Note! We've come to realize that this wasn't the right generalization: copying ("reduplication") is not triggered by the clitic =ʔa per se. We've decided to leave this up as a pedagogical demonstration of how far one can get with the wrong generalization. **Please do not quote or cite!**

Ingush clause chaining as a syntactic phenomenon

Ash Asudeh, Rochester
Oleg Belyaev, Moscow State
Bronwyn Bjorkman, Queen's
Frances Dowle, Oxford
Neil Myler, Boston
Dan Siddiqi, Carleton
Lisa Sullivan, Oklahoma State

November 17
MoMot 2024
lrfg.online

1 Introduction[1]

- In this paper, we examine clause chaining in Ingush (Nichols 2011, Conathan and Good 2001, Peterson 2001):

(1) muusaa-z [kinižka =ʔa d-iiš-aa], ğealie lota + j-a-r.
Musaa-ERG book =ʔa AGR-read-ACV cigarette light AGR-AUX-PAST
'Musa read the book and (then) lit a cigarette.' (Peterson 2001: 146, (7))

- In particular, we examine the 'stem doubling' that occurs in clause-chaining environments, as discussed by Bjorkman (2022).

(2) jett [laq =ʔa + laq-aa] b-el-ar.
cow go.dry =ʔa go.dry-ACV AGR-die-PAST
'The cow stopped giving milk and died.' (Peterson 2001: 147, (11))

- This copying has a pair of properties that make it challenging to explain:

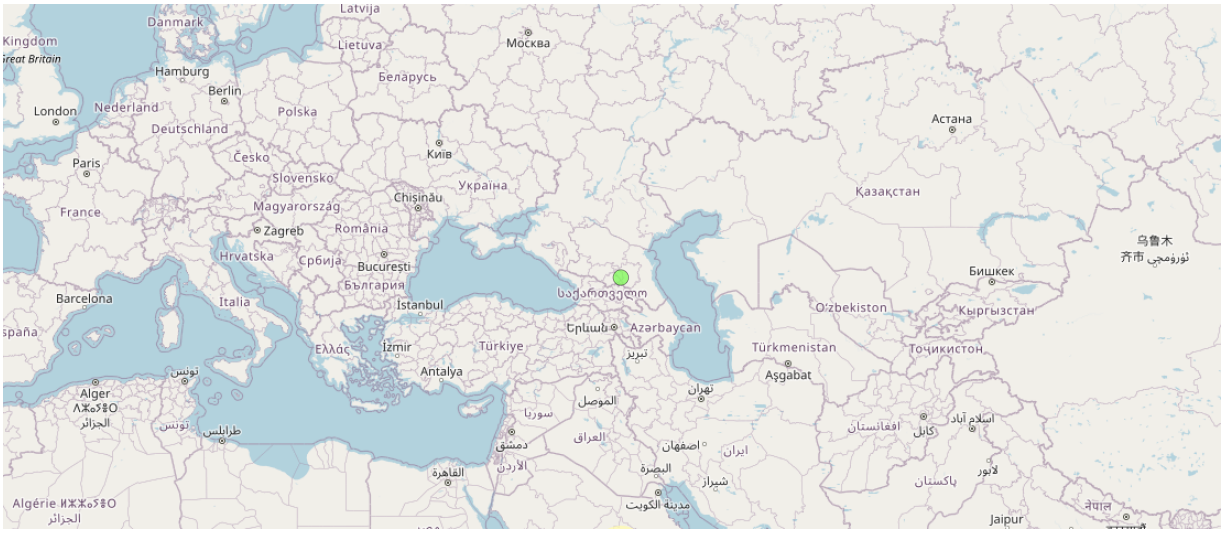
1. Irregular stem allomorphy does not appear in the copy when it appears in the 'base.'
   ∴ not pure phonology
2. The agreement morphology does appear in the copy.
   ∴ not multiple exponence

- We provide an L_RFG analysis that shows that the copying arises from a requirement for concessive clauses to have a constituent in the left-periphery.
  - When this constituent is the verb-agreement complex, it must appear both high and low.
  - This requires an analysis that handles the unique instantiation of the PRED feature, which L_RFG inherits from LFG.

