Marta Ruda* and Nick Huang

Overt pronouns in null subject languages: experimental investigation of Kashubian, Polish, and Silesian

https://doi.org/10.1515/ling-2024-0109
Received June 12, 2024; accepted May 19, 2025; published online September 2, 2025

Abstract: Interpretive differences between overt subject pronouns relative to null subject pronouns are commonly considered a defining property of Consistent Null Subject Languages (CNSLs), in contradistinction to Partial Null Subject Languages (PNSLs). In this article, we present the results of two experimental studies which challenge this view with respect to two environments: variable binding and pronominal coreference (binding) under double embedding. Having shown that given the two phenomena, all four logically possible patterns of binding of overt pronouns are attested in languages with null arguments, we propose that a theoretically meaningful typological *pro*-drop classification should move away from facts surrounding the binding or interpretation of overt pronouns, and should instead focus on what kind of derivational mechanism a language has to yield a null subject structure. Specifically, CNSLs are characterised by a 'rich' inflection-based mechanism, in contrast to PNSLs, for which alternative proposals have been made.

Keywords: Montalbetti effect; null subject language; *pro*-drop typology; pronoun interpretation; binding

1 Introduction

The long line of work on null subject languages within the generative framework (among many others, Alexiadou and Anagnostopoulou 1998; Barbosa 1995, 2009, 2019; Borer 1986, 1989; Chomsky 1981, 1982; Franks 1995; Frascarelli 2007, 2018; Holmberg 2005, 2010a; Holmberg et al. 2009; Huang 1984, 1989; Jaeggli and Safir 1989; Müller 2006; Rizzi 1982, 1986; Roberts 2010, 2019; Roberts and Holmberg 2010; Saab 2020, 2024; Saito 2007; Sigurðsson 2011; Tomioka 2003, and the contributions in Cognola and

***Corresponding author: Marta Ruda**, Institute of English Studies, Jagiellonian University in Kraków, Al. Adama Mickiewicza 9a, 31-120 Kraków, Poland, E-mail: marta.ruda@uj.edu.pl
Nick Huang, Department of English, Linguistics, and Theatre Studies, National University of Singapore, Block AS5, 7 Arts Link, Singapore 117570, Singapore, E-mail: znhuang@nus.edu.sg

Casalicchio 2018a) has led to a distinction between so-called Consistent Null Subject Languages (CNSLs) and Partial Null Subject Languages (PNSLs), among others.

The most prominent features considered to be characteristic of CNSLs include grammatically unconstrained null subjects of all person/number(/gender) combinations and 'rich' subject-verb agreement morphology.[1] In addition to these morphological facts, theoretical accounts of NSLs often prominently feature claims about the information structure properties of overt subject pronouns (see, e.g., Alexiadou and Anagnostopoulou 1998; Barbosa 1995; Barbosa et al. 2005; Camacho 2013; Frascarelli 2007; Frascarelli and Jiménez-Fernández 2019; Roberts 2019). CNSLs and PNSLs differ in how overt subject pronouns are interpreted relative to null subject pronouns. One striking example of this difference (among other examples) is provided by the coreferential preferences assigned to the pronoun *ele* 'he' in (1) in European versus Brazilian Portuguese: barring emphasis, the pronoun is most naturally interpreted as not being bound by (non-coreferential with) the matrix subject *João* in European Portuguese, classified as a CNSL, whereas in Brazilian Portuguese, classified as a PNSL, the bound (coreferential) interpretation can be accessed easily (see Barbosa 2019; Barbosa et al. 2005).

(1) *O João disse que ele comprou um computador.*
the João said that he bought.3SG a computer
'João said that he bought a computer.' [Portuguese, Barbosa 2019: 489]
European Portuguese: João said that someone else not João bought a computer.
Brazilian Portuguese: João said that João himself bought a computer.

In this article, we take a closer look at this claim about overt and null subject pronouns in NSLs with rich verbal morphology. We focus on the two environments that have received much attention in the literature – variable binding and double embedded structures. However, as we discuss in more detail in Section 2, a review of the literature reveals that existing generalizations about binding possibilities for overt and null subjects are somewhat unclear. This lack of clarity calls into question whether these generalizations are a basic difference separating CNSLs and PNSLs. This has typological implications, as it concerns the question of whether these binding facts can be used as a criterion for differentiating between the two types of languages. There are also theoretical implications: if there are truly such basic

1 Some other properties which have been associated with this language type are null expletives, free subject inversion, the lack of the *that*-trace effect, and the unavailability of third person indefinite/arbitrary/generic inclusive null subjects, though these are, relatively speaking, more controversial as criterial properties (for discussions, see, e.g., Barbosa 1995, 2011a,b, 2019; Biberauer 2010; Camacho 2013; Cognola and Casalicchio 2018b; Franks 1995; Genevska-Hanke 2019, 2022; Holmberg 2005, 2010b; Roberts 2010, 2019; Roberts and Holmberg 2010; Ruda 2022, 2024; Sheehan 2010; Willim submitted, and references therein).

differences in binding, our theories of NSLs should strive to explain why such differences exist; otherwise correlations observed in some language groups require alternative explanations.

To strengthen the empirical basis, as well as to better control for potential confounds, such as the lack of clear language-internal baselines for comparison in prior discussions, we conducted two experiments, reported in Sections 3 and 4, to understand how naive native speakers interpret null and overt subject pronouns. We focus on three closely related, West Slavic (Lechitic) languages that meet a number of CNSL criteria: Kashubian, Polish, and Silesian (henceforth K, P, S). K, P, S display *pro*-drop and have largely parallel inflectional systems, allowing for a maximally controlled comparison. We find that overt subject pronouns in both structures vary across the three languages in binding possibilities (relative to null subjects), even though the three languages are otherwise morphosyntactically very similar.

In Section 5, we elaborate upon the implications of this cross-linguistic microvariation, arguing that taking the binding possibilities of overt pronouns as a defining property in our typology of NSLs would require making a potentially arbitrary choice between prioritizing either one or the other of the environments tested. Further supporting this conclusion is the observation that the binding facts for overt subject pronouns in Mandarin Chinese (albeit a discourse *pro*-drop language) differ in yet another way, suggesting that generally, the binding possibilities of overt pronouns cannot be easily integrated into a typology or a theory of null subject languages.

In light of these complications, we argue against making binding possibilities of overt pronouns a core part of our typology and theories of NSLs. Instead, we suggest focusing on null subjects alone; more specifically, classifying languages based on what derivational mechanisms are used to produce or license null subject structures. In this view, CNSLs are languages where a 'rich' inflection-based mechanism is responsible for the distribution of null subjects, in contrast to PNSLs, for which alternative mechanisms have been proposed. Only by making reference to the licensing mechanism can languages that feature a broader distribution of overt pronouns than what is familiar from, say Italian or Polish, be classified in a theoretically meaningful way.

Finally, our study also adds to recent research on microvariation in NSLs by Frascarelli and Jiménez-Fernández (2019), who conducted an acceptability judgment study on Spanish varieties. However, it should be noted that their study and ours differ in two ways: Frascarelli and Jiménez-Fernández are interested in the interpretive preferences of null subjects, while we are interested in the interpretive preferences of overt subject pronouns relative to null subjects. Based on their results, Frascarelli and Jiménez-Fernández also argue that the typological classification of NSLs exists on a continuum, by extension implying that consistent and partial

pro-drop are not discrete categories. Here we will adopt a more traditional viewpoint, although we agree with these authors on the importance of studying the full spectrum of cross-linguistic differences in both the interpretation and the licensing of *pro* (see Section 5 for more discussion).

2 Background

2.1 Variable binding and double embedding

In this section, we review existing claims about how overt subject pronouns differ from null subjects in two environments: variable binding and double embedding.

Whether overt subjects can undergo variable binding is often referred to as the Montalbetti effect, *a.k.a.* the Overt Pronoun Constraint (Montalbetti 1984), stated in (2) and illustrated in (3) for Spanish, a CNSL, and in (4) for Brazilian Portuguese, a PNSL. While in Spanish the bound variable interpretation requires the use of a null subject, in Brazilian Portuguese it is available with an overt pronoun.[2]

(2) Overt Pronoun Constraint
Overt pronouns cannot link to formal variables iff the alternation overt/empty obtains. [Montalbetti 1984: 94]

(3) a. *Muchos estudiantes creen que* **ellas** *son inteligentes.*
many students believe that they are smart
Bound interpretation impossible: **(Many x: x a student) x believes that x is intelligent.*

b. *Muchos estudiantes creen que* ***pro*** *son inteligentes.*
many students believe that are smart
Bound interpretation possible: *(Many x: x a student) x believes that x is intelligent.* [Spanish, Montalbetti 1984: 82–83]

2 As Roberts (2019: 201–202, footnote 5) notes, judgments in this context may be influenced by factors such as the type of quantifier (see also Carminati 2002; Modesto 2000: 82, footnote 71), as well as the language used, with the effect possibly being stronger in Spanish than in Italian, both CNSLs. On the other hand, some speakers of Spanish have been reported to be insensitive to the null/overt distinction in this context (see Camacho 2013; Herbeck 2018, and references therein) and the bound variable interpretation becomes available in CNSLs with overt subjects under focalization or topicalization (see Barbosa 1995; Herbeck 2018; Montalbetti 1984 and references therein), as well as with post-verbal subjects (Barbosa 1995).

(4) *Todo menino acha que* **ele** *é inteligente.*
every boy thinks that he is intelligent
Bound interpretation possible: *(Every x: x a boy) x thinks that x is intelligent.*
[Brazilian Portuguese, adapted from Modesto 2000: 82, footnote 71]

With a non-quantificational matrix subject, as in (5), Montalbetti (1984: 85) reports that the coreferential construal becomes possible, with both null and overt subjects of embedded clauses potentially being able to corefer with the matrix subject. However, Montalbetti also suggests that null subjects might nonetheless be preferred in this context. Although Montalbetti attributes this particular preference for null subjects to pragmatics, Holmberg (2010a: 91) takes the null/overt subject facts in embedded clauses to be a defining property in the distinction between CNSLs and PNSLs.

