Subject Pseudo-incorporation in Laz

Yağmur Sağ, Ömer Demirok, and Muhammet Bal

1. Introduction

Laz, an endangered South Caucasian language, combines features of theoretical interest such as active-ergative case alignment, phi-agreement with objects (besides subjects), and oblique subjects (e.g., Holisky 1991, Lacroix 2009, Öztürk & Pöchtrager 2011, Tuite 2017). Based on novel data from the Pazar (Atina) dialect of Laz, this study reports yet another phenomenon to be available in the language: subject/agent pseudo-incorporation.[1] Documenting how this interacts with the case system and the phi-agreement system in the language, we argue for a unified analysis which postulates a null expletive subject in case of incorporation. Our proposal also resolves a long-standing problem that subject/agent pseudo-incorporation in Turkish posits for a dependent-theoretic account of accusative.

In the following section, we start by demonstrating that Laz allows subject pseudo-incorporation. In Section 3, we discuss the impact of subject pseudo-incorporation on the ϕ-agreement patterns and outline the core puzzle. In Section 4, we propose a syntax for pseudo-incorporation and present the null expletive analysis. Section 5 provides further support based on the unavailability of passivization along with agent/subject pseudo-incorporation. Section 6 is dedicated to concluding remarks and discusses consequences of our analysis for Turkish.

2. Evidence for subject pseudo-incorporation

Case alignment in the Pazar/Atina dialect of Laz is active-ergative in the sense of Woolford (2015) and differentiates external arguments from internal arguments, as exemplified in (1). The subject of a transitive or an unergative verb is marked with the ergative case suffix. The subject of an unaccusative verb and the object of a transitive verb are arguably inflected with a null nominative case marker.

(1) a. laç'i-**k** ts'ari ş-um-s
dog-ERG water.NOM drink.IMPF-PRS.3SG
'The dog is drinking water.' transitive

b. bere-**k** k'i-am-s
child-ERG yell-IMPF-PRS.3SG
'The child is yelling.' unergative

c. ts'ari kor-un
water.NOM get.cold-IMPF.PRES.3SG
'The water is cooling down.' unaccusative

We argue that Laz allows the incorporation of subject arguments, direct evidence of which comes

* Yağmur Sağ, Harvard University, yagmur_sag-parvardeh@fas.harvard.edu. Ömer Demirok, Boğaziçi University, omerfaruk.demirok@bogazici.edu.tr. Muhammet Bal, Boğaziçi University, muhammet.bal@std.bogazici.edu.tr We thank the audiences at Tu+8 and WCCFL 41 for their valuable feedback. We are also grateful to Tanya Bondarenko and Jonathan Bobaljik for their insightful discussions. Any remaining errors are ours.

[1] The third author of this paper is a native speaker of Laz besides being a well-trained linguist. The variety of Laz he speaks, from which the data reported here comes from, is spoken in Pazar. More comprehensive fieldwork is needed to see to what extent subject/agent pseudo-incorporation is attested across different Laz varieties.

from its effect on ergative case.[2] The example in (2-b) contrasts with the one in (2-a) in that the subject lacks the ergative case marker. Furthermore, the caseless (i.e., morphologically unmarked for case) subject needs to occupy the immediately preverbal position, which is evidenced by its inability to be separated from the verb, as shown in (2-c).[3]

(2) a. laç'i-**k** bere-s goyo-k'ap'-u
dog-ERG child-DAT over-attack-PST.3SG
'The dog attacked the child.' canonical subject

b. bere-s laç'i goyo-k'ap'-u
child-DAT dog over-attack-PST.3SG
'One or more dogs attacked the child.' PI'ed subject

c. *laç'i bere-s goyo-k'ap'-u
dog child-DAT over-attack-PST.3SG

Further evidence supporting the presence of subject incorporation in Laz is observed in the interpretation of caseless subjects that directly precede the verb. These subjects exhibit semantic characteristics typical of incorporation, including number neutrality, narrow scope indefinite interpretation, and compliance with the so-called *name-worthiness* requirement (Mithun 1984, Bittner 1994, Chung & Ladusaw 2003, Dayal 2011, a. o.).

To see the case of number neutrality first, let us compare the examples in (2) one more time. The subject noun *laç'i* 'dog' in the regular transitive construction given in (2-a) refers to a unique dog that is familiar (i.e., part of the common ground) and hence yields a definite singular interpretation. In contrast, the caseless subject in (2-b) yields a number-neutral interpretation, referring to one or more dogs, the identity of which are not necessarily part of the common ground. In fact, the best translation that reflects the interpretation of (2-b) is 'The child got dog-attacked', as is commonly done in the incorporation literature.

