst-kleband** ‘durable self-adhesive’; **lediglich-selbst-redend** ‘simply obviously’; *die Jeju Sonder-Selbst-verwaltende* *Provinz* ‘the special self-managing province Jeju’
- d. *das hoch-selbst-gelobte* *Klimaprogramm der Grünen* ‘the highly self-appraised program of the Green (Party)’

As expected, nouns constitute the vast majority of all lexical types: to give a rough impression, among the 147 elements in [Y]_β, 49 are action nouns derived with the suffix *-ung*. Adjectives are less frequent, but their character is quite complex. For instance, the compound *selbst-ständig* ‘independent’, contributing with 28 lexical types (6.2%) and 168 occurrences (18.1% {168/928}), can also be used as a noun as in *der Selbständige* ‘the self-employed (worker)’ found in compounds like *IT-selbst-ständige* ‘[Information] T[echnology]-self-employed worker’. As for participles, the examples in (14c–d) are the only ones found in the working sample. Before turning to nouns, it is worth noting that among the

¹⁷ In this section we will use mainly [X]_α and [Y]_β for referring to the constituents linearly immediately on the left and right of *selbst-* in the hierarchical structure, while we will use <X> and <Y> to refer to the mere strings on the left and the right of <*selbst*>, without any claim about their position in the structure.

¹⁸ Concerning past participles, we have filtered out a couple of examples showing the presence in the corpus of compound verbs with *selbst* treated as part of the verb: *ge-selbst-mordet*, lit. ‘self-killed’ and *mit-selbst-verschuldet* ‘self-accused’. Despite being compelling cases, they fall outside the scope of interest of this paper.

adjectival types we find coordinative structures such as *satirisch-selbst-ironisch* ‘satiric-self-ironic’, or the already mentioned examples *aggressiv-selbst-bewusst* ‘aggressive-self-confident’ and *kriegsfähig-selbst-ständig* ‘war-capable-self-reliant’, as well as subordinate ground compounds like *sprach-selbst-bewusst* ‘language-wise self-confident’. Curiously enough, in a couple of cases the coordinative relation involves an adjective and a *selbst*-compound whose base is the same adjective modifying the *selbst*-compound – *mörderisch-selbst-mörderischer Angreifer* ‘murderous-self-murderous attacker’, *kritisch-selbst-kritisch* ‘critical-self-critical’. In the same track, we can also observe the use of relative left-constituents such as in *super-selbst-bewusst* ‘super self-confident’ and the participle *hoch-selbst-gelobt* ‘highly self-appraised’.

Turning our attention to nouns, we will now give a survey of the structures that occur in our sample roughly following the classification of compounds proposed by Scalise & Bisetto (2009) discussed in §3 above, without aiming at an exhaustive account. On the one hand we limit our ambitions because we encountered a wealth of different structures greatly wider than the general schema for SSCs shown above in (12b). On the other hand, more crucially, we will see below that the more complex the structure, the more ambiguous the interpretation of its structure and of the relation between the constituents. It is worth noting that we have not found examples of exocentric compounds in our sample: so the right-most constituent is always both the grammatical and the semantic head.

In our sample we have found some examples of coordinate compounds also among nouns:

- (15) a. *Selbst-erkenntnis-Selbst-liebe* ‘self-knowledge-self-love’
 b. *Selbst-zweifel-Selbst-reflexion* ‘self-doubt-self-reflection’
 c. *[[Kritik]-[Selbst[kritik]]-Bewegung* ‘critic-self-critic movement’

The examples in (15) show how complex the slot <X> and the slot <Y> in the flat structure can be, often hosting more than one lexeme. On the other hand, coordinate compounds are rather unproblematic, since they usually coordinate binary structures, as in (15a–b) or they are embedded as modifiers in a subordinate compound as in (15c), in which case they classify as exocentric coordinate compounds in Scalise & Bisetto’s (2009) terms.

