The syntax of controllable pronouns: a case study from Uyghur

Madelaine O'Reilly-Brown (Stanford University)

1. Introduction

- This talk: discusses the controllable null subject (henceforth CNS) in Uyghur (Turkic: China)[1]
- The CNS appears in both argument (1a) and adjunct clauses (1b)[2]

(1) a. Ahmet [_ Turpan-gha bër-ish-qa] tirish-t-i
Ahmet Turpan-DAT go-NMLZ-DAT try-PST-3
'Ahmet$_i$ tried *ec*$_{i/*k}$ to go to Turpan'

b. Meryem [_ texse yuy-ghan-da] naxsha
Meryem dish wash-ASP-LOC song
ëyt-i-du
make-NPST-3
'Meryem$_i$ sings while *ec*$_{i/*k}$ washing the dishes.'

- Today, I will restrict my attention to CNS subjects of so-called *-Ish* argument clauses like (1a)
- The CNS exhibits unique interpretive and distributional properties:
  1. Interpretation: two options
     (i) Obligatory coreference: In (1), the null subject must be coreferential with the matrix subject
     (ii) Impersonal interpretations: when not controlled, the CNS gives rise to either existential or generic readings
  2. Distribution:
     (i) The CNS is restricted to caseless subject positions
     (ii) The CNS is incompatible with subject agreement
     – In contrast to (1a), the complement clause in (2) hosts third person subject agreement, and a control interpretation is disallowed

(2) Ahmet$_i$ [Ali-ning/?*pro*$_{k/*i}$ Turpan-gha bër-ish-**i**-ge]
Ahmet Ali-GEN/*pro* Turpan-DAT go-NMLZ-**3**-DAT
tirish-t-i
try-PST-3
'Ahmet$_i$ tried for Ali/?him$_{k/*i}$ to go to Turpan'

- The proposal: the Uyghur CNS is a bare-N pronominal
- This approach provides the tools needed to derive both the distribution and interpretation of the CNS
  1. Interpretation:
     – A bare N introduces a variable ranging over a set of all potential referents in a given context (Ackema and Neeleman 2018)
     – As a consequence, the CNS cannot saturate a predicate; this forces either saturation by a matrix argument (control) or binding by an existential/generic operator (impersonal readings)
  2. Distribution:
     – Due to its drastically reduced size, a bare-N pronoun is restricted to caseless positions (Fenger 2018)
     – Lacking phi-features, the CNS cannot Agree

[1] 'Controllable' instead of 'controlled' is a deliberate terminological choice in view of the emphasis I will place on no-control readings of the CNS.

[2] Unless indicated otherwise, all data are thanks to three native speaker consultants. I thank them for sharing their language with me.

Roadmap:

- Section 2: Properties of *-Ish* clauses
- Section 3: The interpretation of the CNS
- Section 4: The distribution of the CNS
- Section 5: The proposal
- Section 6: Conclusion and next steps

2. The structure of *-Ish* clauses

- *-Ish* clauses: found as the complement to attitude and nonattitude verbs, including *ret qil-maq*, 'refuse', and *qoshul-maq*, 'agree', *tirish-mek*, 'try', and *bashli-maq*, 'start'
- *-Ish* clauses are quite structurally reduced
  - They prohibit negation, aspectual morphology, and tense (Asarina 2011)
- Two properties of *-Ish* clauses with overt subjects:
  1. The subject is obligatorily genitive-marked
  2. Agreement from the possessive agreement paradigm obligatorily suffixes to the *-Ish* morpheme

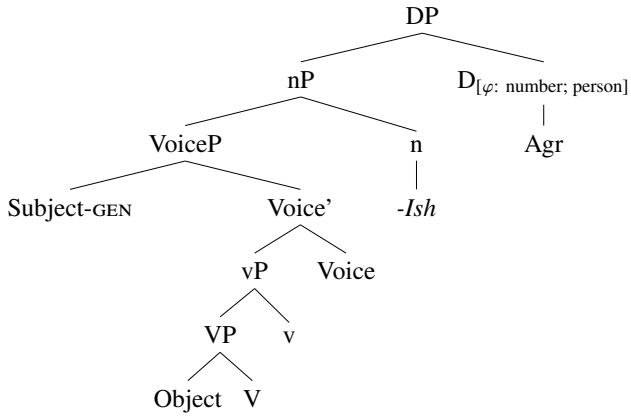
(3) Ahmet [Ali*(-ning) Turpan-gha bër-ish*(-i)-ge]
Ahmet Ali-GEN Turpan-DAT go-NMLZ-3-DAT
tirish-t-i
try-PST-3
'Ahmet tried for Ali to go to Turpan'

