

## Categorial mismatches of pronouns – a diachronic perspective

KRZYSZTOF MIGDALSKI<sup>1</sup> AND HAKYUNG JUNG<sup>2</sup>

<sup>1</sup>*University of Wrocław*

<sup>2</sup>*Seoul National University*

### ABSTRACT

This paper investigates categorial mismatches in Slavic pronouns, focusing on cases where the morphological form of pronominal clitics diverges from their syntactic or prosodic interpretation. We argue that such mismatches are empirically necessary to account for both synchronic variation and diachronic change. The analysis shows that mismatches between PF (phonological form) and LF (logical form) interpretations may give rise to structural ambiguity—one of the key preconditions for language change, as noted by Lightfoot (1979, 1991). Building on accounts of PF/LF mismatches in pronominals proposed by Despić (2011, 2014), Stegovec (2019), and Puškar-Gallien (2022), we argue that diachronic data from Old Russian and synchronic variation in Polish and Macedonian provide strong support for the assumption of categorial mismatches as the most explanatory framework for the observed changes and distributional patterns.

**KEYWORDS** pronouns · clitics · diachronic linguistics · categorial mismatch

### 1 INTRODUCTION

This paper provides an analysis of mismatched pronominal elements in Slavic, that is cases in which their morphological make-up does not correspond to their syntactic interpretation. It contributes to the existing accounts by showing that the categorial mismatches are a factor triggering pronoun reanalysis in the process of language change. For this reason, this paper argues that the assumption of categorial mismatches is an empirical necessity. This paper is organized as follows. §2 provides a motivation for the assumption of categorial mismatches of pronouns, which largely stems from Cardinaletti & Starke's 1999 analysis of structural deficiencies. §3 shows the way pronominal mismatches have been accounted for in Slavic, focusing on analyses developed by Despić (2011), Despić (2014), Stegovec (2019), and Puškar-Gallien (2022), which are contrasted with a recent account by Milosavljević (2023), who argues against the assumption of categorial mismatches. §4 addresses diachronic data from Old Russian as well as discusses synchronic speaker and dialectal variation in the distribution of pronouns in Polish and Macedonian, arguing that the assumption of pronominal mismatches is a necessity account for the presented empirical facts and to explain the process of language change.

### 2 CATEGORIAL LF/PF PRONOUN MISMATCHES

It has been observed in the literature that pronominal elements display systematic variation in their syntactic behavior and possible semantic interpretation depending on their morphophonological strength, which corresponds to their strong pronoun versus clitic status. The first extensive comparative study dates back to Cardinaletti & Starke (1999), though further substantial observations have also been made by Franks (2013), Despić (2011), Despić (2014), and Bošković (2018) on the basis of Slavic data. For example, the Serbian data in (1-a) and (2-a) indicate that only strong pronouns, and not clitics, can be coordinated, whereas (3-a) that clitics are incompatible as complements of pronouns.

Example (4) illustrates a semantic distinction in the distribution, showing that only clitics can be coreferential with an antecedent interpreted as a discourse topic.

(1) a. \*je i ga [Coordination]  
her.CL.ACC and him.CL.ACC

b. nju i njega (BCMS, Milićev 2008: 449-450)  
her.STRONG.ACC and him.STRONG.ACC

(2) a. \*ga i Roka [Coordination]  
him.CL.ACC and Rok.ACC

b. njega i Roka (BCMS, Milićev 2008: 449-450)  
her.STRONG.ACC and Rok.ACC

(3) a. \*na ga [Complement of P]  
on him.CL.ACC

b. na njega (BCMS, Milićev 2008: 449)  
on him.STRONG.ACC

(4) a. What about Kusturica's latest movie? I know directors usually like their most recent movies [The antecedent is a discourse topic]

b. Kusturicini<sub>i</sub> najnoviji film ( $\checkmark g_a_i$ ) je razočarao  
Kusturica's latest movie him.CL.ACC AUX disappoint.PART.M.SG  
(\*njega<sub>i</sub>) him.STRONG.ACC (BCMS)  
'Kusturica's latest movie disappointed him' (see Jovović 2022)

Table 1 below specifies typical morphosyntactic and semantic properties of strong pronouns vs. clitics/weak pronouns, as attested for Slavic and other languages.

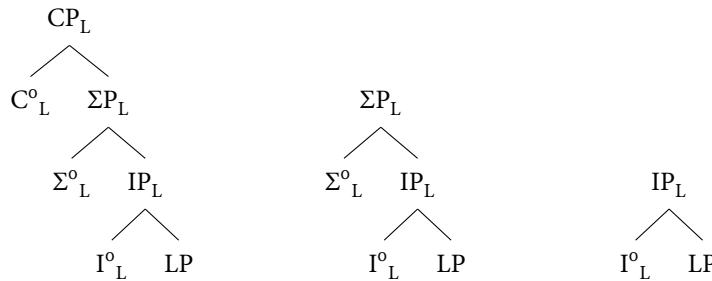
Table 1: Contrasts between strong and clitic pronouns (adapted from Cardinaletti & Starke 1999)

Characteristic	Strong Pronouns	Clitic Pronouns
Non-human reference	✗	✓
Occurrence in coordination	✓	✗
Expletives	✗	✓
Its antecedent must be discourse-topical	✗	✓
Sloppy readings possible	✗	✓
Can act as bound variables	✗	✓

Cardinaletti & Starke (1999) attribute the variation in the behavior of the pronominal elements to the degree of their structural deficiency. As presented in (5), strong pronouns, weak pronouns, and clitics have different syntactic structures, with strong pronouns projecting the highest number of layers, which correspond to their morphological complexity, and which include CP, which encodes referential features, such as range, and humanness, as well as ΣP, which hosts focus, polarity, and is in general taken to be the locus of prosody-related features of a pronominal element. The structure of the weaker elements is poorer: weak pronouns are “peeled” strong pronouns, while clitics are “peeled” weak pronouns.

