English Language and Linguistics 15.2: 303–334. © Cambridge University Press 2011
doi:10.1017/S1360674311000050

English proforms: an alternative account

EVELIEN KEIZER
University of Vienna
(Received 15 June 2010; revised 24 January 2011)

In most theoretical and descriptive treatments of English proforms it seems to be accepted that proforms replace constituents in underlying structure (i.e. phrases or clauses). The aim of the present article is to challenge this assumption. It will be demonstrated that a great many fully acceptable uses of proforms turn out to be quite problematic for the view of proforms as corresponding either to constituents or to semantic and/or syntactic units in underlying representation; nor, it turns out, do proforms necessarily refer to or denote a single (identifiable, retrievable or inferrable) entity. After a brief summary of the relevant literature, the article presents a detailed examination of the actual function and use of English proforms, focusing on a number of frequently used proforms: (i) the indefinite pronoun *one*, (ii) the predicative proform *do so*, (iii) the demonstrative pronouns *that* and *those* and (iv) certain uses of the personal pronouns *we*/*us* and *you.* On the basis of attested examples, it is argued that these proforms do not necessarily express a unit at any level of underlying representation. Instead an alternative account of the use of proforms is suggested, using the theory of Functional Discourse Grammar, which, with its four different levels of analysis (representing pragmatic, semantic, morphosyntactic and phonological information), possesses the kind of flexibility needed to deal with English proforms in a consistent and unified manner. Finally, an attempt is made to explain some of the constraints on the flexible system proposed.

1 Introduction[1]

There seems to be a general consensus among linguists from different theoretical backgrounds that English proforms correspond to semantic or syntactic units of analysis. In formalist theories proforms are typically regarded as replacing constituents in underlying structure (i.e. phrases or clauses; e.g. Burton-Roberts 1986: 55; Radford 1988: 79, 175, 1997: 109; Huddleston & Pullum *et al.* 2002: 1462–3; Aarts & Haegeman 2006: 130; Déchaine & Wiltschko 2003), while non-formalist approaches, including traditional (descriptive) grammars, usually make the somewhat weaker claim that

[1] This article can be considered complementary to an earlier article on the treatment of English pronouns in Functional Discourse Grammar (Keizer to appear) in the sense that it addresses a number of important aspects of this treatment which, due to (understandable) restrictions on length, could not be included in the earlier article. In order for the reader to be able to appreciate the line of reasoning underlying the proposals made in this article, however, many of the data presented in the earlier article have been repeated. The most important additions are section 5.1 (a general introduction to the theory of FDG) and section 6 (constraints on the use of English pronouns). Additions have, however, also been made to other sections (the data set has, for instance, been expanded and further FDG analyses have been added). The two articles together thus provide a more or less complete picture of the particular aspects of pronominalization that form the topic of these articles. At the same time, I have tried to present the material in such a way that both articles can also be read independently.

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

proforms take the place of linguistic (semantic or syntactic) units (including words) (e.g. Wales 1996: 5; Saxena 2006: 131).

The aim of the present article is to challenge this assumption. It will be demonstrated that a great many fully acceptable uses of English proforms turn out to be quite problematic for the generally accepted view of proforms as corresponding either to constituents or to other kinds of semantic and/or syntactic units of analysis. Instead, it will be argued that in order to account for the actual use of pronouns, a flexible system is required, allowing proforms to replace, or correspond to, combinations of elements that do not form a linguistic unit. Authentic examples from a variety of sources will be used to determine the exact circumstances in which the use of a particular English proform is syntactically, semantically and pragmatically acceptable.[2]

Next, it will be argued that the theory of Functional Discourse Grammar (Hengeveld & Mackenzie 2008) offers the kind of flexibility needed to deal with English proforms in a consistent and unified manner, and an FDG treatment will be proposed which is believed to be more in accordance with their actual use. To this end, a brief introduction to the relevant aspect of the theory of FDG will be provided.

Finally, although the flexible system proposed offers a solution for the problems identified, it needs to be constrained in some way, since obviously there are restrictions on which linguistic elements, and which combinations of linguistic elements, can be replaced by a proform. In particular, the fact that in general a verb can be replaced along with its complement(s) needs to be accounted for. An attempt will be made to explain this constraint, as well as some of the other restrictions on the pronouns discussed.

As it will clearly be impossible to present a detailed discussion of all English proforms, the focus in this article will be on a selection of frequently used proforms, including (i) the indefinite pronoun *one*, (ii) the predicative proform *do so*, (iii) the demonstrative pronouns *that* and *those* and (iv) certain uses of the personal pronouns *we*/*us* and *you*.

2 Traditional and theoretical approaches

2.1 Traditional accounts

Traditionally third-person proforms have been regarded as a class of closed items that have 'low semantic content' (Wales 1996: 5), which serve as indicators to the addressee to retrieve the relevant semantic content elsewhere. In addition, proforms are described as 'function words that replace (lexically empty or nearly so) syntactic units of a particular category' (Saxena 2006: 131);[3] they are words 'used instead of a noun

[2] Examples were taken from the British National Corpus (BYU-BNC), the Corpus of Contemporary American English (COCA), and the internet (Google search).

[3] As pointed out by a number of linguists (e.g. Quirk *et al.* 1985: 76, 863; Wales 1996: 1–3), the criterion of substitution is problematic for deictic pronouns, as it is not quite clear what it is they replace. The same problem arises with the generic uses of personal pronouns (Quirk *et al.* 1985: 353). As will become clear in the course of this article, it is important to distinguish the substituting function of proforms (relating to the type of linguistic expression the proform replaces) from their distribution (relating to the position that a proform can take in the construction in which it occurs).

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

or name, to avoid repetition of it' (ibid., quoting Webster's 1913 dictionary). In (1), for instance, the personal pronouns *he* and *it* behave syntactically like NPs,[4] and can be seen as 'replacing' the NPs *John* and *a new car*, while the indefinite pronoun *one* in sentence (2) takes the position of a head noun, 'replacing' the noun *head*:

(1) John has bought a new car. **He** says **he** got **it** cheap.
(2) Can't trade my head in for a new **one**. (COCA)

Proforms are, however, also used outside the nominal domain. In (3a–d), for instance, the proform (*do*) *so* can be used instead of a verbal, adjectival, prepositional and clausal unit, respectively:

(3) (a) They planned to **reach the mountain**, but nobody knows if they **did so**. (Quirk *et al.* 1985: 875)
(b) Prices at present are reasonably **stable**, and will probably remain so. (Quirk *et al.* 1985: 879)
(c) Edward Pitt's name was on his own private list of suspects. But so too was that of Court himself. (BNC)
(d) I think **she'll be there** – I certainly hope so. (Huddleston & Pullum *et al.* 2002: 1463)

It thus seems that proforms can be defined in terms of two basic features (e.g. Huddleston & Pullum *et al.* 2002: 1461):

(a) Proforms are semantically (nearly) empty, relying for their interpretation on information provided elsewhere (the previous discourse or the immediate discourse situation);
(b) Proforms 'replace' ('stand for', 'substitute for') a syntactic unit (the antecedent).

Similar definitions can be found in textbooks on English or linguistic handbooks, including some very recent ones. Some of these emphasize both properties: Leech (2006: 95), for instance, defines a proform as

> a substitute form, that is a word or expression which has no detailed meaning of its own but has the function of 'standing in the place' of another (often more complex) expression.

Other textbooks simply concentrate on the substituting function of proforms. Pearce (2007: 148), for instance, describes the category of proforms as 'items that can stand in for a phrase, clause constituent or clause' (cf. Leech & Svartvik 1994: 56; Finch 2005: 108). A logical next step has been to use replacement by proforms as a test for constituency. Thus, Burton-Roberts (1986: 15) states that

> if you can replace a SEQUENCE OF WORDS in a sentence with a SINGLE WORD without changing the overall structure of the sentence, then that sequence functions as a constituent of the sentence and is therefore a phrase (original emphases).[5]

[4] That is, they function as pro-NPs (or pro-DPs) rather than as pro-nouns (e.g. Quirk *et al.* 1985: 76; Radford 1988: 79; Huddleston & Pullum *et al.* 2002: 1462–3; Déchaine & Wiltschko 2003: 71).

[5] It will be clear from this brief summary that a number of verbs (*replace*, *stand for*, *stand in the place of*, *substitute for*, *used instead of*, etc.) are used as (near) synonyms. In most cases no clear indication is given of what is meant by these verbs (the Webster entry being an exception). It is for this reason that I have placed *replace* in this

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

2.2 *Theoretical approaches*

In generative grammar in particular a great deal of attention has been given to the distribution and analysis of proforms. In most generative approaches, proforms are taken to replace a full phrase (DP, NP, VP, AP, PP, etc.) or a constituent within a phrase (a bar-level or subphrase),[6] resulting, in some cases, in the use of the term 'pro-constituent' (see Radford 1997: 524). Aarts & Haegeman (2006: 130) observe that '[i]n the literature there seems to be a consensus that proforms typically replace constituents, even though this assumption is not always made fully explicit', and proceed to explicitly adopt 'the assumption that a proform replaces a constituent'. Since it is only phrases and bar-levels that form constituents, but not the heads of these phrases and bar-level constituents, the prediction is that proforms can replace a head together with its complement(s) and/or modifier(s), but not the head without its complement(s). Examples like those in (4) and (5) have been used to support this claim:

(4) (a) John **left the office yesterday**; Mary **did so** too.
(b) John **left the office** yesterday; Mary **did so** today.
(c) *John **left** the office, and Mary **did so** the house.
(5) (a) John bought **a big bar of chocolate** yesterday. Mary bought **one** too.
(b) John bought a big **bar of chocolate** yesterday. Mary bought a small **one**.
(c) John bought a bar **of chocolate** yesterday. *Mary bought **one** of soap.

The sequences in (4a, b) and (5a, b) are grammatical: here the predicative proforms *did so* and *one* replace entire phrases (the a-sentences) or bar-level constituents (the b-sentences). In (4c) and (5c), however, the proforms replace just the head, resulting in an ill-formed sentence.

As a consequence of this approach, the idea of using pronominalization as a constituency test also figures prominently in much of the generative literature (see Aarts & Haegeman 2006). Radford (1988: 175), for instance, explicitly states that 'only a unitary constituent can be replaced by a proform – and indeed only a *phrasal* constituent of some sort' (see also Radford 1988: 186, 234, 1997: 109; cf. Van Valin 2001: 121; Déchaine & Wiltschko 2003: 71–89).

3 The use of anaphoric proforms: problematic data

3.1 *Some acknowledged exceptions to the rule*

It is generally acknowledged that there are exceptions to the rule that proforms correspond to units in underlying structure. Some examples are given in (6) (examples

section between quotation marks, and that I prefer the verb *correspond* (indicating a mapping relation between interpersonal (pragmatic) and representational (semantic) units and their morphosyntactic expression).