- Subject genitive case, possessive agreement, and the *-Ish* clause's ability to bear case are all indicators that these clauses are nominalized

Proposal:

- *-Ish* clauses are VoiceP-sized[3]
- The *-Ish* morpheme is in n^{o} and is responsible for nominalizing the VoiceP
- The nominalized complex is headed by a D^{o} with unvalued [person, number] features which Agrees with and assigns genitive case to the subject of the clause

(4) *-Ish* complement clause: overt subject



- Next, I turn to the interpretation of CNS subjects of *-Ish* clauses

[3]For evidence that control variants of *-Ish* clauses are VoiceP-sized and project a syntactic subject (contra a conceivable restructuring (Wurmbrand 2001) analysis), see additional discussion in the Appendix.

3. The interpretation of the CNS

• This section:

1. Demonstrates that coreferent readings of the CNS should be analyzed as control
2. Then, turns to no-control, impersonal readings of the CNS

• I adopt Landau's (2013) Obligatory Control (OC) Signature for identifying control dependencies

(5) The OC Signature
   a. In a control construction [. . . X_i. . . [S PRO_i. . .]. . .], where X controls the PRO subject of the clause S:
   b. The controller(s) X must be (a) co-dependent(s) of S.
   c. PRO (or part of it) must be interpreted as a bound variable. (Landau 2013:29)

• Part (b) of the signature rules out non-c-commanding controllers, which I will use as a diagnostic for control into *-Ish* clauses

• In the absence of an antecedent in the appropriate structural configuration, the null subject is no longer obligatorily coreferent (Mohanan 1983; Clark 1990; Hornstein 1999; Pires 2007; McFadden et al. 2016), and may give rise to arbitrary readings (Landau 2021a)

3.1. *Coreference is control*

• The null subject of (6) must be coreferent with the closest c-commanding DP, *Meryem*; no disjoint construal is possible

(6) Hesen [_ Ekber-ge söz-li-sh-ke] bashli-d-i
Hesen Ekber-DAT talk-RECIP-NMLZ-DAT start-PST-3
'Hesen started to speak with Ekber.'

• Further, only c-commanding antecedents may bind the CNS

(7) Hesen-ning ati-si [_ Ekber-ge
Hesen-GEN father-3SG Ekber-DAT
söz-li-sh-ke] bashli-d-i
talk-RECIP-NMLZ-DAT start-PST-3
'Hesen's$_i$ father$_j$ started $ec_{j/*i}$ to speak with Ekber.'

• The CNS can be shown to be variable-bound by its antecedent, rather than accidentally coreferent with it

• Only a bound-variable interpretation of the CNS is available with a quantified antecedent

(8) her bir oqutquchi [_ mewsümlük imtihan
every one teacher final exam
netije-lir-i-ni waqti-da tapshur-ush-ni] ret
grade-PL-3-ACC time-LOC hand.in-NMLZ-ACC refuse
qil-d-i
do-PST-3
Bound-variable interpretation: $\forall x$ (teacher(x) $\rightarrow$ x refused that x turn in the final exam grades on time))
$*$ Unbound interpretation: $\forall x$ (teacher(x) $\rightarrow$ x refused that (s)he turn in the final exam grades on time))

• Additionally, the CNS may only receive a sloppy reading under VP-ellipsis

(9) Meryem [_ mewsümlük imtihan netije-lir-i-ni
Meryem final exam grade-PL-3-ACC
waqti-da tapshur-ush-ni] ret qil-d-i, Amine mu
time-LOC hand.in-NMLZ-ACC refuse do-PST-3 Amine also
shundaq
such
Strict reading (unavailable):
∗Meryem (λx (x refuses that Meryem turn in the final exam grades on time)) and Amine (λx (x refuses that Meryem turn in the final exam grades on time))
Sloppy reading:
Meryem (λx (x refuses that x turns in the final exam grades on time)) and Amine (λx (x refuses that x turns in the final exam grades on time))

