

VERBAL AGREEMENT SYNCRETISM AND NULL SUBJECTS IN HISTORICAL ICELANDIC

Yiming Liang & Alexandra Simonenko

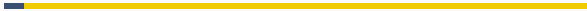
June 18, 2026

Diachronic Generative Syntax (DiGS27), University of Padova



1. Taraldsen's Generalization: why Icelandic is a challenge
2. Measuring agreement syncretism: VASS
3. Historical Icelandic: corpus study and results
4. Beyond Icelandic: cross-linguistic comparison

INTRODUCTION



Taraldsen's Generalization

Languages with sufficiently rich verbal agreement tend to allow referential null subjects.

Rich agreement

Italian, Spanish

⇒ *pro*-drop languages

Poor agreement

Present-Day English, French

⇒ non *pro*-drop languages

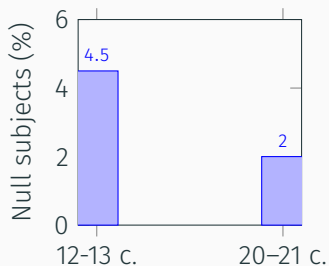
Phi-features of referential *pro* must be identified by **rich** verbal agreement. (Rizzi, 1982; Taraldsen, 1980)

PUZZLE: HISTORICAL ICELANDIC

- Rich verbal inflection over time (typically 5/6 distinct forms)
- Yet null subjects are rare.

	<i>leita</i>	<i>segja</i>	<i>sjá</i>
1sg	leita	segi	sé
2sg	leitar	segir	sér > sérð
3sg	leitar	segir	sér
1pl	leitum	segjum	sjáum
2pl	leitit > -ið	segit > -ið	sjáit > -ið
3pl	leita	segja	sjá
#	4	5	5 > 6

Table 1: Present indicative paradigms of *leita* ‘search’, *segja* ‘say’ and *sjá* ‘see’ in Icelandic (Sigurðsson, 1993).



Null subject rates in IcePaHC
(Kinn et al., 2016).

PREVIOUS ACCOUNTS FOR ICELANDIC NULL SUBJECTS

Sigurðsson (1993):

- Verbal inflection plays no role in licensing null subjects.

Germanic topic drop (diary drop)

- positional and stylistic constraints
- restricted to V1 main clauses
- licensed by an empty prefield (empty SpecCP in main clauses)

- (1) (Ég) þekki það ekki.
(I) recognize that not
'I don't recognize that.'

Genuine *pro*-drop

- licensed by free coindexing with an NP antecedent rather than by agreement

- (2) þá skar Rognvaldr [hár
then cut R. hair
hans]_i, en áðr var ____i
his but before was
úskorit ...
uncut
'Then Rognvaldr cut his hair_i,
but it_i had been uncut before
...'
(Nygaard 1906: 10)

Kinn et al. (2016): empirical study on IcePaHC

- Counterexamples against Sigurðsson (1993): null subjects without NP antecedent are possible in non-V1 main clauses
- Opposite trend of null subjects split by contexts:
 - topic-drop licit: ↑
 - topic-drop illicit: ↓
- 3rd-person null subjects ↓, 1st- and 2nd-person null subjects ↑
- Icelandic underwent *pro-drop* → *topic-drop*
- Shift completed in the early 20th c.

- If the two mechanisms co-exist, the effect of verbal inflection on *genuine pro-drop* may be *blurred*, as topic-drop is not affected by inflection.
→ separate the two licensing mechanisms of null subjects

Research question

Does verbal agreement (syncretism) condition genuine *pro-drop* in Historical Icelandic?

- If the two mechanisms co-exist, the effect of verbal inflection on *genuine pro-drop* may be *blurred*, as topic-drop is not affected by inflection.
 - separate the two licensing mechanisms of null subjects

Research question

Does verbal agreement (syncretism) condition genuine *pro-drop* in Historical Icelandic?