[6] Much depends on the particular generative model used: in X-bar theory, proforms are regarded as substitutes for full phrases (XP) or constituent levels within phrases (X-bar). In later approaches, which distinguish a far greater number of phrases (not only VPs, but, within the VP, also Progressive Phrases and Perfective Phrases), it is claimed that proforms replace maximal projections only (e.g. Radford 1997: 108–9).

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

from Quirk *et al.* 1985: 352–3; see also Huddleston 1984: 273; Wales 1996: 11, 12; Huddleston & Pullum *et al.* 2002: 430):

(6) (a) we doctors, you people
(b) we who have pledged allegiance to the flag
(c) you there, we here
(d) we of the modern age, you in the raincoat
(e) you yourself, he himself
(f) we all, they each, you both

Here the personal pronouns clearly do not function as entire NPs; nor can they straightforwardly be analysed as heads of the entire NPs, modified by what follows. It is true that occasionally proforms can function as heads (e.g. *Is your dog a he or a she?* (Wales 1996: 10); *Democrats have their hes and shes mixed up* (Internet)), in which case they indeed behave as real nominal heads: unlike the proforms in (6), they can be pluralized and co-occur with a determiner.[7]

A similar problem can be found in constructions with the demonstrative pronouns *that* and *those* (examples – including paraphrases – from Quirk *et al.* 1985: 872):

(7) (a) Towards the end of his life, Schubert wrote two remarkable trios: **that** [= the trio] in B-flat, and that [= the trio] in E-flat.
(b) The paintings of Gauguin's Tahiti period are more famous than **those** [= the ones] he painted in France.

Huddleston & Pullum *et al.* (2002: 1504) propose an analysis in which the demonstratives *that* and *those* in these constructions 'function as fused determiner-head in NP structure'. They are therefore not regarded as pronouns, but as determiners that occur in head position. The same kind of analysis is provided for *one* where it is used to replace an entire indefinite NP, as in (8):

(8) I haven't got a pen; can you lend me **one**? (Huddleston & Pullum *et al.* 2002: 412)

Despite these exceptions, however, the traditional idea of a pronoun as 'a word used instead of a noun' has continued to figure prominently in the various definitions of pronouns. According to Huddleston (1984: 274), for instance, 'there can be no doubt that something along the lines of the traditional criterion should figure prominently in a general definition of pronouns', despite the fact that 'a very significant proportion' of the pronouns he discusses do not satisfy this criterion.

3.2 Pronominalization as a constituency test

Some linguists have expressed doubt about the validity of pronominalization as a universal test for constituency. Broadly speaking, two approaches can be found. The

[7] Several linguists have suggested that, given its distribution, the pronoun in (6a) best be analysed as a determiner (e.g. Huddleston 1984: 233; Wales 1996: 2, 11; Radford 1997: 154–5; Huddleston & Pullum *et al.* 2002: 422; Déchaine & Wiltschko 2003). In Keizer (2010), on the other hand, I argue that these constructions would better be regarded as a type of close apposition (see also section 4.4 below).

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

first is to assume that pronominalization is indeed a universally valid test (together with other tests, such as permutation (extraction), ellipsis and coordination). If the tests fail in some languages but not in others, this is taken as evidence that the constituents involved are not universal. Van Valin (2001: 119ff.), for instance, concludes that the 'VP' 'is not universally valid across the languages of the world': it is only in those languages in which a verb and its complements act as one unit in terms of substitution, permutation and coordination, that we need to assume the presence of a VP (Van Valin 2001: 121–2).

The second approach is to assume that pronominalization is not a universally valid constituency test. An exponent of this view is Croft (2001), who, after looking at the use of proforms in various languages concludes that

> The criteria for constituency described in this section [pronominalization, ellipsis, coordination] fail as cross-linguistically valid criteria for a global constituent structure. The tests do not necessarily exist in all languages. Different test constructions give different results ... A decision to choose one construction over another as a 'test' for constituency is opportunistic. (Croft 2001: 189)

It will be clear that if we cannot find any reliable cross-linguistic criteria for constituency structure, it is no use speculating about which units or levels may or may not be present in all languages. Once again, however, this does not mean that these tests cannot be successfully applied to individual languages. The examples in (4) and (5) might still lead to the conclusion that in English pronominalization can be regarded as a reliable test for constituency. The aim of this article is to show that, given the number and the nature of the exceptions, even this claim is unjustified.

4 Additional data

This section will present a detailed discussion of a number of authentic examples from two large corpora (the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA)), providing further evidence for the view that proforms do not serve to replace semantic or syntactic units, and confirming the need for a different general definition of English proforms.

4.1 The indefinite pronoun one

As pointed out before, the predicative pronoun *one* can represent either just the head of an NP (example (9a)), or the head together with its complement(s) (example (9b)) and/or any modifier(s) (example (9c)):[8]

(9) (a) Well how can well how can go without his **keys**? He's got a spare **one**. (BNC)

[8] The pronoun *one* can, of course, also be used referentially, as in *Cecily met a nice guy yesterday. I met* one *too*. Analysis of this use of *one* is far from straightforward and raises a number of interesting questions (see Huddleston & Pullum *et al.* 2002: 1513–14). Unfortunately, the scope of the present article does not allow me to address these questions here.

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

(b) Science, for instance, does not offer a better or more accurate **description of reality** than fiction, only a different **one**. (COCA)
(c) First, because formal **institutionalized procedures** generally attract more publicity than informal **ones**, … (BNC)

As we have seen, it is generally assumed that the pronoun *one* can replace a head without its modifiers (when present), but not a head without its complements. Huddleston & Pullum *et al.* (2002: 441), however, point out that this claim is too strong, as *one* can be used to replace a head noun without its complement(s) if this head noun is a derived noun. The examples they provide are given in (10), where *one* replaces the deverbal nouns *ban* and *proof*, respectively; in both cases, however, the original complement is replaced by a new one.

(10) (a) I support the **ban** on smoking, but not the **one** on alcohol.
(b) The **proof** of Pythagoras' theorem is more clearly formulated than the **one** of Parseval's equality. (Huddleston & Pullum *et al.* 2002: 441)

Some attested examples are provided in (11): in (11a) *one*, on both occurrences, replaces the deverbal head *challenge* without its complement; in (11b) *one*, replacing the head noun *view*, is provided with its own complement:

(11) (a) The **challenge** of bagging a deer with archery gear is a big **one**, but one that is well worth meeting. (COCA)
(b) These notions were easy enough to compare to Indian **views** of prana and Chinese **ones** of qi. (COCA)

Where the head noun is a non-derived relational noun, however, Huddleston & Pullum *et al.* (2002: 1515–16) maintain that *one* cannot take its own PP-complement (e.g. **The one* [= *king*] *of Belgium*, **The one* [= *sleeve*] *of the dress*). Data from the corpus show that this is, indeed, largely correct, though exceptions can again be found. This is illustrated in (12), where *one* is used to substitute basic relational nouns (see also Keizer 2004, 2007: 239–41):

(12) (a) His dark **eyes** meet the watery blue **ones** of the other in the rearview mirror. (COCA)
(b) whether the frail **backs** of the poor, and the callused **ones** of the middle class, bear the nation's heaviest fiscal baggage. (COCA)

A different kind of problem is illustrated in example (13), which shows that *one* can also replace combinations of elements in underlying structure which do not form a syntactic or semantic unit.

(13) So Paul had a **big** blue **felt marker** for days and a red **one** for nights. (BNC)

Clearly, the pronoun *one* in this example, as replacing the head together with a combination of non-adjacent modifiers,[9] does not replace any unit, or level, within the NP.

[9] In (13), *one* may also be interpreted as being codenotational only with *felt marker*, in which case its use is unproblematic.

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

Yet another difficulty is presented by the examples in (14), where the pronoun *one* picks up part of a compound. Although this may lead to questionable results in some cases (*Is that the morning star? ??No, that's the evening one*), the examples in (14) – which all seem to involve a relatively low degree of conventionalization or lexicalization – seem to be acceptable:

(14) (a) Poetry **class** was just about the only **one** Remi really liked. (COCA)
(b) That's when the jokes about the missing bit started, for ring **doughnuts** look just like traditional **ones** with a bit taken out of the middle. (BNC)
(c) I was expecting a science book and got **a history one** instead. (Internet)

A final challenge to existing theories can be found in the following example:

(15) (a) During the late nineteenth century, black Southerners waged campaigns to increase their control over black colleges and to replace white **administrators and teachers** with black **ones**. (COCA)
(b) For every tragic death that occurs, how many close calls go unreported and how many responsible **skiers and drivers** are scared out of their wits by the reckless **ones**?

In both (15a) and (15b) the plural pronoun *ones* replaces a coordinated unit. To account for this use of *ones*, it may therefore be claimed that the pronoun replaces one head, consisting of two coordinated nouns (and as such a syntactic unit). Semantically speaking, however, coordination at the level of the head (or property) is unlikely: the members of the set referred to by the NP *white administrators and teachers* are not at the same time administrators and teachers. Instead, the set as a whole consists of two subsets: a set of administrators and a set of teachers, with the property 'white' applying to the members of both subsets. On this reading, the proform *ones* replaces the heads of both these subsets – a combination of elements that does not form a separate level within the phrase.

4.2 The predicative proform do so

The use of the predicative proform *do so* in (16), as a pro-VP, is unproblematic for generative accounts:

(16) (a) Americans who ignored this warning clearly **did so** at their own risk and on their own responsibility. (COCA)
(b) The Government wanted to create a property-owning democracy, and they **did so**, . . . (BNC)

The same is true for the examples in (17), where *do so* includes modifiers (*as a means of birth control* in (17a), *critically* in (17b)):

(17) (a) Clearly the legislators who **supported abortion as a means of birth control** and the governors who **did so** are going to be held to account in the November elections. (COCA)
(b) In this study, preservice teachers who had the desire and ability **to reflect critically did so** regardless of the type of feedback they received. (COCA)

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

An example like (18), however, does pose a problem: it shows, as in the case of *one*, that the elements captured by *do so* need not be adjacent:

(18) The teacher expressed concern that, although he could **write** quite well **in English**, he only **did so** when writing collaboratively with his friend (who was absent). (BNC)

4.3 The demonstrative pronouns that *and* those

Like *one*/*ones*, the demonstrative pronouns *that* and *those*, when followed by a PP or clause, can replace a determiner and a head, without any complements or modifiers:

(19) (a) Perhaps the best-known **example** is **that** of the Heel Stone at Stonehenge. (BNC)
(b) It is likely that the Scottish **experience** will be exactly the same as **that** in England. (BNC)
(c) **Students**, especially **those** in high school, too often get the message that the contents of the library are more important than the clients. (COCA)

That and *those* may also include a modifier (example (20a)) or a complement (example (20b)). Predictably, it is also possible for *that* and *those* to cover the complement, but not the modifier (example (20c)):