• Together, the c-command requirement on controllers and the obligatory bound-variable interpretations of the CNS show that coreferent readings of the CNS exhibit the OC signature (Landau 2013)

3.2. *No-control readings of the CNS*

• When the CNS is not c-commanded by an antecedent DP, it gives rise to one of two non-referential, impersonal readings:

1. Generic
2. Existential

• Generic readings:

– Also known as 'quasi-universal' (Cinque 1988)

– The domain of quantification of the generic can (but does not have to) include the speaker and addressee

– The reading is 'quasi-universal' in that it applies to all human individuals, but is not falsified by idiosyncratic exceptions

(10) [_ kompyutir-ni etwerka bilen ajrit-ish]
computer-ACC screwdriver with take.apart-NMLZ
aqilanilik
smart
'It's smart (for one/people) to take apart the computer with a screwdriver.'

• Existential readings:

– Assert the existence of a specific individual for whom a given property holds

– Compatible with a specific, bounded time reference (D'Alessandro 2007)

(11) [_ kompyutir-ni etwerka bilen ajrit-ish]
computer-ACC screwdriver with take.apart-NMLZ
aqilanilik i-d-i
smart AUX-PST-3
'It was smart (for someone) to take apart the computer with a screwdriver.' (= someone took apart the computer with a screwdriver)

4. The distribution of the CNS

• The CNS is found in subject position of nonfinite, nominalized complement and adjunct clauses

– This is a familiar distributional pattern and resembles that of, e.g., English PRO

- The CNS alternates with overt subjects
  - Such alternations have been documented in the literature on finite control (Landau 2004) and in non-finite environments in languages like Tamil (McFadden and Sundaresan 2018)
- The novel finding in Uyghur: the CNS is categorically incompatible with subject agreement
  - This generalization holds regardless of whether the embedding verb is an attitude or non-attitude verb

(12) No agreement in complement to attitude verbs

a. men [_ kitap oqu-sh(*-im)-ni]
I book read-NMLZ-1SG-ACC
oyli-wat-i-men
want-PROG-NPST-1SG
'I want to read a book.' (Asarina 2011:48)

b. men [_ sirt-qa chiq-ish(*-im)-qa]
I outside-DAT go.out-NMLZ-1SG-DAT
qoshul-d-um
agree-PST-1SG
'I agreed to go outside.'

c. Hesen [_ oyun körsitish-ke qatnish-ish(*-i)-ni]
Hesen play showing-DAT attend-NMLZ-3-ACC
ret qil-d-i
refuse do-PST-3
'Hesen refused to attend the play.'

(13) No agreement in complement to non-attitude verbs

a. men [_ kitap oqu-sh(*-im)-ni] tirish-t-im
I book read-NMLZ-1SG-ACC try-PST-1SG
'I tried to read a book.' (Asarina 2011:48)

b. Hesen [_ Ekber-ge söz-li-sh(*-i)-ke]
Hesen Ekber-DAT talk-RECIP-NMLZ-3-DAT
bashli-d-i
start-PST-3
'Hesen started to speak with Ekber.'

c. matur-diki mesile mashini-ni [_
engine-LOC.QUAL problem car-ACC
toxtut-ush(*-i)-qa] mejburli-d-i
stop-NMLZ-3-DAT force-PST-3
'The engine problem forced the car to stop.'

- Interactions between the availability of control readings and subject agreement have been noted in, e.g., Słodowicz 2007 and Landau 2015
  - Landau (2015) formalizes this observation as the 'OC-NC Generalization'

(14) [+Agr] blocks control in attitude complements but not in nonattitude complements. (Landau 2015:20)

- However, the generalization in Uyghur is stronger; agreement blocks control in attitude *and* non-attitude complements
- Next, I propose an analysis of the morphosyntax of the CNS that is designed to capture this fact

5. The proposal

- I propose that the Uyghur CNS is a structurally deficient, bare-N pronoun
  - I argue that this approach allows us to straightforwardly derive the distributional and interpretive properties of the CNS

5.1. *Deriving the distribution of the CNS*

- Two facts to account for:
  1. The CNS is restricted to (caseless) subject position
  2. The CNS is incompatible with subject agreement

