Journal of Linguistics (2026), 1–36
doi:10.1017/S0022226726101388

Journal of Linguistics
CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Non-canonical agreement in predicational copular clauses: Evidence from Polish

Sebastian Zawada[1] and Adam Przepiórkowski[1,2]

[1]Institute of Computer Science, Polish Academy of Sciences, Poland
[2]Faculty of Philosophy, University of Warsaw, Poland
Corresponding author: Adam Przepiórkowski; Email: przepiorkowski.adam@gmail.com

Received: 13 August 2025 UTC; **Revised:** 10 July 2026 UTC; **Accepted:** 24 July 2026 UTC

Keywords: agreement; copula; demonstrative; predication; Slavic

Abstract

This paper examines several understudied and unusual types of copular constructions in Polish, in which the copula agrees with the predicate rather than the subject. It argues that this seemingly unexpected agreement pattern follows from a well-known principle: the controller of agreement is the highest accessible phrase in the domain of agreement. This conclusion is defended against appearances, which suggest that the controller of agreement is NOT the highest accessible phrase in the constructions at hand. In the process, the paper advances several theoretical claims, including the proposal that the demonstrative found in the subject position in such constructions is structurally reduced and lacks a case projection, that certain non-agreeing leftmost phrases are base-generated above TP, and that there are two flavours of Pred in Polish, including a highly defective Pred which combines with nominal predicates in the nominative instead of the usual instrumental of predication. The paper also reports the discovery of long-distance agreement in Polish, apparently not noted so far in the context of Slavic.

1. Introduction

Non-canonical agreement in copular clauses is a phenomenon in which the linearly right-hand nominal phrase, KP, controls agreement.[1] It is widely discussed in recent syntactic literature, most frequently in the context of predicate inversion in so-called specificational copular clauses (Moro 1997, den Dikken 2006, Hartmann & Heycock 2023, among many others), as in the Italian example (1), where the left-hand KP is often treated as the underlying predicate. Otherwise, research focuses on so-called "assumed identity" sentences, which in various languages display person-sensitivity effects (Béjar & Kahnemuyipour 2017, Keine et al. 2019,

[1] We take maximal nominal projections to be K(ase)Ps, but almost all empirical discussion in Sections 2–4 (with the exception of some theoretical bits of Section 3.1) is agnostic about the exact structure and category of nominal projections.

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

Coon & Keine 2021, Bhatia & Bhatt 2023), as in (2) from Eastern Armenian, in which the right-hand KP controls agreement because it has higher person than the left-hand KP.

(1) La causa della rivolta {furono / *fu} le foto del muro.
the cause of.the riot were / was the pictures of.the wall
'The cause of the riot was the pictures of the wall.' (Italian, Moro 1997: 28)

(2) Shadi-n yes {et / *er}.
Shadi I was.1SG / was.3SG
'Shadi was me.'
(Eastern Armenian, Béjar & Kahnemuyipour 2017: 483, glosses modified)

In this paper, we contribute to this extensive body of research on copular clauses by introducing and analysing a fundamentally distinct phenomenon: copular constructions in which the right-hand KP controls agreement despite functioning as a standard $\langle e,t\rangle$-type predicate. The constructions under investigation are exemplified in (3)–(5) from Polish.[2]

(3) Z Zosi był dobry człowiek.
of Zosia.SG.F.GEN was.**SG.M** good.SG.M.NOM person.**SG.M**.NOM
'Zosia (lit. of Zosia) was a good person.' (Hentschel 2001: 166)

(4) To był prawdziwy skandal.
TO was.**SG.M** real.SG.M.NOM scandal.**SG.M**.NOM
'It was a real scandal.' (NCP)

(5) Leki to była ostateczność.
medicines.PL.M.NOM TO was.**SG.F** last.resort.**SG.F**.NOM
'Medicines were a last resort.' (CCP)

The structure presented in (3) is an extremely understudied type of copular clause containing a prepositional subject of predication and a nominative predicate. The structure in (4) contains a demonstrative word, *to*, serving as the subject of predication and, again, a nominative predicate. The third structure, shown in (5), has an additional nominal phrase to the left of *to*.[3] Although this last structure has been the topic of numerous analyses within generative linguistics (see, e.g., Rutkowski 2006, Citko 2008, Bondaruk 2013), agreement

[2] All examples come from Polish unless indicated otherwise. The data are drawn from the National Corpus of Polish (NCP, Przepiórkowski et al. 2012), the Corpus of Contemporary Polish (CCP, Kieraś et al. 2025), the Polish Web 2019 Corpus (PW2019) available via Sketch Engine (Kilgarriff et al. 2014) and from the Internet. Morphosyntactic glosses follow the Leipzig Glossing Rules (https://www.eva.mpg.de/lingua/resources/glossing-rules.php), with the additional abbreviation PERS used for personal pronouns. To reduce clutter, only morphosyntactic information pertinent to the phenomenon under investigation is presented. When relevant, matching features of the agreement target and controller are bolded and underlined.

[3] Structures like (4) may also be used for identification (like English *This is John*), and constructions of type (5) may be used to express equative and specificational meaning. Here we focus exclusively on predicational clauses, that is, constructions in which the right-hand NP expresses a property ascribed to the left-hand referential item. See Bondaruk (2013) for a discussion of the structure employed in (5) in light of Higgins's (1973/1979) classification.

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

in such constructions has always been analysed in isolation, as an idiosyncratic phenomenon that stands apart in Polish grammar.[4]

The overarching claim of this paper is that, with respect to agreement, such apparently highly idiosyncratic constructions are all accounted for by a single well-known principle, namely, the one proposed in Bobaljik (2008), with the original formulation given in (6).[5]

(6) The controller of agreement on the finite verbal complex (Infl+V) is the *highest accessible* NP in the *domain* of Infl + V. (Bobaljik 2008: 296, emphasis original)

In our terms, (6) implies that in languages in which only nominative or absolutive KPs are possible agreement controllers, it is the HIGHEST NOMINATIVE/ABSOLUTIVE PHRASE in the domain of T (i.e., within the TP; Bobaljik's Infl + V) that controls agreement, not necessarily the subject. Support for (6) comes from languages such as Icelandic and Hindi/Urdu, which are well-known for the existence of marked – not nominative or absolutive – subjects. When such marked subjects are accompanied by unmarked – i.e., nominative or absolutive – objects, it is the objects that obligatorily control agreement, since they are the highest accessible KPs in the relevant domain.

In contrast to Icelandic and Hindi/Urdu, Bobaljik (2008: 303) classifies Russian – and, one may hypothesise, Slavic languages in general – as having 'rather boring morphological case systems, where m-case [morphological case] tracks GF [grammatical function] fairly neatly', adding that 'the interest comes from languages in which case and GF do not always line up'. This paper shows that evidence for (6) may also be found in such 'boring' languages and, in particular, that – contrary to the common assumption – grammatical function and morphological case do not always line up in the case of Slavic subjects.

Polish constructions exemplified in (3)–(5) provide evidence for Bobaljik's (2008) generalisation only if the actual subjects are inaccessible and the nominative right-hand predicates happen to be the highest accessible KPs. In order to demonstrate this, the paper makes a number of novel empirical observations and defends a number of theoretical claims concerning subjecthood, the internal structure of demonstrative pronouns and the nature of nominal constituents base-generated high in the syntactic tree, above the basic sentential constituent.

In particular, concerning the understudied constructions exemplified in (3), we refute in Section 2 the view expressed in previous literature that the nominative KP predicate is the grammatical subject; such right-hand predicates control agreement not because they are subjects, but because the higher PP, expressing the subject of predication, is inaccessible, as it is NOT A NOMINATIVE KP. Then, in Section 3, we turn to the analysis of constructions exemplified in (4). We present evidence that the word *to* in (4) is a minimally specified demonstrative, consisting of a referential feature (the root of the widely assumed feature geometry of Harley & Ritter 2002) combined with a nominaliser, but not projected any further. Hence, in contrast to the word *to* occurring in typical nominal positions, *to* in (4) is not a complete nominal phrase, that is, it is not embedded in a case projection; in effect, it is NOT A NOMINATIVE KP. This analysis makes it possible to explain intriguing distributional

[4] A notable exception is Hentschel (2001), a non-generative analysis of all three structures. Moreover, Rutkowski (2006), concentrating on constructions such as (5), mentions structures such as (4) as their 'cognitive source'.

[5] Although such an agreement mechanism is probably best known from Bobaljik's (2008) paper, it had been assumed in the literature before. For instance, Bresnan (1994: 118) argues that in English and Chichewa, 'the verb agrees with its most prominent argument (according to the functional hierarchy of SUBJ > OBJ > ...) that lacks indirect case'.

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

differences between the two kinds of *to*. Finally, in Section 4, we show that the left-hand nominative KP in (5) is OUTSIDE THE AGREEMENT DOMAIN, as it is base-generated above the TP. Although this claim has appeared in the prior literature, we provide novel evidence and refute purported counterevidence. This way, the nominative predicates in (3)–(5) remain the only accessible phrases in the agreement domain, so they control agreement even though they are not grammatical subjects.

With the evidence for (6) established, we turn in Section 5 to the formal analysis of constructions exemplified in (3)–(5). First, in Section 5.1, we outline the structural underpinnings of Polish copular clauses and address the important question of why predicates appear in the nominative case in (3)–(5). This fact requires an account, since – in canonical copular clauses – nominal predicates uniformly occur in the instrumental in Polish, as illustrated in (7).

(7) Leki były {ostatecznością / *ostateczność}.
medicines.PL.M.NOM were.PL last.resort.INS / last.resort.NOM
'Medicines were a last resort.'

Building on Citko (2008) and Bondaruk (2013) (a.o.), we postulate two Pred(ication) heads; the default Pred_{INS} assigns the instrumental case to the predicate, while the defective $\text{Pred}_{\varnothing}$ does not assign any case. We show that this proposal accounts for the distributional patterns found.

Then, in Section 5.2, we outline the proposed formalisation of agreement, couched in the interaction/satisfaction model of Agree (Deal 2024, Clem & Deal 2026), and show that it naturally implements Bobaljik's generalisation in (6).

The two ensuing subsections present novel phenomena which strikingly support the proposed analysis. In Section 5.3 we provide additional evidence for the claim that the left-hand KP in constructions exemplified in (5) is outside the agreement domain. Perhaps more interestingly, in Section 5.4 we investigate how the relevant structures interact with raising and control verbs. Embedding sentences such as (3)–(5) under raising verbs results in long-distance agreement (to our knowledge, never discussed in the context of Slavic languages before), whereas embedding them under control verbs results in ungrammaticality. This contrasts with canonical copular clauses, such as (7), which can freely be embedded under both raising and control verbs. We show that this pattern is naturally accounted for by the proposed analysis.

While Section 6 concludes the main part of the paper, Appendix A discusses a potential confound, namely, the role of grammatical person in copular agreement.

2. Sentences with prepositional subjects of predication

Consider examples (8)–(9). The preposition *z* 'of' assigns genitive to its dependent. The predicate, which is obligatorily a nominative KP, expresses a property ascribed to the referent of the genitive KP. The verbal copula must agree with the nominative predicate. The PP is often preverbal, as in (8), but in (9) it remains in the postverbal position, which is a common property of pronominal projections in Polish.

(8) … [z Pauliny]$_{\text{PP}}$ był [niezły gagatek]$_{\text{KP}}$.
of Paulina.SG.F.GEN was.**SG.M** not.bad rascal.**SG.M**.NOM
'Paulina (lit. of Paulina) was quite a rascal.' (PW2019)

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

(9) Kiedyś była [z nas]$_{PP}$ [niezła paczka]$_{KP}$.
once was.**3SG.F** of we.PL.GEN not.bad package.**SG.F**.NOM
'We (lit. of us) used to be a nice bunch.' (CCP)

The fact that the nominative predicate in such structures is the highest accessible phrase is self-evident: neither the prepositional phrase nor the genitive KP within it is a nominative KP. Hence, Bobaljik's generalisation in (6) immediately predicts the agreement facts in (8)–(9). However, in order to distinguish between (6) and the usual generalisation that, in Polish, only subjects control verbal agreement, we need to argue that such predicates are not subjects, contrary to the view expressed in Hentschel (2001) and in apparently the only generative work discussing such constructions so far, Bondaruk (2014).

In traditional Polish linguistics, case and agreement are commonly taken to be THE indicators of subjecthood: nominative constituents agreeing with verbs are subjects by definition. In generative linguistics, control and binding are taken to be cross-linguistically more reliable testing grounds for subjecthood, but they do not give clear results in the case of the construction at hand. As discussed below, both control and binding facts are not fully conclusive. Nevertheless, we take it as the default position that, if there is a subject in such constructions, it is the PP or an empty pronoun coreferential with the genitive KP within it, and not the nominative predicate.[6]

One suggestive argument is that, cross-linguistically, subjects typically denote entities (of type e) or quantifiers (of type $\langle et,t\rangle$), but not predicates (of type $\langle e,t\rangle$). More importantly, in the information-structurally neutral context, the prosodically neutral word order is as in (8)–(9), with the PP adjacent to the verb and the predicate linearly to the right, adhering to the default SVO order in Polish. In contrast, if nominative predicates were subjects, they would reside in Spec,TP, and the default word order would be KP_{NOM} Copula PP, as in (10).

(10) Pechowiec jest z niego.
unlucky.person.SG.M.NOM is of he.SG.M.GEN
'He (lit. of him) is an unlucky person.' (Bondaruk 2014: 44)

In fact, Bondaruk (2014) assumes that (10) illustrates the 'canonical' structure of copular constructions with PP subjects of predication and that left-hand PPs in sentences such as (3), (8) and (9) are topics moved to the CP zone.