(20) (a) However, we have no choice about developing an **educational enterprise** as effective as **that** in Germany. (COCA)
(b) This suspicion was only strengthened by **expressions of cultural imperialism** such as **that** in 1896 by Theodor Herzl in *Der Judenstaat*. (COCA)
(c) These findings suggested that the **MHC susceptibility to PV gene** among Iranians derived from the same ancestor as **that** in the Ashkenazi Jews. (COCA)

As it turns out, however, the complement need not be included: *that* and *those* may take their own complement. This is true for derived heads as well as for relational basic heads (examples (21a, b), respectively), including kinship and body-part nouns (examples (21c, d), respectively):

(21) (a) He said staunch Tory newspapers were calling for the **resignation** of Chancellor Norman Lamont –; to follow **that** of David Mellor. (BNC)
(b) Zulema's **life** and **that** of other female drug lords, (COCA)
(c) Our **children** will suffer -–; but not **those** of rich parents who can afford to send them to public schools. (BNC)
(d) While some of us have sought to help, the real weight of responsibility has fallen on his shoulders and **those** of his wife. (BNC)

Moreover, as the following examples show, it is even possible for *that* and *those* to include head and modifier, but not the complement. In (22), *that* corresponds to the head *image* and its modifier *public* (and also covers the definite article; see section 5.3.2.3), but not the original genitival complement (*Party*):[10]

[10] The following example may be taken as further evidence that the second element of a compound may be taken up by a proform, in this case the proform *that*:

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

(22) Since anti-Semitism was such a prominent feature of **the** Party's **public image**, and **that** of its leader, . . . (BNC)

4.4 Personal pronouns as heads

Although the constructions in (24) have been ignored in most studies on apposition,[11] I have argued elsewhere (Keizer 2010) that, since constructions comply with the general definition of close appositions, they are to be analysed as yet another subtype of close apposition. Like other types of close appositions (e.g. *the poet Burns*, *Burns the poet*, *the name Algernon*), they are therefore provided with an internal head–modifier structure (see also Keizer 2007). On such an analysis, the pronouns in these constructions do not form a separate phrase within the overall construction.

(24) (a) In fact it is how **we Brits** are seen by the French. (BNC, newspaper)
(b) '**You English girls** are so cruel', he was saying. (BNC, prose fiction)
(c) 'In every way, you see, I'm Helen Winslow, based on **me the person**, but synthesised and evolved into my present form by the AIM system. I'd prefer you call me Helen, because that's who I am, but I'll accept Annie if you like.' 'But you are an artificial intelligence system,' said Maxwell. (COCA, fiction)

A completely different use of constructions with a pronominal first element can be found in (25), where they are used as vocatives. Once again, however, the personal pronoun *you* in these examples seems to function as a head rather than an entire NP:[12]

(ia) Furthermore, teaching Portuguese in primary school to children who do not speak it at home further isolates school **experience** from **that** of the home. (BNC)
(ib) State **power**, and consequently *that* of the education authorities, may have been eroded, but it still controls formal education. (BNC)

However, as pointed out by two referees, the status of such N + N sequences as *school experience* and *state power* is a matter of debate. If they are treated as (phrasal) modifier–head constructions, the fact that the second element can be taken up by a proform is of course entirely unproblematic.

[11] See, for instance, Curme (1931), Fries (1989), Lee (1952), Haugen (1953), Hockett (1955), Francis (1958), Sopher (1971), Burton-Roberts (1975, 1994), Quirk *et al.* (1985), Acuña-Fariña (1996, 2009), Langacker (2002) and Keizer (2005, 2007). Although a number of different analyses have been proposed, the general consensus is that close appositions share the following formal properties:

(i) they contain two nominal elements;
(ii) these elements form one intonation unit;
(iii) there is no linking element between the elements (as opposed to such constructions as *the city of Rome*, *the name of Algernon*);
(iv) one element is a proper noun or uniquely defining element, the other a count noun.

[12] Other combinations of pronoun and nominal element, such as those illustrated in examples (i) and (ii), are not relevant for the present discussion:

(ia) Get your heavy metals tested, please. I beg **you guys**. (COCA, spoken)
(ib) Hey, **you all** can put that over the table there. (COCA, spoken)
(ic) Ah, but I get I get embarrassed because **you people** are so cruel! (BNC, meeting)
(iia) I called the authorities, and **they people** told me I was crazy. (COCA, spoken, CNN)
(iib) What are you doing with **them boxes**? (BNC, conversation)
(iic) There were only workhouses and like that for some of them. How did the old ones Eh? How did **they old ones** survive then? (BNC, oral history).

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

(25) (a) '**You fools**, what are you doing?' (COCA, magazine)
(b) Clean it yourself, **you lazy whatsit**! (Internet)

5 A brief introduction to FDG

5.1 General characterization and organization of the model

Functional Discourse Grammar (FDG) is a typologically based model of language use that takes into account only linguistically relevant, i.e. formally expressed, aspects of language (Hengeveld & Mackenzie 2008). As the successor to Simon Dik's Functional Grammar (FG; Dik 1978, 1997), it shares many of FG's underlying principles, in particular its strong emphasis on pragmatics, i.e. on language as a form of verbal interaction between a speaker and an addressee, and the idea that form – however indirectly or inconsistently – reflects communicative function. At the same time FDG is primarily an account of linguistic facts, concerned only with those linguistic phenomena that are encoded in the grammar of a language. This means that within the functional paradigm, FDG takes a moderate position: it is characterized by Hengeveld & Mackenzie (2008: 39) as a form-oriented 'function-to-form' approach to grammar; using Van Valin's (1993) terminology, FDG qualifies as a 'structural-functional' theory of language (see Butler 2003: 30–1).

Compared to other functional frameworks, FDG has a number of distinctive features:

(i) FDG has a top-down organization starting with the formulation of the speaker's intention and then working its way down to articulation.
(ii) FDG takes the Discourse Act as its basic unit of analysis. As such, FDG can accommodate regular clauses, as well as units larger or smaller than the clause.
(iii) FDG analyses Discourse Acts in terms of four independent, but interactive levels, dealing with pragmatics, semantics, morphosyntax and phonology.
(iv) FDG systematically interacts with three other components:

THE CONCEPTUAL COMPONENT is seen as the driving force behind the grammatical component; it contains the prelinguistic conceptual information relevant for linguistic analysis.

THE OUTPUT COMPONENT receives its input from the grammatical component and converts it into acoustic, orthographic or signed output.

THE CONTEXTUAL COMPONENT contains information about the immediate discourse context that affects the form of a linguistic utterance.

Figure 1 gives the general architecture of the model.

(iid) They didn't give it to me. They gave it to **they friends**. (COCA, Spoken, Fox)

As argued in Keizer (2010), the first element in these constructions no longer functions as a proform. In (i) the two elements together are best regarded as fixed (grammaticalized) plural forms of the pronoun *you* (cf. Hengeveld & Mackenzie 2008: 70), while in (ii) the third-person plural pronouns can plausibly be analysed as some kind of all-purpose identifiability markers, whereby the exact interpretation can only be determined on the basis of the context.

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

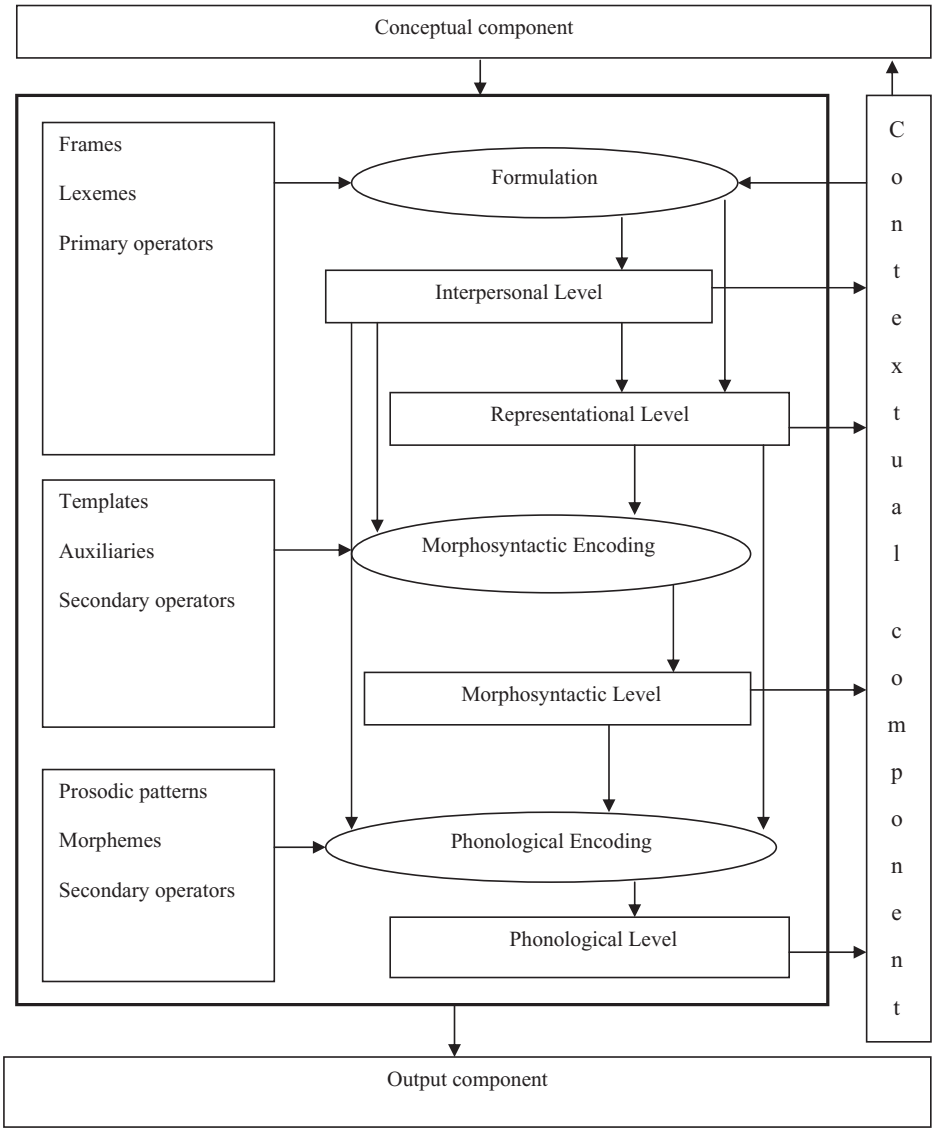


Figure 1. FDG: general layout

This architecture reflects the idea that every linguistic communication starts with some communicative intention at the prelinguistic conceptual level. This information is subsequently fed into the grammatical component where the operation of formulation converts it into interpersonal (pragmatic) and representational (semantic) representations. Next, the operation of encoding translates these into morphosyntactic and phonological representations, which, in turn, feed into the output component.

At each level, the construction of a linguistic expression begins with the selection of a number of language-specific primitives. These primitives come in three kinds. First,

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

there are the structuring primitives, which define the possible combinations of elements at each level (frames or templates). Next, there are the relevant linguistic elements at each level (lexemes, free morphemes and phonemic forms). Finally, each level has its own set of operators for the expression of grammatical information.