5.1.1 The CNS may only be a caseless subject

- (Following Fenger 2018) a bare N is too small to project a KP
  - I assume that a KP layer is a requirement for bearing any non-nominative case
- This restricts the CNS to caseless positions
  - We correctly rule out, for example, a CNS in indirect object position where it gets obligatory dative case
- Next, I show how the CNS' inability to Agree correctly allows it to remain caseless in subject position of *-Ish* clauses

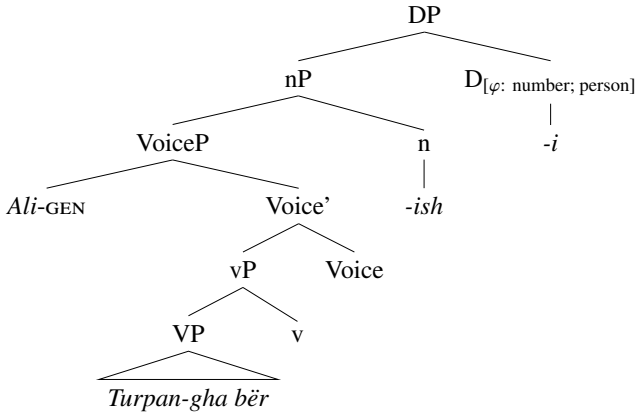
5.1.2 The CNS is incompatible with subject agreement

- We have seen that *-Ish* clauses with overt subjects obligatorily assign them genitive case and Agree with them

(15) Ahmet [Ali*(-ning) Turpan-gha bër-ish*(-i)-ge]
Ahmet Ali-GEN Turpan-DAT go-NMLZ-3-DAT
tirish-t-i
try-PST-3
'Ahmet tried for Ali to go to Turpan'

- I proposed that Dº Agrees with the subject in person and number, and as a consequence assigns it genitive case

(16) *-Ish* complement clause: overt subject

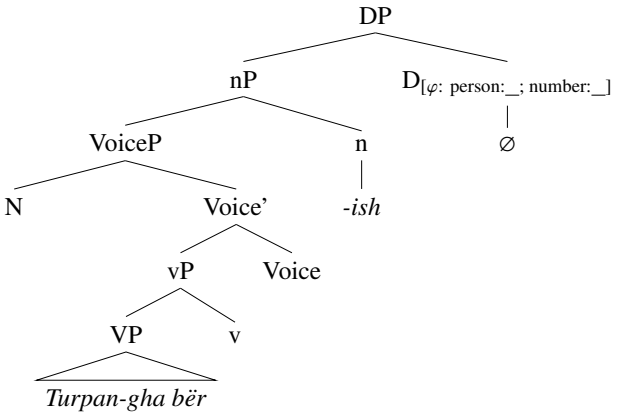


- Alternatively, a CNS subject may be merged (17)

(17) Ahmet [_ Turpan-gha bër-ish-qa] tirish-t-i
Ahmet Turpan-DAT go-NMLZ-DAT try-PST-3
'Ahmet$_i$ tried $ec_{i/*k}$ to go to Turpan'

- The CNS is too small to contain phi-features and thus Agree between the CNS and Dº fails
- Following Preminger (2011), I propose that the result of this failure is not a derivational crash, but morphologically null default agreement

(18) *-Ish* complement clause: CNS subject



5.2. *Deriving the interpretation of the CNS*

• Two types of readings to account for:

1. Obligatory control
2. Impersonal readings

5.2.1 The denotation of bare-N pronouns

• I follow Ackema & Neeleman (2018), for whom the N core of a pronoun denotes the set of all entities in the discourse

– The semantic type of a bare-N pronoun is therefore $\langle e, t \rangle$

– A problem? The CNS is not able to combine with Voice', of type $\langle e, \langle s, t \rangle \rangle$, simply by saturating it

– The solution: The CNS and Voice' combine via Restrict (Chung and Ladusaw 2004)

(19) *Restrict*
If a is of type $\langle e, \langle s, t \rangle \rangle$ and b is of type $\langle e, t \rangle$,
$[[ab]] = \lambda x.\lambda e.[[a]](e, x)\&[b]](x)$.

