Olaf Koeneman (Nijmegen) & Hedde Zeijlstra (Göttingen)

The lack of full pro drop as a consequence of featural overspecification

1. Introduction Apart from incidental contexts in which Germanic varieties allow partial pro-drop (cf. Fuss 2005, Rosenkvist 2007), a categorical difference between Germanic and Romance languages is that none of former allow *full* pro-drop, whereas many of the latter do. This is especially problematic for approaches to null subjects that refer to overall richness (Rizzi 1982; Koeneman 2000; Tamburelli 2006). Romance pro-drop languages are generally rich in the sense that, underlyingly, the paradigm displays all person and number features: for the language learner this means there is clear evidence for [speaker], [addressee] and [singular/plural] features, which, in turn, license pro-drop. However, Icelandic and German are equally rich in this sense, but do not exhibit null subjects. Note, that this cannot have to do with the fact that the Icelandic and German paradigms exhibit a syncretic form; Romanian, also a prototypical null subject language, also displays syncretisms. Icelandic *-ir*, German *-en* and Romanian *-a* all appear in more than one slot and are thus equally rich (taking German 3SG/2PL *-t* to be an accidental similarity). Yet only Romanian exhibits full pro-drop.

(1) Present tenses in rich Romance and Germanic languages

	Romance			*Germanic*	
	Romanian (1st conj.)	*Italian*	*Spanish*	*Icelandic*	*German*
1SG	-∅	-o	-o	-i	-e
2SG	-i	-i	-as	-ir	-st
3SG	-ă	-a	-a	-ir	-t
1PL	-ăm	-iamo	-amos	-jum	-en
2PL	-ați	-ate	-áis	-ið	-t
3PL	-ă	-ano	-an	-a	-en

2. The proposal Following traditional approaches that require pro-drop to be licensed by rich person/number agreement (the person/number features underlying the agreement form should match with those of the corresponding null subject), we argue that this is a necessary, but not a sufficient condition for pro-drop (setting aside so-called 'radical pro-drop', which is a different type of phenomenon, cf. Szendroi & Neeleman 2006). The agreement morpheme licensing *pro* cannot be featurally underspecified (as is the case for English or Dutch, which are poor agreement languages lacking pro-drop (2b, d)), but crucially cannot be overspecified either, as in (2c-d). Overspecification happens when one morpheme expresses not only agreement features, but also another feature, such as tense. We argue (contra Bobaljik & Thraínsson 1998) that this is the case in all those Germanic languages, including those that are rich w.r.t. their φ-feature specification, despite first appearances. Only if the agreement morpheme is neither underspecified nor overspecified (as in (2a)) is full pro-drop possible. The question is then how to determine the status of each language: Why are Icelandic or German not instances of (2a), where pro-drop would be allowed? The answer, here, lies in the past tense.

(2) Language types

	Subject pro	*Morpheme X (on V)*	*Status*	*Prediction*
a.	[person, number]	[person, number]	full specification	+pro drop
b.	[person, number]	[number]	underspecification	-pro drop
c.	[person, number]	[person, number, tense]	overspecification	-pro drop
d.	[person, number]	[number, tense]	under- & overspecification	-pro drop

3. Past tenses We observe that Icelandic and German, unlike their Romance counterparts, have in common that the form appearing in the 3SG present tense never reappears in the past (3). Under a privative feature system and the common assumption that the 3rd person is the result of the absence of person (and number) values (Benveniste 1971), the fact that the 3SG form does not reappear in the past can then not be captured by an impoverishment rule: there are no

feature values that can be impoverished. Crucially, this means that such elsewhere forms do not only compete with other *agreement* forms but also with the *past tense* form. To see this, look at the German paradigm for past tense. At first sight, it seems that *–t(e)-* is a separate past tense morpheme in German, spelling out [T: PAST]. Then, it would follow immediately why, say 2SG.PAST is *–test*: it is a combination of the past tense morpheme *–te* with the [φ: ADDRESSEE] morpheme *-st*. But if that were the case, 3SG.PAST should be *–tet* (*-te* + *-t*), contrary to fact. Since, neither the φ-feature can be said to have been impoverished (as it lacks φ-values to begin with) nor can *–tet* be said to be phonologically ill-formed (witness 2PL.PAST *–tet*), 3SG.PAST *–te* must be a conflated form that is the joint spell-out of [T: PAST] and [φ:], just like *–t* is an elsewhere form for tense *and* agreement, which therefore competes with other agreement and tense markers. Similarly, Icelandic 3SG.PRES *–ir* is an elsewhere form that also competes with other tense markers (such as 3SG.PAST *-ði*) and agreement markers (e.g., 1SG.PRES *-i*).