The narrow scope property is illustrated with the example in (3), where the caseless subject is interpreted under the scope of negation. That is, the sentence in (3) is judged true if no dogs attacked the child and false if some or other dog(s) attacked the child.

(3) bere-s laç'i var goyo-k'ap'-u
child-DAT dog NEG over-attack-PST.3SG
'No dogs attacked the child.' (#some dogs > not)

In short, while an agent NP that is unmarked for number yields a definite singular interpretation when occupying a case-marked argument position, a caseless agent NP that is (necessarily) immediately preverbal is construed number-neutrally and exhibits a narrow scope indefinite behavior.

One additional hallmark of incorporation is the *name-worthiness* requirement, a definedness condition that permits incorporation only when the resulting construction conveys a canonical activity or situation type (see Mithun 1984 and Dayal 2011). The name-worthiness presupposition has a direct impact on the modification of the incorporated noun, restricting it to certain adjectives that contribute to describing a canonical activity type.

In compliance with this requirement, the modification of the incorporated subject *xirsuzi* 'thief' with the adjective *usta* 'master' is felicitous in (4-a). In contrast, the modification with *çuntu* 'fat' is infelicitous in the intended incorporation reading, while the modified subject can instead refer to a unique and familiar thief who is fat, as shown in (4-b).[4] This is because an event where a house is burglarized by master thieves

[2] Object incorporation is also permitted in Laz; however, this paper only focuses on subject incorporation.

[3] Note that the object in (2) is lexically DAT-marked. We call this lexical case, for objects normally appear caseless (i.e., unmarked for case, nominative). These data are important in showing us that the requirement that the caseless subject occupy the immediately preverbal position is not a consequence of two caseless NPs being in the same clause.

[4] The sentence in (4-b) is still grammatical because the verb is unaccusative. The regular subjects of unaccusatives are assumed to be in the null nominative form as stated above. Furthermore, Laz is a scrambling language, which allows the subject to occur in front of a verb for information structural reasons.

could be considered a canonical situation, plausibly a preferred one from the perspective of thieves. On the other hand, the fatness of the thieves breaking into the house does not, in any obvious way, contribute to the burglary. Hence, modification by *çuntu* does not result in a name-worthy situation in (4-b).

(4) a. Ham oxori-şa *usta* xirsuzi ama-xt-u.
this house-ALL master thief IN-go-PST.3s
Lit.: 'Master thief-entering happened to this house.'
'One or more master thieves broke into this house.' PI → number-neutral

b. Ham oxori-şa *çuntu* xirsuzi ama-xt-u.
this house-ALL fat thief.NOM IN-go-PST.3s
'The fat thief broke into this house.'
Not: 'One or more fat thieves broke into this house.' No PI → definite singular

On a final note, we should emphasize that what we are dealing with is an instance of pseudo-incorporation, not canonical head incorporation.[5] The difference between the two phenomena lies in whether the incorporated noun forms a morphological complex with the verb or not.

Recall that an incorporated subject needs to be immediately preverbal in Laz, which still does not exclude the possibility of head incorporation. This point is illustrated once again in (5), with the inability of an adverb to intervene between the subject and the verb:

(5) doktori*(**-k**) ğoma mi-yox-u
doctor-ERG yesterday 1.OBJ-call-PST.3SG
'The doctor called me in yesterday.'
Not: 'I got doctor-called yesterday.'

However, certain particles, such as the additive particle *ti* can cliticize on the incorporated noun, as shown in (6). Since such particles have phrasal attachment, the fact that they do not block incorporation argues against head incorporation. Furthermore, as we have shown in (4-a) above, an incorporated subject noun accepts adjectival modifiers, showing that subject incorporation in Laz is an XP-level process rather than head incorporation. See also Öztürk (2009) for parallel facts on agent incorporation in Turkish.

(6) ğoma doktori **ti** mi-yox-u
yesterday doctor TOO 1.OBJ-call-PST.3SG
'I also got doctor$_F$-called yesterday.'

These properties, which are unexpected of arguments that undergo head incorporation, show that incorporated subjects retain their phrasal status in Laz and hence they are pseudo-incorporated arguments.