• After Restrict, VoiceP is also of type $\langle e, \langle s, t \rangle \rangle$, i.e., a predicate

• Now, this VoiceP is ready to be directly saturated by a matrix argument

5.2.2 Deriving obligatory control

• Consider the subject control example below:

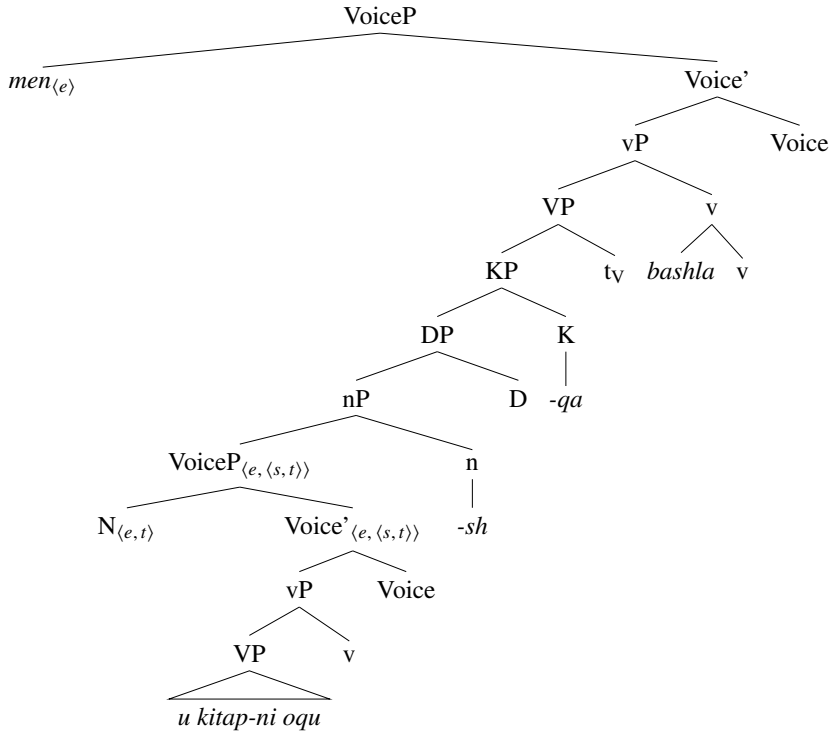
(20) men [__ u kitap-ni oqu-sh-qa] bashli-d-im
I that book-ACC read-NMLZ-DAT start-PST-1SG
'I started to read that book.'

• After merging a bare-N pronoun in Spec,VoiceP, the complement clause denotes $\lambda x.x$ *read that book*, where the reader is (a) member(s) of the set of all potential referents in the context

• (Following Landau 2015) I assume that V-to-v raising occurs in the matrix clause, with the result that the m-command domain of the complement is the matrix vP

• This licenses syntactic predication, and the open property denoted by the embedded clause is applied to the referent of the matrix first person subject

(21) Derivation of subject control into an *-Ish* clause



5.2.3 Deriving impersonal readings

- In the absence of a c-commanding antecedent, the CNS receives an impersonal interpretation
  - Two possible impersonal readings: generic and existential
- Following Ackema and Neeleman 2018, a bare-N pronoun may be bound by either an arbitrary or generic operator
- We observed in section 3.2 that it is matrix tense/aspect which conditions whether an uncontrolled CNS gets an existential or generic reading
- Additionally, an impersonal reading is *only* available if there is no controller in the matrix clause
  - OC readings always supercede impersonal readings

(22) Ahmet [_ Turpan-gha bër-ish-qa] tirish-t-i
Ahmet Turpan-DAT go-NMLZ-DAT try-PST-3
Unavailable: 'Ahmet tried for someone to go to Turpan'

- Proposal: arbitrary and generic operators are hosted in the clausal periphery, above TP[4]
- If *-Ish* clauses are smaller than TP, this correctly predicts that it is not possible to merge an arbitrary or generic operator within the CNS' local clause
  - This prevents merge of an operator inside of the *-Ish* complement clause in (22)
- Additionally, an arbitrary/generic operator in the matrix clause is higher than matrix arguments, so OC via saturation of the embedded VoiceP will take precedence over operator binding

6. Conclusion

- I have analyzed the Uyghur CNS as a bare-N pronominal (drawing heavily from Ackema and Neeleman 2018 and Fenger 2018)

[4]Here I diverge from Ackema and Neeleman 2018, who suggest that these operators merge directly with the pronoun, while acknowledging that an analysis in which they are merged higher in the clause is equally tenable.