(3) Past tenses in Romance and Germanic (imperfects for the Romance languages)

	Romance			*Germanic*	
	Romanian (1st conj.)	*Italian*	*Spanish*	*Icelandic*	*German*
1SG	-a-m	-av-o	-ab-a	-ði	-te
2SG	-a-i	-av-i	-ab-as	-ðir	-test
3SG	-a	-av-a	-ab-a	-ði	-te
1PL	-a-m	-av-amo	-ab-amos	-ðum	-ten
2PL	-a-ți	-av-ate	-ab-ais	-ðuð	-tet
3PL	-a-u	-av-ano	-ab-an	-ðu	-ten

Now, if a morpheme expresses both tense and agreement features, such a morpheme is fully overspecified in the sense of (2). That means that, irrespective of the richness of the Icelandic and German paradigms, such languages cannot be full pro-drop languages. Closer inspection reveals that such overspecification holds not only for rich Germanic languages, but also for Germanic languages with poor agreement (English, Dutch, Faroese). In contrast, the Romance pro-drop languages not only exhibit a paradigm that is rich, these paradigms are also transparent in the sense that no competition between tense and agreement forms takes place. Therefore, tense and agreement are independent morphemes.

4. Analysis Naturally, the question arises as to why overspecification would forbid pro-drop. We assume, much in line with older explanations of pro-drop, that pro enters the derivation value-less but can be valued by the agreement markers. Pace Holmberg (2005), Roberts (2010), but following Pesetsky & Torrego (2007), we allow uninterpretable features on agreement markers to enter the derivation fully valued, which can, in turn, value interpretable features. If and only if, pro can be valued for the same person and number features as the corresponding subject pronouns by the agreement marker by means of feature sharing, can it be interpreted at LF. Such pronouns and pro are even undistinguishable at LF, as the major difference between them is the phonological content (apart from gender features that play no role in pro-drop licensing), so LF has no problem interpreting them. Naturally, if pro were unspecified, LF cannot interpret it in the same vein as an overt pronoun, and pro would be illegible. But, crucially, if pro contained more features than the relevant person/number features, for instance additional tense, aspect or mood features, it would be unrecognizable as a real pronoun, and the derivation would also crash. If pro carries, for instance, tense features, is no longer pronominal in nature. This explains the pattern in (2). Only if a language is rich with respect to its person/number agreement on its finite verbs, and only it has separate morphemes for agreement, tense and any other member of the verbal inflectional paradigm, can it be a full pro-drop language.

Severing case from agreement: Non-finite subjects in Hill Mari[1]

Polina Pleshak (University of Maryland, College Park)

Case is often analyzed as parasitic on agreement (Chomsky 2001, *i.a.*). This includes accounts of subject case in embedded clauses of different types, as pursued by, e.g., Miyagawa 1993 for Japanese, and Kornfilt 2005 for Turkish. For them, apparent case alternations between NOM(inative) and GEN(itive) represent different properties of agreeing heads. In this talk, I present and analyze novel data from Hill Mari (Finno-Ugric, Uralic) arguing that NOM-GEN case alternations are in fact independent of agreement. Case (NOM vs. GEN) is assigned based on the domain where the subject is located, and agreement with NOM and GEN subjects involves two different probes, each in its own domain.

Data. The primary data come from my own fieldwork in the villages of Kuznetsovo and Mikryakovo (Mari El, Russia) in 2016-2019. There is a big variation among speakers, so I concentrate on a particular group of idiolects. Non-finite forms ending in *-mâ/-mä* appear as predicates of relative clauses (RCs) (1), and as nominalizations (2). Only NOM subjects are possible in finite clauses, whereas the subject in *-mâ/-mä* clauses can be either NOM or GEN. Both case forms can occur in the absence of agreement, as shown in (1) and (2).

(1) [***mən'-ən*** / ***mən'*** *mâš-mâ*] *tâgâr-vlä(-**em**)* RC
1SG-GEN 1SG[NOM] wash-PTCP.PASS shirt-PL-POSS.1SG
'the shirts that I washed'

(2) ***tən'-ən*** / ***tən'*** *šâkâr-âm* *näl-m(-**et**)-äm* Nominalization
2SG-GEN 2SG[NOM] bread-ACC take-NMLZ-POSS.2SG-ACC
'[Your mother remembers] that you bought bread'.