Subordinate compounds, which are the most interesting types of compounds for our analysis, are also more problematic from the point of view of their interpretation and can reach pretty complex levels of embedding. Among subordinate compounds we also have synthetic compounds – verbal-nexus subordinate compounds in Scalise & Bisetto’s (2009) terms. Among subordinate compounds we can identify several subtypes, depending on the nature of the relation between the head – in our sample a *selbst*-compound – and its modifier. The criterion for distinguishing these subtypes resides in whether the modifier, the element preceding <*selbst*> in the flat structure, can be interpreted as an argument of the verbal base of the *selbst*-compound or not. Furthermore, if it can be interpreted as an argument, it has to be understood whether the element preceding <*selbst*> constitutes the internal or the external argument of the deverbal head of the compound.

Before considering these different subtypes, let us consider again the general schema proposed in (12b) above, namely the hierarchical structure serving as a common pattern for subordinate compounds, either of the synthetic or of the ground type. The schema works well as long as <X> and <Y> are constituted by a single lexeme, or when they consist of further compounds inserted in the general schema, as reported in (16) in which a compound is inserted in [X]_α.

- (16) a. [[Z [W]]_x [*selbst* [Y]]
- b. *Liebes-kummer-Selbst-therapie*
 [[love [sickness]] [self [therapy]]]
 ‘self-therapy for lovesickness’

Far harder is finding an example of a compound inserted within [Y]_β. Synthetic *Selbst*-compounds usually have as a head a deverbal noun (or seldom an adjective), but even if we include *selbst*-compounds headed by action nouns derived from compound verbs like *Danksagung* ‘acknowledgement’ from *danksagen* ‘to acknowledge’, *Krankschreibung* ‘sick certificate’ from *krankschreiben* ‘to give a sick note’, etc., we scarcely find examples. In (17) and (18) we report the only clear cases of a complex [Y]_β inserted in the general schema in (12b).

- (17) a. [X [*selbst* [[Z [**W**]_v **suff**]_v]]
 b. *Körper·selbst·wahr·nehm·ung*
 [body [self [**true** [**take**]] **ing**]
 ‘self-perception of the body’

The example in (17) has as a head in [Y]_β an action noun derived from the separable complex verb *wahrnehmen* ‘perceive, be aware of’, consisting of *nehmen* ‘to take’ and *wahr* ‘true’. Both the compound verb and its action noun are fairly frequent words: they occur on average between 10 and 100 times every million tokens of the *Dudenkorpus*.

More complex is the case of the example in (18):

- (18) a. [X [*selbst* [**pref** [Z [**W**]_N]_N **suff**]_v **suff**]_v]
 b. *Verabschiedungs·selbst·be·weih·räuch·er·ung*
 [leave_taking [self [**PFX** [**holy** [**smoke**]] **SFX**] **ing**]
 ‘self-incensation/adulation during a leave taking’

The lexeme *weih* is an obsolete adjective meaning ‘votive, holy’ that is nowadays only found in the verb *weihen* ‘to consecrate, sanctify’, also used as a modifier in compounds. In combination with *Rauch* ‘smoke’ it forms the synchronically opaque compound *Weihrauch* ‘incense’ which is subsequently derived into the verb *beweihräuchern* ‘to incense, adulate’ whose action noun forms the [Y]_β of the *selbst*-compound. Therefore, the [Y]_β is derived from a verb like *beweihräuchern* that is not even a compound verb in contrast to *wahrnehmen* above.