(5)

(i) Strong pronouns (with L = any lexical category)      (ii) Weak pronouns      (iii) Clitics



The realization of an actual structure is governed by the “Minimise Structure” principle: given a choice, the most deficient (=clitic) form must be realized (Cardinaletti & Starke 1999: 197). This principle is related to economy of representations, and it captures the insights of other economy principles, such as the “Avoid (lexical) Pronoun” filter (Chomsky 1981), which enforces the choice of null subject pronouns instead of their overt counterparts, as well as Picallo’s 1994 “Avoid Features” principle (“avoid referential features up to identification”). This principle ensures that a form with larger structure (e.g., a strong pronoun) can be selected only if the realization of the most deficient structure is ruled out for independent reasons, such as prosodic restrictions, as in coordination or in Left Dislocation (Cardinaletti & Starke 1999: 201). The impossibility of realizing the most deficient structure may give rise to pronoun mismatches at the interfaces: a pronoun may surface at PF as a regular strong pronoun, but it may display interpretation typical of clitics at LF, as illustrated in the subsequent section.

### 3 CATEGORIAL LF/PF PRONOUN MISMATCHES IN SOUTH SLAVIC

This section overviews instances of South Slavic pronouns whose categorial status as a clitic or a weak pronoun does not correspond to its expected syntactic distribution or interpretation. One way to account for these occurrences is to assume LF/PF mismatches. This section summarizes the main observations made in the literature, for a more comprehensive recent overview (see Milosavljević 2023).

The Slovenian example in (6) due to (Stegovec 2019: ex. 8-9) shows that strong pronouns (such as *njim* in (6)) which are complements of prepositions that assign lexical case may have the sloppy interpretation and refer to inanimate objects. These properties are typical of clitics, so it is unexpected to observe them for strong pronouns. Stegovec (2019) accounts for the mismatch by arguing the pronouns surface as strong forms at PF, given that only strong forms may be complements of prepositions, but they are interpreted as clitics at LF.

(6) Pero se vrača k svojim koreninam in Maja se tudi vrača k  
 Pero REFL returns to POSS.REFL roots and Maja REFL too returns to  
**njim.**  
 them.STRONG.DAT  
 ‘Pero<sub>i</sub> is going back to his<sub>i</sub> root and Maja<sub>k</sub> is going back to [?his<sub>i</sub> roots/her<sub>k</sub> roots]  
 too.’ (Slovenian, Stegovec 2019: ex. 8-9)

Despić (2011), Despić (2014) analyzes other cases of mismatches in Serbo-Croatian, where strong pronouns such as *nju* may unexpectedly refer to non-human objects when followed by the intensifier *sam*, thus they also display a property that is otherwise restricted to clitics (see (7)).

(7) Malo ko obilazi muzeje oko gradske crkve<sub>i</sub>. **Nju<sub>i</sub>** \*(samu),  
 few who visits museums around city church her.STRONG.DAT self  
 opet, dnevno poseti oko 50 turista.  
 again daily visits around 50 tourists  
 'A few people visit museums around the city church. (As for the church itself,) an  
 average of 50 tourists visits it a day.' (BCMS, Despić 2011: 247)

Moreover, Despić (2011, 2014) observes that although only clitics can act as bound variables (Montalbetti 1984; see (8-a)), this restriction is also lifted for strong, focused forms when they are preceded by *samo*, as in (8-b).

(8) a. Svaki predsedniki misli da *ga<sub>i</sub>* /??njega<sub>i</sub> svi vole.  
 every president thinks that him.CL him.STRONG all love.  
 'Every president thinks that everyone loves him.'  
 b. Svaki predsedniki misli da samo **njega<sub>i</sub>** svi vole.  
 every president thinks that only him.STRONG all love  
 'Every president thinks that only he himself is loved by everyone.'  
 (BCMS, Despić 2011: 247; 2014: 67)

Despić argues that in these contexts the focused full pronouns are underlyingly clitics in syntax, but they surface at PF as strong forms to satisfy the prosodic requirements of focus imposed by *samo*, which preclude the realization of prosodically deficient elements.

Puškar-Gallien (2022) develops similar argumentation to Stegovec's 2019 for related cases of strong pronouns that complement prepositions in Serbo-Croatian and assumes PF/LF mismatch in the realization of pronouns, with strong forms realized instead of clitics as last resort due to PF constraints. She shows that the sloppy reading of strong pronouns when they are used as complements of prepositions is also available in Serbo-Croatian.

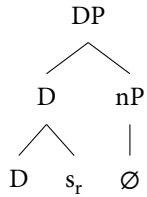
(9) Dok vozi, Ljubica uglavnom koristi svoj telefon<sub>i</sub> za navigaciju a  
 while drives Ljubica mostly uses POSS.REFL phone for navigation but  
 Tamara se dobro snalazi i bez njega<sub>i</sub>.  
 Tamara REFL good manages and without him.STRONG.GEN  
 'While driving, Ljubica mostly uses her phone for navigating and Tamara manages  
 well without Ljubica's phone / Tamara's phone.' (BCMS, Puškar-Gallien 2022)

Puškar-Gallien (2022) assumes that clitics have a different structure than strong pronouns as they lack the nP, which is the locus of humanness/animacy features; therefore, the interpretation of clitics is not restricted with respect to humanness. In line with the "Minimise Structure" principle, she takes the spell-out of a strong pronoun to be a last-resort strategy, which then leads to the realization of nP. In such a scenario, pronouns display inanimate and sloppy readings. Thus, they adopt the form of a strong pronoun, but function as a clitic.

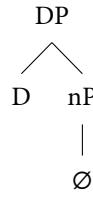
All the analyses presented so far assume that the morphophonological make-up of a pronoun does not necessarily correspond to its anticipated LF interpretation or syntactic behavior. Milosavljević (2023) develops an alternative account, arguing that there is no need for assuming LF/PF mismatches, so in his view morphologically full pronouns are always syntactically stronger than clitics and exhibit the distribution of strong forms. Furthermore, he postulates that the basic difference between pronouns relates to their semantic representation, as strong pronouns, unlike clitics, have a resource situation pronoun (*s<sub>r</sub>*) as an argument of D.

(10)

(i) Strong pronouns



(ii) Clitics



The structure corresponding to the one in (10) is used by Bürig (2004) and Schwarz (2009) for definite descriptions, Arsenijević (2018) for demonstratives, and by Patel-Grosz & Grosz (2017) for strong pronouns. D is assumed to be the locus of *phi*-features. Milosavljević (2023) argues that in combination with the “Minimize Structure” principle, it captures the distribution of strong pronouns and clitics in all environments. He also posits that his analysis is less costly and less demanding in terms of language acquisition and processing than the “mismatch” analyses and that it does not rely on “undesired” theoretical devices. However, given that it is quite common for languages to display mismatches, for example with respect to the position of a syntactic constituent after spell-out and its interpretation (for instance, in *wh*-movement and negative raising), the claim about the simplicity and economy of his alternative analysis, which does not postulate PF/LF mismatches, is debatable.