(5) a. *Juan cree que él es inteligente.*
Juan believes that he is intelligent
b. *Juan cree que* pro *es inteligente.*
Juan believes that is intelligent
'Juan believes that he (=Juan) is intelligent.' [Spanish, Montalbetti 1984: 85]

In particular, Holmberg (2010a) capitalizes on the observation that in CNSLs such as Arabic, Greek, Spanish, Turkish, and Italian, the embedded subject cannot be overt unless it is contrastive or a shifted topic. On the other hand, in PNSLs such as Brazilian Portuguese, Finnish, and Marathi there is no such requirement, with the crucial difference being that the null subject has to be coreferential with the matrix subject and free reference is not possible in this case, unlike for the overt pronoun. However, it has been suggested that Bulgarian actually allows for both an ambiguous interpretation of an embedded null subject and a similarly ambiguous interpretation of an overt pronoun in the same position (6). Nevertheless, Genevska-Hanke (2022), who points out this observation, classifies Bulgarian as a CNSL, treating it as counter-evidence to the claim that CNSLs disallow coreference between overt subjects of embedded clauses and matrix subjects.

(6) *Petar*$_i$ *kaza če* {pro$_{i/j}$ / *toj*$_{i/j}$} *e spal.*
Peter said.3SG that he PRF slept
'Peter$_i$ said that he$_{i/j}$ had slept.' [Bulgarian, from Bojadžiev et al. 1999 via Genevska-Hanke 2022: 188–189]

Furthermore, contrary to a widespread view, such a coreferential reading of overt pronouns has been reported as possible even in Italian, as (7) shows (see also Frascarelli and Jiménez-Fernández 2019 and references therein for other Romance varieties).

(7) *Leo$_i$ ha detto che lui$_i$ ha comprato una casa.*
Leo has.3SG said that he has.3SG bought a house
'Leo$_i$ said that he$_i$ had bought a house.' [Italian, Frascarelli 2018: 225]

Contrasts for null and overt subjects seem to be clearer in the second environment – double embedding structures. Languages such as Brazilian Portuguese (a PNSL) have been reported to require that a null subject be coreferential with the closest c-commanding antecedent (8a) – though see Barbosa et al. (2005); Fernandes et al. (2018) for some counter-evidence – whereas an overt pronoun in this language (see (8b)) can corefer with either of the higher subjects. As for languages such as European Portuguese (a CNSL, see (9)), a null embedded subject can corefer with either of the higher subjects, or can have a discourse antecedent. The literature is less clear about what interpretations overt subject pronouns have in European Portuguese, but given the above discussion about variable binding, a reasonable expectation is that they cannot be coreferential with a local antecedent. Data parallel to (8) are reported for another PNSL, Finnish, as well as for a discourse *pro*-drop language, Mandarin Chinese, by Modesto (2008: 383). Holmberg and Sheehan (2010) propose to account for the PNSL facts by taking them to involve control between the antecedent and the null subject, modeled as syntactic binding by Agree between the unvalued [D]-feature of the null subject pronoun and the antecedent. The unavailability of coreference with the highest subject is a case of defective intervention on their analysis. It is not clear what mechanisms derive the interpretation of the overt pronominal subjects, though.

(8) a. *O Paulo$_1$ disse que o Pedro$_2$ acredita que* pro$_{*1/2/*3}$ *ganhou.*
the Paulo said that the Pedro believes that won
b. *O Paulo$_1$ disse que o Pedro$_2$ acredita que ele$_{1/2/3}$ ganhou.*
the Paulo said that the Pedro believes that he won
'Paulo said that Pedro believes that (he) won.' [Brazilian Portuguese, Modesto 2000: 84]

(9) *O Paulo$_1$ disse que o Pedro$_2$ acredita que pro$_{1/2/3}$ ganhou.*
the Paulo said that the Pedro believes that won
'Paulo said that Pedro believes that (he) won.' [European Portuguese, Modesto 2000: 84]

To sum up, there is a general intuition that CNSLs and PNSLs differ in the binding possibilities for overt and null subjects. To the extent that these differences are systematic, it suggests that binding is a basic differentiator between these two types of languages and raises interesting questions about the information structure and even the morphosyntax of overt pronouns in the two types of languages. However, as can be seen from our review above, (potential) counter-examples have been raised

occasionally. Existing examples are also based on a relatively small sample of languages, with the most detailed discussion centered on Romance languages. We add to the cross-linguistic landscape by seeking to establish relevant generalizations in three West Slavic null subject languages, K, P, and S. To ensure that our generalizations are sound and replicable, our study makes use of formal experiments. But before discussing our experiments, we first review basic facts for these three languages. This is especially important since K and S are less well-studied than P is.

2.2 The null subject parameter in K, P, and S

In contrast with P, which has official status and is widely spoken, K and S are varieties spoken by much smaller communities. K is used mostly in northern Poland (Pomeranian Voivodeship) and S in southern Poland (Silesian and Opole Voivodeships) and northern Czech Republic (Moravian-Silesian Region). In the Polish 2021 Census, 87,600 and 457,900 speakers declared K and S respectively as the main language used at home (although some research shows a much bigger number for K; see Rogowska-Cybulska and Cybulski 2011 and references therein). K is classified as severely endangered in the *Atlas of the World's Languages in Danger* (Salminen 2010) and has the official status of a regional language, while S is excluded from consideration in the *Atlas* and has been classified as a Polish dialect.[3] Both K and S are characterised by a significant degree of dialectal variation. The judgment data presented in what follows come from speakers of Central K (Central Kashubia, Eastern Pomerania region of northern Poland) and Central S (the Katowice metropolitan area in the Upper Silesia region of southern Poland).

While the status of P as a CNSL has been clear (see Barbosa 2019; Franks 1995; Roberts 2019; Ruda 2022; Willim 1989, submitted for this classification, and see also Barski 2013; Bondaruk 2001; McShane 2009; Nillson 1982; Pisarkowa 1969; Wolna et al. 2022 for some previous discussions of null and overt pronominal subjects in P), the situation in K and S needs to be evaluated more carefully, not only because they are relatively understudied, but also because of microvariation in past tense agreement morphology (see Section 2.2.2). As mentioned in the introduction, a number of properties have been used to classify a language as a CNSL; we list six key properties in (10).

(10) Distinguishing features of CNSLs (based on previous literature; for useful overviews see esp. Barbosa 2011a, b; Camacho 2013; Cognola and Casalicchio 2018b; Roberts 2019; Roberts and Holmberg 2010)

3 On April 26, 2024, the lower house of the Polish parliament granted Silesian regional language status, but the bill was vetoed by the president on May 29, 2024.

a. grammatically unconstrained null subjects of all person/number/gender combinations
b. 'rich' subject-verb agreement morphology
c. free subject inversion
d. unavailability of a generic inclusive null subject with plain third person verbal inflection in out-of-the-blue contexts
e. lack of the *that*-trace effect
f. binding differences between null and overt subject pronouns, e.g. null subjects can receive a bound variable interpretation but overt subjects cannot.

In this section, we consider all three languages in light of the properties in (10a–d), showing that they all pattern almost alike, as is expected if they are CNSLs. Property (10e) is difficult to verify due to the dispreference or complete unavailability of long-distance *wh*-extraction in these languages (see Cichocki 1983; Szczegielniak 1999 for discussions of P); but to the extent that such extraction is possible in K and P, the judgments suggest that the *that*-trace effect is not operative in these two languages; the S native speakers whom we have consulted do not accept long-distance *wh*-extraction at all. The final property (10f) is the subject of our experimental study, so we defer our discussion to Sections 3 and 4.

2.2.1 Null subjects of all person/number/gender combinations

Null subjects are available in contemporary K and S varieties, where *pro*-drop is in principle possible with all person/number/gender combinations including third person subjects in matrix clauses, as in (11)–(12) (barring the contexts of agreement drop discussed below) (for references for K, see esp. Nomachi 2014, 2019 and the references therein, as well as the data in Ruda 2024; Ruda et al. 2022a; for S, see the data in Ruda 2024; Ruda et al. 2022b).[4]

(11) a. *Sniég* *òdgarnie* *spòd* *brómë.*
snow shovel.3SG from.in.front.of gate
'He will shovel the snow from in front of the gate.' [K, file: k48.txt]

4 Small K and S corpora of authentic native speaker texts of different genres (blog, newspaper column, news, literary fiction, Facebook post), consisting of 80 Kashubian and 80 Silesian texts of approximately 300 words each have been assembled by Ruda et al. (2022a) and Ruda et al. (2022b). When corpus examples are provided below, the file name refers to the corpus file from which the example has been extracted, available on the OSF platform (https://osf.io/q49y6 and https://osf.io/ypw4v).

b. *Za wiele nie mëslelë, nie òbzérelë sã za jinszima,*
too much not thought.3PL.M not looked.3PL.M SE on others
zdrzelë blós na se, òddelë dwa dobré strzélë a…
looked.3PL.M only on self gave.3PL.M two good shots and
zajãlë czwiôrti plac.
took.3PL.M fourth place
'They didn't think too much, didn't pay attention to others, they only looked at each other, shot two good shots, and…took the fourth place.'
[K, file k61.txt]

(12) a. *Piyrwyj miała cynżko spokopić roztōmajte rzeczy, na*
at.first had.3SG.F difficulty understand different things on
kere miała terozki dować pozōr.
which had.3SG.F now give attention
'At first it was hard for her to understand the different things which she was to look after now.' [S, file: s7.txt]

b. *Lŏcōm po świecie, fajrujōm i nawet jak niyrŏz sie*
fly.3PL around world celebrate.3PL and even if at.times SE
muszōm wspōmōc krykōm, to na starojść niy majōm czasu.
have.to.3PL support stick then on old.age not have.3PL time
'They fly around the world, celebrate, and even if they need to support themselves with a walking stick at times, they don't have the time for the old age.'
[S, file s64.txt]

2.2.2 'Rich' subject-verb agreement morphology

Being closely related languages, K, P, and S have very similar tense inflection paradigms. As Table 1 shows, there are different forms for all person/number combinations in the present tense in all three languages.[5]

In the past tense and the moods where the lexical verb is in the so-called *l*-participle form, which inflects for number and gender, P also requires a person/number agreement marker, which is a clitic that is obligatorily expressed on the verb or a pre-verbal constituent. In K and (in some contexts) S, however, the person/number marker can be absent, a phenomenon we will call 'agreement drop' (Breza 2001; Breza and Treder 1981; Lorentz 1971; Nomachi 2014; Ruda 2017). Agreement drop contexts, exemplified in (13)–(14), require overt subjects, even though K and S

5 The data in Table 1 have been provided by native speakers of Central K, standard P, and Central S.

Table 1: Present tense for ‘to read’.