The remainder of this section will provide some more details about the first three levels of analysis, as these are relevant for the analysis of proforms to be presented in sections 5.2 and 5.3.

The Interpersonal Level

THE INTERPERSONAL LEVEL deals with all the formal aspects of a linguistic unit that reflect its role in the interaction between speaker and addressee. Some of the relevant units at this level are:

- DISCOURSE ACT (A): the smallest identifiable unit of verbal communication (cf. Kroon 1995: 65).
- ILLOCUTION (F): conveying the speaker's communicative intention.
- SPEECH PARTICIPANTS (P_1 and P_2): the functions of Speaker and Addressee, for the expression of linguistically relevant properties of the speech participants.
- COMMUNICATED CONTENT (C): the totality of what the Speaker evokes in the communication.
- ASCRIPTIVE SUBACT (T): Speaker's evocation of a property (e.g. 'tall', 'eat', 'car')
- REFERENTIAL SUBACT (R): Speaker's evocation of an entity.

Units at this level typically have the (simplified) frame in (26):

(26) $(\pi\ A_1\colon [(\pi\ F_1\colon ILL\ (F_1))\ (P_1)_S\ (P_2)_A\ (C_1\colon [\ldots\ (\pi\ T_1)_\Phi\ (\pi\ R_1)_\Phi\ \ldots]\ (C_1)]\ (A_1))$
where: ILL = illocution (declarative, interrogative)
π = operators applying at the different layers, e.g. approximation at the layer of ascription, identifiability at the layer of reference
Φ = pragmatic function, e.g. Topic, Focus, Emphasis, Contrast

A partially instantiated frame is provided in (27b):

(27) (a) (What did the dog chase?) The dog chased a huge wasp.
(b) $(A_I\colon [(F_I\colon DECL\ (F_I))\ (P_I)_S\ (P_J)_A\ (C_I\colon [(T_I)\ (+id\ R_I\colon (T_J))_{Top}\ (-id\ R_J\colon [(T_K)\ (T_L)])_{Foc}]\ (C_I)]\ (A_1))$

Here we have a declarative Discourse Act (A_I) with a Communicated Content consisting of three subacts: (T_I), representing the evocation of the property 'chase'; (+id R_I: (T_J))$_{Top,}$ representing the evocation of an identifiable and topical entity characterized by another Ascriptive Act, evoking the property 'dog'; and (-id R_J: [(T_K) (T_L)])$_{Foc}$, representing non-identifiable referent, focal entity characterized by two Ascriptive Acts, evoking the properties 'huge' and 'wasp'.

The Representational Level

THE REPRESENTATIONAL LEVEL deals with the semantic aspects of a linguistic unit; i.e. with descriptions of entities as they occur in some non-linguistic world. The units

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

at this level differ with respect to the type of entity they denote; each entity type is represented by a different variable. Some of the relevant units are:

Semantic category	*Variable*	*Examples*
Property	f	colour
Individual	x	chair
State-of-affairs	e	meeting
Propositional Content	p	idea
Location	l	top
Time	t	week

A simplified example of a representational frame can be found in (28):

(28) $(\pi\ e_1: [(f_1: lex_V\ (f_1))\ (x_1)_\Phi\ (x_2)_\Phi]\ (e_1))$
where lex_V = a lexeme of the category verb
π = operators applying at the different layers, e.g. tense at layer of e, number at the layer of x
Φ = semantic functions, e.g. Actor, Undergoer, Location

A partially instantiated frame is provided in (29):

(29) (a) The dog chased a huge wasp.
(b) $(past\ e_i: [(f_i: chase_V\ (f_i))\ (1x_i: (f_j: dog))_A\ (1x_j: (f_k: wasp): (f_l: huge))_U]\ (e_i))$

Here we find a SoA (symbolized by the variable e) which took place in the past, in which two participants (both singular individuals, symbolized by the variable x), with the semantic roles of Actor and Undergoer, are involved in the action of chasing (symbolized by the variable f). Both individuals are described lexically; i.e. the lexemes *dog*, *wasp* and *huge* are used to describe properties of the individuals in question.

The Morphosyntactic Level

THE MORPHOSYNTACTIC LEVEL accounts for the linear and formal properties of a linguistic unit. The major units of analysis of this level are Words, Phrases, Clauses and Linguistic Expressions; the general schema is given in (30):

(30) $(Le_1: [(Cl_1: [(Xw_1)\ (Xp_1: [(Xw_2)\ (Xp_2)\ (Cl_2)]\ (Xp_1))\ (Cl_3)]\ (Cl_1))]\ (Le_1))$
Where Le = Linguistic Expression
Cl = Clause
Xp = Phrase (of the type x)
Xw = Word (of the type x) (Hengeveld & Mackenzie 2008: 291)

The largest unit, that of the Linguistic Expression, can contain one or more Clauses, which, in turn, may consist of one or more Words, Phrases and Clauses. Phrases, too, may contain one or more Words, Phrases or Clauses, etc. The different possible combinations lead to a large number of templates, which can be reduced to a limited

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

number of macrotemplates. A sentence like (31a), for instance, would satisfy the macrotemplate in (31b) (Hengeveld & Mackenzie 2008: 295):

(31) (a) The dog chased a huge wasp.
(b) (Cl_i: [(Np_i: [(Gw_i: the (Gw_i)) (Nw_i: dog (Nw_i))] (Np_i)) (Vp_i: (Vw_i: chased (Vw_i)) (Vp_i)) (Np_j: [(Gw_j: a (Gw_j)) (Ap_i: (Aw_i: huge (Aw_i)) (Ap_i)) (Nw_j: wasp (Nw_j))] (Np_j))] (Cl_i))

As we can see in (31b) a distinction is made between lexical Words and grammatical Words: lexical Words correspond to content elements at the Interpersonal and Representational Levels (lexemes), whereas Grammatical Words express lexically empty elements (i.e. operators; e.g. determiners, auxiliaries etc.).

5.2 Proforms in FDG

5.2.1 General characterization

As in most linguistic theories, in FDG, too, proforms are assumed to correspond to pragmatic or semantic units in underlying representation. This view was more or less inherited from FDG's predecessor, Simon Dik's Functional Grammar, where proforms were assumed to correspond to units in underlying representation. Here, however, the units in question were semantic rather than syntactic units (e.g. Dik 1997: 225–6). The personal pronouns and the indefinite pronoun *one*, for instance, were seen as replacing units designating Individuals (persons or objects), the pronouns *it*, *this* or *that* as replacing units designating states-of-affairs, *so* and *that* as replacing units designating propositions, and *one*, *so* and *such* as replacing units designating properties. It was the introduction of the f-variable to represent Properties that made it possible to represent the anaphoric relation in the case of *one* and (*do*) *so*, as illustrated in the following examples (where A is an operator indicating anaphoricity):

(32) (a) John had bought a blue **car** and I will buy a new **one**.
(b) (i 1 x_i: (f_i: car_N (f_i)) (x_i)$_\phi$: (f_j: $blue_A$ (f_j)) (x_i)$_\phi$)
'a blue car'
(i 1 x_j: (Af_i) (x_j)$_\phi$: (f_k: $green_A$ (f_k)) (x_j)$_\phi$)
'a green one' (Hengeveld 1992: 54; cf. Keizer 1992a: 126, Dik 1997: 224–5)
(33) (a) Snoopy is sleeping. **So** is Woodstock.
(b) (Pres e_i: [(Af_i) (x_i: Woodstock)$_\phi$]) (Keizer 1992a: 124–6)

It was soon pointed out that the absence of further internal layers within the predication (i.e. the absence of the equivalent of a VP or V-bar constituent) and the term (e.g. the absence of units consisting of a head and *one*, but not all of its modifiers) formed a problem for this kind of approach (e.g. Keizer 1992a: 127, fn. 8, 1992b: 23, fn. 6; Van der Auwera 1992), but a satisfying solution was never offered. We will return to this issue in the sections to follow.

FDG has adopted the FG analysis of proforms as corresponding to units in underlying representation. However, since FDG distinguishes between an Interpersonal,

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

a Representational and a Morphosyntactic Level, a complete analysis of proforms in FDG will have to take each of these levels of representation into consideration.

5.2.2 Representation at the Interpersonal Level

At the Interpersonal Level, deictic personal pronouns are regarded as 'essentially similar to proper names': they function as a grammatical device for naming oneself, one's Addressee or a third party present in the immediate situation. They are, therefore, regarded as directly reflecting the abstract head of a Referential Subact at IL (Hengeveld & Mackenzie 2008: 118). The English pronominal system allows for the following combinations of features (where [±S] means 'involving the Speaker or not' and [±A] means 'involving the Addressee or not'; note that number distinctions are made at the Representational Level):

(34) First person singular (+id R_1: [+S, –A] (R_1))
First person plural exclusive (+id R_1: [+S, –A] (R_1))
First person plural inclusive (+id R_1: [+S, +A] (R_1))
Second person singular (+id R_1: [–S, +A] (R_1))
Second person plural (+id R_1: [–S, +A] (R_1))
Third person singular (+id R_1: [–S, –A] (R_1))
Third person plural (+id R_1: [–S, –A] (R_1))
(See Hengeveld & Mackenzie 2008: 118–19)

A simple sentence like (35a) will therefore have the IL representation given in (35b):

(35) (a) **I** love **you**.
(b) (A_I: [(F_I: DECL (F_I)) (P_I)$_S$ (P_J)$_A$ (C_I: [(T_I) **(R_I: [+S, -A] (R_I)) (R_J: [-S, +A](R_J))** (C_I))$_\Phi$] (A_I))

Anaphorically used personal pronouns are not analysed as abstract heads of Referential Subacts. At IL they are simply represented as headless Referential Subacts, as demonstrated in (36):

(36) (a) **They** like **him**.
(b) (A_I: [(F_I: DECL (F_I)) (P_I)$_S$ (P_J)$_A$ (C_I: [(T_I) **(R_I) (R_J)** (C_I))$_\Phi$] (A_I))

Anaphoric pronouns can also correspond to Ascriptive Subacts (Hengeveld & Mackenzie 2008: 121). Examples are pronouns functioning as non-verbal predicates and pronouns functioning as the Ascriptive head of a Referential Subact. Examples are given in (37) and (38), respectively:

(37) (a) (John is a pilot.) **So** is Jane.
(b) (A_I: [(F_I: DECL (F_I)) (P_I)$_S$ (P_J)$_A$ (C_I: **[(T_I)** (R_1) (C_I))$_\Phi$] (A_I))

(38) (a) (A: I bought a car yesterday.) B: A new **one**?
(b) (A_I: [(F_I: DECL (F_I)) (P_I)$_S$ (P_J)$_A$ (C_I: [(R_1: **[(T_I)** (T_J)]) (C_I))$_\Phi$] (A_I))

5.2.3 Representation at the Representational Level

At the Representational Level, participants in the SoA (whether arguments or modifiers) typically have a lexical head. It is, however, also possible for the heads of, for instance,

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

Individual-designating units, to be empty or even absent. Some examples are given in (39)–(41) (Hengeveld & Mackenzie 2008: 143, 237–8):

(39) Lexical head:
  (a) **The man** painted the door.
  (b) (1 x_i: [(f_i: man (f_i)) (x_i)$_\Phi$])
(40) Empty head:
  (a) Mary wants a good-looking man but I prefer **an honest one**.
  (b) (1 x_i: [(f_i) (x_i)$_\Phi$]: [(f_j: honest (fj)) (x_i)$_\Phi$])
(41) Absent head:
  (a) The man cleaned the windows and **Ø**/**he** painted the door.
  (b) (1 x_i)

In (39a) the head position of the expression *the man* is occupied by a nominal lexeme, represented as (f_i) in (39b). In (40a), the pronoun *one* is coindexed with the Property *man* (f_i) mentioned earlier in the same sentence (see also Hengeveld & Mackenzie 2008: 167). As can be seen in (40b), there is a head position within the x-layer, but it is filled with only a Property variable (f_i), which itself lacks a head. In (41a) the anaphoric relation is indicated either by the pronoun *he* or by zero-realization; in both cases the Individual-designating unit *the man* functions as the antecedent, as indicated by the presence of a coindexed x-variable in (41b). In this case, the x-variable has no head at all.