→ a gradient measure of agreement syncretism

MEASURING VERBAL AGREEMENT SYNCRETISM

THE MEASUREMENT PROBLEM

Verbal agreement syncretism: categorical distinction “rich” vs. “poor”

- Little consensus on the definition of *richness*:

Source	Strength	Richness criterion
Platzack & Holmberg (1989)	Strong	Person distinctions are morphologically marked on the verb
Roberts (1993: 263–273)	Weak	Overt distinct morphological plural marking
Rohrbacher (1994, 1999)	Strong	In at least one number of at least one tense of the regular verbs, the person features [1st] and [2nd] are both distinctively marked
Vikner (1997)	Strong	Person morphology is found in all tenses
Koenenman (2000: 67–84)	Weak	Three features ([α speaker], [α addressee], [α singular]) are required for the description of a paradigm
Bobaljik (2002)	Weak	Finite verbs may bear multiple distinct inflectional morphemes
Koenenman & Zeijlstra (2014)	Strong	Featural distinctions are expressed in all three dimensions +/–speaker, +/–participant, +/–plural

Figure 1: Theories of richness (Walkden, 2021: 9).

THE MEASUREMENT PROBLEM

Definition	WS. OE	Nth. OE	Sth. ME	Nth. ME	Ear. ModE
Platzack and Holmberg (1989)	Rich	Rich	Rich	Rich	Rich
Roberts (1993: 263–273)	Rich	Poor	Rich	Poor	Poor
Rohrbacher (1994, 1999)	Rich	Poor	Rich	Poor	Poor
Vikner (1997)	Rich	Rich	Rich	Rich	Rich
Koenenman (2000: 67–84)	Rich	Rich	Rich	Rich	Poor
Bobaljik (2002)	Rich	Rich	Rich	Rich	Rich
Koenenman and Zeijlstra (2014)	Rich	Rich	Rich	Rich	Rich?

Table 2: Classifications of richness of Historical English (Walkden, 2021: 14).

- Conflicting richness evaluations even for the same language and the same dataset
- Koenenman and Zeijlstra (2014): counterexamples against RAH disappear using their definition of “richness”.

THE MEASUREMENT PROBLEM

- Many languages, including **Icelandic**, fall between the fully distinctive and fully syncretic extremes.
 - A binary rich vs. poor distinction misses the **gradient** nature of agreement syncretism.
- Corpus data often reveal greater variation and diachronic change in verbal endings.
 - Richness defined from grammar-book paradigms may not reflect the **actual distribution of verbal endings in corpus data**.

THE MEASUREMENT PROBLEM

- Many languages, including **Icelandic**, fall between the fully distinctive and fully syncretic extremes.
 - A binary rich vs. poor distinction misses the **gradient** nature of agreement syncretism.
- Corpus data often reveal greater variation and diachronic change in verbal endings.
 - Richness defined from grammar-book paradigms may not reflect the **actual distribution of verbal endings in corpus data**.

⇒ a **gradient, data-driven** measure of verbal agreement syncretism

→ **VASS**: Verbal Agreement Syncretism Score (Liang & Simonenko, 2025)

QUANTIFY VERBAL AGREEMENT SYNCRETISM

Intuition:

- A less ambiguous (↓) ending reduces **more uncertainty** about the subject distribution (e.g., -s)
- A more ambiguous (↑) ending reduces **less uncertainty** about the subject distribution (e.g., zero)

pers.	sg	pl
1	hear	hear
2	hear	hear
3	hears	hear

Table 3: The present indicative paradigm in Modern English.

⇒ Quantify verbal agreement syncretism by measuring **how much uncertainty remains** about the subject distribution after the verbal ending is observed.