[1]Glosses: ACV: anterior converb; TCV: temporal converb; ALL: allative; ADV: adverbial

1.1 Ingush: Basic characteristics

- Nakh-Daghestanian (glottolog.org)



- Remarkably strict head-finality (SOV)
- Preverbal complexes sometimes intervene between O and V

(3) ʕa= či= j-ill-aa (Peterson 2001: 148, (15))
down.PV= in.PV= AGR-put-ACV

  - Contains maximum two elements
  - Nothing can separate the elements
  - The first element involves deixis: "situates the motion of the event with respect to the speaker" (Peterson 2001: 145).
  - The second element describes the type of motion: "provides more detailed information about the nature of the motion involved" (Peterson 2001: 145).
  - Always have lexicalized meanings with particular roots, similarly to English particles (Peterson 2001: 145)
  - Other than treating the preverbal particles as a constituent (with category PVC), we do not analyze the preverbal particles.

- Highly agglutinative morphology

(4) preverb(s) = nominal = negative = agreement-root-auxiliary-negative-inflection
clitic(s) = clitic = clitic = affix-root-affix-affix-affix
(based on Peterson 2001: 145, (1))

  - We will not look at negation in this paper.

- About a third of verbs have prefixal gender agreement marking (Peterson 2001: 144)
  - Agrees with absolutive argument

1.2 The clitic =ʔa

- The clitic =*ʔa* marks clause chaining, as in (1) above, repeated here:

(1) muusaa-z [kinižka =ʔa d-iiš-aa], ğealie lota + j-a-r.
Musaa-ERG book =ʔa AGR-read-ACV cigarette light AGR-AUX-PAST
'Musa read the book and (then) lit a cigarette.' (Peterson 2001: 146, (7))

- In addition to clause-chaining, =*ʔa* has three other uses:

(5) aaz qa =ʔa b-oaqq hʕuoga. **focus**
1SG.ERG news =ʔa AGR-communicate 2SG.ALL
'And now I'll tell you some NEWS.'
(Peterson 2001: 146 (4a))

(6) je= žʕaalie-z =ʔa je= tsisk-az =ʔa dulx d-iʔ-anz-ar. **emphatic coordination**
or= dog-ERG =ʔa or= cat-ERG =ʔa meat AGR-eat-NEG-PAST
'Neither the dog nor the cat ate the meat.'
(Peterson 2001: 146 (5))

(7) ʕajšiet j-iilx-ača =ʔa muusaa v-ax-anz-ar. **concessive**
Aisha AGR-cry-TCV =ʔa Musa AGR-go-NEG-PAST
'Even when Aisha cried, Musa didn't go.'
(Peterson 2001: 145 (2))

- All uses of =*ʔa* involve "a common pragmatic element of emphasis" (Peterson 2001: 146).
- We don't provide an analysis of the non-clause-chaining uses of =*ʔa*, but the fact that they are all emphatic suggests a shared analysis.
- The copying of the verb stem can also occur in emphatic environments, especially as triggered by negation (Nichols 2011: 308).
  - Importantly, for irregular verbs, the copy has the unmarked form of the stem, whereas the base has the suppletive form of the stem.

(8) **Vala** cy **lezh** vyssaav yz?
die NEG die.SCV remain.PAST(non-witnessed) 3SG
'He didn't die?' (lit. 'He remained not dying?')
(Listener double-checks what narrator said.)

