

Demonstrative TO in Polish – an L_RFG analysis

Sebastian Zawada

Institute of Computer Science,
Polish Academy of Sciences

Proceedings of the LFG'24 Conference

Miriam Butt, Jamie Y. Findlay and Ida Toivonen (Editors)

2024

PubliKon

lfg-proceedings.org

Abstract

This paper offers an L_RFG analysis of the Polish demonstrative word *to*, which may occur in two syntactic environments: (i) in typical nominal positions and (ii) in the unique copular structure TO + BYĆ ‘to be’ + NP, in which the right-hand NP appears in the nominative case and triggers agreement with the copula. In (i), *to* may only refer to antecedents lacking number and gender (e.g., clauses), whereas in (ii) it may refer to any antecedent. *To* is analysed as one underspecified vocabulary item exponing two structures: a noun lacking number and gender in (i) and a demonstrative identifier in (ii). The difference in their anaphoric possibilities follows from blocking: *to* is blocked in (i) by personal pronouns whenever the antecedent has number and gender, due to an anaphoric agreement mechanism and the specification of vocabulary items.

1 Introduction

In typical nominal positions, Polish personal pronouns and the demonstrative pronoun *to* used anaphorically are in a complementary distribution, as shown in (1a)–(1b) and (2a)–(2b).[†] Whenever the antecedent is a nominal phrase bearing number and gender (*nowy komputer* ‘a new computer.SG.M’ in (1)), a corresponding personal pronoun must be used (*on* ‘PERS.SG.M’). Otherwise, for example, in the case of the clausal antecedent in (2), the demonstrative word *to* is used.¹ As will be shown below, *to* in (2a)–(2b) is a noun, and hence this use of *to* is dubbed here TO_N.

- (1) Kupiłem [nowy komputer]_i.
‘I bought a new computer.’
- a. Jest {on_i / #to_i} świetnym narzędziem do nauki.
is.3SG PERS.SG.M / this excellent.INS tool.INS to study
‘It is an excellent tool for studying.’
- b. {Miał on_i / #Miało to_i} świetne parametry.
had.3SG.M PERS.SG.M / had.3SG.N this excellent specifications.ACC
‘It had excellent specifications.’
- (2) [Polska przegrała kolejny mecz]_i.
‘Poland lost another match.’
- a. Było {to_i / #ono_i} prawdziwym skandalem. (TO_N)
was.3SG.N this / PERS.SG.N real.INS scandal.INS
‘It was a real scandal.’

[†]I would like to thank Adam Przepiórkowski for many discussions on the topic of this paper and for providing comments on my draft. I am also grateful to Ash Asudeh and the LFG’24 reviewers for their constructive comments. My thanks also go to Oleg Belyaev for a helpful email exchange. Finally, thanks to the LFG’24 audience for their helpful feedback. All remaining errors are my own. This research was supported by the National Center of Science, Poland (Project No. 2022/45/N/HS2/03109).

¹The class of Polish third-person personal pronouns consists of three singular forms (*on*.SG.M/*ona*.SG.F/*ono*.SG.N) and two plurals: masculine virile (*oni*.PL.MV) and other than masculine virile (*one*.PL.NMV). Personal pronouns are further glossed as PERS. The other morphosyntactic abbreviations follow the Leipzig Glossing Rules (<https://www.eva.mpg.de/lingua/resources/glossing-rules.php>). To avoid clutter, only morphosyntactic information relevant to the phenomenon under investigation is presented. Since personal pronouns in all examples in this paper are in the nominative case, this information is not included in the glosses.

- b. Zaskoczyło mnie $\{to_i / \#ono_i\}$. (TO_N)
 surprised.3SG.N I.ACC this / PERS.SG.N
 ‘It surprised me.’

Sentences (1a) and (2a) are typical Polish copular clauses with instrumental predicates, whereas (1b) and (2b) contain standard non-copular verbs. The anaphoric items in all these sentences can be replaced by lexical nouns. The word order in the sentences above (as well as in (1c) and (2c) below) is driven by information structure and is quite flexible: in principle, in each of these sentences the pronoun can be placed initially, but this would be slightly less natural and could suggest a contrastive interpretation. Note that Polish verbs express gender in the past tense, but not in the present, and hence *Jest* in (1a) is simply glossed as ‘is.3SG’.

Intriguingly, *to* must be used regardless of the syntactic status of the antecedent in the unique copular structure consisting of *to*, the copula BYĆ ‘to be’, and the nominal predicate in the nominative case agreeing with the copula, see (1c) and (2c). This use of *to* is dubbed here TO_D (in §2.1.2, I argue that its syntactic category is demonstrative identifier).²

- (1) Kupiłem [nowy komputer]_i.
 ‘I bought a new computer.’
- c. Był $\{to_i / \#ono_i\}$ świetny zakup. (TO_D)
 was.3SG.M this / PERS.SG.M excellent purchase.SG.M.NOM
 ‘It was a great purchase.’
- (2) [Polska przegrała kolejny mecz]_i.
 ‘Poland lost another match.’
- c. Był $\{to_i / \#ono_i\}$ prawdziwy skandal. (TO_D)
 was.3SG.M this / PERS.SG.N real scandal.SG.M.NOM
 ‘It was a real scandal.’

At this point, it is worth adding that Polish personal pronouns indeed require the presence of some *syntactic* features on their antecedents. Their semantic status does not play a role: they can be inanimate concrete objects (the new computer in (1)), humans (see (3)), and even abstract objects, such as an event expressed by a gerund form (see (4)).³ What is relevant is that they bear agreement features (henceforth: AF) corresponding to the pronoun’s features.⁴

- (3) Ale najbardziej bałam się o dziecko; ono_i musiało
 but most feared.3SG.F REFL about child PERS.SG.N had.3SG.N
 urodzić się zdrowe!
 born.INF REFL healthy
 ‘But what I feared most was for the baby: it had to be born healthy.’ (NCP)

²Note that the labels TO_N and TO_D, which hint at the categorial status discussed later, are only intended to help the reader navigate the data and do not have any theoretical value. Labels such as ITEM1 and ITEM2 could be used instead.

³The sentences marked ‘NCP’ come from the National Corpus of Polish (Przepiórkowski et al. 2012).

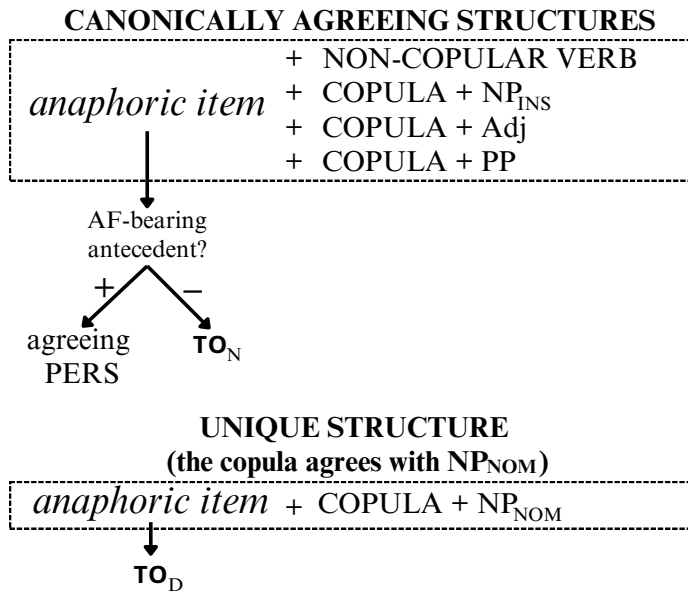
⁴To jointly refer to NUMBER and GENDER, I use the term *agreement features* (AF) instead of *φ-features* – commonly used in the Minimalist and formal semantic literature – to avoid possible confusion with the *φ*-function, mapping c-structure to f-structure. Antecedents will often be described as AF-less and AF-bearing, meaning *agreement-featureless* and *agreement-features-bearing*, respectively.