Example (11) below is included to present the workings of Milosavljević’s 2023 analysis; it concerns focus and bound-variable readings of strong pronouns. Milosavljević argues that strong pronouns differ from clitics at LF universally, including bound-variable environments, so “bound-variable” strong pronouns, unlike clitics, are not genuine bound variables. Strong pronouns and clitics are syntactically licensed in different ways. “Bound-variable” strong pronouns can be coordinated, which is a property of strong pronouns. Thus, it is possible to coordinate “bound variable” pronouns, as in (11), in which case only the strong form can be used, which for Milosavljević indicates that they are syntactically strong.

(11) Svaki predsedniki misli da **njega<sub>i</sub>** i **njegovu<sub>i</sub>** porodicu svi vole.  
 every president thinks that him.STRONG and his family all love  
 ‘Every president thinks that he and his family are loved by everyone.’

(BCMS, Milosavljević 2023)

Regardless of the accuracy of the semantic analysis adopted by Milosavljević, which we leave as an open question here, a problematic aspect of his account is the reduction of the difference between clitics and other types of pronouns to a semantic contrast. It is not clear how his analysis can capture prosodic deficiency, which is certainly the most basic contrast and possibly the most prominent one in language acquisition. Milosavljević reduces prosodic deficiency to focus marking, pointing out that focus has a role in licensing strong pronouns at LF, and it has syntactic and semantic effects. By contrast, for Despić (2011), Despić (2014) and Puškar-Gallien (2022), the role of focus in licensing strong forms is phonological, and strong pronouns must satisfy the prosodic requirements of focus at PF. Due to the reduction of the prosodic requirement of clitics to focus in Milosavljević’s analysis, it is not clear how he can capture the basic, regular and unchangeable prosodic requirements of clitics that apply by default, regardless of focus contexts.

There are two additional problems with Milosavljević’s 2023 proposal, which were pointed out to us by an anonymous reviewer. Namely, as the anonymous reviewer insightfully observes, in Milosavljević analysis, clitics and strong pronouns are categorially identical, and they are both DPs. The D<sup>0</sup> of the strong pronoun DP takes a resource situation as an argument, as suggested for forms such as *njega*, where the *nj-* segment reflects

this semantic dependency. Since on his account resource situations may contain both human and non-human individuals, as well as properties, his analysis in principle allows strong pronouns to refer to non-human entities. To account for the empirical preference for [+human] interpretations of strong pronouns, Milosavljević (2023: 3) introduces the Harmony Scale, which integrates two independently motivated hierarchies:

(12) Pronouns: Strong > Deficient  
 Animacy: Human > Non-human > Inanimate

Under this account, the strong human pronoun occupies the most harmonic position, consistent with the broader linguistic observation that animacy effects tend to be gradual rather than categorical (see Corbett 2006, Arsenijević & Mitić 2016, Aleksić 2019, Arsenijević 2021). However, as the anonymous reviewer observes, Milosavljević leaves unexplained why the Harmony Scale does not apply in specific syntactic configurations, such as when the strong pronoun appears in focus position or as the complement of a preposition. In fact, one might as well expect the opposite distribution under the Harmony Scale analysis: a stronger [+human] preference in focus or PP positions, with no restriction otherwise. Conversely, the categorial mismatch analysis adopted in this paper predicts that the attested facts are the only ones that are possible, whereas the opposite (unattested) situation is not even statable on the categorical mismatch analysis. In our approach, clitics and strong pronouns are categorially different, and strong pronouns appear in focus and as complements of prepositions as a last resort due to independent PF constraints that prevent the realization of clitics in these contexts. Because the underlying interpretation of clitics carries no inherent animacy preference, the insertion of strong pronouns in these syntactic environments preserves the clitic interpretative neutrality. Consequently, the categorial mismatch approach not only captures the attested distributional facts but also excludes unattested patterns, offering a more constrained and predictive analysis.

Furthermore, the anonymous reviewer points out to us that Milosavljević's analysis does not straightforwardly explain why the Person Case Constraint (PCC) affects clitics but not strong pronouns. If, as he proposes, both clitics and strong pronouns are syntactically DPs, then the observed contrast with respect to PCC effects cannot be accounted for via a categorial or syntactic distinction. Instead, the explanation must presumably rest on the resource situation argument of strong pronouns. According to Milosavljević (2023: 20), given that the “different uses of strong pronouns all arise in different syntactic environments, it is plausible to assume that what differentiates them is external rather than internal syntax, i.e. that strong pronouns are uniform syntactically, but their different flavors depend on how the resource situation gets its value.” This assumption may imply that the fact that strong pronouns are not affected by the PCC should arise from the way “the resource situation gets its value” through their resource situation argument. Still, as the anonymous reviewer observes, it is unclear how to extend the resource argument analysis to 1st and 2nd person, as in *me* vs. *mene*. Presumably, *-ne* is the resource situation part, which should somehow provide the coreference possibilities for *mene*. However, it has been shown in the literature (see, for example, Stegovec 2020) that PCC effects are too idiosyncratic to be reduced to a simple pragmatic principle of the speech act; moreover, PCC effects are attested in different variants, such as strong, weak, *me*-first etc., and are clearly a morpho-syntactic phenomenon. Within the categorial mismatch analysis pursued in this paper, the PCC effects can be reduced to a category distinction and, for example, limit its applicability to only heads, on the assumption that clitics are head-like elements. It is not clear how the PCC effects could be independently motivated in Milosavljević's analysis.

Moreover, as will be shown in the subsequent sections, it is also not clear how Milosavljević's analysis can capture PF contrasts in the occurrence of morphologically identical pronominal forms, which are interpreted as clitics or weak pronouns, subject to speaker variation and/or the advancement of language change. Specifically, §4.1 below addresses

the ambiguous contexts of clitic/weak pronouns in Old Russian; §4.2 overviews PF-related speaker variation in the interpretation of weak pronouns in Polish section, while §4.3 discusses Macedonian, in which the distribution of morphologically identical pronouns depends on the type of clitic host. All these contexts are manifestations of language change, which as has been argued in the literature, may be triggered by morphophonological mismatches.