However, it is precisely (10) that is marked with regard to information structure: it is most natural in the context of contrastive topicalisation of the nominative predicate, as in (11a), or predicate fronting expressing surprise, as in (11b).

(11) a. Pechowiec jest z NIEGO, nie z NIEJ.
unlucky.person.NOM is of him.GEN not of her.GEN
'HE is an unlucky person, not HER.'
b. Powiem ci, że PECHOWIEC jest z niego!
tell.1SG you.DAT that unlucky.person.NOM is of him.GEN
'Let me tell you, unlucky is he!'

[6] Since we cannot prove that the PP is the actual grammatical subject, rather than being only coindexed with an empty caseless subject, we refer to this PP as the 'subject of predication'.

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

This suggests that the word order witnessed in (10)–(11) is the result of KP fronting, and the neutral word order of the constructions at hand is PP Copula KP_{NOM} (or Copula PP KP_{NOM} when the PP contains a pronominal KP), as in (8)–(9), which is not immediately compatible with the claim that KP_{NOM} is the subject.

Given these considerations, the need to demonstrate that such nominative KP predicates ARE subjects after all is on the shoulders of advocates of this view. Hentschel (2001) indeed argues for the subjecthood of the nominative KP predicate and against the subjecthood of the PP. Assuming the well-known generalisation that only subjects bind anaphors in Polish, he presents the following Polish examples, with judgements as indicated (Hentschel 2001: 167)[7]:

(12) a. U siebie$_i$ w domu jest on$_i$ nieznośny tyran.
at self in home is.3SG he.NOM unbearable tyrant.NOM
'At home, he is an unbearable tyrant.'
b. ??U siebie$_i$ w domu jest z niego$_i$ nieznośny tyran.
at self in home is.3SG of him.GEN unbearable tyrant.NOM

(13) a. W swoim$_i$ zakładzie jest on$_i$ nieznośnym pedantem.
in self's workshop is he.NOM unbearable pedant.INS
'At his workshop, he is an unbearable pedant.'
b. ??W swoim$_i$ zakładzie jest z niego$_i$ nieznośny pedant.
in self's workshop is of he.GEN unbearable pedant.NOM

Examples (a) are sentences with nominative subjects. (12a) is an infrequent structure (discussed in Section 5.1) with a nominative predicate, i.e., it involves two nominative KPs altogether, and (13a) is a typical Polish copular clause with an instrumental predicate. Examples (b) present analogous structures with PP subjects and nominative predicates. According to Hentschel (2001), the subject-oriented reflexive pronouns *siebie* and *swoim* cannot be bound by the genitive KP in (12b) and (13b) – or, we may add, by the PP (on the assumption that it shares ϕ features with the embedded KP; cf. Řezáč 2008) – although these anaphors can be bound by typical nominative subjects, as shown in (12a) and (13a). Hence, the genitive KP (or the PP) is not a subject, according to this argument.

While we agree that there are acceptability contrasts of the kind illustrated above, they are not related to binding and, hence, they do not say anything about subjecthood. For example, (14a) and (14b) – minimally modified versions of (12b) and (13b) – are similarly degraded, even though they do not involve any binding violations. It seems that the predicates in the structures analysed must express generic, perhaps individual-level properties, with which the locative phrases are not readily compatible (cf. ??*John was intelligent in his room*); hence the drop in acceptability.

(14) a. ??U Marysi$_j$ w domu jest z niego$_i$ nieznośny tyran.
at Mary in home is of he.GEN unbearable tyrant.NOM
b. ??W naszym$_{j+k}$ zakładzie jest z niego$_i$ nieznośny pedant.
in our workshop is of he.GEN unbearable pedant.NOM

[7] Although we follow here the reasoning of Hentschel (2001), in Section 4.2 we introduce important qualifications concerning the use of reflexive binding as a test for syntactic position.

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

Therefore, binding does not seem to provide an argument *against* the subjecthood of PPs. On the contrary, given that acceptability levels of (12b), (13b) and (14a)–(14b) are comparable, i.e., given that binding in (12b)–(13b) does not seem to decrease the acceptability of (14a)–(14b) any further, binding can be argued to provide an argument FOR the subjecthood of the PPs.[8]

Also the complementary part of Hentschel's argument – intended to show that the nominative predicate can bind a reflexive pronoun, and hence is the grammatical subject – is invalid. Consider (15) from Hentschel (2001: 168)[9]:

(15) Jak siadł na konia, zdawało mu się, że z niego prawdziwy
when sit on horse seemed he.DAT REFL COMP of he.GEN real
Piłsudski$_i$ na swojej$_i$ Kasztance.
Piłsudski.SG.M.NOM on self's Kasztanka
'When he got on his horse, he seemed like a real Piłsudski on his Kasztanka.'[10]

Hentschel argues that the fact that the nominative predicate (*prawdziwy Piłsudski* 'real Piłsudski') can bind the reflexive supports its status as the grammatical subject. This, however, is not the case. Consider the attested example (16), where the reflexive pronoun is bound by *Piłsudskiego* 'Piłsudski.ACC', even though the whole KP is in the object position of the verb *widział* 'saw'.[11] Despite the fact that objects cannot control reflexives, KP-internal reflexives can be bound by the head of the nominal phrase, regardless of the syntactic function of the whole KP.

(16) … widział [Piłsudskiego$_i$ [na swojej$_i$ Kasztance]$_{PP}$]$_{KP}$…
saw.3SG.M Piłsudski.ACC on self's.LOC Kasztanka.LOC
'He saw Piłsudski$_i$ on his$_i$ Kasztanka.' (Google)[12]

In short, assuming that it is the subject status of the right-hand KP in (15) that makes binding possible is incorrect.

The other well-known test for subjecthood, control, is inapplicable here. Nominative predicates occurring in such constructions cannot be controlled as the notion of control makes sense only for referential phrases. The opposite configuration – an attempt to create a control structure with a PP (or, more precisely, PRO replacing it) as the controllee – appears to yield a negative result:

(17) *Chcieli być [PRO niezła paczka].
wanted.3PL be.INF not.bad package.SG.F.NOM
Intended: 'They wanted to be a nice bunch.'

[8] However, formal acceptability judgement experiments are needed to confirm the validity of this argument.

[9] Hentschel provides two examples. The second one is omitted here as it is fully analogous. Note that the copula is omitted in this sentence, which is allowed in certain types of copular clauses in the present tense.

[10] Józef Piłsudski (1867–1935) was a prominent Polish statesman and *Kasztanka* is the name of his mare.

[11] Out of context, (16) has two more readings: the reflexive can be part of the KP and bound by the *pro*-dropped subject (X saw Piłsudski when Piłsudski was sitting on X's horse named Kasztanka) or the PP containing the reflexive can be an adjunct at the verbal spine, with the reflexive bound by the *pro*-dropped subject (X was sitting on X's horse named Kasztanka when X saw Piłsudski). However, the context clearly indicates that the reflexive refers to Piłsudski.

[12] http://www.infosokolka.pl/artykul,obchody-swieta-wojska-polskiego.html; accessed on July 9, 2026.

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

However, as discussed in Section 5.1, (17) is ungrammatical for independent reasons, so it does not provide an argument against the subjecthood of the PP.

Bondaruk (2014) also argues that the nominative predicate is the grammatical subject, adducing two arguments. The first is this: 'The fact that the copula always assumes the ϕ features of the nominative predicate clearly indicates that it is the nominative predicate that serves as a subject in sentences of this type' (Bondaruk 2014: 38). This, however, assumes the traditional position that agreement implies subjecthood, rejected in Bobaljik (2008), in the current paper and in much generative literature.

The second piece of evidence proposed in Bondaruk (2014) concerns the control of the phrase *po pijanemu* 'while drunk', which is assumed to be subject-oriented (Dziwirek 1994). Consider (18). This sentence has only one interpretation: the one who was drunk is the referent of the *pro*-dropped subject, not the referent of the feminine pronoun embedded in the PP *z niej* 'of her'. This is meant to demonstrate that *po pijanemu* can only be controlled by subjects.

(18) pro_i Żartował z niej$_j$ po $\text{PRO}_{i/*j}$ pijanemu.
joked.3SG.M of her.GEN while drunk
'He made jokes of her while he was drunk.' (Bondaruk 2014: 41)

According to Bondaruk, when this phrase is within a copular clause with a PP, as in (19), the PRO of *po pijanemu* must be coindexed with the nominative predicate.

(19) Z niego jest niezły zbój$_i$ po PRO_i pijanemu.
of he.GEN is.3SG not.bad thug.SG.M.NOM while drunk
'He is quite a thug while he is drunk.' (Bondaruk 2014: 41; coindexation original)

We find this argument dubious on semantic grounds: only entities may be drunk (as in (18)), not predicates (as claimed about (19)). That is, it must be the referent of the *e*-type (pronoun within the) PP that is interpreted as drunk in (19), rather than the predicate expressed by the $\langle e, t\rangle$-type nominative KP, as the coindexation of *po pijanemu*'s PRO with the predicate would imply.

Moreover, *po pijanemu* is not a reliable test of subjecthood in any case, as the following examples demonstrate:

(20) Wszystko zdarzyło się po pijanemu…
everything happened REFL while drunk
'Everything happened while one was drunk.' (Google)[13]

(21) Mniej szczęścia miał rowerzysta potrącony po pijanemu.
less luck had.3SG.M cyclist.SG.M.NOM hit.PTCP while drunk
'Less fortunate was a cyclist hit by a drunk person.' (Google)[14]

In (20), the subject is *wszystko* 'everything', and the controller of *po pijanemu* is unexpressed. In (21), the context clearly indicates that it was the agent of hitting who was drunk, not the hit cyclist, although it is the cyclist that is expressed by the subject. Moreover, given that *potrącony* 'hit' is a passive form, this agent is at best expressed in syntax as a demoted

[13] https://www.specprawnik.pl/zapytaj-prawnika/problem/art-224-kk/; accessed on July 9, 2026.

[14] https://grodzisk.naszemiasto.pl/nowy-tomysl-kolizja-na-kolejowej-potracenie-pod-boruja-fot/ar/c16-3441809; accessed on July 9, 2026.

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

implicit argument (Collins 2024), if at all, not as a subject. Hence, *po pijanemu* seems to be pragmatically controlled by a semantically prominent agent, not necessarily syntactically projected at all.

In partial summary, none of the arguments in Hentschel (2001) and Bondaruk (2014) for the subjecthood of nominative predicates goes through. Given the suggestive arguments from semantic types and especially word order mentioned at the beginning of this section, as well as binding facts, such nominative predicates should not be considered subjects until convincingly proven otherwise. In the construction at hand, whether the PP is a subject or not, agreement is controlled not by the subject, but by the highest accessible phrase.

Given that this may be the first time Slavic copular clauses with PP subjects have been brought to broader attention, let us also present examples of the relevant structure from Russian, likely overlooked in the generative literature despite extensive research on Russian copular clauses. A key difference between Polish and Russian is that, in Russian, the predicate must be a discontinuous KP consisting of an adjective and a noun. The noun is topicalised and placed in the initial position, while the adjective remains sentence-final. This seems to be the only word order acceptable in Russian, as non-topicalised versions, like those in Polish, are not attested.[15]

(22) Rabotjaga iz menja byl ploxoj…
worker.SG.M.NOM of I.GEN was.SG.M bad.SG.M.NOM
'I wasn't a very good worker. (lit. Worker of me was bad.)' (Russian, RNC)

(23) … no kommerčeskij direktor iz nee byl
but commercial director.**SG.M**.NOM of she.SG.F.GEN was.**SG.M**
neploxoj.
not.bad.SG.M.NOM
'But she was a pretty good commercial director.' (Russian, RNC)

A more detailed analysis of such copular constructions in Russian is left for future research.

3. The word *to*

Let us now turn to the construction exemplified in (4), repeated below as (24).[16] Such a sentence can only be used anaphorically, when – in the case of (24) – *to* must refer to an abstract object, most commonly introduced by a clause.[17]

(24) To {był / *było} prawdziwy skandal. (to_n)
TO was.**SG.M** / was.SG.N real.SG.M.NOM scandal.**SG.M**.NOM
'This was a real scandal.' (NCP)

[15] Examples (22)–(23) are drawn from the Russian National Corpus (RNC, Savchuk et al. 2024).
[16] This section builds on Zawada (2024).
[17] See (i) for an attested example with a context.

(i) Grzegorz Markowski miał 16 lat, kiedy został ojcem. To był prawdziwy skandal!
'Grzegorz Markowski was 16 when he became a father. It was a real scandal!' (Google)

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

In the case of copular constructions discussed in the previous section, it is clear that the PP subject of predication is not an ACCESSIBLE PHRASE in the sense of Bobaljik's (2008) generalisation, as it is not a nominative KP. However, the preverbal word *to* in (24) looks suspiciously like the nominative form of the demonstrative pronoun *to*, so it – rather than the postverbal predicate – should count as the HIGHEST ACCESSIBLE PHRASE and, hence, act as agreement controller.

In this section, we show that the word *to* used in this construction is not the cased pronoun *to*, i.e., that it is not a KP, but rather a maximally reduced demonstrative item consisting solely of the base head responsible for its referential property, just nominalised. As it is not a nominative KP, it is not accessible for agreement purposes, and – hence – agreement is controlled by the nominative predicate. To motivate these claims, we analyse the *to* in (24) in juxtaposition with its full KP counterpart.

3.1. Distributional differences

We show below that, unlike (24), sentences such as (25) and (26) contain the case-bearing *to*. For convenience, we refer to the structurally reduced variant in (24) – analysed below as an *n*P – as to_n and to the KP variant in (25)–(26) as to_{K}.