- (4) Gdy zaczęło się kupowanie_i, miało ono_i gwałtowny
 when started.3SG.N REFL buying.SG.N.NOM had.SG.N PERS.SG.N violent
 charakter.
 character
 ‘When the buying started, it had a violent nature.’ (NCP)

The aim of this paper is to provide an analysis of the observed phenomenon, which is recapitulated in a concise generalisation in (5), and for convenience schematically presented in Figure 1. More specifically, the paper answers the following questions: Why does only TO_N have restricted anaphoric possibilities? What is the relationship between TO_N and TO_D ?

- (5) When used anaphorically, *to* in the copular structure with the nominative nominal predicate (TO_D) may have an antecedent of any type, whereas *to* in any other structure (TO_N) can only have an antecedent lacking number and gender.

Figure 1: Anaphoric items used in Polish syntactic constructions



The proposed analysis adopts the Lexical-Realizational Functional Grammar framework (L_RFG; Melchin et al. 2020; Asudeh et al. 2023, Asudeh & Siddiqi 2023: §5.3; among others). I argue that the realisational character of this approach and a competition mechanism that it employs allow for a parsimonious and insightful explanation of the observed phenomenon. *To* is analysed as an underspecified vocabulary item expounding two structures: an AF-less noun (which gives rise to TO_N) and a demonstrative identifier (TO_D). The AF-less noun TO_N can only be inserted when personal pronouns cannot be expounded because their licensing conditions are not satisfied. This situation arises when the structure they are supposed to expone lacks AF. The presence of AF in the anaphoric nominal structure is, in turn, determined by the form of antecedent.⁵

⁵It should be noted that the paper focuses exclusively on the nominative case of the relevant forms. The complexities of Polish personal pronouns in other cases go far beyond the scope of this paper. Suffice it to say that all non-nominative forms of personal pronouns are suppletive and can have up to four forms:

When the antecedent lacks AF, the nominal structure also lacks them and is expounded by the elsewhere form *to*. The demonstrative identifier (TO_D) can only be realised as *to* because the conditions licensing the personal pronoun stem are in this case never satisfied.

In §2, I present more data and show that TO_N is a noun, but it does not bear gender or number, whereas TO_D is a demonstrative identifier (Diessel 1999). In §3, I introduce the basics of the L_RFG framework and present the analysis briefly sketched in the previous paragraph. §4 discusses some limitations of the proposed analysis and concludes the paper.

2 Analysis

2.1 Categorial status

2.1.1 TO_N is a noun, TO_D is not

The item called here TO_N is a noun, whereas TO_D is not: the former passes all diagnostic tests for nounhood presented below, whereas the latter does not pass any of them.

The first test is modifiability by adjectives. As shown in (6a) and (6b), only TO_N can be modified by the adjective SAM ‘alone’.⁶

- (6) [Polska przegrała kolejny mecz]_i.
 ‘Poland lost another match’
- a. (Samo) *to*_i było skandalem. (TO_N)
 alone this.was.SG.N scandal.SG.M.INS
 ‘This alone was a scandal.’
- b. (*Samo) *to*_i był skandal. (TO_D)
 alone this.was.SG.M scandal.SG.M.NOM
 Intended: ‘This alone was a scandal.’

Another feature that differentiates TO_N from TO_D and indicates the nominal character of the former is the coordination of *to* with another NP. This is possible in the case of TO_N, but not in the case of TO_D, see (7a)–(7b).⁷

- (7) [Polska przegrała kolejny mecz]_i.
 ‘Poland lost another match’
- a. *To*_i i pomeczowa bójka było skandalem. (TO_N)
 this and post-match brawl.SG.F.NOM was.3SG.N scandal.SG.M.INS
 ‘This and the post-match brawl were a scandal.’
- b. *To*_i (*i pomeczowa bójka) był skandal. (TO_D)
 this and post-match brawl.SG.F.NOM was.3SG.M scandal.SG.M.NOM
 Intended: ‘This and the post-match brawl were a scandal.’

long (accented), short (unaccented), a special form occurring after prepositions, and a clitic form attached to prepositions. For instance, consider the genitive singular masculine pronoun in these four versions, respectively: *jego / go / niego / -ń*.

⁶Due to the pronominal semantics, other adjectives do not easily combine with *to*.

⁷The unexpected singular neuter form of the copula in (7a) is briefly discussed in §2.2.1.

Two other empirical facts suggest that TO_N is a noun. Polish nouns inflect for case and, obviously, serve as arguments of verbs and prepositions. *To* also inflects for case (e.g., *tego*.GEN or *temu*.DAT) and can be found in the position of a nominal argument of a verb or preposition (e.g., *Nie lubię tego* ‘I don’t like this.GEN’). Importantly, it will never have an antecedent bearing AF in such situations. This can be treated as further evidence for the nounhood of TO_N .

Given that TO_D can only be found in one particular construction ($TO+BYĆ+NP_{NOM}$), it is not possible to directly check if it inflects for case or serves as an argument of a preposition or a verb (other than the copula *BYĆ* agreeing with the predicate). However, the aforementioned fact that *to* inflected for case or placed in an argument position never refers to an AF-bearing antecedent – and thus always behaves like TO_N – may be treated as indirect evidence against the nominal status of TO_D .

In conclusion, TO_N is a noun, whereas TO_D is not, with the arguments summarised in Table 1.

Table 1: Summary of the diagnostic tests for nounhood

| | TO_N | TO_D |
|------------------------------|--------|--------|
| combines with adjectives | + | – |
| coordinates with NPs | + | – |
| inflects for case | + | (–) |
| serves as a dependent of V/P | + | (–) |

2.1.2 TO_D is a demonstrative identifier

While it is straightforward to demonstrate that TO_D is not a noun, determining its true categorial status is much more challenging. In the Polish tradition, TO_D has been analysed either as a noun (for example by Wiśniewski 1987) or as a quasi-verb – a predicative item formed by the complex unit *TO BYĆ* (see Bogusławski 1988, 2002). Here I argue that TO_D is neither of them, but belongs to the class of demonstrative identifiers (Diessel 1999).⁸

⁸The analysis treating TO_D as a quasi-verb assumes that the form *to* in sentences such as (1c) and (2c) is an integral part of the *verbal* unit *TO BYĆ*, which takes the nominative NP as its sole argument. While there is no space to provide a full critique of this kind of analysis here, it suffices to note that all Polish verbs (and quasi-verbs, such as *TRZEBA* ‘necessary’) attach the negative marker to the left. However, in sentences with TO_D , the negative marker always directly precedes the copula, indicating that this is the only verbal element in this sentence, see (i)–(ii). Employing the attachment of the negation marker as a test for determining verbal character is adopted from Bondaruk (2013: 218).

- (i) a. (Nie) trzeba (*nie) było biegać. (quasi-verb)
NEG necessary.QV NEG was.3SG.N run.INF
 ‘It wasn’t necessary to run.’
- b. (Nie) było (*nie) trzeba biegać. (quasi-verb)
NEG was.3SG.N NEG necessary.QV run.INF
- (ii) a. (*Nie) to (nie) był skandal. (TO_D)
NEG this NEG was.3SG.M scandal.SG.M.NOM
 ‘It wasn’t a scandal.’
- b. (Nie) był (*nie) to skandal. (TO_D)
NEG was.3SG.M NEG this scandal.SG.M.NOM

Also, it is worth noting that adopting the quasi-verb approach would require some reformulation of the final analysis presented in §3, but the main insight would remain intact.