Turning to agreement, Hill Mari has possessive markers on the possessum (3). The same agreement marker can appear on the head noun of the RC (1), on the nominalized verb (2), or, instead of on the head noun of the RC, on its participial verb (4).

(3) *nänä-n* / **nänä pi-*****štä*** [they-GEN / they[NOM] dog-**POSS.3PL**] 'their dog'

(4) [***tən'*** *ro-m-**et***] / [***nänä*** *ro-mâ-**štâ***] *püšängä-m*
you[NOM] cut-PTCP.PASS-POSS.2SG they[NOM] cut-PTCP.PASS-POSS.3PL tree-ACC
'[We saw] the tree cut by you / *by them'.

(5) [***täng-em-vlä(-n)*** *ro-mâ*] *püšängä-**štä**-m* *mä* *už-ân-na.*
friend-POSS.1SG-PL(-GEN) cut-PTCP.PASS tree-POSS.3PL-ACC we see-PRET-1PL
'We saw the tree cut by my friends'.

The participle cannot agree with 3rd person NOM subjects (4), whereas the head noun can (5).

Main question: How do the case & agreement interactions for these non-finite subjects arise?

Analysis: I adopt a configurational theory of case (Marantz 1991; Baker 2015). NOM is *unmarked case* in the clausal domain (TP/CP), and GEN is *unmarked case* in the nominal domain (NP/DP). *Unmarked case* is assigned to DPs that have not received case through other means. Agreement is realized in a probe-goal architecture, subject to locality (Chomsky 2000), with C being a phase head. Finally, the RCs discussed here contain CPs with a relativizing operator in Spec,CP, (6).

I argue that the precise environment for NOM is: containment in the spellout domain of C. Therefore, subjects that are in the specifier of non-finite T show up as NOM. The structure for (5) with a NOM subject is (6).

(6) $[_{DP}\ [_{nP}\ [_{CP}\ Op_j\ [_{TP}$ ***friends*****$_i$**-POSS.1SG e_j *cut*-PTCP.PASS]-$Ø_C$] $[_{nP}$ *pro*$_{i/k}$ *tree*]]-POSS.3PL]

On the other hand, the environment for GEN is: containment in the spellout domain of D. Since RCs are phases whose escape hatch is filled, GEN subjects must be base-generated outside of

[1] The work has been supported by the RFBR grant № 19-012-00627

CP, already in the nominal domain. I assume they are generated in Spec,nP (like bona fide possessors are, see Alexiadou et al. 2007: 562) and can move to Spec,DP. In fact, for my analysis, it is enough that there are two sites for possessors: before and after participial clauses, without specifying exact projections. In support of the uniform nature of possessors and GEN subjects: they cannot co-occur, even though the two realize different semantic roles. (Note the well-formedness of the NOM variant.)

(7) ***Maša(*-n) mâšk-ân šändä-mä män'-ən*** *vârgem-vlä*
Masha(*-GEN) wash-PRET set-PTCP.PASS I-GEN clothes-PL
'my shirts that were washed by Masha'

The structure for (1) with a GEN subject is (8). In those configurations where the GEN DP moves to Spec,DP, it can bind *pro*. (The idea of subjects as possessors is independently proposed in Kornfilt 2005, for Turkish. As the talk will show, the rest of her analysis is incompatible with the Hill Mari facts.) It is only in those cases that the GEN DP can be interpreted as an Agent. If the GEN DP does not move, the presence of a preceding coindexed *pro* is ruled out by independent principles, such as Rizzi's (1986) "recovery" principle or a PF rule (Royer 2020). The agreement pattern with GEN "subjects" now follows; it is the same as possessive agreement, where D probes for a target in its c-command domain. NOM subjects are invisible to D because they are properly contained within the CP phase.

To account for apparent "possessive" agreement of the head noun with a NOM subject, I assume a *pro* in Spec,nP (9), which can co-refer with the NOM subject. Co-reference is possible because *pro* and DP are not in a c-command relation, but the overt DP is a good antecedent for *pro* both for the "recovery" principle and for a PF rule. (As I will show in the talk, *tree* can also show "possessive" agreement that is distinct from the phi-features of the NOM subject, a result of *pro* not co-referring.)