5.2.4 Representation at the Morphosyntactic Level

At the Morphosyntactic Level, a distinction is made between lexical Words, which correspond to lexemes at the Interpersonal or Representational Level, and grammatical Words, which 'either correspond to an operator or a function at the Interpersonal or Representational Level or are introduced as dummies or support elements' (Hengeveld & Mackenzie 2008: 401). In addition, grammatical Words may represent 'units with an abstract head at the Interpersonal or Representational Level' (ibid.). Pronouns, representing abstract heads at the Interpersonal Level and empty or absent heads at the Representational Level, are thus introduced as grammatical Words at the Morphosyntactic Level. The pronoun *one*, for instance, corresponds to the empty head (f_i) of an Individual-designating unit at the Representational Level. As we have seen, it is possible for this empty head to combine with a modifier. At the Morphosyntactic Level, however, Noun phrases cannot occur without a head; the pronoun *one* is therefore introduced at this level to function as the head of the Np (Hengeveld & Mackenzie 2008: 392).

5.2.5 Conclusion

Although the present treatment of proforms in FDG is both innovative and insightful, reflecting the different functions of proforms at the three levels, it nevertheless seems to suffer from the same major flaw as other theories, as it fails to account for the counterexamples provided in sections 3 and 4. Further examination suggests that FDG may even have an additional weakness: as the descriptions of the three levels provided in the previous sections make clear, the verb and its complement(s) do not form a unit at any of these levels. At the Interpersonal Level the Communicated Content consists

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

of a number of unordered and ungrouped Subacts, while at the Representational Level States-of-Affairs are represented as units headed by a Configurational Property consisting of a predicate and all its arguments. At the Morphosyntactic Level the (verbal) predicate tends to correspond to a Verbal phrase (Vp), while the arguments typically appear as Nominal phrases (Nps) and Adpositional phrases (Adps). Together with other (modifying) Phrases and grammatical Words (e.g. conjunctions), these phrases form Clauses. There are, however, no further layers between that of the Phrase and the Clause (there is no higher-order 'superphrase' (Hengeveld & Mackenzie 2008: 299)).

This means that at none of the three levels does FDG distinguish a layer corresponding to what in Generative Grammar is known as the VP, or to any layers within the VP (V-bar levels). Much the same is true for Phrases. These, too, consist of a number of non-hierarchically ordered units (Words, as well as embedded Phrases and Clauses), as illustrated in (42):

(42) (a) John likes that big black French car.
(b) (Np_i: (Gw_i: that (Gw_i)) (Aw_i: big (Aw_i)) (Aw_j: black (Aw_j)) (Aw_k: French (Aw_k)) (Nw_i: car (Nw_i)) (Np_i))

Once again, there are no further layers between the Np and the individual words. This means that the combinations of elements replaced by the pronoun *one* in (43) – *French car*, *black French car* and *big black French car*, respectively – do not form separate layers ('N-bar' levels) within the Np. Nor do they correspond to units at the Interpersonal and Representational Levels.

(43) (a) Jane likes the small red **one**.
(b) Jane likes the small **one**.
(c) Jane likes this **one**.

5.3 Proforms in FDG: a new proposal

5.3.1 The function of proforms

In section 2 it was concluded that most accounts of proforms, be they descriptive or theoretical, describe proforms in terms of two features, one pragmatico-semantic, the other syntactic:

(44) (a) Proforms are semantically (nearly) empty, relying for their interpretation on information provided elsewhere (the previous discourse or the immediate discourse situation);
(b) Proforms replace (stand for, substitute) a syntactic unit (the antecedent).

On the basis of the data discussed in the previous section, however, it seems that the second criterion is no longer tenable: there are too many cases where the proform does not replace a syntactic unit. This leaves us with the semantic property given in (44a).[13]

[13] This is indeed regarded as the one distinctive feature in semantically or cognitively oriented approaches. According to Evans & Green (2006: 490), for instance, 'pronouns can be viewed as a separate category from nouns because they belong to a closed class and because they provide what cognitive linguists call **schematic meaning** rather than **content meaning**' (original emphases).

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

In what follows, it will be assumed that, in an adapted form, this first feature is indeed in itself enough to define the class of proforms. Nevertheless it is believed that a definition in terms of morphosyntactic behaviour is possible as well as useful, provided we take a somewhat different perspective: rather than defining proforms in terms of the syntactic unit (or layer in underlying representation) they replace (clause, phrase, bar-level), proforms can be defined in terms of the position they take in a morphosyntactic template. Instead of the definition in (44), the following FDG definition of proforms will therefore be proposed:

(44′) (a) Proforms correspond to elements at the Interpersonal or Representational Level which do not contain lexical information; as such they are themselves semantically (nearly) empty. For their interpretation proforms therefore rely on information provided elsewhere (the previous discourse or the immediate discourse situation). The elements they correspond to need not form a separate unit (or layer) at the Interpersonal or Representational Level.
(b) At the Morphosyntactic Level proforms are grammatical Words filling an obligatory position in the selected template, either at the level of the Clause (typically Subject) or at the level of the Phrase (typically the head).

Let us consider this second condition in some more detail.[14]

As pointed out before, FDG makes use of morphosyntactic templates (clauses and phrasal) for the ordering of linear elements at the Morphosyntactic Level. Some of the positions in these templates have to be filled. In the case of clausal templates, for instance, English typically requires the Subject position (as well as the Object position in the case of a transitive verb) and the position for the finite verb to be filled. In Phrasal templates, the position of the head must be filled, as well as the positions for any complements of the head. In some cases, however, the Interpersonal and Representational Levels do not provide the lexical material to fill these positions; in that case, proforms can be used. The following example illustrates the use of proforms at the clausal and phrasal levels, respectively:

(45) (a) I bought a new car last week, but **it** wouldn't start this morning.
(b) a red car and a blue **one** (Hengeveld & Mackenzie 2008: 392)

In (45a) the third-person pronoun *it* corresponds to a unit at the Interpersonal Level (a Referential Subact) and Representational Level (an Individual-designating unit). However, since neither of these corresponding units provides the necessary lexical information, the proform *it* is used to fill the Subject position at the Morphosyntactic Level. In (45b) the pronoun *one* also corresponds to a Subact at the Interpersonal Level (in this case an Ascriptive Subact) and to a semantic unit at the Representational Level (a property). Neither the Interpersonal nor the Representational Level, however, provides the lexical information needed to fill the head position of the nominal Phrase

[14] The question of whether is it (pragmatically) appropriate, in a certain discourse context, to use a pronoun (rather than a fully lexical expression) has been discussed in considerable detail (e.g. Cornish 1986, 1999; Ariel 1990; Gundel *et al.* 1993; Wales 1996; Huddleston & Pullum *et al.* 2002: 1470–1; for a summary, see Keizer to appear). I will therefore refrain from discussing the first function here.

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

at the Morphosyntactic Level. Since every Phrase must have a head, the proform *one* is inserted. Nor is it only the indefinite pronoun *one* that can function as the head of the nominal Phrase in the presence of a modifier; this function can also be fulfilled by a personal pronoun (as in *He who has the pen*, *has the power* (BNC)), or as, we have seen before, by a demonstrative pronoun (typically *that* or *those* + clause or PP). The proform *do so*, finally, performs a similar function at the Morphosyntactic Level, the only difference being that in this case the proform functions as the head of a verbal Phrase.[15]

It will be clear that, according to the definition of proforms given in (44′), dummy expressions such as weather-*it*, anticipatory *it* and existential *there* are not regarded as proforms. In example (46), for instance, the element *it* does not express a Subact at the Interpersonal Level. At the Representational Level, the sentence is analysed as involving a zero-place predication frame, which means that *it* does not correspond to any argument either. Thus, neither the Interpersonal nor the Representational Level provides the material to fill the Subject position at the Morphosyntactic Level. Since this position needs to be filled, the dummy element *it* is inserted (Hengeveld & Mackenzie 2008: 347):

(46) **It** is raining.

Moreover, it will be clear that dummy *it* does not function as an instruction to the Addressee to look for lexical information elsewhere. Its sole function is to fill an obligatory position in the morphosyntactic template. Such elements will therefore not be regarded as proforms, but merely as dummies. Similarly, according to the definition of proforms provided, *do so* qualifies as a proform, whereas the auxiliary *do* functions as a dummy. This explains why the two forms can be combined (cf. Quirk *et al.* 1985: 874) and why they display different syntactic behaviour.[16]

[15] There are a number of exceptions to the syntactic criterion, but these are, on the whole, strictly definable and predictable. It seems therefore appropriate to include both functions in the definition of English proforms. For a brief discussion, see Keizer (to appear).

[16] When *do so* appears in a non-finite form, it may combine with the dummy auxiliary *do*, as well as with other modals and auxiliaries:

(ia) And others, even though the EPA report **did not do so**, have pointed to the risk for heart disease. (COCA)

(ib) Those who made the journey to London **might have done so** for a variety of reasons. (BNC)

This clearly shows that the two must be distinguished. The following examples from Huddleston & Pullum *et al.* (2002: 1530) also show that the two dummies do not perform the same function, as they are clearly not interchangeable:

(ii) I didn't invite Kim, but I did Pat/*but I did so Pat.

Additional evidence can be found in the fact that when the dummy auxiliary is used, the main verb can be added, whereas this is not possible in the case of *do so*:

(iiia) Kim didn't come, but Pat did (come).

(iiib) Kim danced, and Pat did so (*dance) too.

These examples clearly show that whereas dummy *do* is indeed an auxiliary, *do so* functions as a proform, filling the position of the main verb.