	Kashubian	Polish	Silesian
1SG	czëtó-m	czyta-m	czytů-m
2SG	czëtô-sz	czyta-sz	czyto-sz
3SG	czëtô-ø	czyta-ø	czyto-ø
1PL	czëtó-më	czyta-my	czytů-my
2PL	czëtô-ta	czyta-cie	czyto-ć́e
3PL	czëta-ją	czyta-ją	czyta-jům

are otherwise *pro*-drop languages. This is presumably because in these contexts, there is no other way of expressing the person feature of the subject.

(13) **(Jô) doch nick nie zrobił — tłomacził sã młodi.*
I PRT nothing not did.SG.M explained.SG.M SE young
‘I didn’t do anything, the young explained himself.’ [K, file: k10.txt]

(14) **(Jo) o nich słyszoł — oni o mie niy.*
I about them heard.SG.M they about me not
‘I heard about them, they about me not.’ [S, file: s27.txt]

When the agreement marker is present, both K and S make null subjects available, as in (15), where the past tense is formed by attaching the agreement marker to the auxiliary BE, and in (16), where the agreement marker is attached to the copula. The overt realization of the agreement marking and agreement drop are available to all S speakers (Ruda 2017); the (non-)realization of agreement in K is subject to dialectal variation (Breza 2001; Jocz forthcoming; Lorentz 1971; Treder 2006).

(15) *Chùtuszkò je-m rozerwôł kùwertã czekawi, jaczé wiadło bënë*
quickly be-1SG tore.SG.M envelope curious what news inside
ni sã tacy […]
not are such
‘I tore the envelope quickly, curious what news could be inside […]’
[K, file: k72.txt]

(16) *Boł-ech niydowno we ślonskij wsi Boronow […]*
was.SG.M-1SG recently in Silesian village Boronow
‘I’ve recently been to the Silesian village Boronow […]’[S, file: s59.txt]

In the future tense, the features of person and number are distinctly expressed for all combinations on the auxiliary (future imperfective) or directly on the lexical verb (future perfective), in accordance with the present tense inflectional pattern. Agreement drop is unavailable in the present and future tenses (but for the present tense of BE, which requires the person/number clitics, like the *l*-participles). In this case, null subjects of all person/number/gender combinations are available in all

languages which are our focus here (see Ruda 2024 and references therein; see also Section 5.2 for some relevant discussion).

2.2.3 Free subject inversion

As (17)–(19) show, free subject inversion is possible in K and S, just as it is in P.[6]

(17) a. *Przyszła Anna.*
came Anna
'Anna came.'
b. *Będzie dzwoniło wielu studentów.*
will call many students
'Many students will call.' [P]

(18) a. *Przëszła Ana.*
came Ana
'Ana came.'
b. *Bãdze zwòniło wiele sztudérów.*
will call many students
'Many students will call.' [K]

(19) a. *Przyszła Ana.*
came Ana
'Ana came.'
b. *Bydźe zwůńůło moc sztudyntůw.*
will call many students
'Many students will call.' [S]

2.2.4 Generic inclusive null subjects and third person inflection

When the generic inclusive interpretation of the null subject is intended with a third person verb form in an out-of-the-blue context, P, S, and contemporary K pattern with CNSLs like Italian and European Portuguese in that they employ the SE-marker *sã*$_K$/*się*$_P$/*sie*$_S$:

(20) *Po ciemku jeździ* **się** *wolniej.*
by dark drive.3SG SE slower
'One drives slower when it's dark.' [P, National Corpus of Polish, Przepiórkowski et al. 2012]

6 The K and S examples and judgments presented here and in what follows were offered by Marika Jocz (Central K) and Bartłůmjej Wanot (Central S).

(21) *Z drëdżi równak stronë, czej wëżi* ***sã*** *wléze, tej barżi*
on second however side when higher SE climb.3SG then more
bãdze bòlec, jak **sã** *zląděje.*
will hurt if SE fall.3SG
'On the other hand though, when one climbs higher, then it hurts more, if one falls down.' [K, file: k24.txt]

(22) *Takowe fotografije* **sie** *dowało do albumu a* **sie** *jich potym*
such photographs SE gave.3SG.N to album and SE them later
oglóndało a pokazowało.
viewed.3SG.N and showed.3SG.N
'Such photographs one would add to an album and one would later view and show them.' [S, file: s72.txt]

An exception here is the generic [3SG] modal construction in K (23), which does not exist in either P or S (see Jocz et al. 2022; Ruda 2022, 2024, and references therein).

(23) a. *Tu ni mòże palëc.*
here not may.3SG smoke
'One may not smoke here.'

One other thing to note is that in some impersonal environments, K and S admit overt pronouns where this is not possible in P, for example with [3PL] arbitrary subjects, illustrated in (24a,b) versus (24c).

(24) a. *Òni pòrisowelë mie aùtół.*
they scratched me car
'They have scratched my car./My car has been scratched.' [K, Jocz et al. 2022: 103]

b. *Uůńi porysowali mi autok.*
they scratched me car
'They have scratched my car./My car has been scratched.' [S, Jocz et al. 2022: 103]

c. *(#Oni) porysowali mi samochód.*
they scratched me car
'They have scratched my car./My car has been scratched.' [P]

However, the availability of arbitrary and generic readings with overt subjects seems to be a point of variation between CNSLs, some of which have, for instance, been reported to admit the generic interpretation not only with null, but also with overt [2SG] subjects (see Camacho 2013: 180 and Genevska-Hanke 2019: 85 respectively for

Spanish and Bulgarian), in contrast to K, P, and S, where the pronoun needs to be null in this context. While the generic and arbitrary interpretations of overt subjects have not been considered as indicative of the *pro*-drop status of a language, it may be worth noticing that for some languages classified as CNSLs interpretive differences between null and overt pronouns are nullified at least in these environments.

2.3 Interim summary: clarifying binding facts for subject pronouns with experiments

Overall, the grammatical features of K, P, and S reviewed above suggest that they fall straightforwardly within the category of CNSLs: null subjects of all person/number/gender combinations are available in matrix and embedded clauses, all person/number combinations are distinguished morphologically in all tenses (barring agreement drop in K and S), all three languages allow free subject inversion and couple third person agreement on the verb with the SE marker to achieve the generic inclusive interpretation in out-of-the-blue contexts (but for modal impersonals in K). These criteria are standard in the NSL literature, even though the latter two may raise some concerns (see the references in Footnote 1 and Section 5 below).

With this background in mind, we move on to consider what the binding facts may be for K, P, and S subject pronouns in two contexts: variable binding and double embedding. Given the above properties, we expect these three languages to largely pattern with each other and other CNSL languages in how overt subject pronouns are interpreted. However, the phenomenon of agreement drop means that K and S are in a sense less consistent in allowing null subjects than P. This opens up a possibility that the languages might actually differ in how null and overt subject pronouns are interpreted, which we explore through formal experiments.

We chose formal experiments over eliciting informal judgments from native speaker consultants, because experiments produce quantifiable data that in principle can reveal more subtle variation between the three languages. Our experiments were presented in an auditory mode, with pre-recorded materials, because K and S are mostly spoken. A further advantage is that pre-recorded materials let us ensure that overt subject pronouns are pronounced consistently in an unstressed manner for all participants. This is important in light of prior observations that the Montalbetti effect can be neutralized in a CNSL like Spanish by stressing or focusing the pronoun.

More specifically, we used a two-alternative forced choice sentence interpretation task, in which participants first heard a sentence with either a null or overt subject pronoun. They then heard a second sentence that makes explicit a particular interpretation of the subject pronoun, and judged whether the second sentence accurately

described the first sentence, by choosing between "True" and "False" options. Hence, while we did not use informal judgments, our experiments were closely modeled upon the informal introspective judgment task standardly used by syntacticians to investigate binding and coreference: for a given sentence that is in principle ambiguous, participants decide whether the coreferential (bound) reading is available.

We should also point out that what our experiments measure is how easy it is to access a particular interpretation. For this reason, for instance, we interpret our results as showing a Montalbetti effect if participants find it harder to obtain a bound variable reading with overt subject pronoun sentences than with null subject sentences, and consequently respond "False" more often to overt subject sentences than to null subject sentences. This departs somewhat from descriptions of the effect in the theoretical literature, in which the Montalbetti effect is the impossibility (ungrammaticality) of a bound variable reading with overt subject pronouns. The departure is necessary because experiments alone cannot definitively show whether a reading is possible or impossible. To illustrate this issue, suppose a participant gives a "False" response to overt subject sentences. This response might well indicate that the interpretation in question is ungrammatical and not available at all for overt subjects, but it might also be the case that the interpretation is available, but difficult to access for processing reasons, for example because it occurs less frequently in naturally occurring speech and texts. Deciding between the two possibilities crucially depends on one's theoretical framework and assumptions about how behavior and grammar are related.

One might wonder whether there are other tasks that could produce results that allow us to make claims about the grammaticality of certain readings. Perhaps with the same pair of sentences, participants could be instead asked to judge the acceptability or possibility of the second sentence as a description of the first. Alternatively, perhaps participants could be presented with a scenario, and then decide whether it is more acceptable to describe the scenario with a sentence containing a null or overt subject pronoun. Or perhaps one could switch to more conventional fieldwork methods, by eliciting informal introspective judgments from native speaker consultants. However, the same limitation is equally applicable to these alternatives: an "unacceptable" or "impossible" response elicited through these tasks might just mean that the bound variable reading is difficult to access. In this respect, then, our judgment task is no worse than these alternatives.

Put differently, judgments obtained for the interpretation of null and overt subjects will always need to be handled carefully. Importantly, in the context of our experiments, we will assume that whatever patterns we observe should be accounted for in terms of grammatical factors (rather than processing ones), as is standardly assumed in the existing literature on null and overt subjects.

3 Experiment 1: variable binding

3.1 Design

Experiment 1 is intended to assess the Montalbetti effect in K, P, and S: whether K, P, and S speakers prefer a null subject (compared with an overt subject) to be bound by a quantificational subject (like a CNSL language like Spanish, as reported by Montalbetti 1984), or whether they are neutral about the binding of null and overt subjects.