(8) $[_{DP}$ *I*-GEN$_j$ $[_{nP}$ $[_{CP}$ Op$_i$ $[_{TP}$ *pro*$_j$ *e*$_i$ *wash*-PTCP.PASS]-Ø$_C$] $[_{nP}$ ~~*I*-GEN~~$_j$ *shirt*]]-POSS.1SG]

(9) $[_{DP}$ $[_{nP}$ $[_{CP}$ Op$_i$ $[_{TP}$ *I e*$_i$ *wash*-PTCP.PASS] -Ø$_C$] $[_{nP}$ *pro*$_i$ *shirt*]]-POSS.1SG]

The restriction against agreement with 3rd person NOMs (4) can also be captured. I argue that T[-fin] is a separate probe, which *never* agrees with 3rd person targets in Hill Mari. Agreement with 3rd person GENs is triggered by an entirely different probe (D). The morphological shape of agreement on D and T[-fin] is an elsewhere form, common to all agreement that is not triggered by T[fin]. However, the two originate on different syntactic heads.

The same analysis also captures subject marking and agreement in (non-RC) nominalizations, as will be shown in the talk.

Conclusions: Difference in non-finite subject case and restrictions on agreement in Hill Mari reflect underlying structural differences. NOM and GEN are each assigned configurationally in their domains: the spellout domain of C for NOM and the spellout of D for GEN. I argue for two separate base positions: clausal subjects (NOM) and possessors of nouns (GEN). The Agent/possessor ambiguity observed in RCs further supports this analysis. Agreement proceeds as each probe (T, D) looks for a goal in its own c-command domain. Taken together, these results show that case-licensing in non-finite clauses is not dependent on agreement, which in turn raises a question whether it is possible to provide a uniform account for both Hill Mari and languages like Japanese and Turkish, where case *is* claimed to depend on agreement.

References: **Baker, M**. (2015). *Case: Its Principles and Parameters.* CUP; **Chomsky, N.** (2000). Minimalist Inquiries. In Rogers et al. (Eds.), *Step by step: in honor of Howard Lasnik*, MIT Press. **Chomsky, N.** (2001). Derivation by Phase. In Kenstowicz (Ed.), *Ken Hale: A life in language*, MIT Press. **Kornfilt, J**. (2005). Agreement and its placement in Turkic non-subject relative clauses. In Cinque & Kayne (Eds.), Handbook of Comparative Syntax, OUP; **Marantz, A**. (1991). Case and licensing. In Westphal et al. (Eds.), *Eastern States Conference on Linguistics*; **Miyagawa, S**. (1993). Case-checking and minimal link condition. In: *Case and Agreement II*, MITWPL 19. **Rizzi, L.** (1986). Null objects in Italian and the theory of pro. *Linguistic inquiry, 17*(3), 501-557. **Royer, J.** (2020). *Binding and the "LOW/HIGH ABS" parameter in Mayan. NELS 51* Abstract.

Late Merger in scrambling – a Mongolian case study

Zhiyu Mia Gong Cornell University

Overview Frank, Lee, and Rambow (1996) observe the following asymmetry in German/Korean local scrambling: a locally scrambled element generally bleeds Condition C, a property typically associated with A-movement. The sole exception to this statement obtains when the binding relation is affected by scrambling involving a subject binder. I present novel data from Mongolian showing that this pattern extends to not only local but also some cases of long-distance (LD) scrambling. These data provide support for Takahashi and Hulsey's (2009) wholesale late merger (WLM), which allows an NP to late merge with D at positions where case can be assigned. I also argue that these data can be explained if we assume accusative case is assigned as dependent case.

Condition C effects in scrambling Frank et al.'s observation holds for Mongolian local scrambling. Condition C reconstruction is obligatory if the binder in the reconstructed sentence is the subject. This holds for both transitive (1) and ditransitive constructions (2).

(1)a. *$\textbf{Ter}_1$ [$\textbf{Čemeg-in}_1$ nom-ii] uragjee
3SG.NOM Č-GEN book-ACC tore
Int. 'She$_1$ tore Čemeg's$_1$ book'
b. *[$\textbf{Čemeg-in}_1$ nom-ii]$_2$ $\textbf{ter}_1$ t_2 uragjee

(2)a. *$\textbf{Ter}_1$ Bat-d [$\textbf{Čemeg-in}_1$ nom-ii] ögsön
3SG.NOM Bat-DAT Č-GEN book-ACC gave
Int. 'She$_1$ gave Bat Čemeg's$_1$ book'
b. *[$\textbf{Čemeg-in}_1$ nom-ii]$_2$ $\textbf{ter}_1$ Bat-d t_2 ögsön

In contrast, Condition C reconstruction does not occur if the binder in the reconstructed sentence is not the subject (3). Condition C reconstruction also does not happen in short scrambling (4).