#### 4 CATEGORIAL MISMATCHES IN LANGUAGE CHANGE

A common assumption, which dates back to Lightfoot (1979, 1991), is that language change may occur during first language acquisition, when due to independent morphological or phonological developments, a structure or a category becomes opaque, and in consequence ambiguous to the language learner. Such a situation may result in language change, which occurs when a learner interprets a linguistic phenomenon in a different way than the parents and other speakers during acquisition and in consequence sets a parameter differently than other language users. According to Lightfoot (1979), the grammaticalization of modal verbs in English, which involved paradigm leveling and the reinterpretation of modal verbs as elements generated in  $T^0$ , arose in such a circumstance. It occurred subsequent to the loss of person/number morphology on the verb in English, which affected all forms except for the 3rd person singular, which retained the  $-s$  morpheme. However, the preterite-present class of verbs, which included modal verbs, did not ever have any person/number marking on the 3rd person singular form, so due to the change, verbs from this class started to stand out as a unique class of morphologically bare verbs in all finite forms. In such a scenario, they became reinterpreted as a separate class of verbs merged under  $T^0$  and started to display special syntactic distribution. Of course, the details of the ways a modified structure can affect language change during language acquisition have been debated since Lightfoot's 1979 original proposal (see, for example, Westergaard 2021 for a recent overview), but what matters for the analysis presented in this paper is a context of opacity or ambiguity that leads to a reanalysis of a given structure and subsequent language change.

##### 4.1 OLD RUSSIAN

We argue in this paper that in Slavic a corresponding case of ambiguous language input concerns structures with prepositions that are complemented by clitics (see also Jung & Migdalski 2022 for an earlier variant of the analysis). In modern Slavic languages, only strong pronouns may act as complements of prepositions; however, Old Russian and Old Polish data include instances of pronominal clitics that follow a preposition. Such cases can be interpreted by the language learner in two ways: (i) prepositions in Old Slavic are proclitics and act as hosts for pronominal enclitics, producing a prosodically independent unit; or (ii) since prepositions cannot be followed by clitics, the pronominal complement of the preposition can be analyzed as a tonic pronoun although it has the morphological form of a clitic. Scenario (ii) led to the reanalysis of clitics as weak pronouns in Old Russian and Old Polish.

(13) *za tę golovy svoi sъkladyvaèmь.*  
for you.ACC head.ACC.PL own.ACC.PL lay-down.1.PL  
'We bow down to you.'

(Old Russian, *Hypatian Chronicle* 1177; Zaliznjak 2008: 36)

(14) *Sam, prawi, przez mię przysiągł jeśm.*  
he say.3.SG.AOR without me swear.PART.M.SG am.AUX  
'He said that he has sworn without me...'

(Old Polish, *Sermon III, On St. Michael's Day*, Migdalski 2016: 303)

Jung & Migdalski (2022) draw the assumption about the reanalysis of the clitic as a weak

pronoun comes on the following observations. Old Russian in the 11th -15th centuries displayed second position cliticization, with pronominal clitics occurring in the dative-accusative order. The occurrence of accusative clitics as complements of prepositions in the earliest texts indicates that they were weak pronouns. Zaliznjak (2008: 36) claims that the combinations of prepositions and pronominal clitics are residues from an earlier stage, in which the pronominal clitics held prosodic independence. However, Jung & Migdalski (2022) point out that a hypothesis that *za tęACC* was an intermediate stage of the change from clitic to pronoun is just as possible as the opposite direction, as illustrated in (15), and that clitics appearing in the PP can be the first hallmark of the (re)strengthening of their prosodic independence.

(15) P + weak pronoun  $\leftrightarrow$  *za tęACC*  $\leftrightarrow$  P + clitic

More evidence for the weak pronoun status of the clitics comes from the fact that pronominal clitics sometimes appear in non-second positions. In (16) the accusative reflexive clitic *se* occupies the initial position, given that the conjunction *a* cannot function as a host.

(16) a *se* ego zapritü  
and REFL.ACC him.ACC shut.3.SG  
'And ... locks himself up'  
(*Gramoty Velikogo Novgoroda i Pskova*, No. 28, 1190-, Jung & Migdalski 2022)

Moreover, the regular ordering of clitics (dative-accusative-auxiliary) could be violated from 13th c. Russian onwards, with the accusative form following the auxiliary. This is a typical property of weak pronouns in Modern Polish (see §4.2 below).

(17) i jela jesmo *se* jemu po ruku  
and take.PART.F.SG am.AUX REFL.ACC him.DAT for hand  
'I promised him...' (BBL No. 731, early 13th c., Zaliznjak 2004: 392)

Janin & Zaliznjak (1993: 289) observe that proclitics could function as hosts for enclitics (for example, *ne li jesi dalí*). Prepositions are proclitics, and the combination of a preposition and a weak pronoun/clitic constitutes a prosodically independent phonetic unit. It is difficult to determine whether the morphologically reduced accusative form in the string *za tęACC* in (15) was prosodically dependent or not, but it is clear that the accusative form occupied an argument position as the object of the preposition. This ambiguity in internal prosodic structure, in Lightfoot's (1979, 1991) sense, provides an adequate condition for the reanalysis of the reduced form as a prosodically independent element.

#### 4.2 POLISH

Migdalski (2016: 302–305) observes that in the oldest Polish text from the late 13th/early 14th, *Holy Cross Sermons* (*Kazania Świętokrzyskie*), pronominal clitics occur in a fixed position after the clause-initial element. This element can be a verb (see (18-a), so they can be either verb-adjacent or second position clitics, but their occurrences are rather infrequent, so it is difficult to determine their cliticization status.

(18) a. a togodla *ji* we złe chustki ogarnęła.  
and therefore him.ACC in bad cloth wrap.PART.F.SG  
'and therefore she wrapped him in bad cloth'  
(Sermon III, *On St. Michael's Day*)

b. Naleźli *ji*, prawi, pieluszkami ogarnienego a w jasłkach  
 find.PART.M.SG him.ACC true nappies.INST wrapped and in cribs  
 położonego.  
 laid  
 'They found him wrapped in nappies and laid in cribs.'  
 (Sermon III, *On St. Michael's Day*)

c. owa *ji* pirzwej widział Habraham  
 so him.ACC first see.PART.M.SG Abraham  
 'So he was first seen by Abraham' (Migdalski 2016: 303)

In subsequent texts, such as *Queen Sophia's Bible/Sárospatak Bible (Biblia Królowej Zofii)*, pronominal elements are located in different positions in the structure, which may indicate that they became strengthened and are increasingly mobile syntactically. Moreover, as shown in (19-a), the clitic forms, such as *ji*, occur together with strong forms, such as *jego* and *jemu* in the same pragmatic and semantic contexts, which suggests that speakers tend to treat both types in a uniform way prosodically in spite of the morphological divergence.