More specifically, we created two conditions that differ in whether the embedded subject was null or an overt pronoun. Since this subject is in principle ambiguous, we made sure to disambiguate by having each sentence be accompanied by a second sentence (a "prompt") that makes explicit the bound variable reading, as shown in (25). Participants were requested to judge whether the second sentence accurately described the first sentence. In the ideal case, if the language in question exhibited the Montalbetti effect, participants should reliably reply "True" if the first sentence had a null embedded subject and reply "False" if that sentence had an overt subject pronoun.

(25) a. Sentence presented
   (i) *Kòżdé dzéwczã rzekło, że (òno) je saja.* [K]
   (ii) *Każda dziewczyna powiedziała, że (ona) jest leworęczna.* [P]
   (iii) *Kożdŏ dzioucha pedziała, iże (ôna) je leworyncznŏ.* [S]
   'Every girl said that (she) is left-handed.'

   b. Prompt
   (i) *Kòżdé z tëch tu dzéwczãt gôdało ò se sami.* [K]
   (ii) *Każda z tych dziewczyn mówiła tu o sobie samej.* [P]
   (iii) *Kożdŏ z tych dziouch gŏdała sam ô siebie samyj.* [S]
   'Each of these girls talked about herself here.'

For thoroughness, we note that we included two other conditions, to study the availability of bound variable readings for overt and null possessors (e.g., in sentences equivalent to *Every girl said that **(her)** mom is left-handed*). We will not discuss this manipulation in this paper, as it does not necessarily concern the Montalbetti effect, which covers overt/null alternations (see (2)). One might choose to analyze null possessors instead as being absent in the syntax altogether, in which case there is no overt/null alternation to speak of.

3.2 Materials

For each language, we created 16 frames, each featuring a distinct combination of lexical items, and corresponding prompts. For each frame we created four versions, one per condition (null/overt subject/possessor), producing a total of 64 sentences. The sentences and prompts (64 pairs total) were sorted into four lists using a Latin Square design, such that each condition appeared four times and every frame appeared only once. All sentences and prompts for all three languages were recorded by native speaker readers, to be presented auditorily.

We then combined the lists with target items from Experiment 2 (see Section 4 below), effectively using one experiment's materials as filler items for the other. We had also explored adding fillers that were not target items from either experiment, but received feedback from native speakers that doing so made the overall study too long and taxing.

3.3 Participants

We recruited 40 adult native speaker participants for each language, by word of mouth. Participants received gifts or vouchers as compensation. Because K and S speakers are almost always also proficient in P, we could not and did not impose any exclusion criterion based on their knowledge of P or any other language.

Our K participants were speakers of the Central K dialect, living in Central Kashubia (Sierakowice and surrounding towns/villages), Eastern Pomerania region of northern Poland. The age of participants ranged from 23 to 70 years old, with the majority being above age 55. In a survey we administered, most participants reported acquiring K from birth, with parents using K at home. A majority of participants (26) reported using K actively at least 50 % of the time, with the other language being P.

Our S participants were speakers of the Central S dialect, living in the Silesian Voivodeship (the Katowice urban area) in the Upper Silesia region of southern Poland. Like K participants, S participants tended to be older adults (45–78 years of age), and generally acquired S from birth, from family and peers. Around half of participants reported using Central S actively at least 50 % of the time, with the other language being P or – in one case – German.

3.4 Procedure

This experiment was conducted in-person with research assistants who are native speakers of K, P, or S. After informed consent was obtained, the experimental session

started with a short informal conversation in the target language, so that participants would feel more comfortable. In addition, doing so helped to mitigate the risk of the intrusion of P in the K and S sessions. Research assistants also explained the instructions in an informal manner in the target language.

We used PCIbex (Zehr and Schwarz 2018) for the display of materials and data collection. Participants were randomly assigned to one of the lists, and listened to recordings played from a laptop via PCIbex. They judged whether the prompt was a "True" or "False" description of the sentences by pressing keys on the keyboard. Participants could listen to each recording as many times as they preferred before giving a judgment.

3.5 Results

The results presented are for 40 P participants, 35 K participants, and 32 S participants; the K and S figures are lower due to unforeseen data loss, where responses were not successfully recorded on PCIbex, for presumably technical reasons. Figure 1 presents the percentage of "True" responses; recall that this measures the availability of a bound variable reading. For statistical analysis, we fitted a mixed-effects logistic regression model for each language's data set, with pronoun type (overt vs. null) as predictors, and with random intercepts by participant. This was the most complex model that converged for all three data sets and did not produce warning messages indicating over-parametrization.

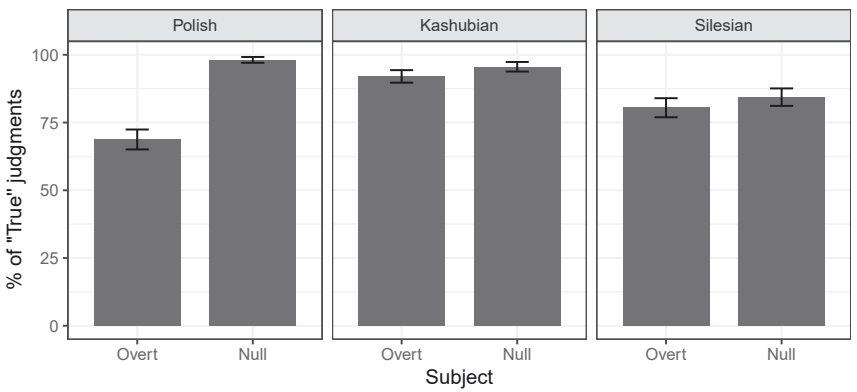

Figure 1: Experiment 1 (variable binding) results; whiskers = standard errors.

P participants obtained the bound variable reading much more reliably with null subjects (98 %) than with overt subjects (69 %) (b = 4.67, z = 5.62, p < 0.01). In contrast, K and S participants obtained the bound variable reading much more easily for both types of subjects at similar rates (about 90 % for K and about 80 % for S). Numerically, there is a slight advantage for null subjects in both languages, but no significant difference was found for K (b = 1.59, z = 1.79, p = 0.07) or S (b = 0.42, z = 1.01, p = 0.31). Given our definition of the Montalbetti effect (Section 2.3), we take these results to indicate that there is cross-linguistic variation between K, P, and S, such that P shows the Montalbetti effect but K and S does not.

3.6 Discussion

Although P shows that bound variable readings are easier to obtain for null subjects than for overt subjects, as expected if P shows the Montalbetti effect, there is a data point that is potentially problematic: the relatively high rate of "True" responses for overt subjects in P (69 %). In existing descriptions of the Montalbetti effect, it is ungrammatical for overt subjects to have a bound variable reading. If so, the rate of "True" responses should have been much lower.

As an anonymous reviewer suggested, this result might indicate something deeper and surprising about binding possibilities in P, in contrast to another CNSL, like Spanish or Italian. Whether that is the case is an empirical question that we will need to leave to future research. But in the meantime, we argue that there might be simpler explanations for why P speakers show a high rate of "True" responses. For instance, we think there is likely to be a "True" bias attributable to our task. Consider a sentence with an embedded overt subject pronoun like "Every girl said that she is left-handed". Participants were requested to judge whether in this sentence, each of the girls talked about herself. Although the grammar of P might not usually allow variable binding for the overt pronoun *ona* 'she', there is no other salient third person female individual that could serve as the pronoun's antecedent. Participants therefore might feel more obliged to give a "True" response. Second, it is also possible that P speakers might make an exception to variable binding if the overt pronoun is stressed, as Montalbetti (1984) reported in the context of Spanish. While the pronoun was not stressed or focused in our stimuli, it is nonetheless possible that some participants perceived otherwise. In our task, participants listened to sentences and interpreted them. In the process of interpretation, participants might have to recall what they heard, but their recollection might not be an accurate reflection of the recordings played. And while participants could listen to the recordings multiple times, doing so would have been effortful and might have introduced further distortions.

Another possibility, suggested by an anonymous reviewer, is the reported tendency for Polish to use overt pronominal subjects more frequently in colloquial spoken language than in written language (Nillson 1982; Pisarkowa 1969). The greater frequency of overt pronominal subjects in spoken P might reflect the disambiguating and emotive functions overt pronouns have, in addition to their different information structural functions (McShane 1999, 2009; Nillson 1982; Pisarkowa 1969 and references therein; see also Pešková 2019 for Spanish and Czech). Setting aside why there is such a difference in use, this general frequency of overt pronouns in spoken P might have made participants more willing to entertain and accept a bound variable reading, even when such a reading is not grammatically licensed.

Similarly, one might also wonder whether high rates of "True" responses and the smaller contrasts in K and S also reflect a more general "True" bias in K and S participants: since there is no standard variety of K and S, participants might be much more willing than P participants to accept any plausible description of the target sentence, even if the description involves an ungrammatical interpretation.

Such a bias seems unlikely to us, especially when we consider participant responses to Experiment 2; recall that each participant completed Experiments 1 and 2 in a single sitting. In Experiment 2, participants also had to give "True" or "False" responses to another set of constructions with overt and null subject pronouns. If K and S participants had a general bias to give "True" responses to overt subject sentences (or even any type of sentence), there should be similarly high rates of "True" responses for overt subject sentences in Experiment 2. That is actually not the case; "True" response rates are generally lower in Experiment 2 for K and S. In fact, contrasts between (certain) K and S conditions in Experiment 2 are also larger and comparable with equivalent contrasts in P (For more details, see Section 4.4.). In other words, even though K and S are not standardized like P is, K and S participants were not indiscriminately giving out "True" responses; they can be as sensitive as P participants are to certain contrasts. In the context of Experiment 1, then, the fact that K and S participants (but not P participants) gave similar responses to the null and overt subject conditions is unlikely to be a null result due to experiment limitations, but rather a genuine case of cross-linguistic variation.

To sum up the results of Experiment 1, a bound variable reading was easier to obtain for null subjects than overt subjects in P. This difference is consistent with P exhibiting a Montalbetti effect and being a CNSL. In contrast, there was no clear difference for both types of subjects in K or S, despite K and S being otherwise typologically very similar and genetically closely related to P.