(3)a. *Bagsh $\textbf{tüün-d}_1$ [$\textbf{Dorj-in}_1$ nom-ii] ögsön
Teacher 3SG-DAT D-GEN book-ACC gave
Int. (The) teacher gave him$_1$ Dorj's$_1$ book
b. [$\textbf{Dorj-in}_1$ nom-ii]$_2$ bagsh $\textbf{tüün-d}_1$ t_2 ögsön

(4)a. *Bagsh $\textbf{tüün-d}_1$ [$\textbf{Dorj-in}_1$ nom-ii] ögsön
Teacher 3SG-DAT D-GEN book-ACC gave
Int. (The) teacher gave him$_1$ Dorj's$_1$ book
b. Bagsh [$\textbf{Dorj-in}_1$ nom-ii]$_2$ $\textbf{tüün-d}_1$ t_2 ögsön

This observation extends to some cases of movement across a clausal boundary. First consider scrambling an ACC-marked (hyperraised) embedded subject. In (5a) the matrix goal argument (*he*-DAT) triggers Condition C violation on the embedded subject. In (5b), scrambling the embedded subject to matrix clause-initial position makes co-indexation between *Bat* and *he*-DAT possible.

(5) a. *Bi $\textbf{tüün-d}_1$ [[$\textbf{Bat-in}_1$ eej -iig] sain xun gej] xel-sen
I he-DAT Bat-GEN mother-ACC good person C say-PST
'I said to him that Bat's mother is a good person.'
b. [$\textbf{Bat-in}_1$ eej-iig]$_2$ bi $\textbf{tüün-d}_1$ [t_2 sain xun gej] xel-sen

Second, consider long-distance (LD) scrambling of ACC-marked embedded objects. I side with Sakamoto (2012) and argue against Fong's (2019) conclusion that Mongolian does not have LD scrambling. LD scrambling is fully possible when the embedded subject is in ACC case. When the embedded subject is in NOM case, LD scrambling is unacceptable (6).

(6) [**Ene em -iig$_2$**], emč [namaig/*bi t_2 uux xeregtei gej] xelsen
this medicine-ACC doctor I.ACC/*I.NOM t_2 drink.INF need C said
'This medicine, the doctor said that I need to drink.'

The LD scrambling exemplified in (6) can bleed Condition C. In (7a), the matrix indirect object pronoun *tüünd* ('he-DAT') may not co-index with *Bat* in the embedded accusative object. However, after the embedded object undergoes LD scrambling, co-indexation becomes possible (7b).

(7) a. *Zaya [bagsh-iig -aa [$\textbf{Bat-in}_1$ esee -g] unsh-san gej] $\textbf{tüün-d}_1$ xel-sen
Zaya teacher-ACC-REFL.POSS Bat-GEN essay-ACC read -PST C he-DAT say-PST
Int. 'Zaya said to him$_1$ that her teacher read Bat's$_1$ essay' (*him ≠ Bat*)
b. ? [$\textbf{Bat-in}_1$ esee-g]$_2$ Zaya [bagsh-iig-aa t_2 unsh-san gej] $\textbf{tüün-d}_1$ xel-sen (*him = Bat*)

The Analysis The facts in (1-7) are accounted for under Takahashi and Hulsey's Wholesale Late Merger (WLM), which allows for determiners to merge and undergo movement alone, with the restrictor merged later. Crucially, whenever WLM applies, the resulting DP must be in a position where it is assigned case. The LD scrambling data examined here provide evidence for the role of

the case assignment mechanism in WLM. Specifically, if it is case that conditions whether the NP restrictor can merge late, we should expect the Condition C effect to be sensitive to the kind of case that the DP is bearing. This prediction is borne out. In Mongolian, *like/love*-class predicates take DAT objects, which can scramble freely just like regular ACC objects. Crucially, in (8) scrambling a DAT object (*Zorig's dog*-DAT) does not bleed Condition C, in contrast to (7) where an ACC object is scrambled. Suppose ACC is assigned as structural case, whereas DAT (in (8)) as lexical case. Since lexical case is assigned in a local configuration with the lexical case assigner, the DP in (8) must be fully spelled out in its base position, triggering a Condition C violation there.