(Nichols 2011: 374, (223), as modified by Bjorkman 2022: (16); emphasis added)

2 Ingush clause chaining

- The properties of Ingush clause chaining are as follows:
  1. Both clauses *must* share a subject, which only appears high in the structure (i.e., in the matrix clause).
  2. Both clauses *can* share an object, which if shared only appears high (i.e., in the matrix clause).
  3. The verb in the subordinate clause always takes a converb inflection, which can be either simultaneous (SCV) or anterior (ACV).
- The role of =*ʔa*:
  1. The subordinator =*ʔa* attaches to the leftmost element of the subordinate clause.
  2. If the subordinate clause does *not* share its object, =*ʔa* attaches to the object.
  3. If there is no object or the object is shared, =*ʔa* attaches to the preverbal complex.
  4. If neither of the previous two conditions holds, =*ʔa* attaches to a copy of the verb-agreement complex, which *also* appears low.
- Here is example (2), repeated from above:

(2) jett [laq =ʔa + laq-aa] b-el-ar.
cow go.dry =ʔa go.dry-ACV AGR-die-PAST
'The cow stopped giving milk and died.' (Peterson 2001: 147, (11))

3 Analysis

- Example (2) suggests a possible reduplication analysis.
- This putative 'reduplication' has several different analyses in the literature: as a Type 5 clitic (Klavans 1985) that triggers actual reduplication (Nichols 2011, Peterson 2001) or as stem doubling triggered by prosodic requirements of the clitic Bjorkman (2022).
- However, fundamental aspects of the data do not support such analyses.
  - The reduplication analysis does not work because an unmarked reduplicant version of the stem appears when the base version of the stem itself is suppletive.

    (9) Muusaa, balkha ga =ʔa gejn, avtovusaa t'eħa-vysar.
    Musa work.ADV delay =ʔa delay.ACV bus miss-PAST
    'Musa was hung up at work and missed the bus.'
    (Conathan and Good 2001: 53, (12a), as modified by Bjorkman 2022: (16))
  - Conathan and Good (2001: 53) note that for irregular verbs, such as *gaa* ('delay'), *gu* ('see'), and *lie* ('die'), the copied verb is not identical to the inflected verb.

- Bjorkman (2022) notes that Conathan and Good do not provide examples for *gu* or *lie*, but notes that Nichols (2011) provides a relevant example, which is our example (8), repeated here:

  (8) **Vala** cy **lezh** vyssaav yz?
  die NEG die.SCV remain.PAST(non-witnessed) 3SG
  'He didn't die?' (lit. 'He remained not dying?')
  (Listener double-checks what narrator said.)
  (Nichols 2011: 374, (223), as modified by Bjorkman 2022: (16); emphasis added)

- This means the process cannot be phonological: it is not a phonological unit that has been copied.
- It also cannot be just stem doubling:
  - The stem-doubling analysis misses that the copy is bimorphemic: both the root and the prefixed agreement morphology occur in both positions.

    (10) k'eank-az vaaza j-uoxa =ʔa j-uoxa j-ea, aara dʕa=qess-ar.
    boy-ERG vase AGR-break =ʔa AGR-break AGR-cause.ACV outside away=throw-PAST
    'The boy broke the vase and threw it outside.'
    (Peterson 2001: 148, (17))

  - A *do*-support style multiple exponence analysis, in which a dummy element is inserted to support a stranded affix, cannot account for the presence of the re-occurring agreement morphology.
    - This problem arises for the stem doubling approach of Bjorkman (2022).

- We instead argue that the content to the left of =*ʔa* is obligatorily displaced material in SpecCP, while =*ʔa* is analyzed as a C.

(11) $\text{CP} \rightarrow \text{X} \quad \overset{\uparrow = \downarrow}{\text{C}'}$