3. Subject pseudo-incorporation and ϕ-agreement

Our core puzzle is how subject pseudo-incorporation (PI, henceforth) affects the transitivity of the clause. In essence, we seek to understand whether a PI'ed subject retains its argument status, or clauses with subject PI are inherently intransitive. The evidence from agreement patterns suggests the presence of a subject in the clause when the agent is PI'ed. Notably, Laz exhibits ϕ-agreement with both objects and subjects, and yet incorporating the subject does *not* eliminate canonical object agreement with theme NPs, indicating that a transitive structure is maintained.

3.1. Background on ϕ-agreement in Laz

The morphological loci of ϕ-agreement are prefixal and suffixal in Laz. Prefixal person agreement, which is tense-invariant, prioritizes *m-set* markers for participant objects, otherwise, hosts *v-set* markers for subjects. Due to reasons of space, we set aside number agreement as well as suffixal person agreement

[5] The term "pseudo-incorporation" is due to Massam (2001). See Mithun (1984) and Baker (1988) for the phenomenon of head incorporation.

in this paper, as they are largely orthogonal to the argumentation. See Atlamaz (2013), Demirok (2013), Blix (2021), and Bondarenko & Zompì (2021) for various proposals on ϕ-agreement in Laz.

(7) m-set agreement

a. **m**- dzir -am -s
1.OBJ- see -IMPF -PRS.3SG.SUBJ
'S/he sees me.'

b. **g**- dzir -am -s
2.OBJ- see -IMPF -PRS.3SG.SUBJ
'S/he sees you.'

(8) v-set agreement

a. **b**- dzir -am -∅
1.SUBJ- see -IMPF -PRS(.1/2.SG.SUBJ)
'I see him/her/it.'

b. ∅- dzir -am -∅
2.SUBJ- see -IMPF -PRS(.1/2.SG.SUBJ)
'You see him/her/it.'

Single-argument verbs, i.e., unaccusatives and unergatives, exclusively exhibit v-set agreement, as shown in (9). In other words, when it comes to ϕ-agreement, single argument verbs and transitive verbs with a non-participant object are on a par with each other.

(9) a. **b**-ğurur, *__m__-ğurur
1.SBJ-die.IMPF, 1.OBJ-die.IMPF
'I am dying.' unaccusative

b. **v**-igzal, *__m__-igzal
1.SBJ-walk.IMPF, 1.OBJ-walk.IMPF
'I am walking.' unergative

To make our assumptions on ϕ-agreement explicit for the purposes of argumentation, we adopt the dependent-theoretic account of object agreement proposed in Bondarenko & Zompì (2021) (henceforth B&Z).[6] We cannot do justice to the intricacies of agreement in Laz or B&Z's proposal. In simplified terms, we borrow from B&Z the idea that the Probe is low in the structure and it searches for a goal within its complement first and then its specifier. We take the Probe to be located in the head that introduces the highest argument. If the Probe successfully copies features from two distinct NPs, we get to see *dependent agreement* (m-set markers) realizing the features of the first target of the Probe (i.e. the theme NP). If the Probe fails to find two NPs, it realizes *unmarked agreement* (i.e. v-set agreement).

This account immediately explains the data in (9). With single-argument verbs, we only observe v-set markers. In such cases, the Probe fails to find a second NP and thus realizes unmarked agreement. It also derives the *dependent agreement* in (7), given that the Probe finds a second NP in its specifier. Notably, assuming that (i) copied person features are the privative [participant] and [speaker] features, and that (ii) first person includes [participant, speaker], second person includes [participant], and third person lacks person features, the unmarked agreement in (8) also follows. The Probe realizes unmarked agreement simply because it fails to copy any features via downward-probing. This is summarized in (10) below.

(10) a. dependent agreement = m-set markers (realizes the set of phi-features copied in downward probing iff the probe finds a second NP in its specifier)

b. unmarked agreement = v-set markers (elsewhere) (after Bondarenko and Zompì 2021)

3.2. Predictions for subject pseudo-incorporation

We now return to the agreement pattern expected under subject PI. If subject PI in essence demotes the agent, resulting in an intransitive construction, we expect the prefixal agreement to host v-set markers on a par with the agreement pattern we see in single-argument verbs shown in (9) above. This is the predicted outcome if the theme ends up being the sole argument in the structure as a result of subject PI.

Contrary to this expectation, in clauses with subject PI, the prefixal agreement with the theme still

[6] See also Marantz (1991) where a dependent-theoretic account of agreement (besides case) was sketched.