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

5.3.2 The representation of proforms

5.3.2.1 The indefinite pronoun one. At the Interpersonal Level, the pronoun *one*, as an anaphoric pronoun, is represented as a headless Ascriptive Subact:

(47) (a) (A: I bought a car yesterday.) B: A new **one**?
(b) (A_I: [(F_I: DECL (F_I)) (P_I)$_S$ (P_J)$_A$ (C_I: [(-id R_1: **[(T_I)** (T_J)]) (C_I))$_\Phi$] (A_I))

At the Representation Level, *one* is represented as an f-variable (designating a property) without a head:

(48) (a) a big car and a small one
(b) (1 x_i: **[(f_i: car$_N$ (f_i))** (x_i)$_\Phi$]: [(f_j: big$_A$ (f_j)) (x_i)$_\Phi$]) & (1 x_j: **[(f_i)** (x_j)$_\Phi$]: [(f_k: small$_A$ (f_k)) (x_j)$_\Phi$])

Note that in this approach it is assumed that the headless unit at the Representational Level is coreferential or codesignating with a single previously introduced semantic unit, as indicated through coindexation on the variables. Thus, in (48b), two different objects are designated (x_i and x_j). Both are described in terms of two properties, one of which, 'car' (f_i), is shared. The second time the property 'car' is designated, it is represented as a co-indexed, headless f-variable (the result being an empty-headed x-variable). In combination with the relevant information from the Interpersonal Level, where the corresponding unit is an Ascriptive Subact, this leads to the insertion of the pronoun *one* in the head position of the corresponding Np at the Morphosyntactic Level. This is illustrated in example (49) (from Hengeveld & Mackenzie 2008: 393):[17]

(49) (Np_i: [(Gw_i: a (Gw_i)) ($Adjp_i$: ($Adjw_i$: big$_A$ ($Adjw_i$)) ($Adjp_i$)) (Nw_i: car-sg (Nw_i))] (Np_i)) &
(Np_j: [(Gw_j: a (Gw_j)) ($Adjp_j$: ($Adjw_j$: small$_A$ ($Adjw_j$)) ($Adjp_j$)) (Nw_j: one-sg (Nw_j))] (Np_i))

As we have seen, however, *one* does not necessarily correspond to a single designating unit (a head) at the Representational Level. Consider in this respect the following examples (cf. examples (42) and (43)):

(50) (a) John likes that big black French car.
(b) Jane likes **the small one**.

In this case the pronoun *one* does not express a single unit at the Representational Level; instead it corresponds to a combination of properties ('black', 'French' and 'car'), as reflected in the following representation:

[17] Since the proform *one* does not correspond to lexical information at the Interpersonal and Representational Level, one might expect it to be analysed as a grammatical Word. As pointed out by Hengeveld & Mackenzie (2008: 401), however, grammatical Words 'often behave in the same way as Words that do correspond to a Lexeme' and can therefore 'be classified according to the analogy between their syntactic distribution and that of Words with lexical content'. Since the pronoun *one* displays much the same distribution as a Lexical noun, it is analysed as a nominal Word, heading a nominal Phrase.

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

(51) (a) (dis 1 x_i: [**(f_i: car$_N$ (f_i))** (x_i)$_\Phi$]: **[(f_j: French$_A$ (f_j))** (x_i)$_\Phi$]: [**(f_k: black$_A$ (f_k))** (x_i)$_\Phi$]: [(f_l: big$_A$ (f_l)) (x_i)$_\Phi$])
(b) (1 x_j: [**(f_i)** (x_j)$_\Phi$]: [**(f_j)** (x_j)$_\Phi$]: [**(f_k)** (x_j)$_\Phi$]: [(f_m: small$_A$ (f_m)) (x_j)$_\Phi$])

In (51), the second object, x_j, is described in terms of four properties, three of which ('black', 'French' and 'car') it shares with the previously designated object x_i. This is captured in (51b) by the use of three headless f-variables, coindexed with the codesignating f-variables in (51a). It is now the combination of these headless, coindexed variables that leads to the insertion of the pronoun *one* at the Morphosyntactic Level. Note that although the pronoun does not replace any particular unit in underlying representation, it still complies with the definition of proforms provided in (44′): (a) it corresponds to non-lexical elements at the Interpersonal and Representational Levels, and serves as an instruction to retrieve lexical information from elsewhere; and (b) it fills an obligatory position at the Morphosyntactic Level.

Note that on this approach it is irrelevant whether the pronoun *one* replaces a whole noun phrase, a subphrase (bar-level), or just a head. As long as the combination of lexical items can be assumed to be accessible to the hearer, *one* can be used as an instruction to retrieve them. This means that constructions such as (52) are no longer problematic either:

(52) (a) white **administrators and teachers** with black **ones**. (COCA) (cf. (15a))
(b) (m x_i: [(f_i: [(m x_j: **(f_j: administrator$_N$ (f_j))** (x_j): (f_k: white$_A$ (f_k)) (x_j)) & (m x_k: **(f_l: teacher$_N$ (f_l))** (x_k): (f_k) (x_k))] (f_i)) (x_i)) &
(m x_l: [(f_m: [(m x_m: **(f_j)** (x_m): (f_n: black$_A$ (f_n)) (x_m)) & (m x_n: **(f_l)** (x_n): (f_n) (x_n))] (f_m))] (x_l))

It is the combination of the two codesignating, headless f-variables, f_j and f_l (representing the properties 'administrator' and 'teacher'), in the second representation in (52b) that triggers the use of the pronoun *ones* at the Morphosyntactic Level, heading the Np corresponding to the unit represented by x_l.

Given the analysis of proforms presented here, FDG can also deal with those cases where *one* picks up part of a compound, as in (53):

(53) Poetry **class** was just about the only **one** Remi really liked. (COCA) (= (14a))

As indicated before, use of *one* in these cases is only allowed when the compound in question involves a relatively low degree of conventionalization or lexicalization. In FDG compounds are analysed as having a complex head, consisting of two separate lexemes at the Representational Level (Hengeveld & Mackenzie 2008: 216). For the compound *poetry class* this results in the representation in (54a), where the complex head f_i consists of the lexeme *class*, restricted by the lexeme *poetry*. Given this analysis, the use of *one* in (53) no longer poses a problem: it can be represented as a headless property f_i, codesignative with the previously mentioned property 'class' (see (54b)):

(54) (a) (x_i: (f_i: (**f_j: class$_N$ (f_j)**: (f_k: poetry$_N$ (f_k)) (f_j)) (f_i)) (x_i)) poetry class
(b) (x_j: (f_l: (**f_j**: (f_m: only$_A$ (f_m)) (f_j)) (f_l)) (x_j)) only one

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

The fact that this is not possible with all compounds suggests that compounds with a very high degree of lexicalization or conventionalization (e.g. *teapot*, *headmaster*) are no longer analysed as having a complex head, but as being restricted by a single lexeme.

5.3.2.2 The predicative proform do so. At the Interpersonal Level, the proform *do so* corresponds minimally to a headless Ascriptive Subact within the Communicated Content, but may also include one or more Referential Subacts. An example is given in (55), where *do so* captures a combination of Ascriptive and Referential Subacts. After all, the properties 'buy' and 'car' are evoked, even if they are not lexically expressed:

(55) (a) A: My sister bought a car. B: **So did I**.
(b) $(A_I: [(F_I: \text{DECL } (F_I))\ (P_I)_S\ (P_J)_A\ (C_I: [\mathbf{(T_I)}\ (R_1)\ \mathbf{(R_J)}])\ (C_I))_\Phi]\ (A_I))$

At the Representational Level, constructions with *do so* can be dealt with in more or less the same way as the pronoun *one*. Consider the simple sentence in (56a):

(56) (a) I **caught a fish** and Bill **did so** too.
(b) $(\text{Past } ep_i: (e_i: (f_i: [\mathbf{(f_j: \text{catch})}\ (x_i)_A\ (1\ x_j: \mathbf{(f_k: \text{fish})}\ (x_j))_U]\ (f_i))\ (e_i))\ (ep_i))) \ \&$
$(\text{Past } ep_j: (e_j: (f_l: [\mathbf{(f_j)}\ (x_k)_A\ (1\ x_l: \mathbf{(f_k)})_U]\ (f_l))\ (e_j)\ (ep_j)))$

Coindexed parts together (i.e. the properties 'catch' f_j and 'fish' f_k) are expressed as *do so*, which expresses the Tense operator. This results in a new Episode ep_j and a new SoA e_j, with different Actor and Undergoer participants (x_k and x_l). This kind of analysis can obviously be applied to any occurrence of *do so*, irrespective of which combination of elements it covers and irrespective of whether these elements appear in adjacent positions in the morphosyntactic template.

5.3.2.3 The demonstrative pronouns that *and* those. Constructions with *that/those* followed by a clause or PP may seem to be more of a challenge, as here the pronoun replaces both the determiner and the head. Note that an analysis in which *that* and *those* function as demonstrative determiners followed by an empty head is not an option. First of all, such an analysis would predict that explicit mention of the head would still be possible; as shown in (57), however, this leads to unacceptable (at best unintended) results:

(57) (a) Perhaps the best-known example is **that** (*?example) of the Heel Stone at Stonehenge. (adapted from (19a))
(b) Zulema's life and **that** (*life) of other female drug lords ... (adapted from (21b))

As we have seen, in FDG anaphoric proforms are represented at the Interpersonal Level as headless subacts. In these particular constructions with *that/those*, however, there is a modifier, which means we also need a head. This is possible if we provide these constructions with the same underlying representation as constructions with the pronoun *one/ones*:

(58) (a) Zulema's life and **that of other female drug lords** ...
(b) $(+\text{id } R_1: [(T_I)\ (T_J: \ldots\ (T_J))])$

Although not lexically expressed, the property 'life' is evoked, and as such represented at the Interpersonal Level as a head (T_I) modified by a complex lexical modifier (T_J).

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

At the Representational Level the relevant units in example (59b) can be represented as follows:

(59) (1 x_i: [(f_i: [(f_j: $life_N$ (f_j)) (x_j)$_{Ref}$] (f_i)) (x_i))
(1 x_k: [(f_k: [(f_j) (x_l – other female drug lords – (x_l))$_{Ref}$] (f_k)) (x_k))

Note that it is only the property 'life' f_j that is shared by the two expressions; all the other designations (that of the unit as a whole and of the Ref element) are replaced by new ones. Moreover, as the examples in (57) already make clear, the pronouns here do not have a demonstrative function, which also explains the absence of the distal/proximate opposition. We can now assume that the forms *that* and *those* are triggered by a combination of information from the Interpersonal Level (an identifiable Referential Subact consisting of a head and a modifier) and the Representational Level (the presence of a headless coindexed f-variable).