4 Experiment 2: double embedding

4.1 Design

This experiment probed a different kind of binding and coreference involving null and overt subject pronouns: when these subjects appear in doubly-embedded contexts, whether they can corefer with the local antecedent or a non-local one. Given existing descriptions, we expect that in a PNSL, null and overt subject pronouns can both corefer with (be bound by) a local antecedent, but only overt subjects can corefer with a non-local antecedent. It is less clear from the literature to what extent overt subjects differ from null subjects in a CNSL, but a reasonable expectation is for null subjects (and not overt subject pronouns) to tend to corefer with a local antecedent, as we discussed in Section 2.

To evaluate these claims about coreference, we again used the same sentence interpretation task: participants listened to a target sentence and a prompt that described that sentence. We used a 2×2 design that crossed two factors: whether the sentence had an overt or null doubly-embedded subject, and whether the prompt named the subject of the main clause or the embedded clause, that is, whether the doubly-embedded subject was intended to be bound by a non-local or local antecedent. To the extent that the subject can be bound by the intended antecedent in the target sentence, participants should respond "True" to the prompt sentence. Example (26) provides an example of these four conditions for P. Again, note that this task is intended to mirror coreference judgments that syntacticians typically use in informal introspection.

(26) Polish example (we mark the position of the null subject with _ for exposition purposes only; participants heard audio recordings)

a. Null doubly-embedded subject / Local antecedent

(i) Sentence presented

Jacek powiedział, że Marcin myśli, że _ dźwignie
Jacek said that Marcin thinks that will.lift
więcej niż 100 kg na siłowni.
more than 100 kg on gym
'Jacek said that Marcin thinks that *pro* will lift more than 100 kg at the gym.'

(ii) Prompt

To Marcin jest tą osobą, która ma dźwignąć więcej niż
PRT Marcin is this person who is.to lift more than
100 kg na siłowni.
100 kg on gym
'It's Marcin who is the person who is to lift more than 100 kg at the gym.'

b. Overt doubly-embedded subject – *on* 'he' / Local antecedent

(i) *Sentence presented*

Jacek powiedział, że Marcin myśli, że on dźwignie
Jacek said that Marcin thinks that he will.lift
więcej niż 100 kg na siłowni.
more than 100 kg on gym
'Jacek said that Marcin thinks that he will lift more than 100 kg at the gym.'

(ii) Prompt

To Marcin jest tą osobą, która ma dźwignąć więcej
PRT Marcin is this person who is.to lift more
niż 100 kg na siłowni.
than 100 kg on gym
'It's Marcin who is the person who is to lift more than 100 kg at the gym.'

c. Null doubly-embedded subject / Non-local antecedent

(i) Sentence presented

Jacek powiedział, że Marcin myśli, że _ dźwignie więcej
Jacek said that Marcin thinks that will.lift more
niż 100 kg na siłowni.
than 100 kg on gym
'Jacek said that Marcin thinks that *pro* will lift more than 100 kg at the gym.'

(ii) Prompt

To Jacek jest tą osobą, która ma dźwignąć więcej niż
PRT Jacek is this person who is.to lift more than
100 kg na siłowni.
100 kg on gym
'It's Jacek who is the person who is to lift more than 100 kg at the gym.'

d. Overt doubly-embedded subject – *on* 'he' / Non-local antecedent

(i) Sentence presented

Jacek powiedział, że Marcin myśli, że on dźwignie więcej niż
Jacek said that Marcin thinks that he will.lift more than
100 kg na siłowni.
100 kg on gym
'Jacek said that Marcin thinks that he will lift more than 100 kg at the gym.'

(ii) Prompt

To Jacek jest tą osobą, która ma dźwignąć więcej niż 100 kg
PRT Jacek is this person who is.to lift more than 100 kg
na siłowni.
on gym
'It's Jacek who is the person who is to lift more than 100 kg at the gym.'

4.2 Materials

Sixteen frames were constructed. For each frame, we created four versions of sentences and prompts, one version per condition. As in Experiment 1, the resulting 64 pairs of sentences and prompts were recorded and sorted into four lists using a Latin Square design, such that each condition appeared four times and each frame appeared only once.

The items for this experiment were combined with those of Experiment 1, as described above.

4.3 Participants and procedure

Since we combined both Experiments 1 and 2 in the same session, the participant recruitment and procedure were the same as those for Experiment 1.

4.4 Results

Figure 2 plots the results for all three languages. Across all three languages, with a non-local antecedent, participants tended to give a "True" response when the doubly-embedded subject was overt, although the rate of "True" responses is generally low (around 50 % for overt, lower for null subjects). With local antecedents, we observe variation between languages. Visually, P and S pattern together: participants are

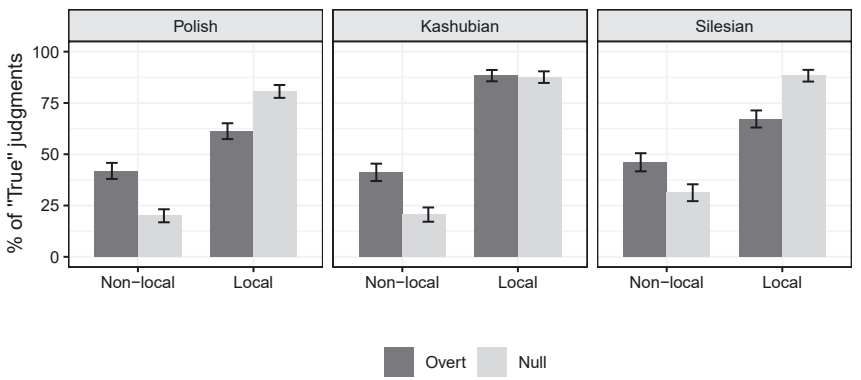

Figure 2: Experiment 2 (double embedding) results; whiskers = standard errors.

more likely to give a "True" response if the doubly-embedded subject was null than if it was overt. This was not the case for K: participants were equally likely to give a "True" response regardless of whether the doubly-embedded subject was null or overt.

Statistical tests confirmed these observations of cross-linguistic differences. More specifically, for each language's data set, we fitted a mixed-effects logistic regression model, with pronoun type (overt vs. null) and antecedent type (non-local vs. local) as predictors, and with random intercepts by frame (again, this was the most complex model that converged for all three data sets). The interaction between pronoun and antecedent types was statistically significant for P ($b = 2.04$, $z = 5.60$, $p < 0.01$) and S ($b = 1.95$, $z = 4.55$, $p < 0.01$), but not K ($b = 0.91$, $z = 1.91$, $p = 0.06$).

4.5 Discussion

We next discuss what these results mean for the classification of K, P, and S. We start by reviewing the responses for the two local antecedent conditions. P and S participants were more likely to allow null subjects than overt subject pronouns to corefer with these antecedents, while K did not show any clear preference. These responses are consistent with a scenario where P is a CNSL – as is usually assumed in the literature – and so is S, while K is a PNSL.

It is harder to draw clear conclusions from the responses for the non-local antecedent conditions. To the extent that any of the three languages are CNSLs, null subjects should be able to be bound by non-local antecedents, and thus one would expect a high rate of "True" responses for the non-local antecedent/null subject condition. However, the actual rate of "True" responses is much lower, around 25 %. Similarly, to the extent that any of the languages is a PNSL, overt subjects should be able to be bound by non-local antecedents, but again, the rate of "True" responses for the non-local antecedent/overt subject condition is much lower, around 50 %.

We suggest that the generally lower rates of "True" responses for non-local antecedents are an artifact of the experiment. For each trial, participants listened to a relatively long sentence with a doubly-embedded structure, processed it, and maintained its representation in memory while judging the prompt (recall that our materials are all auditory). In order to interpret the overt or null subject pronoun, participants would presumably build a dependency between it and what they believe is the intended antecedent. In this dependency-building process, the non-local antecedent arguably has a disadvantage relative to the local antecedent: it appeared earlier in the sentence, before the local antecedent did. As a result, it

might be more difficult to retrieve the non-local antecedent (i.e. have it bind the subject pronoun), due to the relatively longer dependency, memory decay, or both. To the extent this dependency is more difficult to build and/or maintain, this would reduce the percentage of "True" responses for the non-local conditions. That said, this processing account still does not explain why "True" responses are more common for overt subjects than for null subjects. Some other explanation must be sought, perhaps in the differences in the information structural properties of null and overt subjects.

To sum up, while results for the non-local antecedent conditions are tricky to interpret, those for local antecedent conditions nonetheless allow us to form a tentative conclusion: P and S, in allowing null subjects to be bound by local antecedents at a higher rate, have the profile one might expect of a CNSL. K, in allowing both null and overt subjects to be bound by these antecedents at the same rates, provides a clear match with Modesto's (2000) description of Brazilian Portuguese, a PNSL. Setting aside issues relating to interpretation, another noteworthy aspect of the results is the fact that the three languages again do not pattern uniformly with each other, despite being closely related typologically and historically.

5 General discussion

We take the results of our experiments as collectively showing that the properties considered to be characteristic of CNSLs enumerated in (10), repeated for convenience in (27), need to be critically reevaluated.

(27) Distinguishing features of CNSLs (based on previous literature)
- a. grammatically unconstrained null subjects of all person/number/gender combinations
- b. 'rich' subject-verb agreement morphology
- c. free subject inversion
- d. unavailability of a generic inclusive null subject with plain third person verbal inflection in out-of-the-blue contexts
- e. lack of the that-trace effect
- f. binding differences between null and overt subject pronouns.

As reviewed in Section 2, K, P, and S all pretty straightforwardly meet criteria (27a–d). We set aside criterion (27e), which is difficult to test, as long-distance *wh*-extraction might not be productive in these languages. Given that K, P, and S all behave alike so far, a reasonable expectation, then, is for them to all meet

criterion (27f), which is what we tested for with our two experiments. Alternatively, given that K and S exhibit agreement drop and require overt subjects in more environments than P, we might expect K and S to pattern together, to the exclusion of P. However, our results do not accord with either scenario. Of the three languages, only P seems to have overt preverbal subjects whose binding possibilities are different from those of null subjects. Experiment 1, looking at the bound variable reading (the Montalbetti effect), showed that P overt pronouns cannot as easily undergo variable binding as K and S overt pronouns. On the other hand, Experiment 2, looking at whether null subjects can only be bound by a local antecedent, suggests that there is some restriction on the binding of overt subject pronouns by a local antecedent in P and S, but not in K.