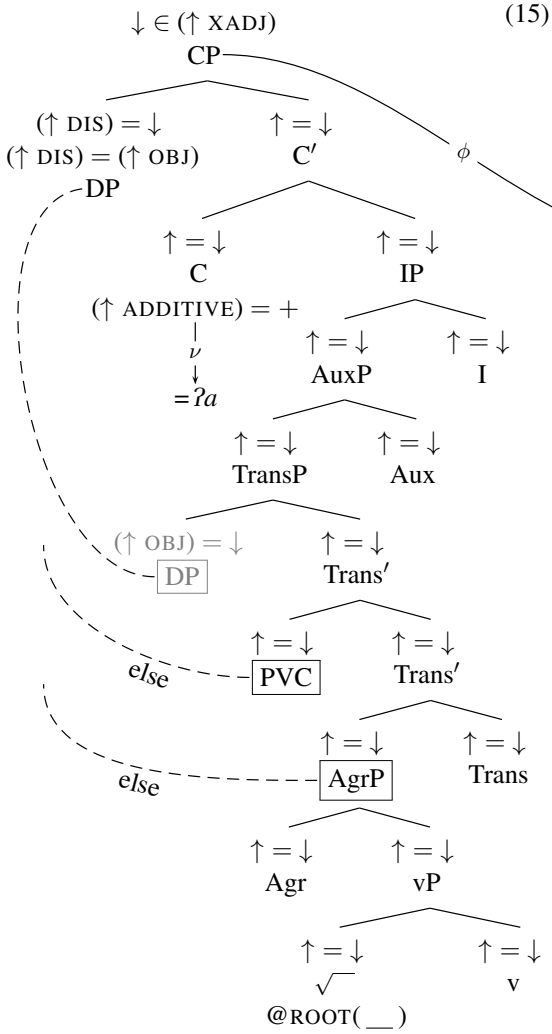
(12) $\text{C}' \rightarrow \overset{\uparrow = \downarrow}{\underset{(\uparrow \text{ADDITIVE}) = +}{\text{C}}} \quad \overset{\uparrow = \downarrow}{\text{IP}}$

(13) $\text{X} \equiv \left\{ \begin{array}{c|c|c} (\uparrow \text{DIS}) = \downarrow & (\uparrow \text{DIS}) = \downarrow & \uparrow = \downarrow \\ (\uparrow \text{DIS}) = (\uparrow \text{OBJ}) & \downarrow \in (\uparrow \text{PVS}) & \text{AgrP} \\ \text{DP} & \text{PVC} & \neg(\rightarrow \{\text{OBJ} \mid \text{PVS}\}) \\ & \neg(\leftarrow \text{OBJ}) & \end{array} \right\}$

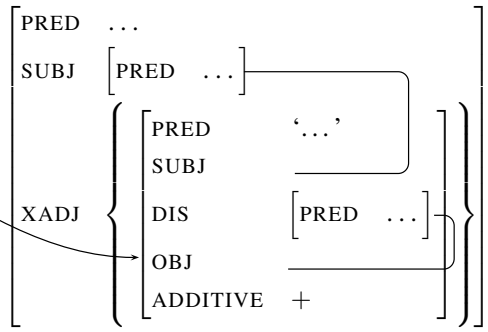
- The grammatical function that the displaced material is related to is always what would otherwise be the highest constituent below =*ʔa*'s C.
  - For example, if the displaced element is an object, its default position would have been immediately to the right of =*ʔa*'s C, as in schematic c-structure (14).
  - In (14), the lower position of the object DP is greyed out to represent where it would otherwise occur (but there is no copy or movement; this is just for presentational purposes).
  - (14) maps to the schematic f-structure (15).

- Rule (13) enforces this through the use of the *off-path* constraints $\neg(\leftarrow \text{OBJ})$ and $\neg(\rightarrow \{\text{OBJ}|\text{PVS}\})$ on the PVC and AgrP nodes, respectively.
  - The notation $(\leftarrow \text{FEAT})$ means that the f-structure that contains the annotated element must contain FEAT. Thus, $\neg(\leftarrow \text{OBJ})$ on the PVC node means that the f-structure that contains the PVC cannot also contain an OBJ.
  - The notation $(\rightarrow \text{FEAT})$ means that the f-structure that contains the annotated element must contain FEAT. Thus, $\neg(\rightarrow \{\text{OBJ}|\text{PVC}\})$ on the AgrP node means that the AgrP's f-structure cannot contain an OBJ or PVS.
  - We return to this at the end of the talk.