However, in many cases the compounds are even more complex; to be sure, <X> and [X]_α, and especially <Y> and [Y]_β do not overlap. As observed in §4 above, the right-most lexemes in complex <Y>s (e.g. *Phase* in [[*Online*[*selbst*[*lern*]]]*phase*] ‘online-self-learning phase’) are also held to be structurally more external than [X]_α, i.e. the lexemes immediately preceding <*selbst*>. This assumption has helped us with the labelling of the data, but does not hold once the data are analyzed more closely. Consider the compounds in (19):

- (19) a. [*Migranten*[*selbst*[*hilfe*]]]
 ‘migrants’ self-help’
 b. [*Migranten*[*selbst*[*organisation*]]]
 ‘migrants’ self-organization’

Both (19a) and (19b) comply with the general structure in which *Migranten-* ‘migrants’ respectively corresponds to $[X]_{\alpha}$ and *-hilfe* ‘help’ and *-organisation* ‘organization’ to $[Y]_{\beta}$. However, when longer <Y>s are found as in (20a) – which is a rather frequent case for compounds containing *Hilfe* ‘help’ – two possible analyses are equally available: one in which *-organisation* is more external than *Migranten-* – see (20b) – and one in which *Migranten-* is more external – (20c).

- (20) a. *Migranten.selbst.hilfe.organisationen*
migrants-self-help-organizations
- b. $[[X [\textit{selbst} [Y]]] Z]$
‘organizations of migrants’ self-help’
- c. $[X [[\textit{selbst} [Y]] Z]]$
‘migrants’ organizations of self-help’

This affects both the interpretation of the meaning of the compound and the type of relation between the constituents. In case $[X]_{\alpha}$ is more internal as in (20b), then it can be interpreted as the subject or the (indirect) object of the underlying verb *helfen* ‘help’ and the resulting synthetic compound modifies *organisation* forming a larger ground compound. On the contrary, in case $[Y]_{\beta}$ is more internal as in (20c), then the *selbst*-compound only modifies *organisation* and we have two successive cycles of ground compounds. In Fig. 1 we exemplify the combinatory richness of the compounds containing *Selbst-hilfe* ‘self-help’:

[X]	[Y]	[Z]
Blinden-		-bewegung
AIDS-		-einrichtung
Sucht-		-fachverband
Suizid-		-forum
Diabetes-		-gruppe
Amputierten-	[selbst [hilfe]]	-netzwerk
Krebs-		-organisation
Jugend-		-projekt
Patienten-		-verein(igung)
Fahrrad-		-werkstatt
Tibeter-		-zentrum

Fig. 1: Combinatory richness of the compounds containing *Selbsthilfe* ‘self-help’

It is worth noting that the lexemes in [Z] – with a couple of exceptions, see below – exhaust the list of lexemes attested in the working sample in this position, while the ones in [X] are far more numerous. Each line in the figure corresponds to an attested compound – e.g. *Blinden·selbst·hilfe·Bewegung* ‘movement for the self-help of blind people’. We can consider the interpretation provided in (20b) to be the more frequent or preferred one even if the other interpretation is equally possible. This is because on the one hand also the compound without the addition of [Z] as external cycle of compounding is attested (e.g. *Diabetes·selbst·hilfe* ‘self-help for diabetis’), and on the other the external cycle [Z] can also consist of other lexemes favoring the interpretation in (20b) (e.g. *AIDS·Selbst·hilfe·gruppe* ‘group for AIDS self-help’, *AIDS·Selbst·hilfe·projekt* ‘project for AIDS self-help’ along with *AIDS·Selbst·hilfe·einrichtung* ‘institution for AIDS self-help’).

The exceptions mentioned above are compounds in which [Z] is clearly more internal than [X] and the compound must be interpreted as having the interpretation in (20c). They are exemplified in Fig. 2:

[X]	[Y]	[Z]
Open-Air-		-tag
Bundes-	[selbst [hilfe]]	-verband

Fig. 2: Ground compounds containing *Selbsthilfe* ‘self-help’

In the compounds in Fig. 2 – *Open-Air-Selbst-hilfe-tag* ‘open-air self-help day’ and *Bundes-selbst-hilfe-verband* ‘federal self-help association’ – the element in [X] is hardly interpreted as an argument of *Hilfe* or as defining the domain in which the action of self-help occurs, but can generally be intended as further specifying the type of *Tag* ‘(commemoration) day’ or of *Verband* ‘association’ already constrained by the compound *Selbsthilfe* ‘self-help’.¹⁹