(8) *[**Zorig-iin$_1$** noxoi-d]$_2$, bi [Zaya-g t$_2$ durtai gej] **tüün-d$_1$** xel-sen (cf. (7))
[Zorig-GEN dog-DAT] I [Zaya-ACC t$_2$ like C] he-DAT say-PST
'I said to him that Zaya likes Zorig's dog' (*him ≠ Zorig*)

Second, T&H adopt the traditional view that case is assigned by designated functional heads via Agree. However, (5)/(7) cannot be straightforwardly captured with this model, since it would require WLM to apply when the scrambled arguments are outside the domain of their case assigners. For this reason I propose that ACC is assigned as dependent case within narrow syntax. I follow Baker and Vinokurova's (2010) hybrid case assignment model, where NOM is assigned by a functional head but ACC is dependent case. Analyzing [Dorji's book] as a DP headed by a covert definite determiner THE -- [THE [Dorji's book]], this model correctly predicts (1-4). In (9) the underlined NP indicates the point where WLM applies. In (9a), [THE] merges with 'gave', then moves to the site between the subject and indirect object. Here WLM of the restrictor [Dorj's book] is licensed, because ACC can be assigned to the whole DP, due to the presence of a higher argument (the subject) in the same domain. The late-merged NP successfully c-commands the pronoun indirect object. In contrast, (9d) is excluded because case can only be licensed at the intermediate landing site, and cannot be licensed when the scrambled object c-commands the pronoun subject.

(9)a. derivation of (4b): Teacher [THE **[Dorji's$_1$ book]**]-ACC he$_1$-DAT [THE] gave
b. (3b): [THE [Dorji's book]]-ACC Teacher [THE **[Dorji's$_1$ book]**]-ACC he$_1$-DAT [THE] gave
c. (1b): *[THE **[Čemeg's$_1$ book]**], she.NOM$_1$ [THE] tore
d. (2b): *[THE [Čemeg's$_1$ book]], she.NOM$_1$ [THE **[Čemeg's$_1$ book]**]-ACC Bat-DAT [THE] gave

Similarly, in order to allow WLM, LD scrambled embedded arguments in (5)/(7) must receive case at a position where they are lower than the matrix subject (for dependent case assignment), but c-command the matrix DAT pronoun (for bleeding Condition C). Although the word order is marginal, the data in (10) indicates that such a position is available.

(10) a. ??Bi **Bat-in$_1$** eej-iig **tüün-d$_1$** [t sain xün gej] xelsen. (cf. (5))
I Bat-GEN mother-ACC he-DAT [t good person C] said (he=Bat)
b. ??Zaya **Baatar-in$_1$** esee-g **tüün-d$_1$** [bagsh-iig-aa t unshsan gej] xelsen. (cf.(7))
Zaya Baatar-GEN essay-ACC he-DAT teacher-ACC-REFL t read C said (he=Bat)

More strikingly, in contrast to Frank et al.'s generalization, it is possible for a scrambled object to bleed Condition C involving a local subject binder, as long as there is a higher case competitor. In (11), the embedded subject 'he.ACC' may not coindex with *Bat* in [*Bat's medicine*] at its base position. However, LD scrambling the ACC object makes the co-indexation possible. Under the current proposal, ACC case is assigned at the intermediate landing site between the matrix subject (*bi* 'I') and the embedded subject (*tüün-iig* 'he.ACC') shown in (12). WLM is licensed at this intermediate site.

(11) [context: in a hospital, nurses are reporting to the doctors about their patients]
Bat-in$_1$ em-iig, bi [**tüün-iig$_1$** önöö öglöö **t** uusan gej] bodoj baina.
Bat.GEN medicine.ACC I he.ACC this morning t drank C think.CVB COP.NPST
'Bat's medicine, I think that he drank this morning.'

(12) [THE[Bat-in$_1$ em-iig]],bi [THE**[Bat-in$_1$ em-iig]**] **tüün-iig$_1$** önöö öglöö [THE] uusan gej bodoj baina.

References

Baker, M. C., & Vinokurova, N. (2010). Two modalities of case assignment: Case in Sakha. *Natural Language & Linguistic Theory*, *28*(3), 593-642.

Fong, S. (2019). Proper movement through Spec-CP: An argument from hyperraising in Mongolian. *Glossa: a journal of general linguistics*, *4*(1).