QUANTIFY VERBAL AGREEMENT SYNCRETISM

Conditional entropy:

$$H(\text{person}|\text{ending}) = - \sum_{x \in \text{person}} \sum_{y \in \text{ending}} P(x, y) \log P(x|y) \quad (1)$$

pers.	sg	pl
1	hear	hear
2	hear	hear
3	hears	hear

$$\begin{aligned} H(\text{person}|\text{ending}) = & - [\log P(1sg|\emptyset) \cdot P(1sg, \emptyset) + \log P(2sg|\emptyset) \cdot P(2sg, \emptyset) \\ & + \log P(3sg|-s) \cdot P(3sg, -s) + \log P(1pl|\emptyset) \cdot P(1pl, \emptyset) \\ & + \log P(2pl|\emptyset) \cdot P(2pl, \emptyset) + \log P(3pl|\emptyset) \cdot P(3pl, \emptyset)] \end{aligned}$$

Entropy $\uparrow \rightarrow$ Uncertainty of the subject distribution $\uparrow \rightarrow$ Level of verbal agreement syncretism $\uparrow$

QUANTIFY VERBAL AGREEMENT SYNCRETISM

New problem: Conditional entropy is highly sensitive to the **unbalanced distribution** of subject's person&number in a dataset.

- e.g., In letters, 1sg subject is predominant. The conditional entropy is low regardless of the richness of agreement system.

Verbal Agreement Syncretism Score (VASS)

$$\text{VASS} = \frac{H(\text{person}|\text{ending})}{H(\text{person})}$$

Low VASS

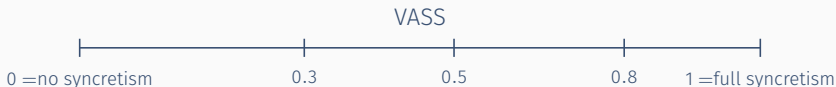
Endings identify the subject well.

Agreement is rich.

High VASS

Endings are more syncretic.

Agreement is poor.



WHAT VASS BUYS US

- gradient rather than categorical
- comparable across languages and periods
- capture subtle variation and change in verbal endings

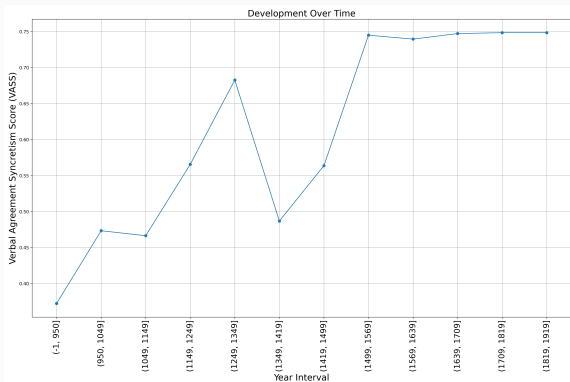


Figure 2: Verb agreement syncretism of present indicative in Historical English (Pearson coef. 0.88, $p < 0.001$, Liang & Simonenko, 2025).

HISTORICAL ICELANDIC: CORPUS STUDY



IcePaHC: parsed Historical Icelandic, 1150–2008 (Wallenberg et al. 2024)

- **Extraction:** Present indicative clauses with a referential pronominal, non-conjoined subject, using CorpusSearch 2 (Randall, 2010)
 - **Exclusion:**
 - Quirky subjects: Verbal agreement does not target the subject.
 - Fusion of a finite verb and the following overt subject clitic
 - Middle voice
- $N = 17,657$ tokens

- VASS was calculated for seven historical periods.
- The periods follow the binning used by Kinn et al. (2016)



Topic-drop licit

- V1 main clauses
 - VFin
 - *pro* Vfin
- V1 conjunct clauses
 - CONJ VFin
 - CONJ *pro* Vfin
- V2 main clauses where the subject occupies the prefield
 - NP-SBJ VFin
 - CONJ NP-SBJ VFin

Topic-drop illicit

- Other contexts (including all subordinate clauses)

Mixed-effects logistic regression:

- **Dependent variable:** null (= 1) vs. overt (= 0) subjects
- **Fixed effects:** VASS, year (log-transformed), topic-drop licitness, VASS $\times$ topic-drop licitness
 - Categorical factors: sum coding
 - Numeric factors: standardized
- **Random intercepts:** text, genre, and verb lemma

Predictions:

- VASS should matter more in topic-drop-illicit contexts (i.e., contexts with genuine *pro*-drop).
- Inverse correlation between VASS and null subjects