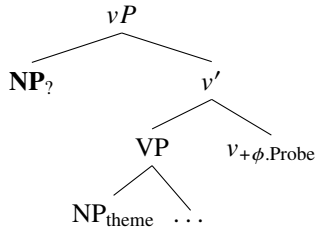
employs m-set markers, representing object agreement, as demonstrated in (11).[7] The variant with v-set agreement in (12) leads to a different interpretation, where *koncolozi* —a witch-like creature in the Anatolian folklore —is understood as the theme of "catching" rather than the agent.

(11) ham oruba-s ma k'oncolozi **m**'-ç'op-um-s
this river-LOC 1.SG *koncolozi* 1.OBJ-catch-IMPF-PRS.3SG
'In this river, I'd get *koncolozi*-caught.'

(12) ham oruba-s ma k'oncolozi **p'**-ç'op-um-∅
this river-LOC 1.SG *koncolozi* 1.SBJ-catch-IMPF-PRS(.NON3SG.SUBJ)
'In this river, I'd do *koncolozi*-catching.'
Not: 'In this river, I'd get *koncolozi*-caught.'

The fact that agreement with the theme is realized via m-set markers shows that the theme NP still counts as an object. Within B&Z's account of ϕ-agreement that we adopt, the availability of m-set agreement markers for the theme NP indicates that the Probe finds a second NP in its specifier, as sketched below. This appears to be inconsistent with the received wisdom on incorporation structures. If the NP$_?$ in (13) is the PI'ed agent itself, how do adjacency requirement, the caselessness requirement, and other semantic hallmarks of incorporation fit the picture? In the next section, we show that adopting a fairly conservative syntax for incorporation is nevertheless possible. In particular, we argue that the NP$_?$ in (13) is not the PI'ed agent but an expletive *pro*.

(13)



4. The syntax of subject pseudo-incorporation in Laz

In this section, we present our analysis to explain the transitive characteristics of a verbal structure encompassing pseudo-incorporation. We propose that a null expletive pronoun occupies the canonical agent position when the agent is incorporated.

Drawing upon the accounts put forth in previous studies by Öztürk (2005) and Sağ (2019, 2022) for Turkish pseudo-incorporation, we adopt a model wherein the verbal structure in Laz comprises two distinct domains: Following Sağ (2019, 2022) we further take pseudo-incorporation to be established via an incorporating head (Inc), which merges with a verb and a θ head (e.g., the Agent head when an agent is incorporated) to create an incorporating verbal complex, represented as θ_{Inc}. This complex verb head takes the PI'ed NP as an argument.

Furthermore, this view, adopting a neo-Davidsonian framework, assumes that both themes and agents are severed from the verb. In other words, a canonical theme argument (the direct object or the subject of an unaccucative) is introduced in the specifier position of a little v head, represented as v_{Th}, projecting above VP, and a canonical agent subject is introduced in the specifier position of a higher little v head, represented as v_{Ag}.

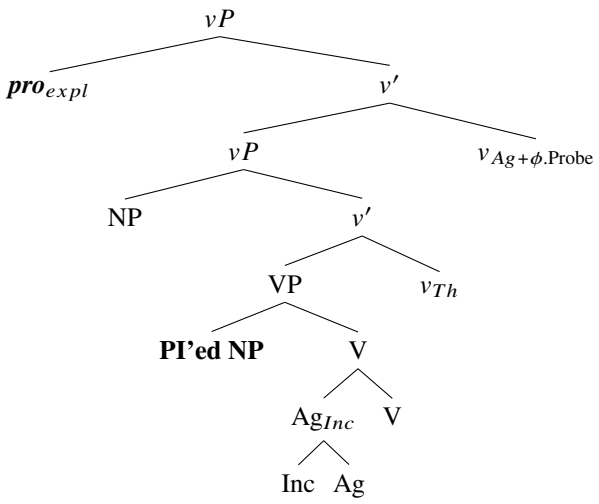
Most importantly, we propose that when subject PI occurs, the v_{Ag} head, bearing the ϕ-Probe is still part of the structure. Although it does not combine with a thematic agent NP, it still has a selectional requirement, which is satisfied by merging a null expletive pronoun in its specifier, as shown in (14).[8]

[7] Note that (11) is only grammatical under the parse where subject PI has occurred. This is because *k'oncolozi*, the agent of a transitive verb, does not bear ERG marking, which is only possible under subject PI.