(19) a. Tegdy wziął Pan Bog człowieka i postawił *ji* w  
 then take.PART.M.GL Lord God man.ACC and put.PART.M.SG him.ACC in  
 raju rozkoszy, aby działał a ostrzegał *jego*.  
 paradise bliss.GEN so-that work.PART.M.SG and protect.PART.M.SG him.ACC  
 I przykazał *jemu...*  
 and command.PART.M.SG him.DAT  
 'The Lord God took the man and put him in the Garden of Eden to work it  
 and take care of it. And the Lord God commanded him...' (Genesis, 2:15-16)

b. uczyńmy *jemu* wspomożenie podobne k *niemu*  
 make him.DAT helper similar to him.DAT  
 '(Let us) make a helper suitable for him' (Genesis, 2:18, Migdalski 2016: 304)

Furthermore, the clitic variants do not need to appear in designated syntactic positions, and they have largely the same distribution as other pronouns. As shown in (20-a), they were not required to be verb-adjacent or in second position.

(20) a. I przywiódł *je* przed Adama, aby *je* opatrzył  
 and bring.PART.M.SG them.ACC before Adam so-that them.ACC see.PART.M.SG  
 a Jimiona *jim* dał  
 and names them.DAT give.PART.M.SG  
 'He brought them to the man to see what he would name them' (Genesis, 2:19)

b. Nazwał *jest* Adam Jimiona *jich* wszelkiemu stworzeniu  
 name.PART.M.SG is.AUX Adam names them.GEN all beings  
 zwierzętemu  
 animal  
 'Adam gave names to all the livestock.' (Genesis, 2:20, Migdalski 2016: 305)

These data seem to indicate that already in Old Polish pronominal clitics started to be analyzed by speakers as weak pronouns, the way they are interpreted in Modern Polish (see Witkoś 1998, Cetnarowska 2003, Migdalski 2006, 2016, Witkoś & Łęska-Bayraktar 2024). Currently, they undergo the process of phonological strengthening. Unlike pronominal clitics in South and West Slavic languages, they are syntactically mobile and may occur in many positions in the clause, although they tend to avoid clause-initial placement (as in (21-a), though see the discussion below). Another difference between

pronominal clitics in South and West Slavic languages and weak pronouns in Polish is that they can also appear in different orders with respect to each other (see (21-b)) and can be separated from each other (see (21-c)). To our knowledge, these types of data were first discussed at length by Rappaport (1988).

(21) a. (\**go*) często (*go*) spotykam (*go*) na ulicy (Pl)  
 him.ACC often him.ACC meet.PRES.1.SG him.ACC on street  
 'I often meet him in the street.'  
 b. Jan chce mu (*go*)/(*go*) mu wypożyczyć  
 Jan want.3.SG him.DAT it.ACC it.ACC him.DAT lend.INF  
 'Jan wants to lend it to him.'  
 c. Jan *mu* chce *go* wypożyczyć a nie sprzedać  
 Jan him.DAT want.3.SG it.ACC lend.INF and not sell.INF  
 'Jan wants to lend it to him rather than sell it.'

As far as the syntactic placement is concerned, the accusative form *go* is excluded from the clause-initial position, but as has been observed in the literature (for example in a corpus study performed by Włodarczyk (2018)), the weak dative form *mi* is increasingly attested clause-initially in texts of various degrees of formality (see (22)). The weak form may even occur in focus or topicalization contexts (see (23-b)) and in co-ordination (see (24-a)).

(22) *Mi* się wydaje, że...  
 me.DAT REFL seems that  
 'It seems to me that...' [attested in parliamentarian speech].  
 (Pl, Włodarczyk 2018: 66)

(23) a. Ja zawsze pięam mocną herbatę z cukrem ale nie  
 I always drink.PART.F.SG strong.ACC tea.ACC with sugar but NEG  
 zawsze *mi* pomagało  
 always me.DAT help.PART.N.SG  
 'I always used to drink strong tea with sugar, but it wasn't always helpful.'  
 b. A *mi* przeważnie pomaga herbata z miodem  
 but me.DAT usually helps tea with honey  
 'And for me tea with honey is usually helpful.' (Pl, Włodarczyk 2018: 65)

(24) a. To miejsce uratowało *mi* i znajomemu wyjazd.  
 this place save.PART.N.SG me.DAT and fellow trip  
 'This place saved the trip for my friend and me.' (Pl, Internet source)  
 b. Ona dała *mi* i jemu/\**mu* pomarańczę  
 she give.PART.F.SG me.DAT and him.STRONG.DAT/him.DAT orange  
 'She gave an orange to him and me.' (Pl, own judgment)

Włodarczyk (2018:75–76) attributes the possibility of clause-initial *mi* placement to "paradigm leveling" that affected pronouns in Polish diachronically and resulted in an unexpected syncretism between dative and accusative forms. Otherwise, case syncretism is attested only between some nominative and accusative as well as some instrumental and locative forms.

(25) Paradigm leveling of pronouns in Polish (OCS, Bartula 1981: 143; Pl, Franks & King 2000: 150)

	Dative		Accusative	
	Strong	Clitic	Strong	Clitic
OCS	<i>mъně</i> ( <i>mъně</i> )	<i>mi</i>	<i>mene</i>	<i>me</i>
Polish	<i>mnie</i>	<i>mi</i>	<i>mnie</i>	<i>mię</i> (now dated; replaced by <i>mnie</i> )

Due to phonetic changes, the former strong dative *mъně* (*mъně*) and accusative *mene* forms attested in Old Church Slavonic became homophonous in Polish, leading to the unexpected syncretism, whereas the former accusative clitic form *mię* fell out of use. This situation could have led to the reinterpretation of *mi* as a strong dative form by some speakers. In consequence, this development has created a categorial mismatch: morphologically identical variants can be interpreted as clitics or strong forms by various speakers, depending on the speaker's interpretation of the pronominal elements.