RESULTS

VASS IN HISTORICAL ICELANDIC

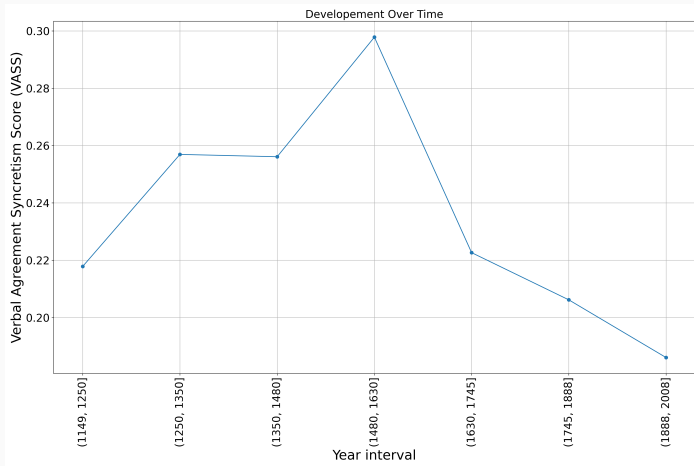


Figure 3: Development of VASS in Historical Icelandic

RESULTS OF THE REGRESSION MODEL

Fixed effect	Est.	SE	<i>z</i>	<i>p</i>	Sig.
Intercept	-3.683	0.207	-17.836	$< 2e - 16$	***
VASS	-0.293	0.152	-1.927	0.054	.
Topic-drop licitness	0.522	0.052	10.009	$< 2e - 16$	***
Year	-0.264	0.140	-1.884	0.060	.
VASS: Topic-drop licitness	0.181	0.058	3.124	0.002	**

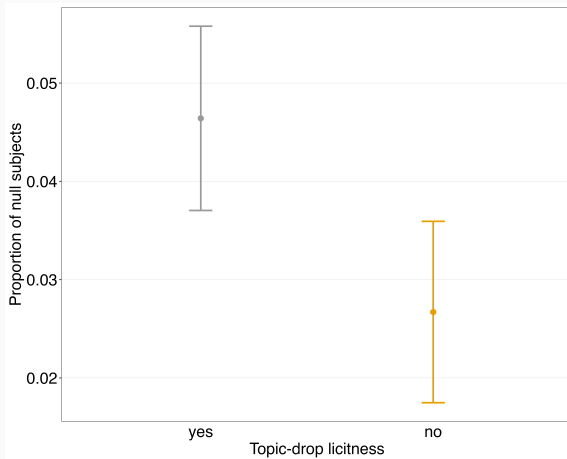
The effect of year does not reach significance:

- $\beta = -0.26$, $p = 0.06$
- pointing in the direction that null subjects decline $\downarrow$ over time.

The independent effect of VASS does not reach significance:

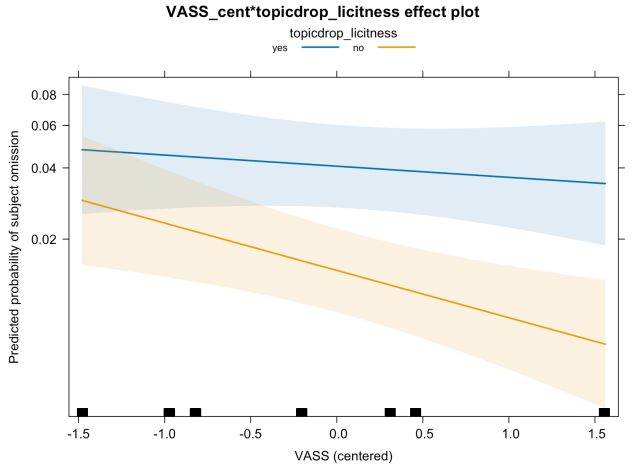
- $\beta = -0.29$, $p = 0.054$
- pointing in the expected direction that a higher VASS disfavours null subjects.