In the survey of the possible structures underlying the <XselbstY> flat structure, we have mentioned different types of relations among the constituents of subordinate compounds. The main criterion resides in the possibility to recognize the constituent in [X] as an argument of the verb from which the constituent in [Y] is derived. In the example of *Migranten-selbst-hilfe*, the situation described by the compound should be interpreted in the sense that migrants autonomously help other migrants and themselves. König (2011) suggests for *Selbsthilfe* an analysis as an example of adverbial reflexive compound highlighting a remarkable agent – it is remarkable that migrants autonomously engage for improving their conditions rather than being helped by some external helper.

An opposite interpretation applies to action nouns such as *Selbst-zerstörung* ‘self-destruction’ in which, contrary to expectation, the patient, instead of being a (inanimate) participant different from the agent, it is the agent itself: someone (or something) performs an action of destruction on itself. Other examples of this can be *Selbst-verwaltung* ‘self-administration’, or *Selbst-messung* ‘self-measurement’. Concerning this last example, in our sample we found the case of *Blutzucker-selbst-messung* ‘blood sugar self-measurement’. The interpretation of *Selbst-messung* as ‘the action of measuring oneself’, theoretically possible for the *selbst*-compound, is made unavailable by the occurrence of *Blutzucker* ‘blood sugar’ as modifier: *selbst* can only refer to the person using the device who takes note of its results measuring the level of blood sugar. In other words, *selbst* refers to the unexpected agent in König’s sense discussed in §3 above performing the measurement and serves as adverbial

¹⁹ An anonymous reviewer points out that “[t]here is nothing in the structure of *Bundes-selbst-hilfe-verband* that forbids an interpretation as *Verband, in dem der Bund sich selbst hilft (statt auf die Länder zu vertrauen)* [‘association in which the Federation helps itself (instead of relying on the individual states of the federation)’]. It is more a question of context and world knowledge”. This case is in fact similar to that of the supposed absence of compounds such as *Selbstscham* ‘self-shame’ or *Selbstrasur* ‘self-shave’ seen above: an adequate context can license the existence of certain compounds or readings, but some readings nonetheless appear as preferred.

focus particle with an exclusive interpretation. On the other hand, the modifier fills the patient argument of the verb: it is clearly not coreferential with the agent and cannot accordingly be (co-)referred to by *selbst*.

However, the shift in interpretation from adnominal to adverbial reflexive compound in König's (2011) terms discussed in §3 above shows that this classification, which should depend on the semantic and pragmatic properties of the predicates underlying the abstract nouns heading the *selbst*-compounds, can be overturned by the addition of further participants, both via syntactic adjunction (see e.g. *Wir empfehlen bei solchen Jugendlichen häufige Selbstmessungen des Blutdrucks* 'we recommend for these young people frequent self-measurements of the blood pressure') and in word-formation. On the other hand, it also shows that *selbst*- does not necessarily saturate arguments of the underlying predicate as a reflexive pronoun would do.

In this vein, among the cases of the adverbial use of the focus particle we can also find examples in which a reflexive interpretation is completely excluded. Consider the examples in (21):

- (21) a. *Dampflokomotiv-selbstbau*
 'self-construction of steam locomotive(s)'
 b. *Hi-Fi-selbstbau*
 'self-construction of Hi-Fi (stereos)'
 c. *Rahmen-selbstbau*
 'self-construction of picture frame(s)'
 d. *der Selbstbau von Radio- und Funkempfängern mit Minilautsprechern*
 'the self-construction of radio-receivers and emitter-receivers with mini loudspeakers'

It is evident that the action of building steam locomotives, stereos, frames and other devices is not performed by the products themselves. At the same time, what we expect from a situation described by the compound *Selbstbau* is that the object that is being built is not produced by a craftsman or in a factory, but by a final user who decides to build the object on his/her own rather than to buy it. In this sense we have an unexpected, remarkable agent, which is however coreferential with a non-argumental beneficiary part of the wider frame of the predicate.