TO_D does not inflect and cannot trigger agreement, and yet it is a strictly referential item that can be used anaphorically and deictically. In his monograph on demonstratives, Diessel (1999: §4:3) calls such peculiar items occurring exactly (and uniquely) in copular clauses demonstrative *identifiers* (in opposition to demonstrative *pronouns*, *determiners* and *adverbs*). The author argues that they can be identified in many typologically unrelated languages based on phonological and morphological clues.⁹ In some languages – such as Supyire, Karanga and Western Bade (Diessel 1999: 80–83) – the stem of the demonstrative in a copular or non-verbal clause is different from the stem of demonstratives used in other contexts.

In other languages, including some Indo-European languages such as French, German, Modern Hebrew, Dutch (den Dikken 2024), and Serbo-Croatian (Browne 1999), the demonstrative identifier can be distinguished based on morphological and syntactic evidence. Namely, it does not agree nor inflect, see the German example in (8).¹⁰

- (8) a. Das ist meine Schwester. (German)
 DEM.NOM/ACC.SG.N is my sister.SG.F
 ‘This is my sister.’ (Diessel 1999: 88)
- b. Das sind meine Freunde. (German)
 DEM.NOM/ACC.SG.N are my friend.PL
 ‘These are my friends.’ (Diessel 1999: 88)

The item dubbed here TO_D perfectly matches the morphological and syntactic criteria for demonstrative identifier, and is therefore classified as belonging to this category.

Of course, postulating a single-element grammatical category may raise some concerns. Diessel’s arguments (having a phonologically distinct stem and/or failing to agree and inflect) do not form conclusive evidence for a distinct category on their own. It seems, however, that we have no alternative – TO_D does not conform to any category usually assumed to exist in Polish. I consider this fact, combined with Diessel’s morphological and syntactic arguments and the frequency of the cross-linguistic pattern, to be sufficient justification for classifying TO_D as a demonstrative identifier. In the c-structure, I will refer to it with the label Dem.

Determining that TO_D is not a noun, but must be analysed as a distinct category (such as Dem adopted in this paper), will turn out sufficient to account for the generalisation presented in (5). This allows us to remain agnostic about the c- and f-structural representation of sentences containing TO_D, such as (1c) and (2c), repeated below as (9) and (10). The only thing relevant for the analysis at hand is that the c-structure rules responsible for building these sentences employ Dem.

- (9) Był to świetny zakup. (TO_D)
 was.3SG.M this excellent purchase.SG.M.NOM
 ‘It was a great purchase.’
- (10) Był to prawdziwy skandal. (TO_D)
 was.3SG.M this real scandal.SG.M.NOM
 ‘It was a real scandal.’

⁹Diessel’s *morphological* evidence for demonstrative identifiers is in fact *morphological and syntactic*: it relies not only on inflection but also on agreement facts.

¹⁰The morphosyntactic glosses in (8) are drawn from Diessel (1999); however, the demonstrative identifier *das* should probably be analysed as caseless and lacking AF, only homonymous to the NOM/ACC.SG.N demonstrative.

Here, we can only point out why it is not straightforward how to analyse such sentences in LFG, leaving a comprehensive analysis for future research. Note that the nominative predicate agrees with the copula, and hence – at least according to what is typically assumed for Polish – it should be analysed as SUBJ. In consequence, *to* must be assigned a different grammatical function, but any that is usually postulated in LFG cannot be adequately motivated for such a strange item: serving as the subject of predication, having pronominal, referential semantics, and being syntactically inert.¹¹ At this point, it can only be concluded that a thorough analysis of sentences such as (9)–(10) can provide valuable insights into the research on copular clauses in LFG (see Dalrymple et al. 2019: §5.4.5 for an overview).

2.2 TO_N does not have agreement features

Consider (11), a modified version of (2b). In the subject position, TO_N triggers what seems to be 3rd person singular neuter agreement.

- (11) To mnie zaskoczyło. (TO_N)
 this I.ACC surprised.3SG.N
 ‘It surprised me.’

At first glance, this may be treated as evidence that TO_N itself is a singular neuter noun. Note, however, that third person singular neuter is also the *default* agreement in Polish (Dziwirek 1990), triggered when the subject is not a nominative NP equipped with AF. The default agreement is observed, for instance, when the subject is an infinitival phrase or an accusative numeral phrase, as illustrated in (12) and (13), respectively.¹²

- (12) Strzelać było zakazane.
 shoot.INF was.3SG.N forbidden
 ‘Shooting (*lit.* to shoot) was forbidden.’ (NCP)
- (13) Pięć osób przyszło.
 five.ACC person.PL.F.GEN came.3SG.N
 ‘Five people came.’ (NCP)

I argue that (11) is also an instance of default agreement triggered exactly for the same reason as in (12)–(13), namely, the absence of an AF-bearing nominative NP in the subject position. *To* in the subject position in (11) is a nominative NP, but lacks AF.

In §2.2.1 and §2.2.2, I present two arguments for AF-less status of TO_N, based on coordination facts and on observations related to situations in which TO_N itself is the antecedent of an anaphoric item.¹³

¹¹This suggests exploring alternative approaches, such as the reductionist approach proposed in Patejuk & Przepiórkowski (2016), which eliminates specific grammatical functions from the f-structure if they cannot be justified by established syntactic tests.

¹²Arguments supporting the analysis assuming the accusative case on Polish numeral subjects can be found, for instance, in Przepiórkowski (1999: §5.3.1.1).

¹³The idea that the demonstrative word *to* in such sentences is AF-less – rather than possessing true singular neuter features – has been proposed in Bartošová (2017) for the Czech demonstrative *to*. However, it was not supported by arguments of the sort presented here.

2.2.1 Coordination

Polish coordinated NPs trigger either plural agreement or, with some restrictions, closest conjunct agreement, see (14). In contrast, coordinated AF-less items always trigger third-person singular neuter (default) agreement, see (15) (such coordinations are slightly degraded because a more natural option – with gerunds instead of infinitives – is available).

- (14) Okno i krzesło {było / były} w tym pokoju
 window.SG.N and chair.SG.N was.SG.N / were.PL in this room
 najładniejsze.
 beautiful.SUP
 ‘The window and the chair were the most beautiful in this room.’
- (15) [Skoczyć na bungee]_{InfP} i [wejść na Kilimandżaro]_{InfP} {?sprawiło
 jump.INF on bungee and climb.INF on Kilimanjaro made.3SG.N
 / *sprawiły} mu najwięcej problemów.
 / made.PL he.DAT most troubles
 ‘Bungee jumping (*lit.* to bungee-jump) and climbing (*lit.* to climb) Kilimanjaro posed the greatest challenges for him.’

Crucially, the coordination of two instances of TO_N cannot trigger plural agreement. Therefore, it aligns with the AF-less infinitival phrases in (15), and not with the singular neuter nouns shown in (14). Note that coordinating two anaphoric uses of *to* is problematic, as it seems impossible to determine which *to* refers to what. However, TO_N can also be used deictically and as a correlate heading a complementiser phrase. Instances of these uses of *to* can be easily coordinated, but they never trigger plural agreement, see (16)–(17).

- (16) To_→ i to_→ mnie {zaskoczyło / *zaskoczyły}.
 this and this I.ACC surprised.3SG.N / surprised.3PL
 ‘This and this surprised me.’ (‘→’ = pointing gesture)
- (17) [To, że Polska przegrała], i [to, że kibice się pobili],
 this COMP Poland lost and this COMP fans REFL brawled
 {było / *były} skandalem.
 was.SG.N / were.PL scandal.SG.M.INS
 ‘[That Poland lost]_{nominalised} and [that fans got into a brawl]_{nominalised} was a scandal.’

Interestingly, the default agreement is also most natural in cases where *to* is coordinated with a lexical noun that serves as the closest conjunct, see (18), which is a modified version of (7a).

- (18) [Polska przegrała kolejny mecz]_i; To_i i pomeczowa bójka
 Poland lost another match this and post-match brawl.SG.F.NOM
 {było / ?była / ??były} skandalem.
 was.SG.N / was.SG.F / were.PL scandal.SG.M.INS
 ‘Poland lost another match. This and the post-match brawl was a scandal.’