#### 4.3 MACEDONIAN

Another case of pronominal mismatch in contemporary Slavic concerns certain contexts of pronominal placement in Macedonian. Pronominal clitics in Macedonian are verb-adjacent, assume proclisis, and may appear in the clause-initial position when their host is a finite verb, such as *raduvame* in (see (26-d)) or the *l*-participle, such as *dale* in (see (27-d)). They may occur low in the structure and do not need to target second position (see (26-d) and (27-d)). They must be verb-adjacent, immediately to the left of their verbal hosts as proclitics, and placement of any intervening material between the verb and the clitics results in ungrammaticality, as in (see (26-d) and (27-d)). Adjacency of a pronominal clitic to an auxiliary is not sufficient, as shown in (27-d). As argued by Migdalski (2025), verb-adjacency is required only with verbs that assign case, such as finite verbs and *l*-participles, thus it is contingent on the case-assigning ability of the host.

(26) a. ( Nie) *si se raduvame na vnučevo*  
           we REFL.DAT REFL.ACC rejoice.PRS.1.SG.DAT to grandson-the.PROX  
           *mnogu.*  
           very-much  
           ‘This grandson of ours is giving us a lot of pleasure.’

b. (Nie) *mnogu si se raduvame na vnučevo.*

c. \**Nie si se mnogu raduvame na vnučevo.*

d. \**Raduvame si se mnogu na vnučevo.* (Mac, Tomić 1999: 10)

(27) a. (Vie) *ste im go dale proektot včera.*  
           you.PL are.AUX him.DAT it.ACC give.PART.PL project-the yesterday  
           ‘As reported, you gave them the project only/already yesterday.’

b. (Vie) *včera ste im go dale proektot.*

c. \**Vie ste im go včera dale proektot.*

d. \**Dale ste im go včera proektot.* (Mac, Tomić 1999: 10)

In structures with predicative nominals, pronominal clitics in Macedonian are excluded clause-initially (see (28-a)). Otherwise, they do not impose any requirements on the categorial status of their preceding host (see Bošković 2001: 255), and some speakers accept them below second position (see (28-e)).

(28) a. \**Mu e tatko (na deteto).*  
           him.DAT is father to child-the  
           ‘[Intended] He is the father (of this child)?’

b. *Tatko mu textite na deteto.*  
           father him.DAT is to child-the  
           ‘He is the father of this child (so he has to take care of him)?’

c. *Toj mu textite tatko.*  
           he him.DAT is father  
           ‘He is his father (and not anyone else)?’ (Mac, Tomić 2000: 295–296)

d. *Petko mi textite tatko.*  
           Petko me.DAT is father  
           ‘Petko is his father.’ (Mac, Franks & King 2000: 83)

e. Petko sekogašmi textite mil.  
 Petko always me.DAT is dear  
 'Petko is always dear to me.' (Mac, Franks & King 2000: 86)

With predicative adjectives, clitics may appear lower than in the second position (see (29-d)). Moreover, they can even occur clause-initially, especially in the presence of more than one clitic, as in (30-a). As was shown earlier, this is increasingly also a typical property of the weak dative pronoun in Polish.

(29) a. Ti *si* ubava žena.  
 you are pretty woman  
 'You are a pretty woman.'  
 b. Ubava žena *si*.  
 c. Ubava *si* žena.  
 d. \*Si ubava žena (Mac, Franks & King 2000: 86-87)

(30) a. ??*Si* *mu* mil.  
 are him.DAT dear  
 'He likes you.'  
 b. Mil *si* *mu*.  
 dear are him.DAT  
 'You are dear to him.'  
 c. Ti *si* *mu* na našion sin mnogu mil.  
 you are him.DAT to our-thehim.PROX son very dear  
 'You are very dear to our son!' (Mac, Tomić 2001: 664)

With predicative passive participles, clitics display a considerable freedom of placement. Some speakers accept them clause-initially, as in (31-a). The clitics do not have to be left-adjacent to the passive participle, such as *rečeno*, (see (31-b)). Moreover, they may occur rather low in the structure, as long as they are located to the left of the participle (see (31-c) and (31-d)).

(31) a. ?*Mu* *e* *rečeno* da bide točen povečke pati.  
 him.DAT is tell.PASS.N to be.SBJV.3.SG punctual more times  
 'He was told to be punctual more than once.' (Mac, Tomić 2000: 296)  
 b. Na Petreta *mu* *e* povečke pati *rečeno* da bide točen.  
 to Peter.DAT him.DAT is more times tell.PASS.N to be.SBJV.3.SG punctual  
 'Peter was told to be punctual more than once.' (Mac, Tomić 2000: 296)  
 c. Na Petreta *mu* *e* od strana na komisijata *mu* *e*  
 to Peter.DAT him.DAT is from side of commission-the him.DAT is  
 povečke pati *mu* *e* *rečeno* da bide točen.  
 more times him.DAT is tell.PASS.N to be.SBJV.3.SG punctual  
 'Peter was more than once told by the commission to be punctual.'  
 (Mac, Tomić 2000: 299)  
 d. Na Petreta *mu* *e* povečke pati *mu* *e* jasno i glasno  
 to Peter.DAT him.DAT is more times him.DAT is clearly and loudly  
*mu* *e* *rečeno* da dojde.  
 him.DAT is tell.PASS.N to come.SBJV.3.SG  
 'Peter was loudly and clearly told to come more than once.'  
 (Mac, Franks & King 2000: 86)

The data involving clitic placement with predicative nominals, adjectives, and passive participles have been analyzed in the literature as instantiating a switch from a complete second position clitic system to a system with verb-adjacent clitic (see Bošković 2001, Tomić 1996, Tomić 1997, 1999), and the non-uniform distribution of the clitics has been attributed to the fact that the change has not been completed. However, the data

discussed above indicate that neither of the systems is consistently at work here, as the clitics do not need to be in second position (except for the examples with predicative nominals), and they also do not need to be adjacent to the passive participle. In fact, Franks & King (2000: 87) show that in structures with the *l*-participle, the clitics may not be separated by any lexical material from the verb (see (32-d)), and they must precede the *l*-participle, as in the grammatical structures presented in (32-d)–(32-d).