(25) To {było / *był} prawdziwym skandalem. (to_{K})
this.NOM was.SG.N / was.SG.M real.SG.M.INS scandal.SG.M.INS
'This was a real scandal.'

(26) To mnie zaskoczyło. (to_{K})
this.NOM I.ACC surprised.SG.N
'This surprised me.'

At first glance, (24) and (25) differ only in the case of the predicate (NOM vs. INS) and in agreement: the right-hand KP obligatorily controls agreement in (24), whereas the copula in (25) takes the third-person singular neuter form, also observed in sentences with *to* being the subject of lexical verbs (see (26)). However, there are several important differences between *to* in (24) on the one hand, and *to* in (25) and (26) on the other.

First, only to_{K} is distributionally equivalent to KPs headed by lexical nouns. That is, one can replace *to* with such a KP in (25) and (26), but – regardless of agreement – not in (24), as shown in (27). (To save space, we do not show examples of *to* occurring with lexical verbs, as it behaves identically to *to* with instrumental predicates.)

(27) a. *Wczorajsze wydarzenie {było / był} prawdziwy (to_n)
yesterday's event.SG.N.NOM was.SG.N / was.SG.M real.SG.M.NOM
skandal.
scandal.SG.M.NOM

b. Wczorajsze wydarzenie było prawdziwym skandalem. (to_{K})
yesterday's event.SG.N.NOM was.SG.N real.SG.M.INS scandal.SG.M.INS
Intended in both: 'Yesterday's event was a real scandal.'

Second, in contrast to to_{K}, to_n cannot be modified by adjectives, see (28).

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

(28) [Polska przegrała kolejny mecz]$_i$. 'Poland lost another match.'
  a. (*Samo) to$_i$ był skandal. (to_n)
     alone TO was.3SG.M scandal.SG.M.NOM
  b. (Samo) to$_i$ było skandalem. (to_K)
     alone this was.3SG.N scandal.SG.M.INS
     'This alone was a scandal.'

Third, only to_K can act as a nominaliser of a complementiser phrase (CP), see (29).

(29) a. *To, co powiedział, był skandal. (to_n)
     TO what said.3SG.M was.SG.M scandal.SG.M.NOM
  b. To, co powiedział, było skandalem. (to_K)
     this.NOM what said.3SG.M was.SG.N scandal.SG.M.INS
     'What he said was a scandal.'

Finally, when to_K is used anaphorically, its antecedent must be a ϕ-less expression, for instance, a clause. In contrast, to_n can refer to any antecedent, including a lexical KP denoting a human; see (30).

(30) Lubiłem Janka$_i$. 'I liked Janek.'
  a. To$_i$ był dobry chłopak. (to_n)
     TO was.SG.M good.SG.M.NOM boy.SG.M.NOM
     'He (lit. This) was a good boy.' (NCP)
  b. #To$_i$ było dobrym chłopakiem. (to_K)
     this.NOM was.SG.N good.SG.M.INS boy.SG.M.INS

The differences are summarised in Table 1.

In brief, to_K behaves like a full KP: it is distributionally equivalent to lexical KPs, it can be modified by adjectives and it can nominalise a CP – similarly to certain nouns taking CP dependents (e.g., *fakt* 'fact', *myśl* 'thought'). Its distributional equivalence to a KP also explains its behaviour with respect to what antecedents it can refer to. Namely, to_K competes with other pronominals having KP distribution, that is, personal pronouns. However, in contrast to to_K, such pronouns have number and gender features (for instance, *on* 'PERS.SG.M', *ono* 'PERS.SG.N'), and therefore outcompete to_K whenever the antecedent bears the corresponding number and gender features. Consequently, to_K is an elsewhere option, occurring only when personal pronouns cannot be used, i.e., when the antecedent is ϕ-less – for

Table 1. Differences between to_n and to_K.

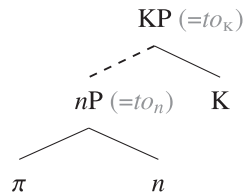
Property	to_n (*to* + NOM predicate)	to_K (e.g., *to* + INS predicate)	example
Distribution of KP	–	+	(27)
Modifiability by adjectives	–	+	(28)
Can nominalise a CP	–	+	(29)
Anaphoric possibilities	any antecedent	ϕ-less antecedents	(30)

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

instance, when it is expressed by a CP or an infinitival TP. The other *to*, to_n, is special: it is a strictly referential item, but it does not behave like a KP – it cannot be replaced by KPs, it cannot be modified by adjectives, nor can it nominalise a CP. As it is not distributionally equivalent to KPs (such as personal pronouns), it does not compete with them, so it can refer to any antecedent.

Given that to_n and to_K share the same morphological form and the same primarily demonstrative function, they should not be analysed as unrelated homonyms. We propose that they are closely related and share a referential feature, namely, the ROOT of the feature geometry (Harley & Ritter 2002), called here π. This root is exponed in Polish as *t-*, as witnessed in various demonstrative forms, such as the adverb *t-ak* 'in this manner' and the numeral *t-yle* 'this much'. Apart from the π feature, to_n and to_K also share the nominal categoriser, *n*, exponed as *-o*; hence the common form *to*. (We will have more to say on this in the following subsection.) However, distributional differences discussed above indicate that only to_K projects to a full KP. Overall, we postulate the following structures for to_n and to_K, with the dashed line signalling our agnosticism about projections between *n*P and KP:

(31) a.



b. $\pi \Leftrightarrow$ *t-*

c. $n \Leftrightarrow$ *-o*

This analysis, treating to_K as a full KP projection, predicts that to_K inflects for case. This is borne out, as shown in (32). Crucially, the case-inflected *to* matches all the characteristics of to_K summarised in Table 1.

(32) a. Nie chcę t-ego.
NEG want.1SG this-GEN
'I don't want this.'

b. Przyglądam się t-emu.
look.1SG REFL this-DAT
'I'm looking at this.'

To summarise, assuming that to_n is the nominalised version of the referential feature, which can be embedded in a case projection (forming to_K), explains (i) why the two – to_K and to_n – are spelled out as the same form, (ii) why to_K competes with other case-bearing pronominals while to_n does not and (iii) why only to_K behaves syntactically like a complete nominal projection: it inflects, combines with adjectives and may nominalise CPs.

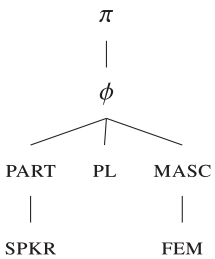
3.2. Feature geometry, π and ϕ

Following Harley & Ritter (2002) and much subsequent work (see, for instance, McGinnis 2005, Béjar & Řezáč 2009), we assume that features relevant for agreement are organised

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

within a hierarchy (a feature geometry). The hierarchy we assume for Polish is presented in (33). Although simplified in some respects that are irrelevant here,[18] it explicitly represents the distinction between the feature π – responsible for referentiality – and the feature ϕ, a strictly grammatical feature present on third-person singular neuter items.

(33)



Distinguishing between $[\pi]$ and $[\pi, \phi]$ makes it possible to capture the difference between ϕ-less referential expressions on the one hand (such as infinitival and clausal phrases in argumental positions, as well as *to* in both its incarnations: to_n and to_K) and third-person singular neuter items on the other (such as the personal pronoun *ono* 'SG.N' and neuter gender nouns such as *okno* 'window.SG.N' or *krzesło* 'chair.SG.N').[19]

There are at least two arguments for the necessity of this distinction in Polish. First, coordination of two singular neuter ($[\pi, \phi]$) items results either in closest conjunct agreement or in plural agreement, as shown in (34). By contrast, coordination of two infinitival phrases or two to_K s in the position of the subject of a lexical verb obligatorily triggers the so-called 'default agreement' (Dziwirek 1990, Preminger 2014), i.e., a 3SG.N form, which signals the lack of an appropriate agreement controller; see (35)–(36). Note that coordinating two anaphoric uses of to_K is difficult, as it is then difficult to determine which *to* refers to which antecedent. However, it is easy to coordinate two deictic uses of to_K (see (36a)) or two CPs nominalised by to_K (see (36b)).[20]

(34) W tym pokoju najładniejsze {było / były} okno i krzesło.
in this room beautiful.SUP was.SG.N / were.PL window.SG.N and chair.SG.N
'The window and the chair were the most beautiful in this room.'

(35) Najwięcej problemów {sprawiło / *sprawiły} mu [skoczyć na bungee]$_{InfP}$ i [wejść na Kilimandżaro]$_{InfP}$.
most troubles.GEN made.3SG.N / made.PL he.DAT jump.INF on bungee and climb.INF on Kilimanjaro
'Bungee jumping (lit. to bungee-jump) and climbing (lit. to climb) Kilimanjaro posed the greatest challenges for him.'

[18] For reasons of space, we do not delve into the complex issue of Polish masculine gender, which is divided into three different subgenders in the Polish structuralist tradition (Mańczak 1956, Saloni & Świdziński 1998).

[19] Note that in nominative singular contexts, singular neuter words (e.g., *krzesł-o* 'chair.SG.N' and the pronoun *on-o* 'SG.N') typically get the suffix *-o*, the same as in *to*. This is expected if the $[\phi]$ feature alone has zero exponence. If so, such singular neuter items only expone the root and – via the same rule that expones *-o* in *to* – the nominaliser *n*; see (31c). Polish singular neuter forms can also have the suffix *-e* in the nominative case, as in (38)–(39), which is mainly driven by the morphophonology of the stem (Orzechowska 1998).

[20] InfP in (35) stands for "infinitival TP".

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

(36) a. {Zaskoczyło / *Zaskoczyły} mnie to$_{\rightarrow}$ i to$_{\rightarrow}$.
surprised.3SG.N / surprised.3PL I.ACC this and this
'This and this surprised me.' ('→' = pointing gesture)

b. Skandalem {było / *były} [to, że Polska przegrała], i
scandal.SG.M.INS was.SG.N / were.PL this COMP Poland lost and
[to, że kibice się pobili].
this COMP fans REFL brawled
'[That Poland lost]$_{\text{nominalised}}$ and [that fans got into a brawl]$_{\text{nominalised}}$ was a scandal.'

This syntactic difference between truly neuter nouns in (34) and ϕ-less constituents in (35)–(36) is accounted for by assuming that infinitival phrases and the word *to* (whether *to*$_n$ or *to*$_{\text{K}}$) are *not* singular neuter, that is, they do not have even the minimal set of ϕ features (the feature [ϕ] itself), in contrast to true singular neuter (pro)nouns. Given this lack of ϕ features on the conjuncts, the usual feature resolution responsible for the plural in (34) cannot take place, so not only particular conjuncts but the whole coordinations in (35)–(36) remain ϕ-less.

The second phenomenon pointing to the same difference – between genuine third-person neuter ([π, ϕ]) and ϕ-less ([π]) items – is based on possible antecedents of pronouns. Polish personal pronouns used anaphorically must agree with their antecedents: they presuppose that their antecedents bear corresponding ϕ features. In this context, note that the singular neuter pronoun *ono* cannot refer to an instance of *to*$_{\text{K}}$, as shown in (37).

(37) A: To$_i$ mnie zaskoczyło.
this I.ACC surprised.3SG.N
'It surprised me.'

B: Mnie też {to$_i$ / #ono$_i$} zaskoczyło.
I.ACC also this / PERS.SG.N surprised.3SG.N
'It surprised me as well.'

The personal pronoun *ono* 'PERS.SG.N' expects its antecedent to have the corresponding features (crucially, [ϕ]), in the same way the pronoun *on* 'PERS.SG.M' requires its antecedent to have the feature [MASC]. The word *to*$_{\text{K}}$ in A's utterance lacks ϕ, and thus *ono* cannot take it as an antecedent. The elsewhere form – *to*$_{\text{K}}$ itself – is used in B's utterance.

This behaviour of infinitival TPs and *to*$_{\text{K}}$ may be contrasted with that of Polish gerund forms, which – just like infinitival TPs – refer to event types, but may be shown to have [π, ϕ]. First, they can trigger plural agreement when coordinated; compare (38) below with (36) above. Second, they may be referred to with the pronoun *ono* 'SG.N'; compare (39) with (37).

(38) Nagradzanie i karanie są ściśle powiązane z
rewarding.SG.N.NOM and punishing.SG.N.NOM are.PL strictly linked with
procesem motywowania pracowników.
process motivating employees
'Rewarding and punishing are closely linked to the process of motivating employees.' (NCP)

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

(39) Jak rozumieć uzdrawianie$_i$ wewnętrznych zranień? Na czym ono$_i$
how understand.INF healing.SG.N inner wounds on what SG.N
polega?
relies
'How do we understand the healing of inner wounds? What does it involve?' (NCP)

Another difference between ϕ-bearing items and *to* is that *to* does not easily bind reflexive pronouns:

(40) a. Ono jest swoim najlepszym obrońcą.
SG.N is REFL.INS best.INS defender.INS
'It is its own best defender.'
(when said about a child, *dziecko*.SG.N in Polish)
b. ??To jest swój najlepszy obrońca.
TO is REFL.NOM best.NOM defender.NOM

While (40a) is fully acceptable, (40b) is considerably worse. This pattern is expected if the reflexive requires a minimal set of ϕ features on its binder: *ono* in (40a) bears the feature [ϕ], while *to* in (40b) lacks it.[21]

Let us finish this section with some speculations on the ϕ-lessness of to_n and to_{K}. The existence of such ϕ-less pronouns should not be particularly surprising. As argued above, personal (*on*-based) pronouns presuppose a specific set of ϕ features on their antecedents – minimally the feature [ϕ] (the singular neuter pronoun *ono*). However, ϕ-less CPs and infinitival TPs also introduce (abstract) objects into the discourse. Accordingly, the existence of ϕ-less pronouns is required to refer to such entities.