Even though the closest conjunct is feminine, the singular neuter form of the verb is the most acceptable. I argue that this is because TO_N , being AF-less, disrupts the standard agreement mechanism of coordinated phrases and imposes default agreement. Other possibilities (closest conjunct and plural agreement) are not fully excluded, suggesting that the grammar exhibits some flexibility in this respect, possibly because such coordinations are not very common.

2.2.2 Reference to TO_N

As previously mentioned, Polish personal pronouns must agree with their antecedents, cf. (1). Note that the singular neuter pronoun *ono* cannot refer to an instance of TO_N , as shown in (19).

- (19) A: To_i mnie zaskoczyło. (TO_N)
 this I.ACC surprised.3SG.N
 ‘It surprised me.’
- B: Mnie też { to_i / # ono_i } zaskoczyło. (TO_N)
 I.ACC also this / PERS.SG.N surprised.3SG.N
 ‘It surprised me as well.’

If we accept the claim that TO_N is AF-less, the phenomenon illustrated in (19) is immediately explained: the pronoun *ono* ‘PERS.SG.N’ expects its antecedent to have singular number and neuter gender. The word *to* in the sentence uttered by A lacks these features, and thus *ono* cannot take it as an antecedent.

Also, this observation justifies the exact formulation of the main generalisation presented in (5). We say that TO_N must refer to AF-less items, although examples (1)–(2) may suggest another interpretation: that TO_N must refer to *non-nominal* items. However, this would undergenerate: TO_N can take as an antecedent another instance of TO_N , which is a noun but lacks AF.

Let us also present evidence that the two tests employed here are indeed sensitive to the AF-bearing vs. AF-less status of some phrase. As already shown in (4), repeated below as (20), Polish gerunds are referred to with the use of the personal pronoun *ono* ‘PERS.SG.N’. Therefore, Polish gerunds pass one of the two tests for having true singular neuter features.

- (20) Gdy zaczęło się kupowanie_i, miało ono_i
 when started.3SG.N REFL buying.SG.N.NOM had.SG.N PERS.SG.N
 gwałtowny charakter.
 violent character
 ‘When the buying started, it was of (*lit.* had) a violent nature.’ (NCP)

This provides a welcome opportunity to validate the tests themselves. Namely, we expect the other test (agreement triggered by coordinated phrases) to yield the result consistent with (20). In other words, we expect coordinated gerunds to be able to trigger plural agreement. If they fail to do so, the tests turn out to be ineffective, as they give conflicting results.

However, coordinated gerunds can indeed trigger plural agreement, as demonstrated in (21)–(22), and thereby they confirm the reliability of the tests.

- (21) Osądzanie i szufladkowanie powodują, że
judging.SG.N.NOM and pigeonholing.SG.N.NOM cause.3PL COMP
rozmówca musi koncentrować się na obronie poczucia własnej wartości.
interviewee must focus.INF REFL on defence feeling own value
‘Judging and pigeonholing cause the interviewee to focus on defending their
self-esteem.’ (NCP)
- (22) Nagradzanie i karanie są ściśle powiązane
rewarding.SG.N.NOM and punishing.SG.N.NOM are.PL strictly linked
z procesem motywowania pracowników.
with process motivating employees
‘Rewarding and punishing are closely linked to the process of motivating em-
ployees.’ (NCP)

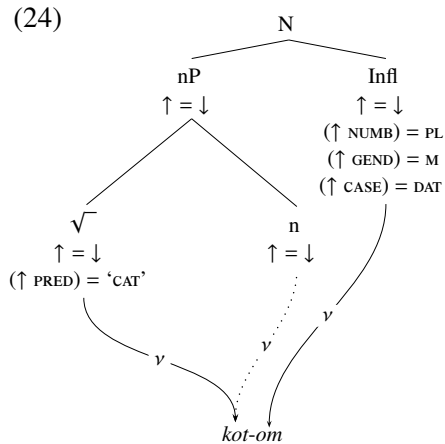
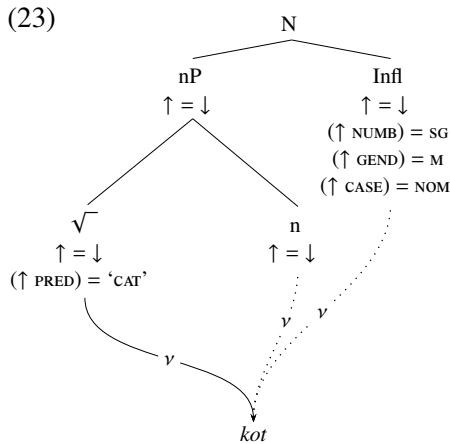
The facts established in this section will serve as a basis for the formal analysis presented in §3. The key points are: (i) TO_N is a noun, but it lacks AF, and (ii) TO_D is not a noun, but it belongs to the class of demonstrative identifiers (Dem) postulated by Diessel (1999).

3 L_RFG formalisation

3.1 Introduction to L_RFG

The analysis presented here adopts Lexical-Realizational Functional Grammar (see Melchin et al. 2020; Asudeh et al. 2023, Asudeh & Siddiqi 2023: §5.3, a.o.), which combines LFG with Distributed Morphology (DM; Halle & Marantz 1993), a realisational and morpheme-based approach to word formation.

In L_RFG, the c-structure terminal nodes are populated not by words coming from the lexicon, but by c-structure labels (categories) equipped with f-descriptions, which are in turn mapped to v(ocabulary)-structures via a separate function (ν). See the examples involving two Polish forms – *kot* ‘cat.SG.M.NOM’ and *kotom* ‘cat.PL.M.DAT’ – in (23)–(24), which are further discussed below.



As shown in the trees above, I assume that the Polish nominal spine consists of a root providing the PRED value ($\sqrt{\quad}$), a morpheme responsible for the category (the nominaliser n), and a nominal suffix (Infl) expressing gender, number and case.¹⁴

Instead of a lexicon, $L_{\text{R}}\text{FG}$ postulates a list of vocabulary items (VIs), that is, a list of mappings from an exponendum (the tuple on the left-hand side of $\overset{\nu}{\mapsto}$ in (25)) to an exponent (the v -structure on the right-hand side of $\overset{\nu}{\mapsto}$ in (25)). The v -structure is further mapped onto the prosodic structure and finally the phonological string. For the purposes of this paper, I present simplified versions of v -structures, containing only the phonological representation, as in (23)–(24).

(25) The structure of a vocabulary item (Asudeh et al. 2023: 23)

$$\langle [C_1, \dots, C_n] , \quad F \cup G \cup I \quad \rangle \overset{\nu}{\mapsto} \left[\quad \right]_{v\text{-structure}}$$

distribution *function/meaning*

The first member of the exponendum is a list of syntactic categories (the labels of c -structure terminal nodes). A vocabulary item can contain more than one category C , as it can *span* multiple adjacent c -structure nodes.¹⁵ The second member ($F \cup G \cup I$) is the union of three possibly empty sets: a set of f -descriptions (F), a set of descriptions of s -structures and Glue meaning constructors (G), and a set of descriptions of i -structures (I).

Consider the Vocabulary Items used in the trees presented above: (26), expounding the stem *kot* ‘cat’, and (27), expounding the plural dative suffix. Given that the suffix *-om* marks the plural dative in all Polish inflecting nouns, regardless of their gender, the VI in (27) is not specified for this feature.