(14)



(15)



- Note that the first two options in the definition of the X metacategory, (13) above, the DP and the PVC cases, the constituent does not also appear low in the tree.
  - However, for the AgrP case, the constituent *does* also appear low in the tree.
  - Since AgrP ultimately contains the root that corresponds to the PRED for the subordinate clause, the doubling of AgrP creates a potential clash of PRED values.
- We must therefore state our vP rule as follows and introduce a type-token distinction between COPY and ROOT's treatments of the PRED feature.

(16) vP → ↑ = ↓ √ { @ROOT(__) | @COPY(__) } ↑ = ↓ v

(17) COPY(X) := (↑ PRED FN) = X

(18) ROOT(X) := (↑ PRED) = 'X'

- COPY specifies the *type* of PRED value by specifying only its general function label, [PRED FN].
- ROOT specifies the type of the [PRED FN] too, but the single quotes mean it is also treated as a *semantic form*, meaning that it is a uniquely instantiated *token* of the given type.

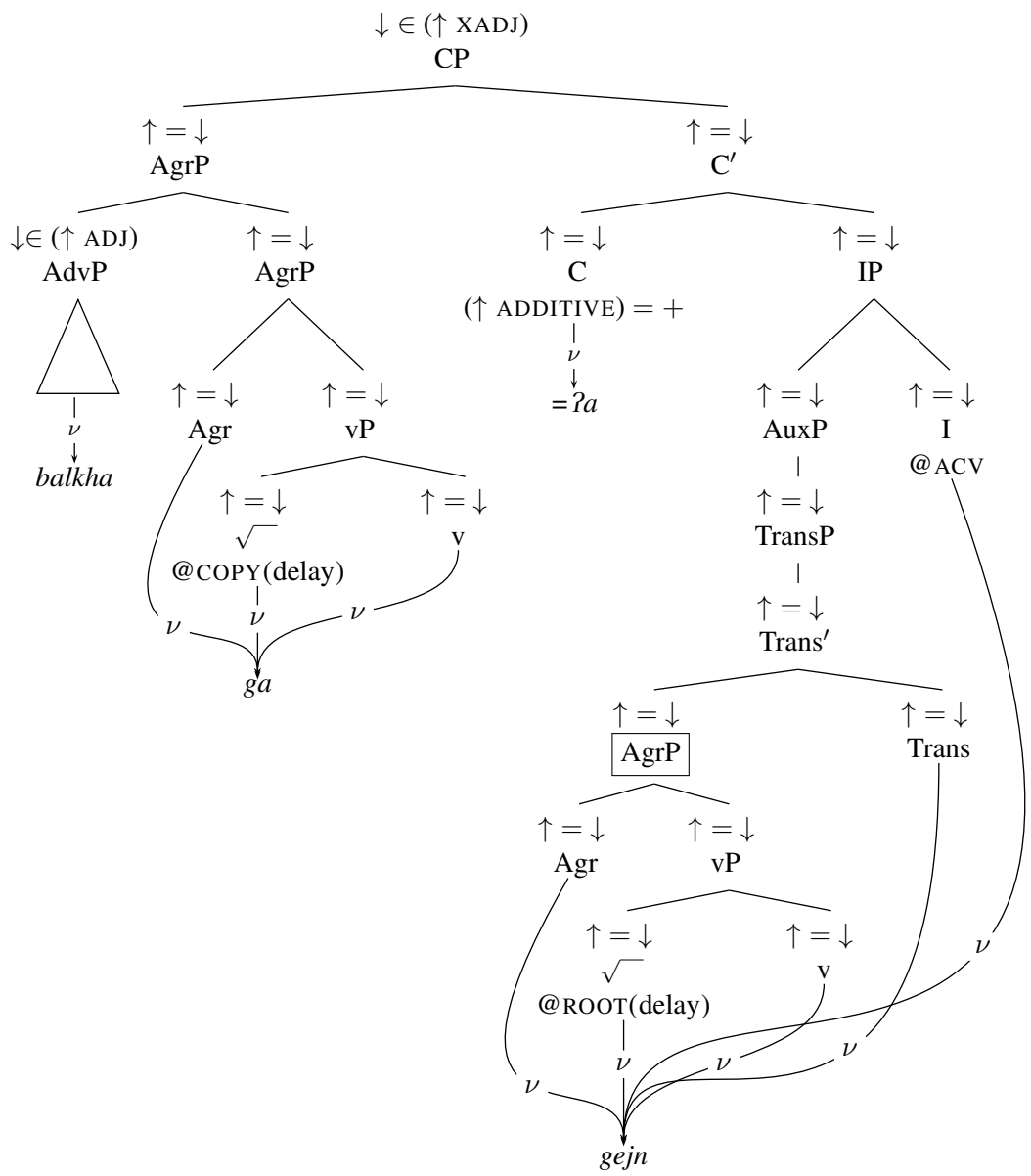
- The vP rule, (16), must only optionally call the PRED-providing ROOT macro, since if both $\sqrt{\ }$s are associated with ROOT, there would be a PRED clash.
  - Although the arguments to COPY and ROOT are underspecified, they must in fact match, because they are ultimately specifications about the same PRED feature.
    - The COPY macro specifies that COPY's argument is the type of function the PRED feature takes as its value, but without unique instantiation.
    - The ROOT macro specifies that ROOT's argument is an instantiated token of its function type.
  - Thus, if the two underspecified arguments don't happen to be the same, there will be a uniqueness violation for the value of PRED.
    - This ensures that COPY is filled by the same verb as ROOT, rather than a dummy element.
- When the highest/closest element to *=ʔa*'s right is the verb/agreement complex, the entire complex appears both high (to the left of *=ʔa*) and low (to the right of *=ʔa*, in its base position).
- We now focus on the verb doubling case exclusively.