This is in line both with Déchaine & Wiltschko (2002), where pronouns are argued to come in different sizes, and with Collins (2024), where implicit arguments (e.g., agents in short passives) are argued to be realised by three kinds of caseless pronouns in English – pro_{gen} (generic), pro_{def} (definite) and pro_{un} (existential) – the last of them also lacking ϕ features. On this approach, the existence of featurally deficient pronouns 'is an immediate consequence of a UG that allows for pronouns defined by various sets of features' (Collins 2024: 58). In fact, as shown in the table in (41), to_{K} fills a gap in the taxonomy of pronouns implicit in Collins (2024).

(41)

	+case	−case
+ϕ	personal pronouns, PRO in PRO-drop languages	pro_{gen} and pro_{def} in English
−ϕ	Polish to_{K}	English pro_{un}, Polish to_n

In summary, there are two variants of the Polish pronoun *to*: a maximal but ϕ-less nominal projection (KP) and a reduced structure (*n*P). We consider this result to be in the spirit of Déchaine & Wiltschko (2002), which proposes that 'it is necessary to recognize (at least) three pronoun types: pro-DP, pro-ϕP, and pro-NP' (p. 409). But while to_{K} might be an

[21] Note that (40b) is the structure with to_n, which can refer to animate objects. Constructing an example with to_{K} and a reflexive is difficult, since to_{K} can only refer to abstract objects, which do not allow for a context with a reflexive.

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

instance of Déchaine & Wiltschko's (2002) pro-DP (given that for them maximal nominal projections are DPs rather than KPs), to_n instantiates a new overt pronoun type, featurally similar to Collins's (2024) phonetically empty pro_{un}.

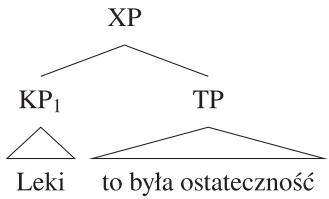
In any case, regardless of its appropriate place in the taxonomy and typology of pronouns, to_n is not a nominative KP, so it is not an accessible agreement controller, which explains why agreement is controlled by the nominative KP predicate lower in the tree.

4. Status of KP_1

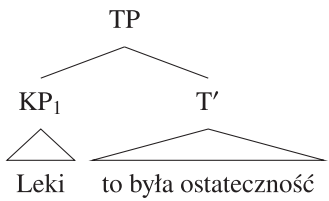
As mentioned in the introduction, the position of the left-hand KP (henceforth, KP_1) in the construction exemplified in (5), repeated below as (42), is a problematic issue. Some analyses, such as Hentschel (2001) and Rutkowski (2006), assume that KP_1 is a left-dislocated topic, so (42) has essentially the same structure as the construction discussed in the previous section, but with an extra constituent in the peripheral position. We schematically present it in (43), but remain agnostic about the category of the relevant projection (hence, the generic XP). Other analyses – including Citko (2008) and Bondaruk (2013) – treat KP_1 as the structural subject, Spec,TP. For both authors, *to* is a special item, distinct from the demonstrative assumed in the previous section: Citko (2008) treats it as a T head, and Bondaruk (2013) as an overt realisation of Pred, a head mediating predication in the small clause. A simplified representation is given in (44).

(42) Leki to była ostateczność.
medicines.PL.M.NOM TO was.**SG.F** last-resort.**SG.F**.NOM
'Medicines were a last resort.' (CCP)

(43) Left dislocation analysis



(44) Subject analysis



https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

In this section, we show that the left dislocation analysis is essentially correct. This, in particular, means that *to* occupies the Spec,TP position, but we additionally assume that it is obligatorily coreferential with KP_1.[22]

In Section 4.1, we present a strong argument that KP_1 occupies an A′ position above TP: it blocks A′-movement. In Section 4.2, we show that known arguments against treating KP_1 as occupying a position above TP are misguided.

4.1. Intervention

The construction at hand does not allow topicalisation and *wh*-movement, as shown in (45)–(46). (The word *to* in parentheses in (45) is an optional topic marker, synchronically unrelated to the demonstrative; the two can co-occur, as in (45b).) In the (a) examples, the right-hand predicate is fronted over KP_1: in (45a), for information-structural purposes, and in (46a), as an instance of *wh*-movement. Both examples are clearly ungrammatical. Compare these with examples in (b), which differ from (a) only in lacking KP_1. Their full acceptability demonstrates that the intervention effect is triggered precisely by the presence of KP_1.[23] To provide a complete picture, examples in (c) are canonical predicational constructions with instrumental predicates and regular subject-verb agreement. These again permit *wh*-movement and topicalisation, since the phrase that the predicate moves over is simply the Spec,TP subject.

(45) a. *Ostateczność$_i$ (to) leki to nie była t_i.
last.resort.**SG.F**.NOM TOP medicines TO NEG was.**SG.F**
Intended: 'Last resort, medicines were not.'

b. Ostateczność$_i$ (to) to nie była t_i.
last.resort.**SG.F**.NOM TOP TO NEG was.**SG.F**
'Last resort, this was not.'

c. Ostatecznością$_i$ (to) leki nie były t_i.
last.resort.SG.F.INS TOP medicines NEG were
'Last resort, medicines were not.'

[22] An anonymous reviewer asks how the coindexation of *to* and KP_1 is established. We assume that this instantiates the standard mechanism of coreference resolution for demonstratives. Accordingly, the coreference relation between KP_1 and *to* in (ii) is established in the same manner as in (i).

(i) Co sądzę o Janku$_i$? Cóż, to$_i$ jest dobry chłopak.
what think.1SG about Janek.LOC well TO is good boy.SG.M.NOM
'What do I think about Janek? Well, he's a good boy.'

(ii) Janek$_i$ to$_i$ jest dobry chłopak.
Janek TO is good boy.SG.M.NOM
'Janek is a good boy.'

[23] As noted by an anonymous reviewer, sentences like (i) are perfectly fine. This, however, does not affect our intervention argument. Note that in (46a), the *wh*-phrase is moved from the very bottom of the tree, where it controls agreement. In contrast, in (i), the deepest, agreement-controlling position is occupied by *Marysia.* Consequently, the *wh*-phrase does not move over it, but is generated higher, in the position of KP_1. Moreover, constructions like (i) are not predicational. They require an answer that IDENTIFIES the referent, as in *Mary was that woman over there*; they cannot be answered with a predicational statement such as *Mary was a good person.*

(i) Kto to {była /*był} Marysia?
who.SG.M.NOM TO was.**SG.F** was.SG.M Marysia.**SG.F**.NOM
'Who was Marysia?'

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

(46) a. *Kto$_i$ Marysia to był z zawodu t_i?
who.**SG.M** Marysia.SG.F TO was.**SG.M** by profession
Intended: 'Who was Marysia by profession?'
b. Kto$_i$ to był (z zawodu) t_i?
who.**SG.M** TO was.**SG.M** by profession
'Who was (s)he (lit. this) by profession?'
c. Kim$_i$ Marysia była z zawodu t_i?
who.INS Marysia.**SG.F** was.**SG.F** by profession?
'Who was Marysia by profession?'

Crucially, in studies specifically aiming at distinguishing between Spec,TP and positions ABOVE Spec,TP, intervening in A′-movement is regarded as an important diagnostic. For instance, Landau (2009), in refuting the 'broad subject' analysis of certain phrases in Hebrew – which posits two specifier positions in TP – appeals to this diagnostic and considers it 'a lethal problem' (Landau 2009: 97) for a TP-internal analysis.[24] The same reasoning also applies to the Polish construction at hand: KP_1 intervenes in A′-movement because it is base-generated in an A′ position itself.[25]

4.2. Refuting counterarguments

Works rejecting the dislocated analysis of Polish KP_1 refer to three phenomena suggesting a TP-internal status: (i) KP_1 may be realised as a quantifier phrase (Citko 2008), (ii) KP_1 can control a reflexive pronoun and (iii) KP_1 can control the subject of an adversative adjunct clause (Bondaruk 2013). We argue that these diagnostics do not provide evidence for a TP-internal status of KP_1, because, as will be shown, (i) cross-linguistically quantifiers are attested in dislocated positions, (ii) Polish reflexive pronouns can be controlled from outside the TP in which they are placed and (iii) subjects of adversative adjunct clauses can be bound by items other than Spec,TP subjects.

Firstly, Citko (2008: 263, fn. 1) dismisses the dislocated analysis on the basis of the observation that quantificational constituents can occur as KP_1 (see (47)), although they are 'impossible in left dislocated positions' (see an attempted hanging topic construction in (48)):

(47) Każdy student to (jest) potencjalny geniusz.
every student TO is potential genius
'Every student is a potential genius.' (Citko 2008: 263)

(48) *Każdy student, on jest potencjalnym geniuszem.
every student he is potential.INS genius.INS
Intended: 'Every student, he is a potential genius.' (Citko 2008: 263)

[24] Similar reasoning can be found in Aoun & Benmamoun (1998), Aoun et al. (2010), Alqarni & Alanazi (2023), Bošković (2024).

[25] This argument is foreshadowed in Rutkowski (2006: 166–168), but without a reference to *wh*-movement. On the other hand, the observation that *wh*-movement over KP_1 is disallowed is present in Citko (2008), but not as an argument for dislocation (Citko's argument AGAINST dislocation is refuted here in Section 4.2).

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

However, quantificational phrases are attested in A′ positions, as in clitic left dislocation constructions best known from Romance languages. Consider the Italian example (49), which contains a left-peripheral phrase co-occurring with a coreferential clitic in the main clause. Similarly, (50) is a Lebanese Arabic sentence containing a left-peripheral quantificational phrase.

(49) Tutti$_i$, non li$_i$ ho visti ancora.
all not them have.1SG seen yet
'All of them, I haven't seem them yet.' (Italian, Cinque 1991: 58, glosses modified)

(50) kəll təlmiize baddun ylabsuw-a maryul
every student.SG.F want.3PL dress.3PL-her uniform
'Every student, they want her to wear a uniform.'
(Lebanese Arabic, Aoun et al. 2010: 195, glosses modified)

Hence, (48) is ungrammatical not because A′ constituents cannot be quantifiers, they can, but because hanging topics must be referential rather than quantificational. Crucially, being in a dislocated position is not equal to being a hanging topic. The latter are best analysed as syntactically unintegrated phrases, simply juxtaposed with the main clause and interpreted as connected to this clause only at the discourse level (Shaer & Frey 2004, Shaer 2009, Villa-García 2023). We argue that KP_1 occupies a dedicated position in the left periphery, not that it is a hanging topic.

Secondly, Bondaruk (2013, 2015) shows that KP_1 can control subject-oriented reflexive pronouns, see (51).

(51) Marek to jest swój najlepszy obrońca.
Marek.SG.M.NOM TO is self's best defender.SG.M.NOM
'Marek is his own best defender.' (Bondaruk 2015: 205)

Importantly, this cannot be attributed to *to* binding the reflexive and – being coreferential with KP_1 – acting as a proxy. As shown in Section 3, such constructions lacking KP_1 are degraded; see (40b), repeated as (52). Therefore, we must conclude that KP_1 directly binds the reflexive.

(52) ??To jest swój najlepszy obrońca.
TO is self's best defender.SG.M.NOM

However, in Section 4.1, we presented syntactic evidence that the left-hand KP behaves as a dislocated phrase, in that it gives rise to intervention effects. How, then, can this be reconciled with the ability of KP_1 to bind reflexives?

We argue that TP does not constitute a locality domain for binding reflexives, and consequently, a TP-external KP_1 can bind a reflexive within that TP. Independent support for the availability of TP-external binders comes from constructions like (53). In this example, the reflexive can be bound not only by PRO – controlled by the matrix object *Piotrowi* 'Piotr.DAT' – but also by the matrix subject, which clearly lacks any syntactic representation within the infinitival TP.

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

(53) Jan$_i$ kazał Piotrowi$_j$ [$_{TP}$ PRO$_j$ przynieść swoje$_{i/j}$ dokumenty].
Jan.NOM ordered Piotr.DAT bring.INF self's documents
'Jan$_i$ ordered Piotr$_j$ to bring his$_{i/j}$ documents.'
(adapted from Przepiórkowski et al. 2002: 178)

The correct generalisation, then, is that the binder of the reflexive must have at least a minimal set of ϕ-features (accounting for the degraded acceptability of (52)) and be located within the same CP – rather than the same TP – as the reflexive (which explains the availability of both interpretations in (53)).[26]

Thirdly, Bondaruk (2013) argues that the dislocated analysis of KP_1 is wrong because KP_1 can control the subject of an adversative adjunct clause (AAC) introduced by *mimo* 'despite':

(54) Mimo PRO częstego słuchania wielu gatunków muzyki, Marek to (jest)
despite frequent listening.GEN many types music Marek TO is
zapalony fan rocka.
great fan.NOM rock.GEN
'Despite frequently listening to many types of music, Marek is a great fan of rock.'
(Bondaruk 2013: 227; glosses modified)

The argument rests on the assumption that only Spec,TP subjects can control the PRO in AAC. This assumption, however, is immediately falsified by attested examples such as (55)–(56). In (55), the subject of the AAC is controlled by a semantically prominent individual, not syntactically represented in the matrix clause at all. In (56), the controller likewise is not represented syntactically (if the precious metals were provided by, say, the ruler's servants) or alternatively, it is realised as the dative complement of the impersonal form *zwrócono* 'returned.IMPS' (if the ruler himself supplied the metals).