(32) a. \*Na Petko *sum mu* poveče pati kažal da dojde  
           to Peter.DAT is.AUX him.DAT more times tell.PART.M.SG to come.SBJV.3.SG  
           ‘Several times, I told Peter to come.’  
   b. Na Petko poveče pati *sum mu* kažal da dojde.  
   c. Poveče pati *sum mu* kažal na Petko da dojde.  
   d. *Sum mu* kažal poveče pati na Petko da dojde

(Mac, Franks & King 2000: 87)

Moreover, in the contexts with passive participles in Bulgarian corresponding to the ones in (31-a), the clitics must be adjacent to the passive participle, and this is the only clitic position that Bulgarian allows (see (33)).

(33) Na Petăr *mu e* kazvano mnogo püti ot strana na komisijata  
           to Peter him.DAT is tell.PASS many times from side of commission-the  
           da büde token.  
           that be.SUBJ.3.SG punctual  
           ‘Peter was more than once told by the commission to be punctual.’ (Bg)

Notably, the acceptability of the structures above is subject to speaker variation. Olga Tomić (p.c.) informs us that clause-initial placement of clitics in the presence of nouns, as in (28-a), is uniformly excluded by all native speakers. As for the other predicative elements, in the Western dialects, clitics may both precede and follow adjectives and passive participles, with the latter distribution being more common in the Eastern dialects. This type of variation indicates that we may observe a process of language change, in which the pronominal forms become strengthened, as is also the case in some other Slavic languages, such as Polish. See also Korubin (1974), Tomić (1997, 2000), Baerman & Billings (1998), and Migdalski (2006, 2025) for more discussion of the diachronic change that seems to be at work here. Importantly, these data indicate that identical pronominal elements may be interpreted in different ways, as pronominal clitics or weak pronouns, depending on the availability of a verbal host that assigns case, and their interpretation is also subject to speaker and dialectal variation. Admittedly, it is difficult to account for these types of data without taking recourse to the assumption of morphosyntactic mismatches.

## 5 CONCLUSION

The data analyzed in this paper demonstrates that categorial mismatches in the syntactic and prosodic interpretations of clitics may have led to instances of language change. Specifically, the change consisted in the strengthening of (former) pronominal clitics in Old Russian, Polish, and clitics that occur with non-verbal predicates in Macedonian. As has been argued for earlier in the literature (Lightfoot 1979, 1991), a condition for language change to take place is a situation of ambiguity. Ambiguity may occur in the case of clitics whose morphological make-up does not correspond to the expected syntactic behavior, as has been shown for Old Russian and Polish. Alternatively, ambiguity may arise in the context of morphologically identical pronominal elements, whose status is contingent on a particular syntactic environment, as in the case of the pronominal forms in Macedonian, which can be interpreted as verb-adjacent clitics or weak pronouns. All these facts provide support for potential PF/LF categorial mismatches of pronouns and in

our view they also indicate that the assumption of categorial mismatches is an empirical necessity.

#### ABBREVIATIONS

1	first person	F	feminine
2	second person	GEN	genitive
3	third person	INF	infinitive
ACC	accusative	M	masculine
AGR	agreement	PART	participle
AOR	aorist	PASS	passive
AUX	auxiliary	PL	plural
BBL	Birch Bark Letters	PROX	proximal
BCMS	Bosnian/Croatian/ Montenegrin/Serbian	PRS	present
CL	clitic	REFL	reflexive
DAT	dative	SBJV	subjunctive
		SG	singular

#### ACKNOWLEDGEMENTS

We would like to thank Wayles Browne, Alberto Frasson, Steven Franks, Miloje Despić, Adrian Stegovec, Neda Todorović, and the FASL audience for discussion and helpful feedback. We also thank Olga Tomić for her comments on the Macedonian data we received earlier. We are grateful to the anonymous reviewer for the very detailed comments which have strengthened the claims made in this paper. All errors remain our responsibility.

#### CONTACT

KRZYSZTOF MIGDALSKI — [krzysztof.migdalski@uwr.edu.pl](mailto:krzysztof.migdalski@uwr.edu.pl)  
HAKYUNG JUNG — [hakyung@snu.ac.kr](mailto:hakyung@snu.ac.kr)

#### REFERENCES

Aleksić, Danilo. 2019. *Kategorija animatnosti i forma akuzativa u srpskom jeziku*: University of Belgrade PhD dissertation.

Arsenijević, Boban. 2018. Atypical demonstratives in an articleless language. In Marco Coniglio, Andrew Murphy, Eva Schlachter & Tonjes Veenstra (eds.), *Atypical demonstratives: Syntax, semantics and pragmatics*, 161–195. Berlin: De Gruyter.

Arsenijević, Boban. 2021. No gender in ‘gender agreement’: On declension classes and gender in Serbo-Croatian. *Balcania et Slavia* 1(1). 11–46. doi: 10.30687/BES/0/2021/01/001.

Arsenijević, Boban & Ivana Mitić. 2016. Effect of animacy and agentivity on the processing of agreement in Serbo-Croatian. In Sabina Halupka-Rešetar & Silvia Martínez-Ferreiro (eds.), *Studies in languages and mind. Selected papers from third Novi Sad workshop on psycholinguistics, neurolinguistic and clinical linguistic research*, 41–77. University of Novi Sad: Faculty of Philosophy.

Baerman, Matthew & Loren Billings. 1998. Macedonian clitics and the trisyllabic stress Window. *University of Trondheim Working Papers in Linguistics* 31. 13–32.

Bartula, Czesław. 1981. *Podstawowe wiadomości z gramatyki staro-cerkiewno-słowiańskiej na tle porównawczym*. Warszawa: Państwowe Wydawnictwo Naukowe.

Bošković, Željko. 2001. *On the nature of the syntax–phonology interface. Cliticization and related phenomena*. Amsterdam: Elsevier.

Bošković, Željko. 2018. On pronouns, clitic doubling, and argument ellipsis: Argument ellipsis as predicate ellipsis. *English Linguistics* 35(1). 1–37.

Bürig, Daniel. 2004. Crossover situations. *Natural Language Semantics* 12(1). 23–62.

Cardinaletti, Anna & Michal Starke. 1999. The typology of structural deficiency: A case study of the three classes of pronouns. In Henk van Riemsdijk (ed.), *Clitics in the languages of Europe*, 185–234. Berlin: Mouton de Gruyter.