(26) $\langle [\sqrt{\quad}], (\uparrow \text{ PRED}) = \text{‘CAT’} \quad \rangle \overset{\nu}{\mapsto} \textit{kot}$
cat : $(\uparrow_{\sigma} \text{ VAR}) \multimap (\uparrow_{\sigma} \text{ RESTR})$

¹⁴This is a departure from what is typically assumed in Distributed Morphology, where gender, number and case are usually represented on separate nodes (see Norris 2022 for an overview). However, as noted by Belyaev (2024), there are no good arguments for such a large structure in Russian, where gender, number and case are marked with a single fusional suffix. This also holds true for Polish. Note that the gist of the analysis presented here does not hinge on the size of the nominal spine and can be easily reformulated to adhere to typical DM standards.

¹⁵On spanning, see for example Haugen & Siddiqi (2016) and references therein.

$$(27) \quad \langle [\text{Infl}], (\uparrow \text{NUMB}) = \text{PL} \quad \rangle \xrightarrow{\nu} -om \\ (\uparrow \text{CASE}) = \text{DAT}$$

We have not postulated vocabulary items for the nodes *n* and the nominative version of *Infl* (specified as in (23)). The mapping of these nodes in (23)–(24) represent instances of the so-called Pac-Man spanning (expressed by dotted line): if a terminal node would be left unexponed due to the absence of an appropriate vocabulary item, it is mapped to a neighbouring exponent.¹⁶

The mapping from *c*-structure to *v*-structure must maximally satisfy the set of **MostInformative** functions, which are an $L_{\text{R}}\text{FG}$ implementation of the Subset Principle postulated in Distributed Morphology. We will concentrate on **MostInformative**_{*f*}, as it will serve as the primary tool in our analysis. A VI to be inserted must contain a subset (possibly a perfect match) of the features located at the relevant part of the *c*-structure. The **MostInformative**_{*f*} chooses the VI that is best specified in this respect. It takes two VIs as input and returns the VI which defines an *f*-structure containing the greater set of features, as shown in (28).¹⁷ The function π_1 returns the VI's exponentum, π_2 extracts the second coordinate out of it (the ‘ $F \cup G \cup I$ ’ part), and Φ (‘big phi’) maps *f*-descriptions to the minimal *f*-structures that satisfy them. The notation $g \sqsubset f$ indicates that the *f*-structure *g* properly subsumes the *f*-structure *f* (see Bresnan et al. 2016: Ch. 5).

(28) Given two VIs, α and β ,

$$\text{MostInformative}_f(\alpha, \beta) = \begin{cases} \alpha & \text{if } \exists f.f \in \Phi(\pi_2(\pi_1(\alpha))) \wedge \forall g.g \in \Phi(\pi_2(\pi_1(\beta))) \rightarrow g \sqsubset f \\ \beta & \text{if } \exists f.f \in \Phi(\pi_2(\pi_1(\beta))) \wedge \forall g.g \in \Phi(\pi_2(\pi_1(\alpha))) \rightarrow g \sqsubset f \\ \perp & \text{otherwise} \end{cases}$$

Having provided this brief introduction, we can now turn our attention to the analysis of Polish pronouns.

3.2 Analysis

A straightforward, “naive” analysis of the presented phenomena could posit two lexical entries – one for TO_N and one for TO_D . The former would include a restriction forcing TO_N to refer solely to AF-less antecedents. Such a solution, however, would be nothing more than a description of the empirical data in formal terms. It postulates mere homonymy of TO_N and TO_D , neglecting their formal and semantic affinity, and just stipulates the difference in their anaphoric possibilities (AF-less vs. any antecedent), without providing any insights about where it comes from.

The $L_{\text{R}}\text{FG}$ framework allows for a more insightful analysis, in which a single vocabulary item related to the form *to* expones both TO_N and TO_D . Crucially, in the proposed analysis, the aforementioned difference (AF-less vs. any antecedent) follows from an independent mechanism of competition between vocabulary items (**MostInformative**_{*f*}).

¹⁶Note that (23)–(27) present just a toy example ignoring inflectional classes, affected by factors such as morphophonology of nominal stems. For a full-fledged analysis of part of a declension system in $L_{\text{R}}\text{FG}$, see Asudeh et al. (2024), devoted to Latin.

¹⁷The formalisation of **MostInformative**_{*f*} presented here differs from the one assumed in previous $L_{\text{R}}\text{FG}$ works. I would like to thank Adam Przepiórkowski for suggesting it and Ash Asudeh for discussion.

The first ingredient of the analysis are the relevant vocabulary items. The demonstrative stem *to* is the least specified form, as it only expones the pronominal root providing the PRED value, see (29).^{18,19}

$$(29) \quad \langle [\sqrt{\quad}], (\uparrow \text{PRED}) = \text{'pro'} \rangle \xrightarrow{\nu} to \\ x_i : \uparrow_{\sigma}$$

The personal pronoun stem (*on*) can only be used in the presence of AF (regardless of their value), manifesting the phenomenon known as secondary exponence (see Noyer 1997; Asudeh et al. 2024). It means that (30) requires that GEND and NUMB be present in the f-structure to which the relevant portion of c-structure ($[\sqrt{\quad}, n]$) is mapped. However, these features and their values cannot be directly defined on $\sqrt{\quad}$ or *n* (in fact, they are specified on Infl, as in (23)–(24)). This is modelled by the existential constraints (\uparrow GEND) and (\uparrow NUMB).

$$(30) \quad \langle [\sqrt{\quad}, n], (\uparrow \text{PRED}) = \text{'pro'} \rangle \xrightarrow{\nu} on \\ (\uparrow \text{NUMB}) \\ (\uparrow \text{GEND}) \\ x_i : \uparrow_{\sigma}$$

Note that the f-descriptions in the VIs are sufficient for *on* to outcompete *to* (via **MostInformative_f**), whenever the former is applicable. At first glance, this makes the distinction between the two spans ($[\sqrt{\quad}, n]$ vs. $[\sqrt{\quad}]$) redundant. However, this difference is intended to capture the fact that *on* occurs exclusively in the pronoun, whereas *to* appears in various other forms, whose detailed analysis is beyond the scope of this paper. Crucially, it is present in the demonstrative adjective (see *to dziecko* ‘this child.SG.N’), which seemingly does have AF, as a result of adjective-noun concord. Without the difference in spans, the analysis would incorrectly predict that the stem *on* should also occur in the adjective.²⁰

Let us now consider the suffixes. The singular feminine suffix that can be found in *on-a* ‘PERS.SG.N’ is presented in (31), and the singular neuter suffix occurring in *on-o* ‘PERS.SG.N’ is shown in (32). The singular masculine nominative (*on*) lacks an overt suffix, and as a result such a specification of Infl will be mapped to the stem through Pac-Man spanning (as in (23)).

¹⁸The semantics of anaphoric items is presented here simply as a variable (x_i). A more comprehensive semantic account, including deictic uses, should take into consideration an alternative approach that treats the pronominal root as an operator forming a definite description with a hidden argument (see Ahn 2022 and references therein).

¹⁹In Distributed Morphology, pronouns are often argued to lack roots (see, for instance, Moskal 2015), which is also generally adopted in L_RFG. However, if we conceptualise roots just as category-neutral nodes providing the PRED value, it makes sense to have them in pronouns. This also seems to be the most straightforward way (i) to have one VI expounding two items of different categories (the form *to* expounding just the root common for TO_N and TO_D, and possibly for other demonstrative items mentioned in fn. 20), and (ii) to enable the competition between *on* and *to*: they must have something in common, which is here the same root providing the pronominal PRED value.