Frank, R., Lee, Y. S., & Rambow, O. (1996). Scrambling, reconstruction and subject binding. *Rivista di Grammatica Generativa, vol. 21*, 67-106.

Sakamoto, Y. (2012). *A study of sluicing and cleft in Mongolian: A comparison with Japanese.* MA thesis, Tohoku University.

Takahashi, S., & Hulsey, S. (2009). Wholesale late merger: Beyond the A/Ā distinction. *Linguistic Inquiry*, *40*(3), 387-426.

Against the Argument/Adjunct Distinction

Andrew McInnerney – University of Michigan

BACKGROUND: Most syntacticians today seem to assume that the argument/adjunct distinction (A/AD) has some syntactic implementation. Many authors continue to assume a more-or-less traditional (Government-and-Binding-like) approach to the A/AD. However, as noted by Harley (2007, 2014), the GB approach is difficult to implement in Bare Phrase Structure (because of the elimination of bar-level information). Consequently, some authors now adopt one of a variety of novel proposals for distinguishing arguments from adjuncts (e.g. Chomsky's "Pair Merge").

At the same time, a few authors have advanced syntactic analysis that neutralize the traditional A/AD. E.g., Larson (1988, 1990) proposed that VP-shells made it natural to analyze canonical adjuncts (like locative PPs) as "innermost complements." Jackendoff (1990) criticized Larson's proposal for "neutraliz[ing] the structural distinction between arguments and modifiers," but later Culicover & Jackendoff (2005:289) also claimed that "there is no syntactic difference between arguments and adjuncts." Despite the precedent for skepticism towards the A/AD, an attempt at a complete, empirically-based refutation of it has not yet been undertaken. That is the goal of this paper. I argue the evidence supports eliminating the A/AD in favor of alternative explanations for effects traditionally attributed to it.

EVIDENCE FOR THE A/AD: The evidence for the syntactic A/AD comes in the form of syntactic "argumenthood diagnostics." I have identified at least seven of these (from a variety of sources, including Huddleston & Pullum 2002, Goldberg 2006, Hedberg & DeArmond 2009, Needham & Toivonen 2011, Schütze 2012, Zyman 2020, *inter alia*): **[A]** *Omisssibility*: Adjuncts are (always) omissible, while arguments are (generally) non-omissible. **[B]** *Exemptability*: If the projection of a verb is targeted by a syntactic operation (like *do so*-substitution or VP-fronting), the adjuncts of that verb may optionally be excluded from the effects of that operation, while arguments may not. **[C]** *Islandhood*: Adjuncts but not arguments are islands for extraction. **[D]** *Iterability*: Adjuncts can be freely iterated, but arguments cannot. **[E]** *Condition C reconstruction*: Arguments but not adjuncts reconstruct for Condition C. **[F]** *Permutability*: Adjuncts but not arguments can be freely permuted. **[G]** *Extractability*: Arguments but not adjuncts can be extracted out of weak islands.

I evaluate each of these diagnostics by asking two questions: **[1]** Do the results of the diagnostic align with canonical categorizations? (e.g. are canonical adjuncts like locatives/temporals/ benefactives treated one way, and canonical arguments like goals/recipients treated another way?) **[2]** Are the distinctions drawn by the diagnostic predicted by independently motivated grammatical properties, in the absence of the A/AD? (By Occam's Razor, if a diagnostic motivates the A/AD, it must identify a distinction *not already expected* by properties independent of the A/AD.)

EVALUATING THE DIAGNOSTICS: Space does not allow detailed discussion of all the diagnostics in the abstract, so I will focus on what I take to be the most important ones. The remaining diagnostics are discussed briefly at the end.

Omissibility [A] is well-known to provide poor evidence for the A/AD (e.g. Culicover 2009:69, Needham & Toivonen 2011:9). First, many canonical arguments are omissible. Focusing on PPs, a corpus search (CoCA) reveals that, in fact, *most* PPs canonically considered arguments (e.g. *talk* ***to NP***, *look* ***at NP***, etc.) are omissible. There are also many canonical adjuncts which are non-omissible (e.g. *behave* ***nicely/rudely*** *to someone*, adverbials in middles, see Sailor & Schütze 2013 for further examples). Thus, if there is an A/AD, *omissibility* is a poor heuristic for it. Additionally, whether or not a given constituent is omissible depends on principles entirely independent of the A/AD; there is no need to connect the (im)possibility of omission to the A/AD. Instead, constituents are (non-)omissible for a variety of reasons. One reason is pragmatic felicity (as argued by Ernst 1984, 2002; McConnell-Ginet 1994; Goldberg & Ackerman 2001). Another is idiomaticity. Any time a verb "selects" a particular preposition (which is semantically bleached

and invariable, like *on* in *depend **on***, or *to* in *belong **to***), that preposition forms an idiom with the verb (Pesetsky 1995:135), and hence the PP it heads is non-omissible.