[8] It is crucial for our analysis that what syntactically introduces the agent NP and what semantically introduces the thematic function looking for an agent can be disassociated in languages that can incorporate its agents. Although the details of the semantic composition need to be worked out, this presupposes voice-splitting languages in the sense of Harley (2017) to be available.

Essentially, the presence of an expletive pro enables us to retain a transitive structure when the subject undergoes pseudo-incorporation. This, in turn, allows us to explain the fact that, under subject PI, the theme NP continues to display *dependent* agreement through m-set markers. This follows from the fact that the Probe finds both the theme NP via downward-probing (i.e., in its complement) and the expletive pro in upward-probing (i.e. in its specifier) and realizes the first set of phi features it finds using dependent agreement markers, i.e., m-set markers.

(14)



5. Further support from passivization

In this section, we bring evidence in favor of the expletive analysis through passivization patterns. In a canonical passive form, the main morphosyntactic reflex of passivization is the pre-root vowel *i-* appearing on the verbal complex. Furthermore, agreement with the theme NP is no longer via m-set markers as the theme NP fails to trigger *dependent agreement* but exhibits *unmarked agreement*. Compare the passive construction in (15-b) with the active construction in (15-a).

(15) a. ma **m**-dzir-am-s
1.SG 1.OBJ-see-IMPF-PRS.3SG.SUBJ
'S/he is seeing me.' active: dependent agreement with the theme NP

b. ma **v**-**i**-dzir-er
1.SG 1.SBJ-PASS-see-PASS.IMPF.PRS.NON3SG.SUBJ
'I am being seen.' passive: unmarked agreement with the theme NP

Passivization is even possible with single-argument verbs, resulting in impersonal passive constructions. For example, the only argument in (16-a), which is *berepe* 'children' is demoted in (16-b) through passivization, resulting in an existential interpretation.

(16) a. bere-pe-k germa-pe-s k'i-am-an
child-PL-ERG mountain-PL-LOC yell-IMPF-PRS.3PL
'Children scream in mountains.'

b. Germa-pe-s **i**-k'i-en
mountain-PL-LOC PASS-yell-PASS.IMPF.PRS.3SG
'(People) scream in mountains.'

Based on these facts and the general perspective on the semantics of passivization in the literature, we take the passive prefix *i-* to signal that the (highest) argument slot is existentially saturated (cf. Taylan & Öztürk 2014, Eren 2021).

Our analysis predicts that passivization should not be available in clauses with subject PI since the

highest argument slot is occupied by a non-thematic expletive, which cannot be demoted by existential closure. This prediction is borne out, as evidenced by the ungrammaticality of the following example:

(17) *ham oruba-s ma k'oncolozi **v-i**-ç'op-er
this river-loc 1.SG *koncolozi* 1.SBJ-PASS-catch-PASS.IMPF
Intended: 'In this river, I would be *koncolozi*-caught.'

6. Concluding remarks

In this paper, we have explored verbal structure of Laz and showed that it exhibits subject pseudo-incorporation, with consequences for ϕ-agreement patterns. We have proposed that a null expletive *pro* is merged in the canonical subject position when subject pseudo-incorporation takes place.

Our study, although focused on data from Laz, carries implications for the broader cross-linguistic architecture of pseudo-incorporation. Turkish, another language renowned for allowing subject PI, poses a challenge to the Dependent Case Theory, which postulates the accusative case to be dependent on another c-commanding NP (e.g., Baker & Vinokurova 2010, Baker 2015).

(18) Dependent Case Assignment (Baker 2015: 48-49)
If there are two distinct NPs in the same spell-out domain such that NP1 c-commands NP2, then value the case feature of NP2 as *accusative* unless NP1 has already been marked for case.

To be more specific, in Turkish, when subject PI takes place, the direct object retains its accusative marking, as seen in (19-b). This is unexpected if no NP c-commands the theme argument.[9]

(19) a. Köpek Ali-**yi** ısır-dı.
dog.NOM Ali-ACC bite-PST
'The dog bit Ali.' no PI

b. Ali-**yi** köpek ısır-dı.
Ali-ACC dog bite-PST
'Ali got dog-bitten.' subject PI

The fact that the accusative marker does appear along with subject PI is fully predicted under our proposal that the syntax of subject PI crucially involves an expletive in the c-commanding spec, vP position. This null expletive ensures that the accusative case on the theme NP is still licensed even when the subject undergoes PI in Turkish.

In summary, our investigation, focusing on two cross-linguistic reflections of transitive syntax, has demonstrated that subject PI maintains a transitive structure due to the presence of an expletive occupying the canonical argument position of the PI'ed subject.