Even more striking is the case of the example in (22):

- (22) *Heu-selbst-entzündung*
 hay-self-ignition
 ‘self-ignition of the hay’

The verb *entzünden* ‘to ignite’ is a verb with a typically anti-causative alternation between a transitive argument structure (23a) and an intransitive (anti-causative) argument structure marked by the reflexive marker *sich* (23b) (examples from the *Dudenreferenzkorpus*):

- (23) a. *Viele verschiedene Menschen [...] entzündeten ihre Kerze am Altar.*
 ‘a lot of different people lit their candle at the altar’
 b. *das Heu hat sich [von selbst] entzündet*
 ‘the hay ignited [by itself]’

In *Heu-selbst-entzündung* ‘self-ignition’, *selbst-* can’t be interpreted as reflexive, because the hay does not qualify as a remarkable agent, but only as the patient of the anti-causative reading of the verb. In this case *selbst-* can only function as a focalizing adverb which stresses the spontaneity of the phenomenon and the absence of any external volitional cause that can be made responsible for the ignition.

6. Towards a constructional account of *selbst*-compounds

On the basis of the wide empirical exemplification provided in §5 above, it seems straightforward to conclude that the role of *selbst* in compounds does not consist in binding a variable, as maintained by König, but rather in focusing on an unexpected responsible of an activity. Accordingly, *selbst* in compounds can be treated as a context-bound instantiation of its main function of focus particle. The alleged reflexive interpretation is therefore to be seen as a side-effect of its value of contrastive co-reference marker which picks up the available potential referents and fulfills in this way its general function of focus particle.

In a constructional framework, we might tentatively represent the general value of *selbst* in compounds in the following terms which adopt the standard compound schema in (24a) as elaborated by Booij (2016):

- (24) a. $\langle [N_i N_j]_{Nk} \leftrightarrow [SEM_j \text{ with relation R to } SEM_i]_k \rangle$
 b. $\langle [N_i [V\dots]_{Nj}]_{Nj} \leftrightarrow [SEM_j \text{ with relation R to } SEM_i]_j \rangle$

In (24b) Booij's standard compound schema is further specified for subordinate synthetic compounds (see Gaeta & Zeldes 2017). In the latter, the relation R is normally interpreted as argumental: this can be expressed by referring to semantic roles like agent, patient, and the like, or to syntactic relations like subject, object, etc. as already discussed in §3 above. Remind the synthetic compound *Taxifahrer* 'taxi driver' (in contrast to the non-synthetic *Sonntagsfahrer* 'Sunday driver') in which the relation R profiles a patient/object. Note that in (24b) $[V\dots]_{Nj}$ represents a deverbal noun derived by means of affixation (suffixation and/or transfixation: see respectively *Selbstentzündung* 'self-ignition' and *Selbsthilfe* 'self-help', or conversion: *Selbstbau* 'self-construction') and provides the head of the synthetic compound. In (25a) the synthetic compound schema of (24b) is combined with the constructional schema of *selbst* (25b):

- (25) a. $\langle [N_i\text{-selbst-}[V\dots]_{Nj}]_{Nj} \leftrightarrow [\text{Contrastive interpretation of } SEM_i \text{ with relation R to } SEM_j]_j \rangle$
 b. $\langle [X_i \dots \text{selbst}]_{XPi} \leftrightarrow [\text{Contrastive interpretation of } SEM_i]_i \rangle$

The constructional schema of *selbst* in (25b) accounts for both the adverbial and the adnominal interpretation of *selbst* as discussed with regard to the examples in (5) and (7) above. In particular, especially when *selbst* is postposed to the modified constituent (see fn. 3 above), a contrastive or exclusive interpretation obtains which generally focuses on an unexpected responsible of an activity. In the example (2) above and reported in (26) below, Helga is portrayed as the unexpected and exclusive responsible of the activity of washing her grandmother's hair, which by inference is normally carried out by someone else, and in first place by her grandmother:

- (26) *Helga hat die Haare ihrer Großmutter selbst gewaschen.*
 'Helga herself washed her grandma's hair.'