5.1 Are variable binding and double embedding reliable bases for generalizations about pronoun interpretation?

One possible response to the inconsistency in the results is to maintain that overt subject pronouns in CNSLs do have marked interpretations relative to null subjects, but variable binding and/or double embedding does not provide a reliable environment for determining the binding possibilities of overt pronouns and null subjects. If that is the case, the associated experiment is confounded, and its results should be set aside.

Indeed, there are some factors that could make it difficult to reliably detect the Montalbetti effect in practice. As we mentioned in the introduction, Montalbetti (1984), among others, observes that for Spanish speakers, an overt subject pronoun can be bound by a quantificational main clause subject if the overt pronoun is focused (e.g., if stressed). In addition, when the binder is not quantificational, languages seem to vary as to whether an overt pronoun can be bound or not (Camacho 2013; Frascarelli 2018; Genevska-Hanke 2022; Montalbetti 1984). That said, we proactively took steps to eliminate these potential issues in Experiment 1: our binders were exclusively quantificational, and we made use of audio recordings, which let us make sure that the overt pronoun was not pronounced with stress that could be misinterpreted as contrastive focus. While we obviously still could not prevent our participants from perceiving that the overt pronoun received focus, we believe that the results of Experiment 1 are as reliable as they can be. Rejecting the variable binding results in favor of the double embedding results – in which case one might classify P and S, but not K, as CNSLs – would then be arbitrary.

5.2 A purely morphosyntactic approach toward the classification of CNSLs and PNSLs

Another response would be to drop the binding possibilities of overt pronouns as a criterion for classifying languages as CNSLs. Instead, the classification criteria would consist of only morphosyntactic ones like (27a–e), or even (27a–b) exclusively (if the issues raised in the references in Footnote 1 with respect to (27c–e) are insurmountable, which is beyond the scope of this paper to determine). In this analysis, K, P, S would all be classified as CNSLs.

There are several arguments in favor of this view. First, research on pronouns has suggested that overt pronouns in general can be further divided into smaller classes across and within languages, with different structural and interpretational properties. To pick just two influential proposals in this literature, Cardinaletti and Starke (1999), for instance, distinguish between 'strong' and 'weak' pronouns (which are in turn different from pronominal clitics), while Déchaine and Wiltschko (2002) propose that overt pronouns can be DPs, ϕPs, and NPs. While both proposals differ in technical details, they agree that overt pronouns can vary in prosody, structure, and importantly, semantics (see also, among others, Noguchi 1997; Patel-Grosz 2020; Patel-Grosz and Grosz 2017).[7] From this perspective, it is unsurprising to find that overt pronouns do not behave uniformly in K, P, S, although we must leave a detailed analysis of overt pronouns in these languages to future research.[8] We only note here that in light of structure-based constraints such as Cardinaletti and Starke's (1999) *Minimise Structure* or Patel-Grosz and Grosz's (2017) *Minimize DP!*, a language in which null and overt pronouns have equally economical structures (e.g., the same number of syntactic nodes) could be expected to evade (27f) irrespective of its *pro*-drop type.

Relatedly, there appears to be an implicit assumption in the literature on Consistent and Partial NSLs that overt subject pronouns are such that they must either allow variable binding and also corefer freely with the higher subjects in double embedding (in the case of PNSLs) or not do so at all (in the case of CNSLs). However, this assumption is called into question when we consider a wider range of languages. Mandarin Chinese overt subject pronouns, for instance, can be bound by

7 A question which arises in this context with respect to the analysis which takes agreement to be pronominal in CNSLs (see Alexiadou and Anagnostopoulou 1998; Barbosa 1995, 2011a and references therein) is whether we would not expect similar types of differences in the composition of pronominal agreement, resulting in different interpretive properties of subjects in these languages.

8 While we can note that just like P, K and S have both full and reduced forms in some cells of their personal pronoun paradigms (e.g., 3 SG.M.DAT: *jemu/mu*$_P$, *jemù/mù*$_K$, *łónymu/mu*$_S$), to the best of our knowledge in-depth descriptions of their syntactic and semantic properties are not available at present.

(corefer with) either of the higher subjects in double embedding; in this respect Mandarin behaves like a PNSL (28). But Mandarin has also been reported to show a Montalbetti effect, in that an overt subject pronoun cannot undergo variable binding, much like in a CNSL (29) (Huang 1991; see also Aoun and Li 1990 for a more nuanced discussion). While we also do not have much to say about overt pronouns in Mandarin, we believe that these observations provide further evidence that overt pronouns across languages are a heterogeneous class of lexical items in their own right. Given their diversity, overt pronouns might not always stand in clear contrast with null subjects, contrary to what is usually assumed in the CNSL/PNSL literature.

(28) *Zhangsan*$_1$ *shuo* *Lisi*$_2$ *juede* *ta*$_{1/2/3}$ *bu* *neng* *lai.*
Zhangsan say Lisi feel 3SG not can come
'Zhangsan said that Lisi feels that s/he (Zhangsan, Lisi, or someone else) can't come.' [Mandarin Chinese]

(29) *Meiyou* *ren*$_1$ *shuo* *(*ta*$_1$*)* *hen* *xihuan* *Lisi.*
no person say 3SG very like Lisi
'No one says that s/he likes Lisi.' [Mandarin Chinese, Huang 1991, ex. (47)]

Finally, there is also a conceptual argument for dropping the interpretation of overt pronouns from the list of criteria in (27): while the properties in (27a)–(27d) clearly reflect the formal, morphosyntactic features of the language, the interpretation of overt pronouns is a matter of the interface between syntax and information structure, where the alignments clearly vary across CNSLs (see Jasinskaja and Simík forthcoming for a discussion of perspectives from Slavic) and where judgments and preferences are especially susceptible to influences from language processing (see, among many others, Fernandes et al. 2018; Filiaci et al. 2014; Sorace 2005; Sorace et al. 2009; Wolna et al. 2022). In this domain, we note that CNSLs can differ in the interpretation of overt subjects, even non-pronominal ones. The quantifier interaction data in (30) provide a concrete example: these Polish sentences are natural with a narrow scope interpretation of the subjects, i.e. the intended, common-sense interpretations involving different handsome men, policemen, and lottery winners (this interpretation is facilitated by placing stress on *jakiś* 'some'/*ktoś* 'someone'). In contrast, parallel structures are deviant in Greek, another CNSL, and are offered to support an A-bar analysis of preverbal subjects, in which such subjects always receive an obligatory wide scope interpretation (Alexiadou and Anagnostopoulou 1998: 505–506). Thus, if the logic of the argument holds, Polish (but not Greek) preverbal subjects can in fact be in an A-position, despite the *pro*-drop status of both languages, or, alternatively, the effect of A-bar movement on the scopal properties of the subject observed in Greek, as well as in Romance varieties (Barbosa 1995), is not universal for all CNSLs. In either case, overt subjects are not interpreted uniformly

even across the CNSLs which (on the surface) pattern alike with respect to the property in (27f).

(30) a. *Jakiś przystojny mężczyzna ożenił się z każdą z moich koleżanek w zeszłym roku.*
some handsome man married SE with every from my friends in last year
'Some handsome man married each friend of mine last year.' ("each" > "some")

b. *Jakiś policjant stał przed każdym domem zeszłej nocy.*
some policeman stood in.front.of every house last night
'Some policeman stood in front of every house last night.' ("every" > "some")

c. *Ktoś wygrał każdą loterię w zeszłym roku.*
someone won every lottery in last year
'Someone won every lottery last year.' ("every" > "some")

While investigating data of this type goes beyond the scope of this paper, they further underscore the need for much more detailed comparative research into subject interpretation in CNSLs before the correlations observed in the current literature can be lifted to the status of reliable generalizations expected to apply beyond the language types investigated so far. Clearly, various independent aspects of the grammars of particular languages can be expected to influence the interpretive options available to both overt and null pronouns, diminishing the categorizing potential of pronoun interpretation in diagnosing null subject language types (in addition to the structure of the pronouns themselves, potentially, e.g., free/fixed word order, expression of information structure, presence of CLLD, presence of articles; see, e.g., Barbosa 2019 and references therein for a discussion of interpretive differences between subjects in discourse *pro*-drop languages).

To the extent that typological classifications such as the one in (31) from Barbosa (2019: 487–488) can be theoretically informative, we suggest that if a language fulfills the description specified in (31) for CNSLs, this should be sufficient to classify it as such, regardless of its patterns and frequencies of use of overt pronouns (see also Genevska-Hanke 2022). In other words, a CNSL is a language in which *pro*-drop is licensed by rich agreement (in one way or another, see Barbosa 1995, 2011a, 2019; Camacho 2013; Cole 2010; Holmberg 2005, 2010a; Roberts 2010, 2019; Saab 2020, 2024; Sigurðsson 2011).

(31) Null subject (pro-drop) languages
– **Consistent null subject languages:** languages with rich subject agreement morphology, such as Italian, Greek, among others; subjects are freely dropped under the appropriate discourse conditions.

- **Partial null subject languages:** languages with agreement and referential null subjects whose distribution is restricted, such as Hebrew, Finnish, Russian, Brazilian Portuguese.
- **Semi-pro-drop languages:** languages that only have impersonal and quasi-argumental null subjects (Icelandic, Faroese, a range of creoles).
- **Discourse pro-drop languages:** languages that lack agreement, such as Chinese, Japanese and Korean. These have been described as allowing any argument to be dropped, not just subjects.

A consequence of this definition is that drawing the fine line between consistent and partial *pro*-drop can now be based on purely formal considerations: languages in which verbal inflection is derivationally responsible for the distribution of null subjects are CNSLs. This sets them apart from PNSLs, where several alternative derivational mechanisms have been proposed: null NP anaphora (Barbosa 2013, 2019) or control (Holmberg 2005, 2010a; Holmberg et al. 2009; Holmberg and Sheehan 2010); and see Frascarelli (2018); Frascarelli and Jiménez-Fernández (2019) for another view where CNSLs and PNSLs exist on a continuum.