(55) Waga nie spada, mimo PRO przyjmowania drastycznie niskich
weight.NOM NEG drop despite consuming drastically low
ilości kalorii…
amounts calories.GEN
'(One's) weight isn't going down despite consuming drastically low amounts of calories.' (Google)[27]

(56) Mimo PRO dostarczenia kilkunastu ton złota i srebra, władcy
despite delivering dozen.or.so tons gold and silver ruler.DAT
Inków nie zwrócono wolności.
Incas.GEN NEG returned.IMPS freedom
'Despite delivering a dozen tons of gold and silver, the ruler of the Incas was not granted his freedom. (lit. to the ruler of Incas was not granted of freedom)' (Google)[28]

[26] Note that we refer to the position of KP_1 as an A′ position, even though anaphor binding is traditionally taken as a diagnostic of A positions. Together with intervention facts, the binding data indicate that the position of KP_1 exhibits mixed A/A′ properties: a situation that is not without precedent (see, for instance, Webelhuth 1992: 207–209 and van Urk 2015, Bošković 2024). See also Bošković (2025) for an analysis of binding insensitive to the A/A′ distinction.

[27] https://dieta-sportowca.pl/2018/04/metabolicdamage/; accessed on July 9, 2026.

[28] https://czaskultury.pl/felietony/szesc-i-pol-tony-konkwisty/; accessed on July 9, 2026.

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

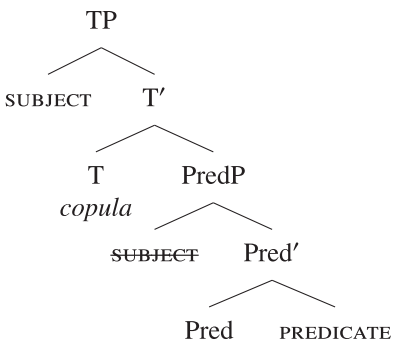
In summary, the assumption that KP_1 is base-generated above the TP agreement domain immediately explains intervention effects. Together with the status of *to*$_n$ discussed in the previous section, this assumption also explains why the nominative predicate is the highest accessible phrase and, hence, controls verbal agreement. Known arguments against this assumption either fail entirely (as in the case of quantificational phrases and adversative adjunct clauses) or can be reconciled with the TP-external analysis of KP_1 by granting it mixed A/A′ status or by adopting an analysis of binding insensitive to this distinction (as in the case of reflexive binding).

5. Formalisation

5.1. PredP(s) and nominative KP predicates

We adopt the currently standard view on the structure of copular clauses, which holds that there is a functional head mediating nonverbal predication, Pred (see, for instance, Bowers 1993, 2001, Baker 2003: Chapter 2, den Dikken 2006 for general discussion and Citko 2008, Bondaruk 2013 for its application to Polish copular clauses).[29] Its maximal projection, PredP, is selected by the copula – whether it is a lexical verb (V), a light verb (*v*) or simply the Tense head. This last minimal approach, on which the copula (*być* 'be') is an exponent of T in the context of PredP, is also adopted in this work (following, for instance, Baker 2003 and Roy 2013). In information-structurally neutral contexts, the constituent closest to T, i.e., the subject of predication in Spec,PredP, moves to Spec,TP to satisfy EPP. The schematic structure of such copular clauses is presented in (57).

(57)



The first question to be addressed is why predicates in constructions under investigation appear in the nominative case. From the perspective of Polish grammar, this is unexpected. Nominal predicates in canonical clauses – when the subject of predication is a nominative KP – typically take the instrumental case, as shown in (7), repeated below as (58).

(58) Leki były {ostatecznością / *ostateczność}.
medicines.PL.M.NOM were.PL last.resort.INS / last.resort.NOM
'Medicines were a last resort.'

[29] For a dissenting view, see Matushansky (2019).

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

Relatedly, why is it the case that non-canonical subjects occurring in the constructions analysed in this paper – PP[*z*] and *to*$_n$ – cannot occur with instrumental predicates? It is straightforward to show that they cannot in the case of PP[*z*], see (59). As for *to*$_n$, see Section 3, where we show that the item occurring with instrumental predicates is *to*$_{\text{K}}$, i.e., a canonical KP subject, rather than *to*$_n$.

(59) *Z Marysi było dobrym człowiekiem.
of Marysia.SG.F.GEN was.SG.N good person.SG.M.INS
Intended: 'Marysia (lit. of Marysia) was a good person.'

At its core, our proposal follows Citko (2008) and Bondaruk (2013): we assume there are two Pred heads – a 'canonical' one, which assigns the instrumental case to its complement (Pred$_{\text{INS}}$), and a 'defective' one, which does not assign case (Pred$_{\varnothing}$).[30] The two differ with regard to selectional restrictions, as shown in (60). The canonical Pred$_{\text{INS}}$ takes an instrumental predicate and combines with the same kind of specifiers as lexical verbs – typically KPs. By contrast, Pred$_{\varnothing}$ is 'defective' in the sense that it places no selectional restrictions on its complement and specifier.[31]

(60)

head	specifier	complement
Pred$_{\text{INS}}$	KP*	KP$_{\text{INS}}$
Pred$_{\varnothing}$	—	—

KP* in the table refers to 'canonical' subjects, i.e., ones that lexical verbs normally combine with. Besides nominative nominal phrases, these include infinitival phrases and CPs, as shown in the context of Pred$_{\text{INS}}$ in (61).

(61) a. [Wprowadzić ją w błąd]$_{\text{InfP}}$ było rzeczą bardzo trudną.
lead.INF she.ACC in mistake was.SG.N thing.INS very difficult
'To mislead her was a very difficult thing (to do).' (NCP).

b. [Że się od razu nie pozabijali]$_{\text{CP}}$ było zasługą matki…
COMP REFL immediately NEG killed was.SG.N achievement.INS mother.GEN
'That they didn't kill each other right away was thanks to the mother (lit. the mother's achievement).' (NCP)

Given that Pred$_{\varnothing}$ has no restrictions on its complement and specifier, the next question is why only non-canonical subjects – PP[*z*] and *to*$_n$ – can occur with it; for example, why is the Pred$_{\varnothing}$ version of (58), with the nominative predicate, unacceptable with the canonical subject? We propose that this is straightforwardly explained in terms of Pāṇinian blocking:

[30] The idea that there can be a family of Pred heads in a language is also present in Adger & Ramchand (2003), who argue for two Preds – eventive and non-eventive – in Scottish Gaelic. See also Citko (2008) and Bondaruk (2013) for discussion of semantic differences between Polish Preds.

[31] Note that we focus exclusively on nominal predicates here, ignoring the notorious problem of the grammatical case of adjectival predicates; for discussion, see e.g., Przepiórkowski (2004), Witkoś (2008), Lindert (2016).

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

since KP subjects have the dedicated Pred_{INS} selecting for such items, $\text{Pred}_{\varnothing}$ (the 'elsewhere' Pred) is blocked, as it is less specified for such subjects.

Interestingly, canonical subjects CAN occasionally occur with nominative predicates, but this structure is rare and requires an expressive predicate, as in (62)–(63).

(62) Ale król był straszna świnia, zepsuł mu gitarę!
but king.**SG.M**.NOM was.**3SG.M** awful pig.SG.F.NOM broke he.DAT guitar
'But the king was a terrible pig, he broke his guitar!' (Polish comedy song by W. Sikora)[32]

(63) No, dziadek był kawał drania – potaknął
well grandfather.SG.M.NOM was.3SG.M piece.SG.M.NOM bastard.GEN nodded
Izzat.
Izzat
'Well, grandpa was a bit of a bastard – nodded Izzat.' (CCP)

Paradoxically, this phenomenon confirms the analysis in terms of Pred_{INS} and $\text{Pred}_{\varnothing}$. It has been reported that violating economy conditions does not necessarily lead to ungrammaticality if it induces a pragmatic effect (Schlenker 2005, Patel-Grosz & Grosz 2017). This is exactly the case here: using $\text{Pred}_{\varnothing}$ with a nominative subject, instead of the better match, Pred_{INS}, is only licensed when the expression is intended to convey a special interpretive effect, namely, marking the speaker's expressive attitude towards the assertion.

Importantly, (62) also shows the workings of agreement when two nominative phrases are present in the agreement domain: the HIGHEST one controls agreement, precisely as predicted by Bobaljik's generalisation.

Before we proceed to the formal analysis of agreement, let us show how our minimal inventory of Preds interacts with control facts mentioned in Section 2; see (17), repeated as (64).

(64) *Chcieli być [PRO niezła paczka].
wanted.3PL be.INF not.bad package.SG.F.NOM
Intended: 'They wanted to be a nice bunch.'

The predicate *niezła paczka* 'a nice bunch' is fine with both *to* and PP in the finite matrix structure (as illustrated for PP in (9)). However, a control structure cannot be formed. This follows naturally under the assumption that PRO bears case (see Chomsky & Lasnik 1993, Sigurðsson 2008). Given that case-bearing items have a more specialised Pred available, the defective $\text{Pred}_{\varnothing}$ is blocked.

Crucially, as shown in (65), strongly expressive nominative predicates can occur in control structures. This directly follows from the analysis proposed for constructions like (62)–(63): once the predicate is sufficiently expressive, a violation of the economy conditions described above is pragmatically licensed. As a result, PRO can occur with $\text{Pred}_{\varnothing}$ despite the availability of the more specified alternative, Pred_{INS}.

[32] Available on YouTube: https://www.youtube.com/watch?v=4kzpeHsUdW0; accessed on July 9, 2026.

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

(65) a. Gdyby-m chciał być [PRO kawał dziada]…
if-1SG want be.INF piece.NOM bastard.GEN
'If I wanted to be a piece of bastard…' (Google)[33]

b. Ludzie myślą, że ja zwyczajnie [PRO próbuję być kozak]…
people think COMP I simply try.1SG be.MG daredevil.NOM
'People think I simply try to be a daredevil…' (Google)[34]

Having laid out our assumptions regarding the basic structure of copular clauses, we now turn to the formal implementation of the agreement mechanism.

5.2. Agreement

Bobaljik (2008) does not provide a formal derivational analysis of case-dependent Agree. To implement Bobaljik's generalisation formally, we require a mechanism involving (i) case-discriminating agreement probes which (ii) copy the ϕ features of their goal. Under a certain set of assumptions, these can be achieved by any model of Agree, including the classic formulation in Chomsky (2000, 2001). A central feature of Chomsky's approach, however, is the claim that agreement and case-valuation proceed in tandem, and the eventually nominative phrase must be 'active' – bear an unvalued case (i.e., the [ucase:] feature) – in order to participate in agreement. This assumption, in turn, runs counter to one of the core insights of Bobaljik (2008): that agreement is dependent on case (see also Preminger 2014).[35]

For concreteness, we adopt the interaction/satisfaction model of Agree (Deal 2015, 2024, Clem & Deal 2026, a.o.), as it allows us to present the analysis in the most perspicuous way. We modify this approach in a way directly implementing case-dependent Agree. As specified below, the agreement probe postulated here searches for the nominative case. In this respect, we depart from Clem & Deal (2026), where case marking is treated as a by-product of agreement.

In the interaction/satisfaction model, agreement probes are specified by two parameters: INTERACTION (which features the probe interacts with) and SATISFACTION (which features, if any, cause the probe to stop probing further). We propose that Polish T both interacts with and is satisfied by nominative KPs, thus implementing Bobaljik's generalisation, rather than by some set of ϕ features, as is more frequently assumed. That is, the probe on T is specified as [INT: NOM, SAT: NOM].

Given this, we adopt the 'coarse' approach to feature copying (see, for instance, Grishin 2024, Deal & Royer 2025), which assumes that '*all* features of the goal are copied onto the probe, not just those that entail the interaction feature' (Grishin 2024: 19).[36] That is, by saying that the probe T interacts with NOM, we imply that it copies all the features present on a phrase bearing NOM, including the ϕ features, which results in observed agreement.

Let us illustrate this analysis with a structure involving to_n (but not KP_1), such as (4), repeated again as (66).

[33] https://forum.audio.com.pl/topic/99381-czy-prezentujecie-sw%C3%B3j-sprz%C4%99t-znajomymnie-audiofilom/page/2/; accessed on October 29, 2026.

[34] https://belgianasznowydom.blogspot.com/2024/01/o-mojej-empatii-normalnej-inaczej.html; accessed on October 29, 2026.

[35] For a critical discussion of the activity condition, see Deal (2025).

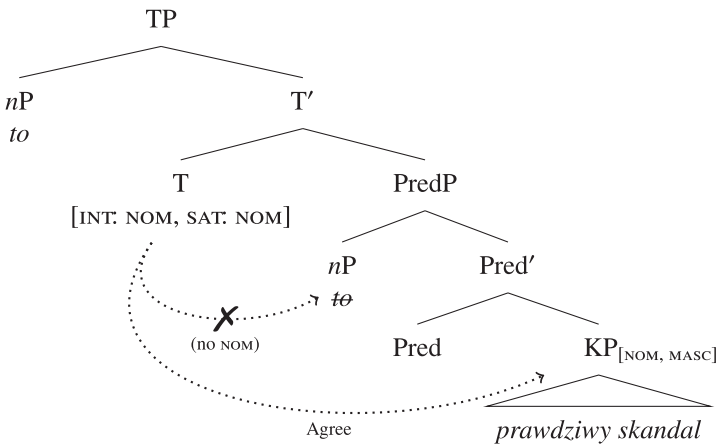
[36] An alternative approach, adopted, for example, in Clem & Deal (2026), posits that a probe specified as [INT: F] copies only the feature F and all features that dominate F in a feature geometry.

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

(66) To {był / *było} prawdziwy skandal. (to_n)
TO was.**SG.M** / was.SG.N real.SG.M.NOM scandal.**SG.M**.NOM
'This was a real scandal.' (NCP)

The right-hand KP predicate occurs in the nominative case and obligatorily controls agreement. The word *to* is a reduced demonstrative lacking a case projection, which makes it invisible to the probe, see (67).[37] The word *to* moves to Spec,TP for EPP reasons (crucially, detached from agreement, as is argued for Slavic in, for instance, Lavine & Freidin 2002, Bailyn 2004).