²⁰Other forms presumably containing the stem *to* include the manner adverb *tak* ‘this way’, and the locative adverbs *tu* ‘here’ and *tam* ‘there’. It is not clear whether *-o* is part of the stem *to* and is subsequently deleted due to morphophonological constraints, or whether the stem itself should be analysed as *t-*. For simplicity, the current analysis assumes the former, though it should not be considered a definitive solution.

$$(31) \quad \langle [\text{Infl}], (\uparrow \text{ NUMB}) = \text{SG} \rangle \xrightarrow{\nu} -a$$

$$(\uparrow \text{ GEND}) = \text{F}$$

$$(\uparrow \text{ CASE}) = \text{NOM}$$

$$(32) \quad \langle [\text{Infl}], (\uparrow \text{ NUMB}) = \text{SG} \rangle \xrightarrow{\nu} -o$$

$$(\uparrow \text{ GEND}) = \text{N}$$

$$(\uparrow \text{ CASE}) = \text{NOM}$$

The second ingredient of the analysis are the c-structure rules. Let us first present the rule forming the demonstrative identifier.

$$(33) \quad \text{Dem} \rightarrow \begin{array}{ccc} \sqrt{\quad} & & \text{dem} \\ \uparrow = \downarrow & & \uparrow = \downarrow \\ (\uparrow \text{ PRED}) = \text{'pro'}$$

Rule (33) simply posits the categoriser *dem*, which gives the demonstrative identifier its category and is restricted to combining with the pronominal root only.

The rules that form nouns are more complex. The lowest part of a noun (nP) consists of an acategorial root ($\sqrt{\quad}$) and the nominaliser (n). The template used in (34) is a metarule saying that any PRED value possible in a language can satisfy it. Its definition is given in (35) (Asudeh et al. 2024).²¹

$$(34) \quad \text{nP} \rightarrow \begin{array}{ccc} \sqrt{\quad} & & \text{n} \\ \uparrow = \downarrow & & \uparrow = \downarrow \\ @\text{ROOT}(_) \end{array} \quad (\text{nP for both lexical nouns and pronouns})$$

$$(35) \quad \text{ROOT}(X) := (\uparrow \text{ PRED}) = \text{'X'}$$

While both lexical nouns and pronouns share the rule presented in (34), they differ at the higher level of noun formation, related to their inflection. The rule forming lexical nouns, previously used in (23) and (24), is shown in (36). It states that nP combines with a node carrying gender, number and case. The templates ending in “!” are special types of templates called *bang macros* (Asudeh et al. 2024), which simply enumerate possible values of a given attribute. The definitions of the bang macros used here are given in (37), with only part of the definition of @CASE! shown to minimise clutter.

$$(36) \quad \text{N} \rightarrow \begin{array}{ccc} \text{nP} & & \text{Infl} \\ \uparrow = \downarrow & & \uparrow = \downarrow \\ & & @\text{NUMBER!} \\ & & @\text{GENDER!} \\ & & @\text{CASE!} \end{array} \quad (\text{N for lexical nouns})$$

$$(37) \quad \begin{array}{ll} \text{NUMBER!} & := (\uparrow \text{ NUMB}) = \text{SG} \mid (\uparrow \text{ NUMB}) = \text{PL} \\ \text{GENDER!} & := (\uparrow \text{ GEND}) = \text{M} \mid (\uparrow \text{ GEND}) = \text{F} \mid (\uparrow \text{ GEND}) = \text{N} \\ \text{CASE!} & := (\uparrow \text{ CASE}) = \text{NOM} \mid (\uparrow \text{ CASE}) = \text{GEN} \mid (\uparrow \text{ CASE}) = \text{DAT} \mid \dots \end{array}$$

²¹Note that the system must be restricted to prevent generating non-existent nouns. This problem is approached in Asudeh et al. (2024) by postulating that nominalisers, specific for particular inflection classes, restrict which PRED values license them. Similarly, the v-structures of the suffixes must specify their hosts. However, since this paper focuses on the distribution of Polish anaphoric items, and not on Polish declension, these issues are not addressed here.

As demonstrated in the introductory section of this paper, TO_N is in a complementary distribution with personal pronouns, which must agree with their antecedents. Given that they are all nouns, they are built by the same rule given in (38). To capture the pronoun’s dependency on its antecedent, the rule for pronouns replaces @GENDER! and @NUMBER! with the restriction @ANT-AGR, which ensures that either the pronoun and the antecedent match in gender and number or both lack these features. Such an agreement mechanism is present exclusively in pronouns, and hence the @ANT-AGR-bearing Infl is equipped with the constraining equation which requires the presence of the *pro* PRED value in the f-structure it is mapped to.

$$(38) \quad N \rightarrow \begin{array}{c} \text{nP} \\ \uparrow = \downarrow \end{array} \quad \begin{array}{c} \text{Infl} \\ \uparrow = \downarrow \\ (\uparrow \text{ PRED FN}) =_c \textit{pro} \\ \text{@ANT-AGR} \\ \text{@CASE!} \end{array} \quad (\text{N for pronouns})$$

Formally, @ANT-AGR is a template comprising a few equations, as shown in (39).

$$(39) \quad \text{@ANT-AGR} := \begin{array}{l} \text{(i)} \quad (\text{GF}^+ \uparrow) \text{ GF} = \% \text{ANT} \\ \text{(ii)} \quad \mathcal{R}((\uparrow_\sigma \text{ INDEX})) = ((\% \text{ANT})_\sigma \text{ INDEX}) \\ \text{(iii)} \quad (\% \text{ANT NUMB}) \implies (\% \text{ANT NUMB}) = (\uparrow \text{ NUMB}) \\ \text{(iv)} \quad (\% \text{ANT GEND}) \implies (\% \text{ANT GEND}) = (\uparrow \text{ GEND}) \end{array}$$

Line (i) in @ANT-AGR establishes a pathway to an argument and assigns it the local name ANT(ecedent). In line (ii), this argument is declared to be an antecedent of the pronoun by the function \mathcal{R} taken from PCDRT (Haug 2014; Dalrymple et al. 2018).²² Lines (iii)–(iv) indicate that if the antecedent bears number and gender, the pronoun will have the same values of these features. The equations in (iii)–(iv) employ implications as defined in Bresnan et al. (2016: 60–61). If the left-hand side of the implication sign is satisfied, then the right-hand side (the consequent) will hold as a defining equation. Otherwise, the consequent is not treated as a defining equation. Consequently, if the antecedent lacks AF, the pronoun will lack them as well, given that they are not defined anywhere else in the nominal structure.

As a result, a pronoun having an AF-less antecedent will also be AF-less, while a pronoun referring to an AF-bearing antecedent will match its agreement features. The next section demonstrates how this analysis works in practice.

3.3 Analysis at work

Recall the generalisation presented in (5) and repeated below:

- (5) When used anaphorically, *to* in the copular structure with the nominative nominal predicate (TO_D) may have an antecedent of any type, whereas *to* in any other

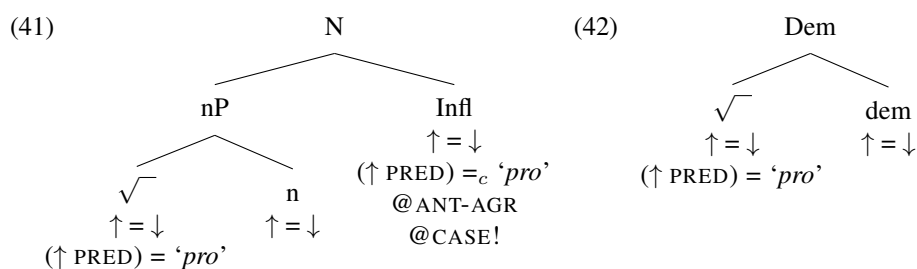
²²Note that (i) and (ii) require the antecedent to be present in the same f-structure as the pronoun. Therefore, @ANT-AGR is unable to capture intersentential anaphora. For the purposes of this paper, I assume that @ANT-AGR operates across f-structures, although formalising such a mechanism poses a significant challenge for LFG, where access to previous f-structures is denied. This issue is briefly discussed in §4.

structure (TO_N) can only have an antecedent lacking number and gender.