[9] Dikmen et al. (2023) identify this case-theoretic puzzle on accusative under subject PI and offer a distinct solution. We leave a comparison to future work.

References

Atlamaz, Ümit. 2013. Cyclic agreement and empty slots in pazar laz. In *Proceedings of bls37*.

Baker, Mark. 1988. *Incorporation: a theory of grammatical function changing*. Chicago: University of Chicago Press.

Baker, Mark. 2015. *Case: its principles and its parameters*. Cambridge: Cambridge University Press.

Baker, Mark C. & Nadya Vinokurova. 2010. Two modalities of case assignment: case in sakha. *Natural Language and Linguistic Theory* 28. 593–642. https://doi.org/https://doi.org/10.1007/s11049-010-9105-1.

Bittner, Maria. 1994. Case, Scope, and Binding. In *Case, Scope, and Binding*, 1–48. Dordrecht: Springer Netherlands.

Blix, Hagen. 2021. Spans in South Caucasian agreement. *Natural Language Linguistic Theory* 39. 1–55. https://doi.org/10.1007/s11049-020-09475-x.

Bondarenko, Tanya & Stanislao Zompì. 2021. *Leftover Agreement*. Ms. MIT. https://lingbuzz.net/lingbuzz/006215.

Chung, Sandra & William A. Ladusaw. 2003. *Restriction and saturation*. Cambridge, Massachusetts: The MIT Press.

Dayal, Veneeta. 2011. Hindi pseudo-incorporation. *Natural Language Linguistic Theory* 29(1). 123–167. (20 July, 2023).

Demirok, Ömer. 2013. *Agree as a Unidirectional Operation: Evidence from Laz*. M.A. thesis, Boğaziçi University.

Dikmen, Furkan, Ömer Demirok & Ümit Atlamaz. 2023. *Revisiting Agent Pseudo-Incorporation in Turkish: A Dependent Case Theoretic Perspective*. Accepted for publication in *The Linguistic Review*. https://ling.auf.net/lingbuzz/007507.

Eren, Ömer. 2021. Fake arguments as apparent valency changers: Evidence from Laz. In *Proceedings of WCCFL38*.

Harley, Heidi. 2017. The 'bundling' hypothesis and the disparate functions of little v. In R. D'Alessandro, I. Franco & Gallego Á. (eds.), *The Verbal Domain*, 3–28. Oxford: OUP.

Holisky, Dee Ann. 1991. Laz. In A. C. Greppin & A. C. Harris (eds.), *The Indigenous Languages of the Caucasus, Volume 1: The Kartvelian Languages*, 395–472. Delmar/New York: Caravan Books.

Lacroix, René. 2009. *Description du dialecte laze d'arhavi*. University of Lumiere Lyon 2 dissertation.

Marantz, Alec. 1991. Case and licensing. In *Proceedings of ESCOL'91*.

Massam, Diane. 2001. Pseudo noun incorporation in niuean. *Natural Language Linguistic Theory* 19(1). 153–197.

Mithun, Marianne. 1984. The evolution of noun incorporation. *Language* 60(4). 847–894. (20 July, 2023).

Öztürk, Balkız. 2005. *Case, referentiality, and phrase structure*. Amsterdam: John Benjamins Publishing Company.

Öztürk, Balkız. 2009. Incorporating agents. *Lingua* 119(2). 334–358.

Öztürk, Balkız & Markus Pöchtrager. 2011. *Pazar laz*. LINCOM.

Sağ, Yağmur. 2019. *The semantics of number marking: reference to kinds, counting, and optional classifiers*. Rutgers University dissertation.

Sağ, Yağmur. 2022. Bare singulars and singularity in turkish. *Linguistics and Philosophy* 45(4). 741–793. https://doi.org/10.1007/s10988-021-09323-0.

Taylan, Eser Erguvanlı & Balkız Öztürk. 2014. Transitivity in pazar laz. *Acta Linguistica Hungarica* 61(3). 271–296. http://www.jstor.org/stable/26191758.

Tuite, Kevin. 2017. Alignment and orientation in Kartvelian (South Caucasian). In J. Coon, D. Massam & L. Travis (eds.), *Oxford Handbook of Ergativity*, 1114–1138. Oxford University Press.

Woolford, Ellen. 2015. Ergativity and Transitivity. *Linguistic Inquiry* 46(3). 489–531.