Cetnarowska, Bożena. 2003. On pronominal clusters in Polish. In Peter Kosta, Joanna Błaszczałk, Jens Frasek, Ljudmila Geist & Marzena Zygis (eds.), *Investigations into formal Slavic linguistics: Contributions of the fourth European conference on Formal Description of Slavic Languages – FDSL IV*, 13–30. Frankfurt am Main: Peter Lang.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Corbett, Greville G. 2006. *Agreement*. Cambridge: Cambridge University Press.

Despić, Miloje. 2011. *Syntax in the absence of aeterminer phrase*. Storrs, CT: University of Connecticut dissertation.

Despić, Miloje. 2014. Intensifiers, focus, and clitics: Is pronoun position truly an argument for D in SC? In Lilia Schürcks, Anastasia Giannakidou & Urtzi Etxeberria (eds.), *The nominal structure in Slavic and beyond*, 39–74. Berlin/Boston: De Gruyter Mouton.

Franks, Steven. 2013. Orphans, doubling, coordination, and phases: On nominal structure in Slovenian. *Slovene Linguistic Studies* 9. 55–92.

Franks, Steven & Tracy Holloway King. 2000. *A handbook of Slavic clitics*. New York: Oxford University Press.

Janin, Valentin Lavrentievich & Andrej Anatolévič Zaliznjak. 1993. *Novgorodskie gramoty na bereste: iz raskopok 1984–1989 gg.* Moscow: Nauka.

Jovović, Ivana. 2022. Condition B and other conditions on pronominal licensing in Serbo-Croatian. *Linguistic Inquiry* 55(2). 402–421.

Jung, Hakyung & Krzysztof Migdalski. 2022. Toward a four-way pronoun hierarchy: A view from Slavic. *Journal of Slavic Linguistics* 30. 1–11.

Korubin, Blagoja. 1974. Mesto glagola sum kak vspomogatel'nogo i kak svjazki v makedonskom literaturnom jazyke. In N. Švedova et al. (ed.), *Grammatiáskoe opisanie slavjanskix jazykov: koncepcii i metody*, 244–250. Moskva: Nauka.

Lightfoot, David. 1979. *Principles of diachronic syntax*. Cambridge: Cambridge University Press.

Lightfoot, David. 1991. *How to set parameters: Arguments from language change*. Cambridge, MA: MIT Press.

Migdalski, Krzysztof. 2006. *The Syntax of compound tenses in Slavic*. Utrecht: Tilburg University dissertation.

Migdalski, Krzysztof. 2016. *Second position effects in the syntax of Germanic and Slavic languages*. Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.

Migdalski, Krzysztof. 2025. Case contingency of Macedonian cliticization. *Journal of Slavic Linguistics* Forthcoming.

Milićev, Tania. 2008. How strong are full pronouns in Serbian? In Gerhild Zybatow, Luka Szucsich, Uwe Junghanns & Roland Meyer (eds.), *Formal Description of Slavic Languages*, 448–460. Frankfurt am Main: Peter Lang.

Milosavljević, Stefan. 2023. What you see is what you get: On the uniformity of morpho-syntactic and morpho-phonological properties of pronouns. In *GLOW 46 (Workshop Mismatched Pronouns)*, Graz.

Montalbetti, Mario. 1984. *After Binding: On the interpretation of pronouns*: MIT dissertation.

Patel-Grosz, Pritty & Patrick Grosz. 2017. Revisiting pronominal typology. *Linguistic Inquiry* 48(2). 259–297.

Picallo, M. Carme. 1994. Catalan possessive pronouns: The avoid pronoun principle revisited. *Natural Language & Linguistic Theory* 12(2). 259–299.

Puškar-Gallien, Zorica. 2022. A theory of (pro)nominal structure and its consequences on resolving morphosemantic mismatches. In *Formal Description of Slavic Languages 15*, Humboldt Universität Zu Berlin.

Rappaport, Gilbert. 1988. On the relationship between prosodic and syntactic properties of pronouns in the Slavic languages. In Alexander M. Schenker (ed.), *American contributions to the tenth international congress of Slavists*, 301–327. Bloomington: Slavica.

Schwarz, Florian. 2009. *Two types of definites in natural language*: University of Massachusetts, Amherst dissertation.

Stegovec, Adrian. 2019. Crop to fit: Pronoun size and its relation to strict/sloppy identity and animacy. In *LSA 2019*, New York City.

Stegovec, Adrian. 2020. Taking case out of the person-case constraint. *Natural Language & Linguistic Theory* 38(1). 261–311. doi:10.1007/s11049-019-09443-0.

Tomić, Olga. 1997. Non-First as a default clitic position. *Journal of Slavic Linguistics* 5. 301–323.

Tomić, Olga. 1999. Cliticization and clitichood. In István Kenesei (ed.), *Crossing boundaries: Advances in the theory of Central and Eastern European languages*, 9–32. Amsterdam: Benjamins.

Tomić, Olga. 2000. On clitic cites. In Frits Beukema & Marcel Den Dikken (eds.), *Clitic phenomena in European languages*, 293–316. Amsterdam: Benjamins.

Tomić, Olga. 2001. The Macedonian negation operator and cliticization. *Natural Language & Linguistic Theory* 19. 647–682.

Tomić, Olga. 1996. The Balkan Slavic clausal clitics. *Natural Language & Linguistic Theory* 14. 811–872.

Westergaard, Marit. 2021. Language acquisition, microcues, parameters, and morphosyntactic change. In Richard D. Janda, Brian D. Joseph & Barbara S. Vance (eds.), *The handbook of historical linguistics, Volume II*, 357–374. Hoboken, NJ: Wiley Blackwell.

Witkoś, Jacek. 1998. *The syntax of clitics: Steps towards a minimalist account*. Poznań: Motivex.

Witkoś, Jacek & Paulina Łeska-Bayraktar. 2024. The person-case constraint, person ordering constraint and pronominal clusters in polish. *Glossa: a journal of general linguistics* 9(1). doi:10.16995/glossa.10122.

Włodarczyk, Hélène. 2018. Mnie czy mi? O użyciu zaimka pierwszej osoby w celowniku. *Poradnik Językowy* 9. 64–80.

Zaliznjak, A. A. 2004. *Drevnenovgorodskij dialect*. Moscow: Jazyki slavjanskoj kul'tury.

Zaliznjak, Andrej Anatolévič. 2008. *Drevnerusskie ènklitiki*. Moscow: Jazyki slavjanskoj kul'tury.