Consider examples (1b)–(1c), partly illustrating this phenomenon, which are repeated below as (40a)–(40b).²³

- (40) Kupiłem [nowy komputer]_i.
 ‘I bought a new computer.’
- a. {Miał on_i / #Miało to_i} świetne parametry.
 had.3SG.M PERS.SG.M / had.3SG.N this excellent specifications.ACC
 ‘It had excellent specifications.’
- b. Był {to_i / #on_i} świetny zakup. (TO_D)
 was.3SG.M this / PERS.SG.M excellent purchase.SG.M.NOM
 ‘It was a great purchase.’

The anaphoric item in (40a) is a noun, and hence it is built by rules (34) and (38) presented in §3.2, see (41). The demonstrative identifier in (40b) is formed by rule (33), and its structure is presented in (42).



The vocabulary items that can potentially expone these structures are given in (43) and (44).

- (43) $\langle [\sqrt{\quad}], (\uparrow \text{PRED}) = \text{'pro'} \rangle \xrightarrow{\nu} to$
 $x_i : \uparrow_{\sigma}$
- (44) $\langle [\sqrt{\quad}, \text{n}], (\uparrow \text{PRED}) = \text{'pro'} \rangle \xrightarrow{\nu} on$
 $(\uparrow \text{NUMB})$
 $(\uparrow \text{GEND})$
 $x_i : \uparrow_{\sigma}$

Let us first focus on the nominal version, that is, (41). Consider the @ANT-AGR restriction again:

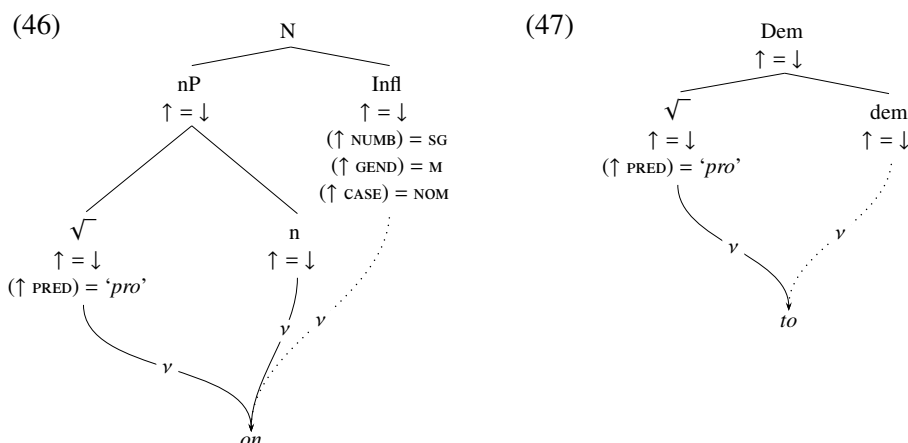
- (45) @ANT-AGR :=
- (i) $(\text{GF}^+ \uparrow) \text{GF} = \% \text{ANT}$
 - (ii) $\mathcal{R}((\uparrow_{\sigma} \text{INDEX})) = ((\% \text{ANT})_{\sigma} \text{INDEX})$
 - (iii) $(\% \text{ANT NUMB}) \implies (\% \text{ANT NUMB}) = (\uparrow \text{NUMB})$
 - (iv) $(\% \text{ANT GEND}) \implies (\% \text{ANT GEND}) = (\uparrow \text{GEND})$

The antecedent (*nowy komputer* ‘new computer.SG.M’) has NUMBER and GENDER, and

²³Example (1a) is omitted here, as it will behave exactly like (1b). Analogously, (2a) will be omitted in the later part of this section.

hence lines (iii) and (iv) force the pronoun to have matching values of these features. This, in turn, makes the VI for *on* available. Because it is better specified than the VI for *to*, the latter will be blocked by **MostInformative_f**, see (40a). In contrast, the demonstrative identifier in (40b) cannot be expounded by *on* – it lacks gender and number, and also does not contain the n node (see (42)); therefore, it is realised by *to*. Note that the structure in (42) is indifferent to the AF of the antecedent.

The structures discussed above together with ν -mappings are presented in (46) and (47). In the case of (46), I present only those parts of the f-description attached to Infl which are relevant for exponence, that is, number and gender (in their unpacked form – originally they are contained in the @ANT-AGR macro).



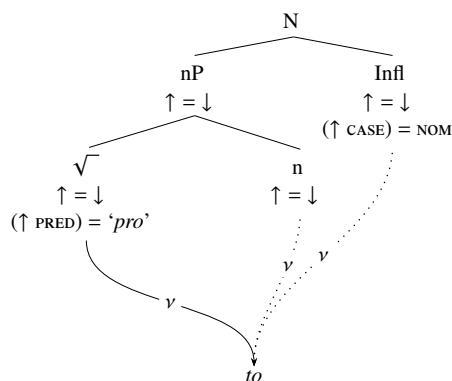
Now, consider examples (2b) and (2c), repeated below as (48a) and (48b).

- (48) [Polska przegrała kolejny mecz]_i.
 ‘Poland lost another match.’
- a. Zaskoczyło mnie {to_i / #ono_i}. (TO_N)
 surprised.3SG.N I.ACC this / PERS.SG.N
 ‘It surprised me.’
- b. Był {to_i / #ono_i} prawdziwy skandal. (TO_D)
 was.3SG.M this / PERS.SG.N real scandal.SG.M.NOM
 ‘It was a real scandal.’

The antecedent lacks AF, and *to* is present in both types of structures: as a noun (TO_N), and as a demonstrative identifier (TO_D). Again, two c-structures are available, (41) and (42). In the nominal structure, @ANT-AGR checks if the antecedent has AF. It does not, so the pronoun cannot have them either: the implications in lines (iii) and (iv) are not satisfied. In consequence, the nominal structure is AF-less, as number and gender are not specified anywhere else.

The VI for the personal pronoun stem (see (44)) requires AF in the structure, so it cannot be used. Instead, *to* is used in (48a), the only anaphoric item able to expone the resulting structure, which is presented in (49). Analogously to (46), I only include the equations which are directly relevant to exponence. The Infl node in (49) does not contain number and gender, since lines (iii) and (iv) in @ANT-AGR have not been satisfied.

(49)



The demonstrative identifier in (48b) is expounded by *to* as usual, see (47).

To summarize, this subsection has demonstrated how the proposed analysis predicts the difference in anaphoric possibilities between the two uses of *to*. The nominal use (TO_N) can only refer to antecedents lacking number and gender, because otherwise it is blocked by personal pronouns. In contrast, the item present in the copular clause with a nominative nominal predicate is a demonstrative identifier. It can never meet the conditions required for personal pronouns to be expounded, and hence is always realised as *to* (TO_D).

4 Open problems and conclusion

Let us begin this summary by discussing the limitations of the proposed analysis and point to potential directions for future research. At least three issues require further investigation to achieve a comprehensive analysis of the Polish pronominal system.

The first issue, as noted in fn. 22, is intersentential anaphora. Polish personal pronouns bear the same *syntactic* features as their antecedents. This property seems to be encoded in their lexical entries (similarly to the practice of encoding the fact that English *he* requires a *male* antecedent, see Heim & Kratzer 1998: 244–245, Dalrymple et al. 2019: 533). This has been implemented here using a set of constraints comprising a template called @ANT-AGR. However, within the current LFG architecture, it seems impossible to enforce this constraint intersententially, across f-structures. Let us only hint at a possible solution here, which is to shift anaphoric agreement to the module explicitly designed to track information throughout the discourse, that is, to the discourse structures from the Discourse Representation Theory (e.g., PCDRT, as already integrated with LFG in Dalrymple et al. 2018, Dalrymple et al. 2019: Ch: 14–15). It would assume that lexical items such as nouns and pronouns introduce predicates to the discourse structure carrying information such as *expressed by a masculine noun*. A formalisation of such an approach and an exploration of its consequences are left for future research.