5.3.2.4 Personal pronouns as heads. Unlike the proforms discussed so far, the first elements of a pronominal apposition like *we teachers* are always deictic (first- or second-person) pronouns. As we saw in section 5.2, deictic pronouns are represented at the Interpersonal Level as the abstract head of a Referential Subact, consisting of a combination of the features [±S] and [±A]. English first and second pronouns are thus represented as follows:

(60) First person singular/plural exclusive (+id R_1: [+S,-A] (R_1))
First person plural inclusive (+id R_1: [+S,+A] (R_1))
Second person singular/plural (+id R_1: [-S,+A] (R_1))

The difference between simple deictic pronouns and pronominal appositions is that in the latter the Referential Subact consists not only of an abstract head, but also contains an Ascriptive Subact. As in the case of *that/those*, these constructions must therefore be analysed as containing a head and a modifier at the Interpersonal Level. This may lead to the following possible respresentations:[18]

(61) (a) we women
(b) (R_I: [[+S,-A] (T_I)] (R_I))
(62) (a) we the women
(b) (R_I: [[+S,-A] (+id T_I)] (R_I))

At the Representational Level the constructions in (61) and (62) are provided with the same underlying representation:[19]

(63) (x_i: [(mx_j): (mx_k)] (x_i))

[18] For a more complete treatment of these constructions, see Keizer (2010).
[19] Note that plurality is specified at the level of the component elements; this is to account for those cases where the two elements do not (formally) agree in number, e.g. *we the council*.

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

Vocatives like *you fools!*, which fulfil a completely different function, are provided with a different representation. In FDG Vocatives are analysed as a special subclass of Interactive Acts (Hengeveld & Mackenzie 2008: 78). If the Vocative is used to gain the Addressee's attention, this may be done by means of an invariant lexical item, such as *hey* as in (64a). If the Vocative consists of the pronoun *you*, this pronoun will be analysed as restricting the variable representing the Addressee. In that case, the F-variable has an abstract head (Interpellative) (example (64b)):

(64) (a) Hey!
(a′) $(A_I: [(F_I: \text{hey } (F_I))\ (P_I)_S\ (P_J)_A]\ (A_I))$
(b) You!
(b′) $(A_1: [(F_1: \text{INTERP } (F_1))\ (P_1)_S\ (P_J: [\text{-S,+A}]\ (P_J))_A]\ (A_1))$

The construction in (65a) will be analysed as in (64b), but now expanded with a Communicated Content consisting of one Ascriptive Subact (corresponding to the property 'fool'):

(65) (a) You fool!
(b) $(A_I: [(F_I: \text{INTERP } (F_I))\ (P_I)_S\ (P_J: [\text{-S,+A}]\ (P_J))_A\ (C_I: [(T_I)]\ (C_I)))\ (A_1))$

Vocatives are regarded as fulfilling a purely interpersonal function: they do not refer to or designate an extralinguistic entity. As such, they lack a representation at the Representational Level. All information needed to trigger the right form of the proform at the Morphosyntactic Level is provided at the Interpersonal Level.

6 Constraints on the use of English proforms

So far it has been argued that, contrary to what has been claimed in most of the literature on the subject, English proforms do not necessarily replace constituents in underlying representation. Instead, it has been demonstrated that pronouns can capture many different combinations of elements that do not form a unit at any level of analysis. It was then concluded that, in a sufficiently flexible system like FDG, this need not be a problem, and that all the uses of English pronouns mentioned in this article can quite easily be dealt with.

At the same time, however, it will be clear that pronouns cannot just replace any combination of linguistic elements. In other words, although we need a flexible system, there must also be constraints to prevent any unacceptable or ungrammatical uses of proforms. In this section some of these constraints will be briefly discussed.[20]

Since the proforms *one* and *do so* seem to perform very similar functions in the nominal and verbal domain, respectively, they will be dealt with in one section. We

[20] This section will only deal with some of the morphosyntactic restrictions on the use of proforms. In addition, the use of a proform may be inappropriate for pragmatic or semantic reasons, i.e. a hearer may not be able to retrieve the lexical information the proform is supposed to correspond to (see the definition of proforms given in section 5.1). This kind of violation will not be discussed here (but see e.g. Cornish 1986, 1999; Ariel 1990; Gundel *et al.* 1993; Wales 1996; Huddleston & Pullum *et al.* 2002: 1470–1).

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

will start by looking at the predicative proform *do so*, as it is in the verbal domain that the constraints on the use of proforms seem to be strongest.

Although, as demonstrated in section 4.2, there are many exceptions to the claim that *do so* must replace a constituent ('VP' or 'V-bar') in underlying representation, it cannot be denied that in most cases replacing a verb without its complement(s) leads to an ungrammatical construction. One quite obvious reason for this constraint is that – as pointed out many times before – the verb and its complement(s) form one semantic unit, and that the relation between the two is, indeed, very strong, the result being that the meaning of the verb is incomplete without the presence of the complement(s). This is, of course, largely due to the function of the verb in the predication: a two- or three-place verb serves to indicate a relation between a first ('external') argument and one or more other ('internal') arguments; in order to be able to do so, these other arguments need to be mentioned explicitly. This strong semantic relation between verb and (internal) argument is reflected in morphosyntactic behaviour. Thus, cross-linguistically, the internal arguments (typically realized as objects) tend to immediately precede or follow the verb, while omitting them usually results in an ungrammatical construction.[21] This has led many linguists to believe that morphosyntactically, too, the verb and its complement must form a unit. The fact that, on the whole, the proform *do so* could only be used to replace this unit as a whole, and not just the verb, has been seen as further confirmation of this view.

Subsequently, this approach was extended to the nominal domain. Several studies have emphasized the parallels between the internal structure of verb phrases and noun phrases (e.g. Abney 1987; Rijkhoff 2002; Broekhuis & Keizer forthcoming). When we compare the syntactic behaviour of the pronoun *one* with that of the proform *do so*, such a unified approach seems justified: both can be used to replace the head together with its complement(s) and modifier, whereby the complements, if present, have to be covered by the proform. The examples in section 4.1, however, convincingly show that, unlike in the verbal domain, the pronoun *one* need not include the complement of the noun it replaces, and may be used in combination with a new complement. Now, one may of course conclude that in those cases the element in question is not a complement but a modifier of the noun – a conclusion that may not be entirely unjustified. As demonstrated in Keizer (2004), it is not always easy to distinguish between prepositional complements and modifiers within the noun phrase. Thus it turns out that application of the various tests defined to determine the status of a postnominal PP (e.g. Huddleston 1984; Radford 1988; Huddleston & Pullum *et al.* 2002), including replacement by the pronoun *one*, yields equivocal results, and that only very few nouns can be qualified as truly relational. Instead, the evidence presented suggests that it is only PPs designating concepts that are intrinsically and strongly

[21] Note that, as with the use of proforms, these are indeed tendencies rather than rules: in English, for instance, objects can appear in other positions (can be extracted or displaced), while arguments can be omitted under certain circumstances (e.g. in generic statements).

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

connected to the concept designated by the head noun that have to be included in the reference of the proform *one* (see Keizer 2004: 346).

It therefore seems justified to conclude that there is no reason to assume that the proform *one* must replace a noun along with its complement. In some cases this may definitely be preferred, due to the strong intrinsic relation between noun and complement. But what we see is a tendency rather than a hard and fast syntactic rule.

So why is it that so many linguists, for so many years, have held on to the idea that the proform *one* must replace a head noun and its complement(s)? It may well be that this view is largely inspired by the wish to find parallels between the verbal and the nominal domain, as well as by a (usually theory-specific) wish to provide all phrases with the same basic underlying structure. In other words, if in the verbal domain the proform *do so* must replace a verb and its complement(s), then its nominal equivalent *one* must have the same function in the nominal domain.

We can, however, also turn this argument around. If there is no solid basis for the assumption that the proform *one* must correspond to a noun and its complement(s), then there is no reason to assume that proforms in general, including *do so*, substitute constituents. What the data tell us is that when there is a very strong, intrinsic semantic relation between a head and its complement, it is likely, and sometimes necessary, for both the head and the complement to be included in the reference of the proform. This is virtually always the case in the verbal domain, as here the bond between head and complement is presumably the strongest. The same may be true for some relational nouns, but certainly not for all of them.

It will be clear that once we let go of the requirement that proforms correspond to units in underlying structure, all other problematic uses of the proforms *do so* and *one* described in section 4, in particular those in which the proform replaces two or more elements that do not form a constituent, can easily be accounted for – as was indeed shown in section 5.

So far, we have looked solely at the role of the intrinsicness of the semantic relation between head and complement to account for possible constraints on the use of proforms. Other semantic relations, however, also apply. Let us concentrate on some of these further restrictions within the nominal domain.

When we compare the use of *one* with the use of *that*, we find that there is a certain amount of overlap between the two pronouns. In the following examples, for instance, the pronouns *the one*(*s*) can be replaced by *that/those*:

(66) (a) I support the **ban** on smoking, but not **the one**/**that** on alcohol (cf. (10a)).
(b) They were open stairs like **the ones**/**those** in the house his Mum and Dad were buying (BNC, fiction prose; adapted). (Huddleston & Pullum *et al.* 2002: 441)

Unlike *that*, however, *one* can only be used with countable nouns (see Huddleston & Pullum *et al.* 2002: 1508):

(67) (a) Their political knowledge surpassed **that**/*__the one__** of the members of the party cell. (BNC, adapted)

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

(b) Normal lung tissue and **that**/*__the one__ from patients with focal fibrosis expressed very little ET-1. (BNC; adapted)

Moreover, *that* seems to be the preferred choice with countable abstract nouns, as in example (68):

(68) (a) Perhaps the best-known **example** is **that**/??**the one** of the Heel Stone at Stonehenge. (BNC; adapted) (cf. (19a))
(b) It was also weighing on Bernard's mind that they were opening more and more shops in Europe and yet the European **commercial scene** was not as stable as **that**/?**the one** in the UK. (BNC; adapted)

The pronoun *one*, on the other hand, can be used to denote animate entities. This is not possible with *that*, although the plural form, *those*, does allow for this use:

(69) (a) By March, he'd won the right to remove incompetent principals –; like **the one**/*__that__ in the Bronx who had been convicted of possessing crack cocaine. (BNC; adapted)
(b) Two men –; including **the one**/*__that__ in the photograph –; were cleared of her murder. (BNC; adapted)
(c) **Students**, especially **the ones**/**those** in high school, too often get the message that the contents of the library are more important than the clients. (COCA; adapted) (cf. (19c))

Finally, when we consider appositional constructions with a pronominal first element (section 4.4), we see that these constructions only freely allow first- and second-person plural pronouns (*we*/*us* and *you*). These are typically followed by a bare plural noun (example (70), here repeated for convenience), but may also be followed by a definite NP. In the latter case, the determiner can only be the definite article or a possessive determiner (example (71)):

(70) (a) In fact it is how **we Brits** are seen by the French. (BNC, newspaper)
(b) '**You English girls** are so cruel,' he was saying. (BNC, prose fiction)
(c) He never lay a hand on any of **us children**. (BNC, oral history)
(71) (a) This follows the American tradition: '**We the people of the United States** . . . do ordain and establish this constitution'. (BNC, broadsheet newspaper)
(b) Marshal, here are we before you, O savior of France. **We your little buddies** swear to follow where you advance. (COCA, fiction)

The first-person singular pronoun does occur in these constructions, but only sporadically, and only when followed by a definite NP (example (72a)); the second-person singular pronoun seems to occur somewhat more frequently (example (72b)):

(72) (a) . . . 'In every way, you see, I'm Helen Winslow, based on **me the person**, but synthesized and evolved into my present form by the AIM system. I'd prefer you call me Helen, because that's who I am, but I'll accept Annie if you like.' 'But you are an artificial intelligence system,' said Maxwell. (COCA, fiction)
(b) You will not feel the guilt you have felt these almost eleven years, **you the man who broke up a family**. (COCA, fiction)

No examples, however, could be found with any of the third-person singular pronouns (**she/he* (*the*) *doctor*; **it* (*the*) *child*); the third-person plural pronouns *them* and *they*

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

do occur, but, as argued in Keizer (2010), in those cases the pronoun does not function as the head of the construction, but rather as a kind of all-purpose definiteness marker (see also note 10).