- I argue that my account is able to derive characteristic properties of the CNS
    1. Distribution: restricted to caseless subject position, cannot Agree
    2. Interpretation: bound-variable OC readings and impersonal (existential and generic) readings
- This approach raises two higher-level points regarding how we think about and analyze control
    1. Significant empirical gains follow from exploring the morphosyntax of controllable pronouns as an additional parameter
    2. Thinking about PRO as a *category of referentially deficient pronouns* offers us the flexibility to derive much of the crosslinguistic variation that has been problematic for earlier theories

6.1. *Deriving properties of control from the morphosyntax of the controlled pronoun*

- Classical theories of control, as in Chomsky 1981, postulate that all control constructions involve PRO
- Some proposals for PRO:
    - PRO is a null DP which is obligatorily caseless (Chomsky 1981)/bears null case (Chomsky and Lasnik 1993) and is restricted to subject position of nonfinite clauses
    - PRO is a DP minimal pronoun with unvalued phi-features (Landau 2015, Landau 2021a)
- These approaches range with respect to how stipulative they are forced to be about caselessness, nullness, and distributional restrictions
- A persistent problem: adequately describing and accounting for no-control readings
- What happens if we take no-control readings as central to our theories of control?
    - No-control readings of the Uyghur CNS are virtually identical to readings of Germanic *men*-type impersonal pronouns documented in Egerland 2003, Fenger 2018
- Unlike the control literature, the literature on impersonal and referential pronouns puts much more explanatory burden on morphosyntax of the pronoun (see, e.g., Cardinaletti and Starke 1999, Déchaine and Wiltschko 2002, Holmberg 2005, Shlonsky 2009, a.o.)
- Many dedicated impersonal pronouns share properties with controlled pronouns:
    - Non-referential readings
    - Caselessness
    - Restricted to subject position
    - Inability to Agree/default agreement
- Many studies derive these properties from *defective/reduced* pronominal structure (Cinque 1988, Egerland 2003, Fenger 2018, Ackema and Neeleman 2018)
- Reconsidering whether PRO is *necessarily* a DP could provide another avenue for deriving its characteristic properties

6.2. *PRO as a category of pronouns*

- Should our theories simply dispense with DP PRO and replace it with a bare-N pronominal?
  - No!
- Crosslinguistic work on control has revealed extensive variation (see Landau 2024 for an overview):
  - Null vs. overt controllable pronouns
  - (Un)availability of inflected controlled complement clauses
  - Case of controllable pronouns
  - Types of no-control interpretations
  - (Un)availability of control into finite clauses
- Control properties may differ crosslinguistically in part because of variation in morphosyntax of controllable pronouns
  - I.e., the English controllable pronoun may well be a (minimal) DP
- Not all controllable pronouns need be identical![5]

6.2.1 Final remarks

- My claim: future work on control should shift more of the explanatory burden to the morphosyntax of the controllable pronoun(s)
- This move raises a new set of questions:
  1. What properties do all controllable pronouns share in common?
     - An idea: controllable pronouns are obligatorily referentially defective
  2. How should these shared properties be represented structurally?
  3. What hypothetical controllable pronouns are predicted to be possible/impossible and are attested/unattested?

[5] This claim is anticipated in Baker and Ikawa 2024, who suggest that PRO is part of a class of pronouns.

References

Ackema, Peter, and Ad Neeleman. 2018. *Features of Person: From the Inventory of Persons to Their Morphological Realization*, volume 78. Cambridge, MA: Linguistic Inquiry Monographs.

Asarina, Alevtina. 2011. Case in Uyghur and Beyond. Doctoral Dissertation, Massachusetts Institute of Technology.

Baker, Mark, and Shiori Ikawa. 2024. Control theory and the relationship between logophoric pronouns and logophoric uses of anaphors. *Natural Language and Linguistic Theory* 42:897—-954.

Cardinaletti, Anna, and Michal Starke. 1999. The typology of structural deficiency: A case study of the three classes of pronouns. In *Clitics in the Languages of Europe*, ed. Henk Van Riemsdijk. Berlin: Mouton de Gruyter.

Chomsky, Noam. 1981. *Lectures on Government and Binding: The Pisa Lectures*. 9. Walter de Gruyter.

Chomsky, Noam, and Howard Lasnik. 1993. 'The Theory of Principles and Parameters'. In *Syntax: An International Handbook*

of Contemporary Research, ed. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann, 506–569. De Gruyter Mouton.