(67)



Constructions with PP subjects of predication are fully analogous, with PP[*z*] in place of to_n in (67) (cf. (71) in Section 5.4).

Moreover, (67) also illustrates how agreement operates in structures such as (5), repeated below as (68). KP_1 is base-generated above TP (and, hence, not shown in (67)) and is therefore fully irrelevant for the T probe.

(68) [$_{XP}$ Leki [$_{TP}$ to była ostateczność]].
medicines.PL.M.NOM TO was.**SG.F** last.resort.**SG.F**.NOM
'Medicines were a last resort.' (CCP)

Note that under the view that PredP is a phase (see Citko 2014: §5.1), a locality issue arises, since the probe would have to enter the phase, thereby violating the Phase Impenetrability Constraint (PIC; Chomsky 2000). However, issues concerning the interaction between agreement and PIC are not unique to our analysis and have prompted a range of proposals addressing how such configurations should be handled. One possibility is to weaken the PIC, as proposed in Chomsky (2001), who argues that a phase is sent to

[37] Although we subscribe to the view that ϕ features are organised hierarchically (as in (33)), in the trees we present only the features crucial for agreement (we omit ϕ and π).

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

Spell-Out only after another phase head has been merged. On this view, the structures analysed here do not involve a PIC violation: T probes into PredP immediately, before the next phase head, C, is merged. Other possibilities are that agreement is not restricted by phases (only movement is), as argued in Bošković (2007), or that Pred (at least $\text{Pred}_{\varnothing}$) is a 'defective' phase head – i.e., one lacking a full set of uninterpretable ϕ features – and is therefore transparent for agreement (as proposed in Richards 2012: 151 to account for apparent PIC-violating agreement).

Due to space constraints, we do not attempt to adjudicate between these approaches here, focusing instead on additional data in the following two subsections that confirm the analysis postulated in this paper.

5.3. KP_1 outside agreement domain

Given the interaction/satisfaction model, especially as presented in Clem & Deal (2026), one could claim that KP_1 is within the agreement domain after all, but the 'insatiable' probe interacts with BOTH KP_1 and the KP_2 predicate, and only expones the features of the last – 'outermost' (Grishin 2024, Przepiórkowski 2025) – item it interacts with, i.e., the features of KP_2. This would be compatible with an analysis on which KP_1 is base-generated within TP, but not with the analysis proposed above. Dually, if it could be shown that the probe does not interact with KP_1 at all, this would be a confirmation of the analysis proposed here.

In this subsection, we demonstrate that KP_1 indeed does not interact with the probe: KP_1's features fail to be exponed on the copula even when it is the ONLY nominative phrase equipped with ϕ features in the sentence. Consider (69), where the predicate is a ϕ-less adverbial phrase.[38]

(69) a. […] jedna tabletka to {było / *była} za mało.
one pill.SG.F.NOM TO was.SG.N was.SG.F too little
'One pill was not enough (lit. was too little).' (PW2019)

b. Przeszło miesięczna męka to {było / *była} za dużo…
over month.long ordeal.SG.F.NOM TO was.SG.N was.SG.F too much
'The ordeal of more than a month was too much.' (PW2019)

There are no nominative KPs in these sentences other than KP_1,[39] so – on the 'insatiable' Agree account sketched above – KP_1s should control agreement here. But it is clear that they do not: if they did, the copula would expone the feminine gender, which it does not. Instead, only the singular neuter form of the verb is allowed here, resulting from 'default agreement' (Dziwirek 1990, Preminger 2014), i.e., the lack of an actual agreement controller. This confirms that KP_1 is base-generated outside the agreement domain.

[38] Such sentences also contradict the commonly expressed empirical claim that in such constructions the left-hand constituent and the predicate must have the same grammatical category (Citko 2008, Bondaruk 2013).

[39] There exist broadly nominal – actually, numeral – phrases such as *za mało lekarstwa* 'too little medicine.GEN' and *za dużo kłopotu* 'too much trouble.GEN', but it may be demonstrated that such numeral phrases in Polish do not occur in the nominative (Franks 1995, Przepiórkowski 1999), and the nouns within them are in the genitive, so even on an elliptical analysis of the apparently adverbial predicates in (69) as reduced numeral phrases, there are no nominative KPs there apart from the KP_1s.

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

5.4. Long-distance agreement

Consider (70), with copular constructions embedded under raising verbs.

(70) a. Marysia musiała [być dobrym człowiekiem].
Marysia.**SG.F**.NOM must.**SG.F** be.INF good person.SG.M.INS
b. Z Marysi musiał [być dobry człowiek].
of Marysia.SG.F.GEN must.**SG.M** be.INF good person.**SG.M**.NOM
c. Marysia to musiał [być dobry człowiek].
Marysia.SG.F.NOM TO must.**SG.M** be.INF good person.**SG.M**.NOM
'Marysia must have been a good person.'

In these sentences, the finite modal verb expresses the features of the only nominative phrase in (70a) (canonical copular clause) and (70b) (non-canonical copular clause with a prepositional subject), or of the right-hand nominative phrase in (70c) (non-canonical copular clause with to_n and KP_1), as might be expected. However, unlike all previous examples, (70b) and (70c) are instances of long-distance agreement, a phenomenon in which a matrix verb exhibits agreement with an argument of an embedded verb. The literature related to such phenomena is abundant,[40] with various theoretical proposals: from those arguing that such phenomena involve a sequence of cyclic, strictly local agreement processes, with the effect that the non-locality is only apparent (e.g. Haug & Nikitina 2016), to those which suggest that agreement is not limited by virtually any locality constraints (Bošković 2007). Although a detailed analysis of how the phenomenon presented in (70b)–(70c) can be understood within the broader theoretical research on long-distance agreement is beyond the scope of the paper, we would like to note that the account offered above captures these patterns immediately.

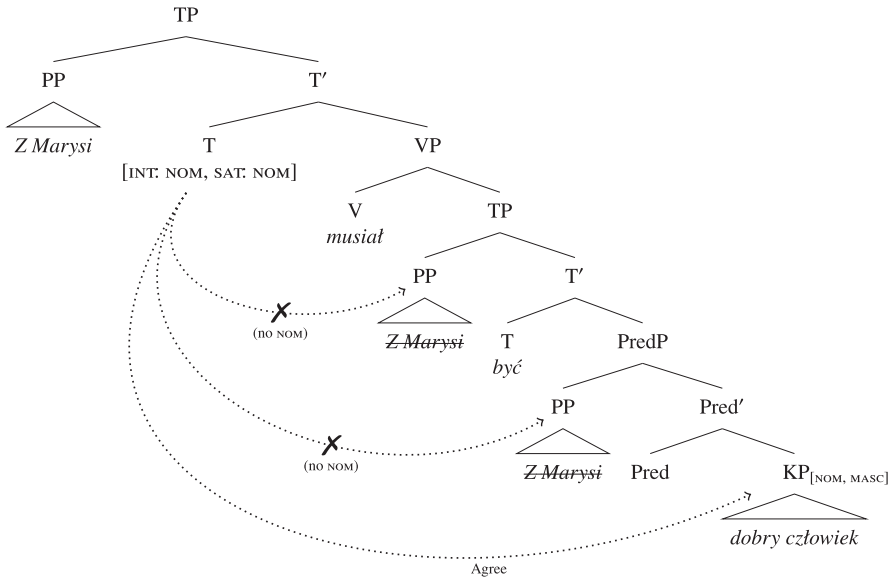
From the perspective of Bobaljik's generalisation adopted here, there is nothing special about agreement in (70b)–(70c): the highest accessible phrase controls agreement with the copula, see (71) corresponding to (70b) above.[41] Once the embedded TP is built and selected by the raising modal verb, the main sentential T merges and probes for a nominative KP, finding it at the very bottom of the tree.

[40] See Bhatt & Keine (2017) for an overview.

[41] Analogously to *to* in (67), the PP in (71) moves to satisfy EPP. The assumption that EPP can be satisfied by a PP is not new; see, for instance, Babyonyshev's (1996) analysis of Russian locative inversion and Bailyn's (2012) reformulation of EPP in Russian in terms of argument prominence.

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

(71)



As expected, control verbs – such as *chcieć* 'want' – cannot occur with such non-canonical structures, although they can occur with a canonical copular clause with an instrumental predicate:

(72) a. Marysia chciała być dobrym człowiekiem.
Marysia.**SG.F**.NOM wanted.**SG.F** be.INF good person.SG.M.INS
b. *Z Marysi chciał być dobry człowiek.
of Marysia.SG.F.GEN wanted.**SG.M** be.INF good person.**SG.M**.NOM
c. *Marysia to chciał być dobry człowiek.
Marysia.SG.F.NOM TO wanted.**SG.M** be.INF good person.**SG.M**.NOM
'Marysia wanted to be a good person.'

The crucial factor here is that control verbs project their own subject (in Spec,*v*P), which – as in the case of all other lexical verbs – cannot be 'non-canonical' (*n*P, PP[*z*]). As postulated in Section 5.1, lexical verbs select for KP* – the shorthand for 'canonical' subjects. By contrast, raising verbs do not project their Spec,*v*P and raise their (possibly non-canonical) subjects from the embedded TP.

Another relevant piece of evidence pointing in the same direction is related to available interpretations of modal verbs. Note that the canonical copular sentences embedded under *móc* 'can' allow for both epistemic and non-epistemic interpretations (cf. Kratzer 1991). That is, on the one hand, (73a) can be interpreted so that Janek has the ability or is permitted to be 'a terrible pig' (non-epistemic). On the other, it can be interpreted epistemically, i.e., as saying that Janek being 'a terrible pig' is compatible with the evidence. In contrast, the non-canonically agreeing sentence (73b) can only be interpreted epistemically: the sentence simply asserts that it is compatible with the evidence that Janek was a terrible pig. For example, it cannot be interpreted deontically, as saying that somebody permitted Janek to misbehave. (Similar considerations apply to non-canonical copular constructions with to_n.) This is in line with the common view that epistemic modals are expressed by raising verbs,

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

while non-epistemic modals are typically expressed by control verbs (e.g., Jackendoff 1972, Kratzer 1991, Brennan 1993).

(73) a. Janek mógł być straszną świnią.
Janek.**SG.M**.NOM could.**SG.M** be.INF terrible pig.SG.F.NOM
'Janek could be a terrible pig.' (*epistemic*/*non-epistemic*)

b. Z Janka mogła być straszna świnia.
of Janek.SG.M.GEN could.**SG.F** be.INF terrible pig.**SG.F**.NOM
'Janek could be a terrible pig.' (*epistemic*/*#non-epistemic*)

In summary, when non-canonical copular constructions discussed in this paper are complements of epistemic raising verbs, the nominative KP predicate still controls agreement, but this time it is long-distance agreement between a verbal head (the epistemic modal) and a phrase embedded within its (infinitival) TP complement. That is, what the finite verb agrees with in such constructions cannot reasonably be analysed as this verb's subject.

This is apparently the first report of long-distance agreement in Slavic in the literature; it strikingly confirms the view of agreement expressed in Bobaljik's generalisation (6) and disconfirms the view common in traditional Slavic linguistics that verbs only agree with their subjects.[42]

6. Conclusion

Agreement in Polish follows the generalisation proposed in Bobaljik (2008), and does so in a nontrivial way. The traditional view claiming that the subject always controls agreement is merely an approximation that breaks down in certain contexts. A more accurate generalisation is that agreement is controlled by the highest accessible phrase within the relevant domain. This paper demonstrated this on the basis of copular clauses, which seem to be the only environment where the difference between these two approaches to agreement becomes transparent in Polish.

The case-driven analysis of agreement is strongly supported by the various facts discussed in this paper: the prepositional subject of predication and to_n cannot control agreement, as they are not nominative KPs. The left-hand nominative KP in "extended" sentences with to_n is, in turn, outside the agreement domain. Thanks to the fact that the predicates in such sentences are in the nominative case, we can observe the otherwise unexpected non-canonical agreement.

[42] Although the discussion above focuses on modal verbs, the phenomenon is not restricted to them, but includes raising verbs in general. See (ia), involving the verb *zacząć* 'to begin', and (ib) with *przestać* 'to stop'.

(i) a. Nazwy rond to zaczął być problem…
names.PL.F.NOM roundabouts.GEN TO started.**SG.M** be.INF problem.**SG.M**.NOM
'Names of roundabouts started to be a problem.'
(Google; https://mobile.nowiny.pl/egazeta/tygodnik-rybnicki/2012-06-26/82631-solidarnosc-na-rondzie.html; accessed on October 19, 2026.)

b. Uszka na wigilię to przestał być problem…
dumplings.PL.N.NOM for Christmas.Eve TO stopped.**SG.M** be.INF problem.**SG.M**.NOM
'Dumplings for Christmas Eve are no longer a problem.'
(Google, A consumer's comment on a food processor)

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

While the above account of agreement, of the word(s) *to* and of the position of KP_1 are amply supported by a variety of phenomena, the overall syntax of copular constructions assumed here remains more tentative. Despite extensive research on the topic, the structure of nonverbal predication is still poorly understood. The proposed analysis assumes that nonverbal predication is mediated by a functional head typically called Pred, which aligns with the mainstream view, but a Pred-less analysis, following Matushansky (2019), may ultimately prove more compelling. We hope that future research will explore this possibility and will attempt to incorporate the findings of this paper into a Pred-less analysis of copular clauses.