Exemptability [B] is best illustrated by *do so*-substitution. It has been believed since Ross & Lakoff (1976) that *do so*-substitution can strand adjuncts, but not arguments, on the basis of contrasts like *Lee poured water {onto the floor/on Monday} and Terry did so {*into the sink/on Tuesday}*. However, subsequent work on *do so* (e.g. Przepiórkowski 1999, Culicover & Jackendoff 2005, Houser 2010) has revealed that many canonical arguments can systematically be stranded under *do so* substitution. E.g., "patients" can be stranded using *to*-phrases (as in *John kicked Mary and Peter **did so to Ann*** from Przepiórkowski 1999). Whether or not a given constituent can be stranded under *do so*-substitution is better understood without reference to the A/AD. Strandability of a given PP under *do so*-substitution depends on the compatibility of that PP with the lexical verb *do*. PPs that are compatible with lexical *do* can be stranded, while those that are incompatible with *do* cannot be stranded. (Consider **Lee did something into the sink, Lee did something to Ann, Lee did something on Tuesday*). *Do so*-substitution thus does not provide good evidence for the A/AD. VP-fronting is traditionally analyzed as evidence for the A/AD by similar argumentation, but again, previous work shows that strandability-under-VP-fronting is orthogonal to the A/AD (unfortunately space does not permit discussion in the abstract, see data in, e.g., Phillips 2003, Landau 2007, Culicover & Winkler 2019).

Finally, ***islandhood*** [C] has sometimes been taken as evidence for the A/AD, e.g. by Chomsky et al. (2019), but I argue this is incorrect. First, despite certain claims to the contrary (to be addressed shortly), most canonical adjunct PPs are *not* islands. E.g., benefactive *for*-phrases, instrumental and comitative *with*-phrases, locatives, and durationals (all canonical adjuncts), are all transparent for extraction (examples can be found in Takami 1988, 1992). The sentence *What did John decide on* is sometimes cited as evidence that locatives are islands (e.g. Chomsky et al. 2019:249). However, Newmeyer (2005:116) and others have pointed out that extraction from locative *on*-phrases is perfectly acceptable; in this context, it happens to be odd for pragmatic reasons (extracting *what* from a locative is odd whether the locative is a canonical argument or a canonical adjunct. Compare **What does Lee live in*, **What did Lee drive on*). Additionally, I demonstrate that adjuncthood is irrelevant to canonical "adjunct islands": the PPs that are "adjunct islands" (like clausal *after*-phrases) are islands whether they are canonical arguments or canonical adjuncts. E.g., temporal PPs must be considered arguments, not adjuncts, in certain contexts (e.g. *The party lasted **[until Lee insulted Terry]***, *The detective placed the time of the murder **[before Lee insulted Terry]***). Extraction is unacceptable there, just as when the PPs are canonically considered adjuncts (e.g. **Who did the party last until Lee insulted __?*, **Who did the detective place the time of the murder before Lee insulted __?*).

The remaining diagnostics fare similarly. Prior work has shown the A/AD to be orthogonal to *Condition C reconstruction* [E] (Adger et al. 2017, Bruening & Al Khalaf 2019) and *extractability* [G] (Hukari & Levine 1995, Szabolcsi 2006). On *permutability* [F], the relevant permutation classes cross-cut the A/AD (i.e. many canonical arguments and adjuncts can be freely permuted, as in *Terry went <with Lee> to the store <with Lee>, Terry argued <for hours> about politics <for hours>*). Finally, *iterability* [D] applies to canonical arguments in exactly the same way as it applies to canonical adjuncts: iteration is possible when it can be construed as "concentric" or "nested," (see Ernst 2002:135 and Goldberg 2006:37fn15, who reach this conclusion when looking just at locative "*adjuncts*." I show the same principle applies to "arguments," e.g. *Lee put the book [in the library][on the shelf][next to* Aspects*]*).