Chung, Sandra, and William A. Ladusaw. 2004. *Linguistic inquiry monographs*, chapter Restriction and Saturation. Cambridge, MA: MIT Press.

Cinque, Guglielmo. 1988. On Si Constructions and the Theory of Arb. *Linguistic Inquiry* 19:521–581.

Clark, Robin. 2016. *Thematic theory in syntax and interpretation*. Routledge.

D'Alessandro, Roberta. 2007. *Impersonal si constructions: Agreement and interpretation*, volume 90. Berlin, New York: Walter de Gruyter.

Déchaine, Rose-Marie, and Martina Wiltschko. 2002. Decomposing Pronouns. *Linguistic Inquiry* 33:409–442.

Egerland, Verner. 2003. Impersonal pronouns in Scandinavian and Romance. *Working papers in Scandinavian syntax* 71:75–102.

Fenger, Paula. 2018. How impersonal does *one* get?: A study of *man*-pronouns in Germanic. *The Journal of Comparative Germanic Linguistics* 21:291–325.

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

Hornstein, Norbert. 1999. Movement and control. *Linguistic Inquiry* 30:69–96.

Landau, Idan. 2004. The scale of finiteness and the calculus of control. *Natural Language & Linguistic Theory* 22:811–877.

Landau, Idan. 2013. *Control in generative grammar: A research companion*. Cambridge: Cambridge University Press.

Landau, Idan. 2015. *A Two-Tiered Theory of Control*. Cambridge, MA: MIT Press.

Landau, Idan. 2021a. *A Selectional Theory of Adjunct Control*. Cambridge, MA: MIT Press.

Landau, Idan. 2024. Noncanonical obligatory control. *Language and Linguistics Compass* 18:e12515.

McFadden, Thomas, and Sandhya Sundaresan. 2018. Reducing pro and PRO to a single source. *The Linguistic Review* 35(3):463–518.

McFadden, Thomas, Sandhya Sundaresan, Christopher Hammerly, and Brandon Prickett. 2016. Failure to control is not a failure: it's pro. In *Proceedings of NELS*, volume 46, no.3, 1–10.

Mohanan, Karuvannur P. 1983. Functional and anaphoric control. *Linguistic Inquiry* 14:641–674.

Pires, Acrisio. 2007. The derivation of clausal gerunds. *Syntax* 10:165–203.

Preminger, Omer. 2011. Agreement as a fallible operation. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Shlonsky, Ur. 2009. Hebrew as a partial null-subject language. *Studia Linguistica* 63:133–157.

Słodowicz, Szymon. 2007. Complement control in Turkish. *ZAS Papers in Linguistics* 47.

Wurmbrand, Susanne. 2001. *Infinitives: Restructuring and clause structure*. Walter de Gruyter.

Appendix

Evidence for a syntactically projected subject

- I argue that null-subject *-Ish* clauses project a syntactic subject on the basis of reciprocal licensing
- Reciprocals in this language require a clausemate, syntactically projected coordinated or semantically plural antecedent, as the contrast between (23a) and (23b) shows

(23) a. Meryem bilen Zöhre birbiri toghrisidiki
Meryem and Zöhre each.other about
maqali-lar-ni oqu-d-i
article-PL-ACC read-PST-3
'Meryem and Zöhre read the articles about each other.'

b. ??Meryem bilen Zöhre [Amine-ning yaz-ghan
Meryem and Zöhre Amine-GEN write-ASP
birbiri toghrisidiki maqali-lir-i-ni] oqu-d-i
each.other about article-PL-3-ACC read-PST-3
Intended: 'Meryem and Zöhre read the articles about each other that Amine wrote.'

- If a reciprocal is licensed in a null-subject clause, we can assume that the clause projects a (null) antecedent which binds the reciprocal
- Reciprocals are licensed in null-subject *-Ish* clauses

(24) [_ birbiri-ni këchir-ish] yaxshi
each.other-ACC forgive-NMLZ good
'It's good to forgive each other.'

- Example (24) also suggests that the CNS can be semantically plural
  - A (still speculative) thought: partial control seems to be widely available in Uyghur, including in non-attitude complements (contra Landau 2015)
  - It would be interesting to probe whether semantic plurality of the CNS could be leveraged derive partial control