Acknowledgements. For helpful feedback on the material presented in this paper, we would like to thank the audiences at the Slavic Linguistics Society (Verona, 10–12 September 2025), Formal Description of Slavic Languages (Wrocław, 22–24 September 2025) and Generative Linguistics in Poland (Warsaw, 6–7 November 2025), as well as the participants of the Formal Linguistics Seminar at the Institute of Computer Science, Polish Academy of Sciences. We are also grateful to Oleg Belyaev for a helpful discussion of the Russian data presented in Section 2. Finally, we thank the anonymous reviewers for their generous comments. All remaining errors are our own.

Funding statement. This research was supported by the National Center of Science, Poland (NCN PRELUDIUM, 2022/45/N/HS2/03109).

References

Adger, David & Gillian Ramchand. 2003. Predication and equation. *Linguistic Inquiry.* 34(3). 325–359. https://doi.org/10.1162/002438903322247515.

Alqarni, Muteb & Mohammad Alanazi. 2023. There are no broad subjects in Standard Arabic. *Linguistics.* 61(5). 1131–1164. https://doi.org/10.1515/ling-2021-0125.

Aoun, Joseph & Elabbas Benmamoun. 1998. Minimality, reconstruction, and PF movement. *Linguistic Inquiry.* 29(4). 569–597.

Aoun, Joseph E., Elabbas Benmamoun & Lina Choueiri. 2010. *The syntax of Arabic*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511691775.

Babyonyshev, Maria A. 1996. *Structural connections in syntax and processing: Studies in Russian and Japanese.* Cambridge, MA: MIT dissertation.

Bailyn, John Frederick. 2004. Generalized inversion. *Natural Language & Linguistic Theory.* 22(1). 1–49. https://doi.org/10.1023/B:NALA.0000005556.40898.a5.

Bailyn, John Frederick. 2012. *The syntax of Russian*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511984686.

Baker, Mark C. 2003. *Lexical categories: Verbs, nouns and adjectives*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511615047.

Béjar, Susana & Arsalan Kahnemuyipour. 2017. Non-canonical agreement in copular clauses. *Journal of Linguistics.* 53(3). 463–499. https://doi.org/10.1017/S002222671700010X.

Béjar, Susana & Milan Řezáč. 2009. Cyclic Agree. *Linguistic Inquiry.* 40(1). 35–73. https://doi.org/10.1162/ling.2009.40.1.35.

Bhatia, Sakshi & Rajesh Bhatt. 2023. Copular agreement in Hindi-Urdu. *Glossa: A Journal of General Linguistics.* 8(1). https://doi.org/10.16995/glossa.5777.

Bhatt, Rajesh & Stefan Keine. 2017. Long-distance agreement. In Martin Everaert & Henk C. Riemsdijk (eds.), *The Wiley Blackwell companion to syntax*, 1–30. Hoboken, NJ: Wiley-Blackwell 2nd edn. https://doi.org/10.1002/9781118358733.wbsyncom083.

Bobaljik, Jonathan David. 2008. Where's phi? Agreement as a postsyntactic operation. In Harbour et al. (2008) 295–328.

Bogusławski, Andrzej. 2002. *To jest* – 'das ist'. Unteilbare Spracheinheiten? Doch! Teil II. *Die Welt der Slaven.* 47(2). 251–274.

Bondaruk, Anna. 2012. Person Case Constraint effects in Polish copular constructions. *Acta Linguistica Hungarica.* 59(1–2). 49–84. https://doi.org/10.1556/aling.59.2012.1-2.3.

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

Bondaruk, Anna. 2013. *Copular clauses in English and Polish. Structure, derivation and interpretation*. Lublin: Wydawnictwo KUL.
Bondaruk, Anna. 2014. Polish predicational clauses with the preposition *z* 'of'. *Roczniki Humanistyczne.* 62. 33–48.
Bondaruk, Anna. 2015. Case and agreement in Polish predicational clauses. *Lingue e Linguaggio.* 14(2). 197–215.
Bondaruk, Anna. 2019. Agreement with the post-verbal DP in Polish dual copula clauses. In María J. Arche, Antonio Fábregas & Rafael Marín (eds.), *The grammar of copulas across languages*, 107–129. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198829850.003.0006.
Bošković, Željko. 2007. On the locality and motivation of Move and Agree: An even more minimal theory. *Linguistic Inquiry.* 38(4). 589–644. https://doi.org/10.1162/ling.2007.38.4.589.
Bošković, Željko. 2024. On wh and subject positions, the EPP, and contextuality of syntax. *The Linguistic Review.* 41(1). 7–58. https://doi.org/10.1515/tlr-2024-2002.
Bošković, Željko. 2025. Contextuality of syntax and adieu to the A/A′-distinction. Unpublished manuscript. https://ling.auf.net/lingbuzz/008345.
Bowers, John. 1993. The syntax of predication. *Linguistic Inquiry.* 24(4). 591–656.
Bowers, John. 2001. Predication. In Mark Baltin & Chris Collins (eds.), *The handbook of contemporary syntactic theory*, 299–333. Oxford: Blackwell.
Brennan, Virginia M. 1993. *Root and epistemic modal auxiliary verbs*. Amherst, MA: University of Massachusetts dissertation.
Bresnan, Joan. 1994. Locative inversion and the architecture of Universal Grammar. *Language.* 70(1). 72–131. https://doi.org/10.2307/416741.
Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–156. Cambridge, MA: The MIT Press.
Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), Ken Hale*: A life in language*, 1–52. Cambridge, MA: The MIT Press.
Chomsky, Noam & Howard Lasnik. 1993. The theory of principles and parameters. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann (eds.), *Syntax: An international handbook of contemporary research*, vol. 1, 506–569. Berlin: Walter de Gruyter.
Cinque, Guglielmo. 1991. *Types of A′-dependencies* Cambridge, MA: The MIT Press.
Citko, Barbara. 2008. Small clauses reconsidered: Not so small and not all alike. *Lingua.* 118(3). 261–295. https://doi.org/10.1016/j.lingua.2007.05.009.
Citko, Barbara. 2014. *Phase Theory: An introduction*. Cambridge: Cambridge University Press.
Clem, Emily & Amy Rose Deal. 2026. Dependent case by Agree: Ergative in Shawi. *Linguistic Inquiry.* 57(2). 267–313. https://doi.org/10.1162/ling_a_00529.
Collins, Chris. 2024. *Principles of argument structure: A Merge-based approach*. Cambridge, MA: The MIT Press. https://doi.org/10.7551/mitpress/15436.001.0001.
Coon, Jessica & Stefan Keine. 2021. Feature gluttony. *Linguistic Inquiry.* 52(4). 655–710. https://doi.org/10.1162/ling_a_00386.
Deal, Amy Rose. 2015. Interaction and satisfaction in ϕ-agreement. In Thuy Bui & Deniz Özyldz (eds.), *Proceedings of the Forty-Fifth Annual Meeting of the North East Linguistic Society*, 179–192. Amherst, MA: GLSA.
Deal, Amy Rose. 2024. Interaction, satisfaction, and the PCC. *Linguistic Inquiry.* 55(1). 39–94. https://doi.org/10.1162/ling_a_00455.
Deal, Amy Rose. 2025. Current models of Agree. *Berkeley Papers in Formal Linguistics.* 4(1). https://doi.org/10.5070/BF2.48662.
Deal, Amy Rose & Justin Royer. 2025. Mayan animacy hierarchy effects and the dynamics of Agree. *Natural Language & Linguistic Theory.* 43. 1607–1663. https://doi.org/10.1007/s11049-024-09648-y.
Déchaine, Rose-Marie & Martina Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry.* 33(3). 409–442. https://doi.org/10.1162/002438902760168554.
den Dikken, Marcel. 2006. *Relators and linkers: The syntax of predication, predicate inversion, and copulas*. Cambridge, MA: The MIT Press. https://doi.org/10.7551/mitpress/5873.001.0001.
Dziwirek, Katarzyna. 1990. Default agreement in Polish. In Katarzyna Dziwirek, Patrick Farrell & Errapel Mejías-Bikandi (eds.), *Grammatical relations: A cross-theoretical perspective*, 147–161. Stanford, CA: CSLI Publications.
Dziwirek, Katarzyna. 1994. *Polish subjects*. New York & London: Garland.
Franks, Steven. 1995. *Parameters of Slavic morphosyntax*. New York, NY: Oxford University Press.

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

Grishin, Peter. 2024. How to agree with the lowest DP. Manuscript. https://ling.auf.net/lingbuzz/007112.
Harbour, Daniel, David Adger & Susana Béjar (eds.). 2008. *Phi theory: Phi-features across modules and interfaces*. Oxford: Oxford University Press.
Harley, Heidi & Elizabeth Ritter. 2002. Person and number in pronouns: A feature-geometric analysis. *Language.* 78(3). 482–526. https://doi.org/10.1353/lan.2002.0158.
Hartmann, Jutta M. & Caroline Heycock. 2023. Person effects in agreement with Icelandic low nominatives: An experimental investigation. *Natural Language & Linguistic Theory.* 41(3). 1029–1090. https://doi.org/10.1007/s11049-022-09564-z.
Haug, Dag Trygve Truslew & Tatiana Nikitina. 2016. Feature sharing in agreement. *Natural Language & Linguistic Theory.* 34(3). 865–910. https://doi.org/10.1007/s11049-015-9321-9.
Hentschel, Gerd. 2001. On the perspectivisation of noun phrases in copula sentences, mainly in Polish: *(Y) to (jest) X* and similar phenomena. In Viktor S. Chrakovskij, Maciej Grochowski & Gerd Hentschel (eds.), *Studies on the syntax and semantics of Slavonic languages*, 161–213. Oldenburg: Bibliotheks- und Informationssystem der Universität Oldenburg.
Higgins, Francis Roger. 1973. *The pseudo-cleft construction in English*. Bloomington, IN: Indiana University Linguistics Club.
Higgins, Francis Roger. 1979. *The pseudo-cleft construction in English*. New York, NY: Garland.
Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar.* Cambridge, MA: The MIT Press.
Keine, Stefan, Michael Wagner & Jessica Coon. 2019. Hierarchy effects in copula constructions. *Canadian Journal of Linguistics/Revue canadienne de linguistique.* 64(4). 617–648. https://doi.org/10.1017/cnj.2019.28.
Kieraś, Witold, Małgorzata Marciniak, Marek Łaziński, Marcin Woliński, Krystyna Bojałkowska, Wiktor Eźlakowski, Łukasz Kobyliński, Dorota Komosińska, Katarzyna Krasnowska-Kieraś, Michał Rudolf, Aleksandra Tomaszewska, Joanna Wołoszyn & Natalia Zawadzka-Paluektau. 2025. Korpus Współczesnego Języka Polskiego. Dekada 2011–2020. *Język Polski.* 105(2). https://doi.org/10.31286/JP.001055.
Kilgarriff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý & Vít Suchomel. 2014. The Sketch Engine: Ten years on. *Lexicography ASIALEX.* 1. 7–36. https://doi.org/10.1007/s40607-014-0009-9.
Kratzer, Angelika. 1991. Modality. In Arnim von Stechow & Dieter Wunderlich (eds.), *Semantics: An international handbook of contemporary research*, 639–650. Berlin: de Gruyter.
Landau, Idan. 2009. Against broad subjects in Hebrew. *Lingua.* 119(1). 89–101. https://doi.org/10.1016/j.lingua.2008.08.004.
Lavine, James E. & Robert Freidin. 2002. The subject of defective T(ense) in Slavic. *Journal of Slavic Linguistics.* 10(1/2). 251–287.
Lindert, Patrick. 2016. Default case, NP-ellipsis, and the theory of control in Polish. *Questions and Answers in Linguistics.* 3(1). 1–26. https://doi.org/10.1515/qal-2016-0002.
Matushansky, Ora. 2019. Against the PredP theory of small clauses. *Linguistic Inquiry.* 50(1). 63–104. https://doi.org/10.1162/ling_a_00301.
Mańczak, Witold. 1956. Ile rodzajów jest w polskim? *Język Polski.* XXXVI(2). 116–121.
McGinnis, Martha. 2005. On markedness asymmetries in person and number. *Language.* 81(3). 699–718. https://doi.org/10.1353/lan.2005.0141.
Moro, Andrea. 1997. *The raising of predicates: Predicative noun phrases and the theory of clause structure*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511519956.
Orzechowska, Alicja. 1998. Rzeczownik. In Renata Grzegorczykowa, Roman Laskowski & Henryk Wróbel (eds.), *Gramatyka współczesnego języka polskiego: Morfologia*, 270–332. Warsaw: Wydawnictwo Naukowe PWN 2nd edn.
Patel-Grosz, Pritty & Patrick G. Grosz. 2017. Revisiting pronominal typology. *Linguistic Inquiry.* 48(2). 259–297. https://doi.org/10.1162/LING_a_00243.
Preminger, Omer. 2014. *Agreement and its failures*. Cambridge, MA: The MIT Press.
Przepiórkowski, Adam. 1999. *Case assignment and the complement-adjunct dichotomy: A non-configurational constraint-based approach*. Tübingen: Universität Tübingen dissertation.
Przepiórkowski, Adam. 2004. On case transmission in Polish control and raising constructions. *Poznań Studies in Contemporary Linguistics.* 39. 103–123.
Przepiórkowski, Adam. 2025. Concord is insatiable Agree: Heterogeneous case and phi patterns in Slavic and Finnic. Unpublished manuscript. https://ling.auf.net/lingbuzz/009491.
Przepiórkowski, Adam, Mirosław Bańko, Rafał L. Górski & Barbara Lewandowska-Tomaszczyk (eds.). 2012. *Narodowy Korpus Języka Polskiego*. Warsaw: Wydawnictwo Naukowe PWN.