Focusing on the mechanism rather than the surface distribution of null and overt subjects may in some cases be crucial for categorizing a language one way or another, as this provides flexibility to impose additional formal restrictions on *pro*-drop even in CNSLs. To illustrate this point, suppose we use only surface distribution as a criterion (contrary to what we propose). If that is the case, we might conclude that Italian, the generative poster child of CNSLs, is not a CNSL, since a null [2SG] subject cannot be freely dropped – it is blocked in the subjunctive in the language due to inflectional syncretism (Cardinaletti 1997; Cognola and Casalicchio 2018b; Roberts 2019). In contrast, taking into account the derivational mechanism generally available in a language lets us accommodate additional restrictions on *pro*-drop (such as inflectional syncretism), and in doing so, avoid an intuitively incorrect shift in the classification of languages like Italian. On this approach, inflection-based constraints can in fact provide evidence for the role of agreement in licensing *pro*-drop in a language, supporting its classification as a CNSL.

Importantly, viewing cross-linguistic differences in *pro*-drop in terms of derivational mechanisms rather than purely surface properties is a crucial component of classifying languages such as K and S. As Section 2 noted, there is a correlation between *pro*-drop and agreement drop. Since this correlation basically boils down to inflectional syncretism arising from agreement drop, it supports the view on which the derivation of null subjects in K and S is agreement-based (in addition to the availability of null subjects of all person/number/gender combinations in both matrix and embedded clauses), implying that they belong to the class of CNSLs.

6 Conclusions

In this paper, we have considered coreference data from K, P, and S in light of the defining features of CNSLs postulated in some previous literature, showing that the binding of overt subject pronouns relative to null pronouns does not necessarily pattern together with the purely grammatical criteria. Jointly with the observations about Mandarin Chinese from the literature, the two experimental studies presented here demonstrate that, given the two phenomena under investigation (variable binding and pronominal coreference under double embedding), all four logically possible patterns of interpretation of overt subject pronouns are attested, as summarized in Table 2.

Table 2: Overt subject pronoun interpretation.

	Kashubian	Polish	Silesian	Mandarin
Variable binding	OK	*	OK	*
Binding by local antecedent	OK	*	*	OK

This result is consistent with an independent body of work showing that overt pronouns cross-linguistically vary in structure and interpretation. In the context of null subjects, this result further suggests that the interpretations of overt subject pronouns might not necessarily stand in opposition to those of null subjects, and so cannot be considered as a reliable diagnostic in *pro*-drop typology. Faced with this conclusion, it might be empirically and conceptually simpler if morphosyntactic criteria – specifically, what derivational mechanism yields null subject structures – were to form the backbone of language classification in this domain.

Acknowledgments: We are grateful to our RAs and translators Marika Jocz, Grzegorz Kulik, Magdalena Piosek, Bartłůmjej Wanot, and Paulina Węsierska, without whose help the experimental studies presented here would not have been possible.
Research funding: This research was supported by National Science Centre, Poland, grant 2018/31/D/HS2/00130 and by a National University of Singapore faculty start-up grant.
Data availability: The data generated and analyzed during this study are available on the OSF platform at https://osf.io/52cx8/.

References

Alexiadou, Artemis & Elena Anagnostopoulou. 1998. Parametrizing Agr: Word order, verb-movement and EPP-checking. *Natural Language and Linguistic Theory* 16. 491–539.

Aoun, Joseph & Yen-hui Audrey Li. 1990. Minimal disjointness. *Linguistics* 28. 189–203.

Barbosa, Pilar. 1995. *Null subjects*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Barbosa, Pilar. 2009. Two kinds of subject *pro*. *Studia Linguistica* 63. 2–58.

Barbosa, Pilar. 2011a. *Pro*-drop and theories of *pro* in the minimalist program. Part 1: Consistent null subject languages and the pronominal-Agr hypothesis. *Language and Linguistics Compass* 5. 551–570.

Barbosa, Pilar. 2011b. *Pro*-drop and theories of *pro* in the minimalist program. Part 2: Pronoun deletion analyses of null subjects and partial, discourse and semi *pro*-drop. *Language and Linguistics Compass* 5. 571–587.

Barbosa, Pilar. 2013. Partial *pro*-drop as null NP anaphora. In Yelena Fainleib, Nicholas LaCara & Yangsook Park (eds.), *Proceedings of the 41st Annual Meeting of the North East Linguistic Society*, vol. 2, 71–85. Amherst, MA: Graduate Linguistic Students' Association.

Barbosa, Pilar. 2019. *Pro* as a minimal nP: Toward a unified approach to *pro*-drop. *Linguistic Inquiry* 50. 487–526.

Barbosa, Pilar, Maria E. L. Duarte & Mary Aizawa Kato. 2005. Null subjects in European and Brazilian Portuguese. *Journal of Portuguese Linguistics* 4. 11–52.

Barski, Ewelina. 2013. *Spanish and Polish heritage speakers in Canada: The overt pronoun constraint*. London, ON: University of Western Ontario dissertation.

Biberauer, Theresa. 2010. Semi null-subject languages, expletives and expletive *pro* reconsidered. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric variation: Null subjects in minimalist theory*, 153–199. Cambridge: Cambridge University Press.

Bojadžiev, Todor, Ivan Kuzarov & Jordan Penčev. 1999. *Săvremenen bălgarsky ezik* [Modern Bulgarian language]. Sofia: Petar Beron.

Bondaruk, Anna. 2001. Pro-drop in Polish and Irish. *Poznań Studies in Contemporary Linguistics* 37. 35–52.

Borer, Hagit. 1986. I-Subjects. *Linguistic Inquiry* 17. 375–416.

Borer, Hagit. 1989. Anaphoric AGR. In Osvaldo A. Jaeggli & Kenneth J. Safir (eds.), *The null subject parameter*, 69–109. Dordrecht: Springer.

Breza, Edward. 2001. Podstawowe wiadomości z morfologii [The basics of morphology]. In Edward Breza (ed.), *Kaszubszczyzna. Kaszëbizna* [Kashubian], 125–184. Opole: Uniwersytet Opolski.

Breza, Edward & Jerzy Treder. 1981. *Gramatyka kaszubska: Zarys popularny [Kashubian grammar: A popular sketch]*. Gdańsk: Zrzeszenie Kaszubsko-Pomorskie.

Camacho, José A. 2013. *Null subjects*. Cambridge: Cambridge University Press.

Cardinaletti, Anna. 1997. Subjects and clause structure. In Liliane Haegeman (ed.), *The new comparative syntax*, 33–63. London: Longman.

Cardinaletti, Anna & Michal Starke. 1999. The typology of structural deficiency: A case study of the three classes of pronouns. In Henk van Riemsdijk (ed.), *Eurotyp*, vol. 5, pt. 1: *Clitics in the languages of Europe*, 145–234. Berlin: De Gruyter Mouton.

Carminati, Maria Nella. 2002. *The processing of Italian subject pronouns*. Amherst, MA: University of Massachusetts Amherst dissertation.

Chomsky, Noam. 1981. *Lectures on government and binding: The Pisa lectures*. Dordrecht: Foris Publications.

Chomsky, Noam. 1982. *Some concepts and consequences of the Theory of Government and Binding*. Cambridge, MA: The MIT Press.

Cichocki, Władysław. 1983. Multiple WH-questions in Polish: A two-COMP analysis. *Toronto Working Papers in Linguistics* 4. 53–71.

Cognola, Federica & Jan Casalicchio (eds.). 2018a. *Null subjects in generative grammar: A synchronic and diachronic perspective*. Oxford: Oxford University Press.

Cognola, Federica & Jan Casalicchio. 2018b. On the null-subject phenomenon: An example of successful linguistic research. In Federica Cognola & Jan Casalicchio (eds.), *Null subjects in generative grammar: A synchronic and diachronic perspective*, 1–28. Oxford: Oxford University Press.

Cole, Melvyn Douglas. 2010. Thematic null subjects and accessibility. *Studia Linguistica* 64(3). 271–320.

Déchaine, Rose-Marie & Martina Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry* 33. 409–422.

Fernandes, Eunice G., Paula Luegi, Eduardo Correa Soares, Israel de la Fuente & Barbara Hemforth. 2018. Adaptation in pronoun resolution: Evidence from Brazilian and European Portuguese. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 44. 1986–2008.

Filiaci, Francesca, Antonella Sorace & Manuel Carreiras. 2014. Anaphoric biases of null and overt subjects in Italian and Spanish: A cross-linguistic comparison. *Language, Cognition and Neuroscience* 29. 825–843.

Franks, Steven. 1995. *Parameters of Slavic morphosyntax*. New York: Oxford University Press.

Frascarelli, Mara. 2007. Subjects, topics and the interpretation of referential *pro*: An interface approach to the linking of (null) pronouns. *Natural Language and Linguistic Theory* 25. 691–734.

Frascarelli, Mara. 2018. The interpretation of *pro* in consistent and partial null subject languages: A comparative interface analysis. In Federica Cognola & Jan Casalicchio (eds.), *Null subjects in generative grammar: A synchronic and diachronic perspective*, 211–239. Oxford: Oxford University Press.

Frascarelli, Mara & Ángel L. Jiménez-Fernández. 2019. Understanding partiality in *pro*-drop languages: An information-structure approach. *Syntax* 22(2–3). 162–198.

Genevska-Hanke, Dobrinka. 2019. *Overt and null subjects in Bulgarian and in L1 Bulgarian-L2 German interlanguage*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Genevska-Hanke, Dobrinka. 2022. Subject realization in Bulgarian, a consistent null subject language: Theoretical issues and empirical facts. In Gréte Dalmi, Egor Tsedryk & Piotr Cegłowski (eds.), *Null subjects in Slavic and Finno-Ugric: Licensing, structure and typology*, 177–203. Berlin: De Gruyter Mouton.

Herbeck, Peter. 2018. Deriving null, strong and emphatic pronouns in Romance *pro*-drop languages. In Pritty Patel-Grosz, Patrick Georg Grosz & Sarah Zobel (eds.), *Pronouns in embedded contexts at the syntax-semantics interface*, 171–213. Dordrecht: Springer.

Holmberg, Anders. 2005. Is there a little pro? Evidence from Finnish. *Linguistic Inquiry* 36. 533–564.

Holmberg, Anders. 2010a. Null subject parameters. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric variation: Null Subjects in minimalist theory*, 88–124. Cambridge: Cambridge University Press.