The same is true of the example mentioned in fn. 3 above – *Selbst gebackener Kuchen schmeckt am besten* 'Homemade cake tastes best' – in which the arbitrary interpretation of *selbst* focuses on the unexpected and exclusive responsible of the activity of baking cakes

which is normally carried out by professional bakers. The concrete interpretation of *selbst*, i.e. its value of contrastive coreference marker for unexpected agents or patients, has to be contextually specified as already discussed above for König's adverbial or adnominal interpretation – see respectively (5) and (7) above. The inheritance of these syntactic properties into the partially specified morphological schema for subordinate synthetic compounding in (24b) provides evidence for considering the schema in (25a) in terms of a constructional idiom, in which the interpretation is held to take place at an intermediate level between word-formation and syntax (see Booij 2002, 2016). Accordingly, the relation R profiles an unexpected, contrastive agent or patient, in dependence of the verb underlying the deverbal head.

7. Conclusion

To sum up, with the help of a significant sample of concrete examples, we hope to have shown the richness and the variety of German synthetic *selbst*-compounds. They cannot simply be treated as cases of compounds incorporating a reflexive pronoun. Instead, they are likely to be constructional idioms in which the focus particle *selbst* conveys contrastive coreference. The latter can be interpreted in different ways in dependence of the general environment scoping within the compound and outside of it, in the general syntactic context of occurrence. On the other hand, *selbst*-compounds as constructional idioms also show clear word-formation properties, as shown by the possibility of picking up modifiers of an argumental or of a grounding nature. These mixed properties are condensed into their general typologically relevant profile of constructions expressing contrastive coreference at the word-level.

References

- Altmann, Hans. 2011. *Prüfungswissen Wortbildung*. 3. ed. Göttingen: Vandenhoeck & Ruprecht.
- Angster, Marco. to appear. *At the boundaries of word-formation. Contrastive coreference in 30 European languages*.
- Baroni Marco, Vito Pirrelli & Emiliano Guevara. 2007. NN Compounds in Italian: Modelling category, induction and analogical extension. *Lingue e Linguaggio* 6(2). 263–290.
- Becker, Karl Ferdinand. 1836. *Ausführliche deutsche Grammatik als Kommentar der Schulgrammatik. Erster Band*. Frankfurt am Main: Kettembeil.