As for the vocative constructions discussed in section 4.4, it will not come as a surprise that in these constructions we only find the (singular or plural) pronoun *you*.

7 Conclusions and implications

The aim of this article has been to provide a definition of English proforms that captures all their various uses and functions. It has been argued that proforms can be defined in terms of two basic functions, one pragmatico-semantic, the other syntactic in nature. Prototypically, proforms exhibit both functions; under clearly definable circumstances, however, the second requirement need not be fulfilled.

What distinguishes this definition from previous definitions of proforms is that proforms are no longer defined in terms of substitution of a pragmatic, semantic or syntactic unit. It is true that in many cases anaphoric proforms share their reference or designation with a phrasal or clausal antecedent. This tendency, together with the fact that proforms appear as one form at the Morphosyntactic Level, filling one position in a morphosyntactic template, has prompted the belief that this is a necessary feature of proforms. As it turns out, however, this is merely an accidental feature, not an absolute requirement.

Subsequently, an alternative analysis, in which proforms can replace units smaller and larger than constituents, was shown to be possible within the theory of Functional Discourse Grammar. The analysis proposed was based on the idea that, if proforms can be defined morphosyntactically at all, this should not be done in terms of substitution, but in terms of the position they take in a morphosyntactic template: in the absence of a lexical element, they fill an obligatory position at clause or phrase level. An obvious further advantage of this approach is that by replacing the substitution criterion by the positional criterion it has become possible to suggest a treatment of proforms that can be applied to all proforms: definite and indefinite, referential and predicative, phoric and deictic.

Finally, the proposal also has an important broader implication. As we have seen, in many linguistic theories substitution by means of proforms (or pronominalization) is regarded as one of the most reliable tests for determining constituency. The data presented in this article show this assumption to be ill-founded. This does not mean that substitution cannot be used as a constituency test at all; it does mean that it has to be applied with caution and always in combination with other tests.

Author's address:
Institut für Anglistik und Amerikanistik / Department of English
Universität Wien
Campus d. Universität Wien
Spitalgasse 2–4/Hof 8.3

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

1090 Vienna
Austria
evelien.keizer@univie.ac.at

Corpora

Davies, Mark. 2004–. BYU-BNC: The British National Corpus. http://corpus.byu.edu/bnc.
Davies, Mark. 2008−. The Corpus of Contemporary American English (COCA). www.americancorpus.org.

References

Aarts, Bas & Liliane Haegeman. 2006. English word classes and phrases. In Bas Aarts & April MacMahon (eds.), *The handbook of English linguistics*, 117–45. Oxford: Blackwell.
Abney, Steven. 1987. The English noun phrase in its sentential aspect. PhD dissertation, MIT.
Acuña-Fariña, Juan Carlos. 1996. The puzzle of apposition: On so-called appositive structures in English. *Monografías da Universidade de Santiago de Compostela* 190, Universidade de Santiago de Compostela.
Acuña-Fariña, Juan Carlos. 2009. Aspects of the grammar of close apposition and the structure of the noun phrase. *English Language and Linguistics* 13(3), 453–81.
Ariel, Mira. 1990. *Accessing noun phrase antecedents*. New York: Routledge.
Broekhuis, Hans & M. Evelien Keizer. Forthcoming. *Modern grammar of Dutch: Nouns and noun phrases*, part 1.
Burton-Roberts, Noel. 1975. Nominal apposition. *Foundations of Language* 13, 391–419.
Burton-Roberts, Noel. 1986. *Analysing sentences: An introduction to English syntax*. London: Longman.
Burton-Roberts, Noel. 1994. Apposition. In R. E. Asher & J. M. Y. Simpson (eds.), *The encyclopaedia of language and linguistics*, vol. 1, 184–7. Oxford: Pergamon Press.
Butler, Christopher S. 2003. *Structure and function: A guide to three major structural–functional theories*. Amsterdam: John Benjamins.
Cornish, Francis. 1986. *Anaphoric relations in English and French: A discourse perspective*. London: Croom Helm.
Cornish, Francis. 1999. *Anaphora, discourse, and understanding: Evidence from English and French*. Oxford: Clarendon Press.
Croft, William. 2001. *Radical Construction Grammar*. Oxford: Oxford University Press.
Curme, George O. 1931. *Syntax*. Boston: Heath and Co.
Déchaine, Rose-Marie & Martina Wiltschko. 2003. On pro-nouns and other 'pronouns'. In Martine Coene & Yves D'hulst (eds.), *From NP to DP*, vol 1: *The syntax and semantics of noun phrases*, 71–98. Amsterdam: John Benjamins.
Dik, Simon C. 1978. *Functional Grammar*. Amsterdam: North Holland.
Dik, Simon C. 1997. *The theory of Functional Grammar*, vol. 2: *Complex and derived constructions*. Berlin: Mouton de Gruyter.
Evans, Vyvyan & Melanie Green. 2006. *Cognitive linguistics: An introduction*. Edinburgh: Edinburgh University Press.
Finch, Geoffrey. 2005. *Key concepts in language and linguistics*, 2nd edn. London: Palgrave.
Fortescue, Michael, Peter Harder & Lars Kristoffersen (eds.). 1992. *Layered structure and reference in a functional perspective*. Amsterdam: John Benjamins.
Francis, W. Nelson. 1958. *The structure of American English*. New York: Ronald.

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

Fries, Udo. 1989. The crew have abandoned the ship: Concord with collective nouns revisited. *Arbeiten aus Anglistik und Amerikanistik* 13, 98–104.
Gundel, Jeanette K., Nancy Hedberg & Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69, 274–307.
Haugen, Einar. 1953. On resolving the close apposition. *American Speech* 28, 165–70.
Hengeveld, Kees. 1992. *Non-verbal predication: Theory, typology, diachrony*. Berlin: Mouton de Gruyter.
Hengeveld, Kees. 2008. Prototypical and non-prototypical noun phrases in Functional Discourse Grammar. In Jan Rijkhoff & Daniel García-Velasco (eds.), *The noun phrase in F(D)G*, 43–62. Berlin: Mouton de Gruyter.
Hengeveld, Kees & J. Lachlan Mackenzie. 2008. *Functional Discourse Grammar*. Oxford: Oxford University Press.
Hockett, Charles F. 1955. Attribution and apposition in English. *American Speech* 30, 99–102.
Huddleston, Rodney. 1984. *An introduction to the grammar of English*. Cambridge: Cambridge University Press.
Huddleston, Rodney & Geoffrey K. Pullum *et al.* 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
Keizer, M. Evelien. 1992a. Reference, predication and (in)definiteness in Functional Grammar: A functional approach to English copular sentences. PhD dissertation, Vrije Universiteit Amsterdam.
Keizer, M. Evelien. 1992b. Predicates as referring expressions. In Fortescue *et al.* (eds.), 1–27. Reprinted in Matthew Anstey & J. Lachlan Mackenzie (eds.). 2005. *Crucial readings in Functional Grammar*, 109–36. Amsterdam: John Benjamins.
Keizer, M. Evelien. 2004. Postnominal PP complements and modifiers: A cognitive distinction. *English Language and Linguistics* 8(2), 1–28.
Keizer, M. Evelien. 2005. The discourse function of close appositions. *Neophilologus* 89, 447–67.
Keizer, M. Evelien. 2007. *The English noun phrase: The nature of linguistic categorization*. Cambridge: Cambridge University Press.
Keizer, M. Evelien. 2010. We teachers, you fools: Pronouns in close appositions. Presented at the International Conference on Functional Discourse Grammar (IC-FDG-2010), Centro Científico e Cultural de Macau, Lisbon.
Keizer, M. Evelien. To appear. English pronouns in Functional Discourse Grammar. *Language Sciences* (special issue).
Kroon, Caroline. 1995. *Discourse particles in Latin*. Amsterdam: Gieben.
Langacker, Ronald W. 2002. *Concept, image, and symbol: The cognitive basis of grammar*, 2nd edn. Berlin: Mouton de Gruyter.
Lee, D. W. 1952. Close apposition: An unresolved problem. *American Speech* 27, 268–75.
Leech, Geoffrey. 2006. *A glossary of English grammar*. Edinburgh: Edinburgh University Press.
Leech, Geoffrey & Jan Svartvik. 1994. *A communicative grammar of English*, 2nd edn. London: Longman.
Pearce, Michael. 2007. *The Routledge dictionary of English language studies*. London: Routledge.
Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.
Radford, Andrew. 1988. *Transformational Grammar: A first course*. Cambridge: Cambridge University Press.
Radford, Andrew. 1997. *Syntactic theory and the structure of English: A minimalist approach*. Cambridge: Cambridge University Press.

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press

Rijkhoff, Jan. 2002. *The noun phrase*. Oxford: Oxford University Press.
Saxena, Anju. 2006. Pronouns. In Keith Brown (ed.), *Encyclopedia of language and linguistics*, 2nd edn. 131–3. Oxford: Elsevier.
Sopher, H. 1971. Apposition. *English Studies* 52, 401–12.
Van der Auwera, Johan. 1992. Free relatives. In Fortescue *et al.* (eds.), 329–54.
Van Valin, Robert D. 1993. A synopsis of Role and Reference Grammar. In Robert D. Van Valin (ed.), *Advances in Role and Reference Grammar*, 1–164. Amsterdam: John Benjamins.
Van Valin, Robert D. 2001. *An introduction to syntax*. Cambridge: Cambridge University Press.
Van Valin, Robert D. & Randy J. LaPolla. 1997. *Syntax. structure, meaning and function*. Cambridge: Cambridge University Press.
Wales, Katie. 1996. *Personal pronouns in present-day English*. Cambridge: Cambridge University Press.

https://doi.org/10.1017/S1360674311000050 Published online by Cambridge University Press