Holmberg, Anders. 2010b. The null generic subject pronoun in Finnish: A case of incorporation in T. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric variation: Null subjects in minimalist theory*, 200–230. Cambridge: Cambridge University Press.

Holmberg, Anders, Aarti Nayudu & Michelle Sheehan. 2009. Three partial null-subject languages: A comparison of Brazilian Portuguese, Finnish and Marathi. *Studia Linguistica* 63(1). 59–97.

Holmberg, Anders & Michelle Sheehan. 2010. Control into finite clauses in partial null-subject languages. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric variation: Null subjects in minimalist theory*, 125–152. Cambridge: Cambridge University Press.

Huang, C.-T. James. 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15. 531–574.

Huang, C.-T. James. 1989. Pro-drop in Chinese: A generalized control theory. In Osvaldo Jaeggli & Kenneth J. Safir (eds.), *The null subject parameter*, 185–214. Dordrecht: Kluwer.

Huang, C.-T. James. 1991. Remarks on the status of null objects. In Robert Freidin (ed.), *Principles and parameters in comparative grammar*, 56–76. Cambridge, MA: The MIT Press.

Jaeggli, Osvaldo & Kenneth J. Safir. 1989. The null subject parameter and parametric theory. In Osvaldo Jaeggli & Kenneth J. Safir (eds.), *The null subject parameter*, 1–44. Dordrecht: Kluwer.

Jasinskaja, Katja & Radek Simík. Forthcoming. Slavonic free word order. In Jan Fellerer & Neil Bermel (eds.), *The Oxford guide to Slavonic languages*. Oxford: Oxford University Press.

Jocz, Lechosław. Forthcoming. Central and Western Lechitic: Kashubian, Slovincian and Polabian. In Jan Fellerer & Neil Bermel (eds.), *The Oxford guide to the Slavonic languages*. Oxford: Oxford University Press.

Jocz, Marika, Marta Ruda & Bartłůmjej Wanot. 2022. Konstrukcje bezosobowe z podmiotem arbitralnym i generycznym w gramatyce kaszubskiej i śląskiej na tle gramatyki polskiej [Impersonal constructions with arbitrary and generic subjects in the Kashubian and Silesian grammar against the background of Polish grammar]. *Polonica* 49. 95–113.

Lorentz, Friedrich. 1971. *Kaschubische Grammatik* [Kashubian grammar]. Hildesheim: Verlag Dr. H. A. Gerstenberg.

McShane, Marjorie J. 1999. The ellipsis of accusative direct objects in Russian, Polish and Czech. *Journal of Slavic Linguistics* 7. 45–88.

McShane, Marjorie J. 2009. Subject ellipsis in Russian and Polish. *Studia Linguistica* 63. 98–132.

Modesto, Marcello. 2000. *On the identification of null arguments*. Los Angeles: University of Southern California dissertation.

Modesto, Marcello. 2008. Topic prominence and null subjects. In Theresa Biberauer (ed.), *The limits of syntactic variation*, 375–409. Amsterdam: John Benjamins.

Montalbetti, Mario M. 1984. *After binding*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Müller, Gereon. 2006. Pro-drop and impoverishment. In Patrick Brandt & Eric Fuß (eds.), *Form, structure, and grammar. A Feststschrift presented to Günther Grewendorf on occasion of his 60th birthday*, 93–115. Berlin: Akademie Verlag.

Nillson, Barbro. 1982. *Personal pronouns in Russian and Polish: A study of their communicative function and placement in the sentence*. Stockholm: Almqvist & Wiksell International.

Noguchi, Tohru. 1997. Two types of pronouns and variable binding. *Language* 73. 770–797.

Nomachi, Motoki. 2014. On the Kashubian past tense form *jô bël* 'I was' from a language contact perspective. In Elżbieta Kaczmarska & Motoki Nomachi (eds.), *Slavic and German in contact: Studies from areal and contrastive linguistics*, 27–57. Sapporo: Slavic Research Center, Hokkaido University.

Nomachi, Motoki. 2019. Placing Kashubian on the language map of Europe. In Andrii Danylenko & Motoki Nomachi (eds.), *Slavic on the language map of Europe: Historical and areal-typological dimensions*, 453–490. Berlin: De Gruyter Mouton.

Patel-Grosz, Pritty. 2020. Pronominal typology and the *de se/de re* distinction. *Linguistics and Philosophy* 43. 537–587.

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

Pešková, Andrea. 2019. Slavic and Romance *pro*-drop in contrast. *Languages in Contrast* 19. 310–333.

Pisarkowa, Krystyna. 1969. *Funkcje składniowe polskich zaimków odmiennych* [Syntactic functions of the Polish inflecting pronouns]. Wrocław: Zakład Narowowy Imienia Ossolińskich/Wydawnictwo Polskiej Akademii Nauk.

Przepiórkowski, Adam, Mirosław Bańko, Rafał L. Górski & Barbara Lewandowska-Tomaszczyk (eds.). 2012. *Narodowy Korpus Języka Polskiego* [National Corpus of Polish]. Warsaw: Wydawnictwo Naukowe PWN.

Rizzi, Luigi. 1982. *Issues in Italian syntax*. Cinnaminson, NJ: Foris.

Rizzi, Luigi. 1986. Null objects in Italian and the theory of *pro*. *Linguistic Inquiry* 17. 501–557.

Roberts, Ian. 2010. A deletion analysis of null subjects. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric variation: Null subjects in minimalist theory*, 58–87. Cambridge: Cambridge University Press.

Roberts, Ian. 2019. *Parameter hierarchies and Universal Grammar*. Oxford: Oxford University Press.

Roberts, Ian & Anders Holmberg. 2010. Introduction: Parameters in minimalist theory. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric variation: Null subjects in minimalist theory*, 1–57. Cambridge: Cambridge University Press.

Rogowska-Cybulska, Ewa & Marek Cybulski. 2011. Je̜zyk kaszubski [Kashubian]. In Barbara Oczko, Elżbieta Szczepańska & Tomasz Kwoka (eds.), *Słowiańskie je̜zyki literackie: Rys historyczny*, 29–42. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.

Ruda, Marta. 2017. Rich agreement and dropping patterns: *Pro*-drop, agreement drop, no drop. In Yohei Oseki, Masha Esipova & Stephanie Harves (eds.), *Formal Approaches to Slavic Linguistics 24, The New York University Meeting*, 321–341. Ann Arbor, MI: Michigan Slavic Publications.

Ruda, Marta. 2022. Interpreting null subjects in Polish: Against left-peripheral linking. In Gréte Dalmi, Egor Tsedryk & Piotr Cegłowski (eds.), *Null subjects in Slavic and Finno-Ugric: Licensing, structure and typology*, 105–132. Berlin: De Gruyter Mouton.

Ruda, Marta. 2024. On impersonals in Kashubian, Polish, and Silesian: Generic modals and [3sg] in null subject languages. *Journal of Slavic Linguistics* 31. 1–20.

Ruda, Marta, Bartłůmjej Wanot & Marika Jocz. 2022a. Teksty kaszubskie [Kashubian texts]. OSF. Available at: https://osf.io/q49y6.

Ruda, Marta, Bartłůmjej Wanot & Marika Jocz. 2022b. Teksty śląskie [Silesian texts]. OSF. Available at: https://osf.io/ypw4v.

Saab, Andrés. 2020. Remarks on Oku's generalization: Anti-agreement and subject deletion in Spanish and Japanese. *Temas em Estudos Formais da Linguagem* 6(1). 14–40.

Saab, Andrés. 2024. Agree, agreement dissociation and subject ellipsis: Towards a new characterization of the Null Subject Parameter. *Probus* 36. 217–250.

Saito, Mamoru. 2007. Notes on East Asian argument ellipsis. *Language Research* 43. 203–227.

Salminen, Tapani. 2010. Europe and the Caucasus. In Christopher Moseley (ed.), *Atlas of the world's languages in danger*, 32–42. Paris: United Nations Educational, Scientific & Cultural Organization.

Sheehan, Michelle. 2010. Free' inversion in Romance and the null subject parameter. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric variation: Null subjects in minimalist theory*, 231–262. Cambridge: Cambridge University Press.

Sigurðsson, Halldór Ármann. 2011. Conditions on argument drop. *Linguistic Inquiry* 42. 267–304.

Sorace, Antonella. 2005. Syntactic optionality at interfaces. In Leonie Elise Alexandra Cornips & Karen P. Corrigan (eds.), *Syntax and variation: Reconciling the biological and the social*, 46–111. Amsterdam: John Benjamins.

Sorace, Antonella, Ludovica Serratrice, Francesca Filiaci & Michela Baldo. 2009. Discourse conditions on subject pronoun realization: Testing the linguistic intuitions of older bilingual children. *Lingua* 119(3). 460–477.

Szczegielniak, Adam. 1999. That-trace effects' cross-linguistically and successive cyclic movement. In Karlos Arregi, Benjamin Bruening, Cornella Krause & Vivian Lin (eds.), *Papers on morphology and syntax, cycle one*, 369–393. Cambridge, MA: MIT Working Papers in Linguistics.

Tomioka, Satoshi. 2003. The semantics of Japanese null pronouns and its cross-linguistic implications. In Kerstin Schwabe & Susanne Winkler (eds.), *The interfaces: Deriving and interpreting omitted structures*, 321–339. Amsterdam: John Benjamins.

Treder, Jerzy (ed.). 2006. *Język kaszubski. Poradnik encyklopedyczny* [Kashubian. An encyclopedic handbook]. Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego & Oficyna Czec.

Willim, Ewa. 1989. *On word order: A government-binding study of English and Polish*. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.

Willim, Ewa. Submitted. A dedicated null impersonal pronoun in Polish and the Null Subject Parameter. Ms., Jagiellonian University in Kraków.

Wolna, Agata, Joanna Durlik & Zofia Wodniecka. 2022. Pronominal anaphora resolution in Polish: Investigating online sentence interpretation using eye-tracking. *PLoS One* 17. 1–28.

Zehr, Jeremy & Florian Schwarz. 2018. PennController for Internet Based Experiments (IBEX). Available at: https://doi.org/10.17605/OSF.IO/MD832.

Supplementary Material: This article contains supplementary material (https://doi.org/10.1515/ling-2024-0109).