- Becker, Karl Ferdinand. 1852. *Schulgrammatik der deutschen Sprache*. 7. ed. Frankfurt am Main: Kettembeil.
- Booij, Geert. 2002. Constructional Idioms, Morphology, and the Dutch Lexicon. *Journal of Germanic Linguistics* 14(4). 301–329.
- Booij, Geert. 2016. Construction Morphology. In Andrew Hippisley & Gregory Stump (eds.), *The Cambridge Handbook of Morphology*, 424–448. Cambridge: Cambridge University Press.
- Duden. 2022. = Angelika Wollstein (ed.). 2022. *Duden. Die Grammatik*. 10. ed. Berlin: Dudenverlag.
- Edmondson, Jerold A. & Frans Plank. 1978. Great Expectations: An Intensive Self Analysis. *Linguistics and Philosophy* 2(3). 373–413.
- Eichinger, Ludwig M. 2000. *Deutsche Wortbildung. Eine Einführung*. Tübingen: Narr.
- Eisenberg, Peter. 2020. *Grundriss der deutschen Grammatik. Der Satz*. 5. ed. Berlin: Metzler.
- Elsen, Hilke. 2014. *Grundzüge der Morphologie des Deutschen*. 2. ed. Berlin & Boston: De Gruyter.
- Fleischer, Wolfgang & Irmhild Barz. 2012. *Wortbildung der deutschen Gegenwartssprache*. 4. ed. Berlin & Boston: De Gruyter.
- Gaeta, Livio & Amir Zeldes. 2012. Deutsche Komposita zwischen Syntax und Morphologie: Ein korpusbasierter Ansatz. In Livio Gaeta & Barbara Schlücker (eds.), *Das Deutsche als kompositionsfreudige Sprache: Strukturelle Eigenschaften und systembezogene Aspekte*, 197–217. Berlin & New York: De Gruyter.
- Gaeta, Livio & Amir Zeldes. 2017. Between VP and NN: On the Constructional Types of German *-er* Compounds. *Constructions & Frames* 9(1). 1–40.
- Harnisch, Rüdiger & Igor Trost. 2009. Adjektiv. In Elke Hentschel & Petra M. Vogel (eds.), *Deutsche Morphologie*, 17–37. Berlin & New York: De Gruyter.
- Iacobini, Claudio. 2020. Parasyntesis in Morphology. In *Oxford Research Encyclopedia in Linguistics*. Oxford: Oxford University Press.
- Jakubiček, Miloš, Adam Kilgarriff, Vojtěch Kovář, Pavel Rychlý, & Vít Suchomel. 2013. The Ten-Ten Corpus Family. In *7th International Corpus Linguistics Conference CL*, 125–127.
- König, Ekkehard. 2011. Reflexive nominal compounds. *Studies in Language* 35(1). 112–127.
- König, Ekkehard & Peter Siemund. 1999. Intensifiers and reflexives: a typological perspective. In: Zygmunt Frajzyngier & Traci S. Curl (eds.), *Reflexives: Forms and functions*, 41–74. Amsterdam & Philadelphia: Benjamins.
- Leo-Wortgrammatik*. <https://dict.leo.org/grammatik/deutsch/Wortbildungsregeln/Komposition/N-Comp/Pron+N.html?lang=de> (accessed 10 September 2023).
- Leys, Odo. 1973. Bemerkungen zum Reflexivpronomen. In Hugo Moser (ed.), *Linguistische Studien IV. Festgabe für Paul Grebe zum 65. Geburtstag*, 152–157. Düsseldorf: Schwann.
- Lohde, Michael. 2006. *Wortbildung des modernen Deutschen. Ein Lehr- und Übungsbuch*. Tübingen: Narr.
- Metrich, René & Eugène Faucher. 2009. *Wörterbuch deutscher Partikeln. Unter Berücksichtigung ihrer französischen Äquivalente. In Zusammenarbeit mit Jörn Albrecht*. Berlin & New York: De Gruyter.
- Motsch, Wolfgang. 2004. *Deutsche Wortbildung in Grundzügen*. 2. ed. Berlin & New York: De Gruyter.
- Ortner, Lorelies, Elgin Müller-Bollhagen, Hanspeter Ortner, Hans Wellmann, Maria Pümpel-Mader & Hildegard Gärtner. 1991. *Substantivkomposita. Komposita und kompositionsähnliche Strukturen 1*. Berlin & New York: De Gruyter.

Scalise, Sergio & Antonietta Bisetto. 2009. The classification of compounds. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford Handbook of Compounding*, 34–53. Oxford: Oxford University Press.

Siemund, Peter. 2007. Reflexivum. In Ludger Hoffmann (ed.), *Handbuch der deutschen Wortarten*, 707–725. Berlin & New York: De Gruyter.

Marco Angster
University of Zadar
Department of Linguistics
Trg kneza Višeslava 9
HR–23000 Zadar, Croatia
mangster@unizd.hr

Livio Gaeta
University of Turin
Department of Humanities
Via Sant’Ottavio, 20
I–10124 Torino, Italy
livio.gaeta@unito.it



This is an open access publication. This work is licensed under a Creative Commons Attribution CC-BY 4.0 license. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>