1. Recall from above that our c-structure rules create two locations for [PRED FN], one high and one low.
   - The high instance does not add a unique index, which makes it deficient.
   - The combination of both [PRED FN] features requires that they be identical, otherwise there will be a uniqueness violation, since [PRED FN] would have two distinct values.
   - This means the two $\sqrt{\ }$s ultimately must be exponed by the same verb.
2. The system creates a situation where there are two instances of $\sqrt{\ }$, in different morphological environments.
   - One is adjacent to I and thus can be subject to a *vocabulary span*.
   - The other is not adjacent to I and thus cannot be vocabulary spanned with I.
   - The former will be exponed with an 'irregular' form, while the latter will be 'regular'.
   - This is illustrated in example (9) above, repeated here.

   (9) Muusaa, balkha ga =ʔa gejn, avtovusaa t'eħa-vysar
   Musa work.ADV delay =ʔa delay.ACV bus miss-PAST
   'Musa was hung up at work and missed the bus.'
   (Conathan and Good 2001: 53, (12a), as modified by Bjorkman 2022: (16))

   - The tree for the concessive clause in (9) is shown in (19).

(19)



- Here are the regular, (20), and irregular, (21), VIs for the relevant verbs in (9):

(20) $\langle$ [Agr, $\sqrt{}$, v], (↑ PRED FN) = delay $\rangle \xrightarrow{\nu}$ *ga*

(21) $\langle$ [Agr, $\sqrt{}$, v, Trans, I], (↑ PRED FN) = delay $\rangle \xrightarrow{\nu}$ *gejn*
@ACV

4 Conclusion

- We have modelled the copying of the verb in Ingush clause chaining environments.
- We have also shown how the agreement morphology appears in both places, which precluded a stem-doubling analysis.
- We have also shown that the suppletion facts fall naturally out of the positions of the two roots relative to the I node; these suppletion facts precluded a reduplication analysis.
- A few questions remain:

1. Why doesn't the DP copy?
   - If the DP copied, there would have to be some procedure that goes through and resolves all possible PRED clashes within the copy. Moreover, the procedure would have to be fully recursive, because DP's can have complex syntax, including embedded clauses.
   - Our solution instead is to treat the DP as an item that has been displaced in some kind of focus construction. This aligns with Peterson's claim that the clitic *=ʔa*'s host always involves "a common pragmatic element of emphasis" (Peterson 2001: 146).
2. Why doesn't the PVC copy?
   - Again, if the PVC copied, there would have to be some procedure to resolve all possible PRED clashes within the copy.
   - Again, our solution is to treat the PVC as displaced in some kind of focus construction.
3. If neither of these two other options is possible, why *must* we copy the AgrP, rather than just displacing it?
   - One possible answer is given by the framework: it is pretty trivial to copy an AgrP node and deal with its PRED clash, because it perforce contains only a single root.
     - In the case of the PVC, it potentially contains more than one root, and the DP can contain an unlimited number of roots.
       - It would be computationally much less tractable for the system/speaker/learner to employ a generalized PRED-clash resolving mechanism in the DP and PVC cases.
     - But why is AgrP copied? Ultimately this has to do with the prosodic requirements of the clitic *=ʔa* which requires something to lean on to its left, which is crucially in its own prosodic domain (precluding the matrix subject or object).
     - However, the lower copy of the AgrP must also be present for *other* affixes to attach to, per the grammar.
4. The c-structure rule in (13) accurately models but does not seem to explain why the thing that is displaced (DP or PVC) is always the thing that *would* have been closest to the low AgrP.
   - A possibly more explanatory approach is sketched in the appendix.

A Appendix

(22)
$$\left\{ \begin{array}{c|c|c} (\uparrow \text{DIS}) = \downarrow & (\uparrow \text{DIS}) = \downarrow & \\ (\uparrow \text{DIS}) = (\uparrow \text{OBJ}) & \downarrow \in (\uparrow \text{PVS}) & \uparrow = \downarrow \\ \text{DP} & \text{PVC} & \text{AgrP} \\ & \neg@\text{CLOSEST}(\text{DP}, \text{AgrP}) & \neg@\text{CLOSEST}(\text{PVC}, \text{AgrP}) \end{array} \right\}$$

(23) $\text{CLOSEST}(\text{X},\text{Y}) := n \in \phi^{-1}(\uparrow) \wedge \lambda(n) = \text{X} \wedge n' \in \phi^{-1}(\uparrow) \wedge \lambda(n') = \text{Y} \wedge \textbf{closest}(n, n')$

(24) For all c-structure nodes, n, n', n'', in the set of c-structure terminal nodes T for some c-structure, $\textbf{closest}(n, n') \Leftrightarrow \textbf{c-command}(n, n') \ \wedge \neg[\textbf{c-command}(n, n'') \ \wedge \textbf{c-command}(n'', n')] \wedge n \neq n'$

(Asudeh, Bögel, and Siddiqi 2023: 31, (8))

References

Asudeh, Ash, Tina Bögel, and Daniel Siddiqi. 2023. Modelling exponents. In Miriam Butt, Jamie Y. Findlay, and Ida Toivonen, eds., *Proceedings of the LFG23 Conference*, 23–44. Konstanz: PubliKon.

Bjorkman, Bronwyn. 2022. Verb doubling and the architecture of realization. Presented at Yale University, April 25.

Conathan, Lisa, and Jeff Good. 2001. Morphosyntactic reduplication in Chechen and Ingush. In Arika Okrent and John Boyle, eds., *The Proceedings from the Panels of CLS36*, 49–61. Chicago: Chicago Linguistic Society.

Klavans, Judith L. 1985. The independence of syntax and phonology in cliticization. *Language* 61(1): 95–120.

Nichols, Johanna. 2011. *Ingush Grammar*. No. 143 in University of California Publications in Linguistics. University of California Press. Retrieved from https://escholarship.org/uc/item/3nn7z6w5.

Peterson, David A. 2001. Ingush ʔa: The Elusive Type 5 Clitic? *Language* 77(1): 144–155.