The other two limitations follow from the empirical focus of this paper, which resulted in leaving other intriguing issues related to the presented data unexplored. The aim of this work was to analyse the distribution of *to* in two types of sentences. Consequently, the syntactic structure of these sentences – particularly the peculiar copular clause with TO_D , where the predicate occurs in the nominative case and triggers agree-

ment – has not been analysed. The numerous issues related to the c- and f-structural representation of such sentences, as well as their semantic composition, undoubtedly require a separate study.

The final issue, noted in fn. 5, is that the intricate system of Polish personal pronouns in non-nominative cases has been neglected. This opens up an interesting direction for exploration, particularly in the context of the morphemic and realisational framework of L_RFG.

Notwithstanding these limitations, this paper offers a novel perspective on an intriguing phenomenon concerning Polish pronouns, aiming to explain the distribution of the word *to* and personal pronouns used anaphorically. A key observation that led to this research was the fact that *to*, when used in typical nominal positions, can only refer to AF-less antecedents (TO_N), whereas *to* in the unique copular structure, containing a nominal predicate in the nominative case, can have an antecedent of any type (TO_D). The closer examination of the data allowed for a principled explanation of this phenomenon: the difference follows from the mechanism of anaphoric agreement, the specification of the relevant vocabulary items, and the rules of exponence postulated in L_RFG.

More precisely, the personal pronoun stem (*on*) can only be expounded in the presence of AF. These features, in turn, can (and must) be present on the pronoun when the antecedent possesses AF. If the antecedent is AF-less, the anaphoric pronoun also is and cannot be realised by *on*. *To* is inserted then, forming the AF-less pronoun TO_N. Being underspecified for grammatical category, it also expounds the demonstrative identifier (TO_D), present in the aforementioned copular structure with the nominative predicate.

References

- Ahn, Dorothy. 2022. Indirectly direct: An account of demonstratives and pointing. *Linguistics and Philosophy* 45(6). 1345–1393. <https://doi.org/10.1007/s10988-022-09350-5>.
- Asudeh, Ash, Bronwyn Bjorkman, Frances Dowle, Neil Myler, Dan Siddiqi & Lisa Sullivan. 2024. Fusional morphology, metasyncretism, and secondary exponence: A morphemic, realizational approach to Latin declension. This volume.
- Asudeh, Ash, Tina Bögel & Dan Siddiqi. 2023. Modelling exponents. In Miriam Butt, Jamie Y. Findlay & Ida Toivonen (eds.), *The Proceedings of the LFG'23 conference*, 23–44. PubliKon. <https://lfg-proceedings.org/lfg/index.php/main/article/view/38>.
- Asudeh, Ash & Daniel Siddiqi. 2023. Morphology in LFG. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*, 855–901. Berlin: Language Science Press. <https://doi.org/10.5281/zenodo.10185976>.
- Bartošová, Jitka. 2017. *Topics in copular clauses*. Ph.D. thesis, McMaster University.
- Belyaev, Oleg. 2024. Sublexical nodes and affix ordering in LrFG. Talk presented at the LFG'24 Conference, Accra, Ghana.

- Bogusławski, Andrzej. 1988. *To jest*. *Rozprawy Komisji Językowej Łódzkiego Towarzystwa Naukowego* XXXIV. 30–35.
- Bogusławski, Andrzej. 2002. *To jest – ‘das ist’*. *Unteilbare Spracheinheiten? Doch! Teil I. Die Welt der Slaven* 47. 131–154.
- Bondaruk, Anna. 2013. *Copular clauses in English and Polish. Structure, derivation and interpretation*. Lublin: Wydawnictwo KUL.
- Bresnan, Joan, Ash Asudeh, Ida Toivonen & Stephen Wechsler. 2016. *Lexical-Functional Syntax* (2nd edn.). Chichester, UK: Wiley-Blackwell. <https://doi.org/10.1002/9781119105664>.
- Browne, Wayles. 1999. *Vsjako Načalo e Trudno: Traps in Lesson I in the South Slavic Languages*. In Mila Dimitrova-Vulchanova & Lars Hellan (eds.), *Topics in South Slavic syntax and semantics*, vol. 172, 193–204. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/cilt.172.09bro>.
- Dalrymple, Mary, Dag Haug & John Lowe. 2018. Integrating LFG’s binding theory with PCDR. *Journal of Language Modelling* 6(1). 87–129. <https://doi.org/10.15398/jlm.v6i1.204>.
- Dalrymple, Mary, John J. Lowe & Louise Mycock. 2019. *The Oxford reference guide to Lexical Functional Grammar*. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198733300.001.0001>.
- Diessel, Holger. 1999. *Demonstratives: Form, function and grammaticalization*. Amsterdam: Benjamins. <https://doi.org/10.1075/tsl.42>.
- den Dikken, Marcel. 2024. Agreeable predicate nominals. In Matteo Greco & Davide Mocci (eds.), *A Cartesian dream: A geometrical account of syntax. In honor of Andrea Moro*, Lingbuzz Press.
- Dziwirek, Katarzyna. 1990. Default agreement in Polish. In Katarzyna Dziwirek, Patrick Farrell & Errapel Mejías-Bikandi (eds.), *Grammatical relations: A cross-theoretical perspective*, Stanford, CA: CSLI Publications.
- Halle, Morris & Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from Building 20*, 111–176. Cambridge: MIT Press.
- Haug, Dag Trygve Truslew. 2014. Partial dynamic semantics for anaphora: Compositionality without syntactic coindexation. *Journal of Semantics* 31(4). 457–511. <https://doi.org/10.1093/jos/fft008>.
- Haugen, Jason D. & Daniel Siddiqi. 2016. Towards a restricted realization theory: Multimorphemic monolistemicity, portmanteaux, and post-linearization spanning. In Daniel Siddiqi & Heidi Harley (eds.), *Morphological metatheory*, 343–386. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.229.12hau>.

- Heim, Irene & Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell.
- Melchin, Paul B., Ash Asudeh & Dan Siddiqi. 2020. Ojibwe agreement in Lexical-Realizational Functional Grammar. In Miriam Butt & Ida Toivonen (eds.), *Proceedings of the LFG'20 conference*, 268–288. Stanford, CA: CSLI Publications. <https://csli-publications.stanford.edu/LFG/LFG-2020/lfg2020-mas.pdf>.
- Moskal, Beata. 2015. Limits on allomorphy: A case study in nominal suppletion. *Linguistic Inquiry* 46(2). 363–376.
- Norris, Mark. 2022. Nominal inflection in Distributed Morphology. Manuscript. <https://lingbuzz.net/lingbuzz/006516>.
- Noyer, Rolf. 1997. *Features, positions and affixes in autonomous morphological structure*. New York: Garland Publishing.
- Patejuk, Agnieszka & Adam Przepiórkowski. 2016. Reducing grammatical functions in LFG. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway King & Stefan Müller (eds.), *Proceedings of the joint 2016 conference on Head-Driven Phrase Structure Grammar and Lexical Functional Grammar*, 541–559. Stanford, CA: CSLI Publications. <https://doi.org/10.21248/hpsg.2016.28>.
- Przepiórkowski, Adam. 1999. *Case assignment and the complement-adjunct dichotomy: A non-configurational constraint-based approach*. Ph.D. thesis, Universität Tübingen.
- Przepiórkowski, Adam, Mirosław Bańko, Rafał L. Górski & Barbara Lewandowska-Tomaszczyk (eds.). 2012. *Narodowy Korpus Języka Polskiego*. Warsaw: Wydawnictwo Naukowe PWN.
- Wiśniewski, Marek. 1987. Formalnogramatyczny opis leksemów TO. 1. Słowo *to* w funkcji rzeczownika lub przymiotnika. *Acta Universitatis Nicolai Copernici. Filologia Polska* XXIX. 27–42.