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

Przepiórkowski, Adam, Anna Kupść, Małgorzata Marciniak & Agnieszka Mykowiecka. 2002. *Formalny opis języka polskiego: Teoria i implementacja*. Warsaw: Akademicka Oficyna Wydawnicza EXIT.
Řezáč, Milan. 2008. Phi-agree and theta-related case. In Harbour et al. (2008) 83–129.
Richards, Marc. 2012. Probing the past: On reconciling long-distance agreement with PIC. In Artemis Alexiadou, Tibor Kiss & Gereon Müller (eds.), *Local modelling of non-local dependencies in syntax*, 135–154. Berlin: de Gruyter.
Roy, Isabelle. 2013. *Non-verbal predication: Copular sentences at the syntax–semantics interface*. Oxford: Oxford University Press.
Rutkowski, Paweł. 2006. From demonstratives to copulas: A cross-linguistic perspective and the case of Polish. *Journal of Universal Language.* 7. 147–175. http://doi.org/10.22425/jul.2006.7.2.147.
Saloni, Zygmunt & Marek Świdziński. 1998. *Składnia współczesnego języka polskiego*. Warsaw: Wydawnictwo Naukowe PWN.
Savchuk, Svetlana O., Timofey Arkhangelskiy, Anastasiya A. Bonch-Osmolovskaya, Ol'ga V. Donina, Yuliya N. Kuznetsova, Ol'ga N. Lyashevskaya, Boris V. Orekhov & Mariya V. Podryadchikova. 2024. Russian National Corpus 2.0: New opportunities and development prospects. *Voprosy Jazykoznanija.* 2. https://doi.org/10.31857/0373-658X.2024.2.7-34.
Schlenker, Philippe. 2005. *Minimize Restrictors!* (Notes on definite descriptions, Condition C and epithets). *Proceedings of Sinn und Bedeutung.* 9. 385–416. https://doi.org/10.18148/sub/2005.v9i0.776.
Shaer, Benjamin. 2009. German and English left-peripheral elements and the orphan analysis of non-integration. In Benjamin Shaer, Philippa Cook, Werner Frey & Claudia Maienborn (eds.), *Dislocated elements in discourse: Syntactic, semantic and pragmatic perspectives*, 366–397. New York, NY: Routledge.
Shaer, Benjamin & Werner Frey. 2004. 'Integrated' and 'non-integrated' left-peripheral elements in German and English. *ZAS Papers in Linguistics.* 35(2). 465–502. https://doi.org/10.21248/zaspil.35.2004.238.
Sigurðsson, Halldór Ármann. 2008. The case of PRO. *Natural Language & Linguistic Theory.* 26(2). 403–450. https://doi.org/10.1007/s11049-008-9040-6.
van Urk, Coppe. 2015. *A uniform syntax for phrasal movement: A case study of Dinka Bor*. Cambridge, MA: MIT dissertation.
Villa-García, Julio. 2023. Hanging Topic Left Dislocations as extrasentential constituents: Toward a paratactic account. Evidence from English and Spanish. *The Linguistic Review.* 40(2). 265–310. https://doi.org/10.1515/tlr-2023-2003.
Webelhuth, Gert. 1992. *Principles and parameters of syntactic saturation*. New York, NY, Oxford: Oxford University Press.
Witkoś, Jacek. 2008. Control and predicative adjectives in Polish. In Jacek Witkoś & Gisbert Fanselow (eds.), *Elements of Slavic and Germanic grammars: A comparative view*, 255–277. Frankfurt: Peter Lang.
Wiśniewski, Marek. 1990. Formalnogramatyczny opis leksemów TO. 2. Słowo *to* w funkcji spójnika, partykuły lub czasownika niewłaściwego. *Acta Universitatis Nicolai Copernici. Filologia polska.* XXXI. 91–119.
Zawada, Sebastian. 2024. Demonstrative TO in Polish – an lfg analysis. In Miriam Butt, Jamie Y. Findlay & Ida Toivonen (eds.), *The Proceedings of the LFG'24 Conference*, 264–284. PubliKon. https://doi.org/10.18148/lfg/2024.v29i.52.

Appendix A. Unsettled empirical issue: person and KP_1

This appendix discusses an empirical issue which remains unsettled due to conflicting claims in the extant literature. Namely, the question is whether KP_1 in sentences such as (5), repeated again as (74), can take the form of a first- or second-person pronoun.

(74) Leki to była ostateczność.
medicines.PL.M.NOM TO was.**SG.F** last.resort.**SG.F**.NOM
'Medicines were a last resort.' (CCP)

Many researchers claimed that first- and second-person pronouns cannot occur as KP_1 in predicational constructions with *to* (Wiśniewski 1990, Hentschel 2001, Bogusławski 2002, Bondaruk 2012, 2013, 2015). However, to the extent that KP_1 can be a first- or second-person pronoun at all, the copula certainly does not agree with the right-hand KP_2 predicate, as it does in (74), but with KP_1, see (75).

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

(75) Ja to {?jestem / *jest} dobry lekarz.
I.SG.NOM TO am.1SG / is.3SG good doctor.SG.M.NOM
Intended: 'I am a good doctor.'

But if agreement with KP_1 is possible, then this provides an argument against the analysis in the main text, on which KP_1 is base-generated above the agreement domain TP.

In brief, this appendix argues that examples such as (75), with first- or second-person KP_1, are NOT instances of the non-canonical copular construction discussed in Section 4, so they do not invalidate the analysis offered in the main text.

In making this argument, we follow Bondaruk (2013: 246), where *to* in examples such as (75) is treated as a 'marker of emphasis', also known as 'topic particle' or 'topic marker'; i.e., it is neither to_n nor to_K.[43] Indeed, the form *to* is extremely multifunctional: besides functioning as a demonstrative (and having many other uses not relevant here), it is also used as a particle occurring between the left-hand topic and the remainder of the sentence, see (76a). As such, it can also be used in copular clauses, as in the canonical copular sentence with an instrumental predicate in (76b).[44]

(76) a. Skrzydła to miały rozpiętość co najmniej dwóch metrów!
wings.NOM TOP had.3PL span what least two meters
'As for the wings, they had a span of at least two meters!' (CCP)

b. Ty to jesteś szczęśliwą kobietą…
you.SG.NOM TOP are.2SG happy woman.SG.F.INS
'You are a happy woman.' (CCP)

In consequence, Bondaruk (2013) analyses the marginally acceptable versions of sentences such as (75) as a bare binominative structure (KP_{NOM} – BE – KP_{NOM}), the one we discussed in Section 5.1. On such a view – which we fully adopt here and further motivate below – the more acceptable version of (75), with *jestem* 'am.1SG', exhibits canonical agreement with the TP-internal subject and involves the emphatic marker *to*, in full analogy to (76b).

Importantly, Bondaruk (2019) dissents from such an analysis. She treats sentences such as (75) as representing the same structure as (74). Consequently, she analyses agreement in such sentences as person-sensitive – exhibiting KP_2-agreement UNLESS KP_1 is a first- or second-person pronoun. She does not discuss it in the context of her previous proposal based on the 'topic particle' interpretation of *to*, but proposes attested data instead, presented in (77a)–(77d) (Bondaruk 2019: 113), which she uniformly analyses as involving the 'pronominal copula', i.e., in our terms, the non-canonical construction with KP_1 and to_n.

(77) a. Ja to jestem patriota, nie jestem nacjonalista.
I.**SG.NOM** (?) am.**1SG** patriot.SG.NOM NEG am.1SG nationalist.SG.NOM
'I am a patriot, I am not a nationalist.'

b. My to jesteśmy złodzieje.
we.**PL.NOM** (?) are.**1PL** thief.PL.M.NOM
'We are thieves.'

c. Ty to jesteś szczęściarz.
you.**SG.NOM** (?) are.**2SG** lucky.man.SG.M.NOM
'You are a lucky man.'

d. Wy to jesteście sadyści.
you.**PL.NOM** (?) are.**2PL** sadist.PL.M.NOM
'You are sadists.'

However, we dispute the adequacy of Bondaruk's (2019) account of this data. Let us take a closer look at the status of *to* in these examples – for now glossed as (?), but ultimately argued to be topic markers, TOP. Each sentence in (77) lends itself perfectly to the analysis treating *to* as an emphatic marker, with the underlying bare binominative structure. Recall that, in Section 5.1, we noted that the bare binominative structure requires an expressive predicate.

[43] It is the same topic marker that occurs in (45) in Section 4.1.

[44] Sometimes *to* in such sentences does not directly suggest a contrastive topic interpretation, but signals some kind of pragmatic effect. For example, (76b) can be interpreted as a wish that the speaker were as happy as the hearer, or as a regret that she is not as happy the hearer.

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

Indeed, each sentence in (77) contains one and is also acceptable without *to*.[45] For example, in the case of (77b), *to* is readily removable, and if retained, the predicate can be changed into its verbal analogue *kradniemy* 'steal (1PL)', which is shown below. (78) presents the broader context from which (77b) is excerpted, and (79) demonstrates that it is equally natural with *my to kradniemy* 'we TOP steal (1PL)'. This shows that *to* is not a part of the copular construction (in our terms, it is not to_n), but rather a topic marker.

(78) …powiedzieli o nas, że *my* *to* *jesteśmy* *złodzieje* i
said.3PL about us COMP we.**PL.NOM** TOP are.**1PL** thief.PL.M.NOM and
trzeba się od nas trzymać z daleka!
necessary REFL from us keep of distance
'They said of us that we are thieves and one should stay away from us!' (NCP)

(79) …powiedzieli o nas, że *my* *to* *kradniemy*…
said.3PL about us COMP we.**PL.NOM** TOP steal.**1PL**
'They said of us that we steal…'

Regarding (77c), the original attested version does not end with a period, but with an exclamation mark. It is clearly an expressive sentence, like *Ty to masz szczęście!* 'You TOP have luck'. In the case of (77d), it is quite unnatural without *jesteście* (??*Wy to sadyści*), which is surprising given that the verbal part can be freely omitted in sentences with to_n in the present tense. By contrast, it is fine without *to* (*Wy jesteście sadyści*), which again shows that *to* in this sentence should be analysed as the topic particle. In summary, the data provided in (77) do not instantiate the non-canonical copular construction investigated in this paper.

In order to conclusively exclude the possibility of an agreeing first- or second-person KP_1 in this construction, it is necessary to examine constructions in which *to* is immune to the confusing topic marker interpretation. To this end, we employ structures with corrective focus stress on the left-hand NP, which is incompatible with the topical *to* (TOP) – the latter marks its left-hand neighbour as topical, which consequently cannot be focused, see (80).

(80) a. A: Janek to już tu nie pracuje, prawda?
Janek TOP already here NEG work.3SG right
'Janek doesn't work here anymore, right?'

b. B: #Nie, MARYSIA to już tu nie pracuje.
no Marysia TOP already here NEG work.3SG
Intended: 'No, it is Marysia who no longer works here.'

In contrast, KP_1 in non-canonical copular constructions with to_n can be focused, as in (81), which shows that to_n is a different element from the topic marker *to* in (80).

(81) a. A: Marysia to jest nowy przewodniczący, prawda?
Marysia TO is.3SG new chairman.SG.M.NOM right
'Marysia is the new chairman, right?'

b. B: Nie, MAREK to jest nowy przewodniczący.
No Marek TO is.3SG new chairman.SG.M.NOM
'No, Marek is the new chairman.'

Crucially, first- or second-person pronouns placed in structures such as (81) are strongly unnatural, showing that the apparent copula in sentences such as (77) is in fact the topic marker, as suggested above; see (82). The infelicity of B's answer has nothing to do with focusing the first-person pronoun itself, as demonstrated by the perfectly natural example with a canonical structure involving an instrumental predicate in B′'s utterance.

[45] Moreover, sentence (77a) is most likely misinterpreted. An examination of the context in the corpus reveals that the author of this sentence did not use Polish diacritics. It is only the diacritic on the letter *a* (*ą*, representing the nasal [õ] sound) that distinguishes the nominative *patriota*/*nacjonalista* from the instrumental *patriotą*/*nacjonalistą*. With the instrumental forms *patriotą* and *nacjonalistą*, (77a) sounds even more natural in the given context. Hence, (77a) is most probably an example of the canonical copular construction with an instrumental KP predicate.

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press

(82) A: Marysia to jest nowy przewodniczący, prawda?
Marysia TO is.3SG new chairman.SG.M.NOM right
'Marysia is the new chairman, right?'

B: #Nie, JA to jestem nowy przewodniczący.
no I TO am.1SG new chairman.SG.M.NOM

B′: Nie, JA jestem nowym przewodniczącym.
no I am.1SG new chairman.SG.M.INS
'No, I am the new chairman.'

This reasoning shows that, in predicational sentences, first- and second-person pronouns indeed cannot occur as a left-dislocated KP_1 next to to_n. Once the topic interpretation becomes impossible, such sentences are much degraded. This supports the claim that the alleged cases of such first- or second-person KP_1 occurring with *to* and controlling agreement actually represent the bare binominative structure with the topic marker *to*, in line with Bondaruk (2013).

To conclude, this appendix shows that non-canonical copular sentences with to_n do not allow for first- or second-person pronouns as KP_1. Sentences in (77), which at first seem to contradict this claim, are examples of the binominative structure in which KP_1 is generated low and controls agreement. Hence, such examples are not problematic for the analysis in the main text.

Cite this article: Zawada, Sebastian and Przepiórkowski, Adam. 2026. "Non-canonical agreement in predicational copular clauses: Evidence from Polish." *Journal of Linguistics*. doi:10.1017/S0022226726101388

https://doi.org/10.1017/S0022226726101388 Published online by Cambridge University Press