THE EFFECT OF TOPIC-DROP LICITNESS



Topic-drop-licit contexts favor null subjects compared to illicit contexts ($\beta = 0.52, p < 0.001$).

KEY RESULT: VASS INTERACTS WITH TOPIC-DROP LICITNESS



Increasing VASS **strongly** reduces null subjects in **topic-drop-illicit** contexts, but only **weakly** affects null subjects in topic-drop-licit contexts ($\beta = 0.18, p < 0.01$).

Topic-drop illicit

- Only genuine *pro*-drop is possible.
- Steeper slope ($= -0.47$).
- Higher syncretism $\rightarrow$ fewer null subjects.

Topic-drop licit

- Both topic-drop and genuine *pro*-drop are possible.
- Flatter slope ($= -0.11$).
- Weaker effect of verbal inflection.

Topic-drop illicit

- Only genuine *pro*-drop is possible.
- Steeper slope ($= -0.47$).
- Higher syncretism $\rightarrow$ fewer null subjects.

Topic-drop licit

- Both topic-drop and genuine *pro*-drop are possible.
- Flatter slope ($= -0.11$).
- Weaker effect of verbal inflection.

Post hoc analysis: run a separate regression model on each context $\rightarrow$ richer agreement favors null subjects **only in topic-drop-illicit** (i.e., genuine *pro*-drop) contexts ($\beta = -0.6, p < 0.05$).

Topic-drop illicit

- Only genuine *pro*-drop is possible.
- Steeper slope ($= -0.47$).
- Higher syncretism $\rightarrow$ fewer null subjects.

Topic-drop licit

- Both topic-drop and genuine *pro*-drop are possible.
- Flatter slope ($= -0.11$).
- Weaker effect of verbal inflection.

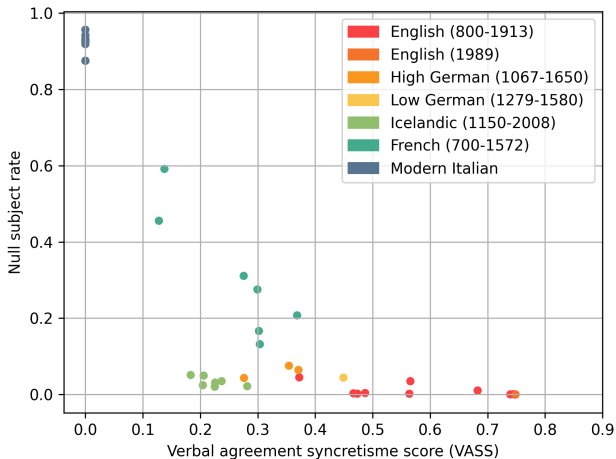
Post hoc analysis: run a separate regression model on each context $\rightarrow$ richer agreement favors null subjects **only in topic-drop-illicit** (i.e., genuine *pro*-drop) contexts ($\beta = -0.6, p < 0.05$).

Interim conclusion

Historical Icelandic genuine *pro*-drop is affected by verbal agreement syncretism, in line with Taraldsen's generalisation

CROSS-LINGUISTIC PERSPECTIVE

VASS AND NULL-SUBJECT RATES ACROSS LANGUAGES



Null-subject rates are inversely correlated with VASS ($r = -0.77$, $p < 0.001$): higher verbal syncretism, fewer null subjects.

WHY DOES ICELANDIC DIVERGE FROM FRENCH?

Icelandic and High German: comparable VASS with Historical French (0.2 – 0.35), but substantially lower null-subject rates.

Proposed explanation:

- Icelandic and High German are stronger V2 systems than Old French.

Icelandic

XVS in about 47% of V2 matrix clauses.

Old French

XVS in about 26% of comparable clauses.

- In V2 languages, the preverbal position is closely tied to topichood, both in High German (Speyer, 2019) and Icelandic (Faarlund, 2004).
- Dropping topics is restricted by conditions distinct from *pro*-drop licensing.

MAIN TAKEAWAYS

- Historical Icelandic **supports Taraldsen's Generalization**, when topic-drop contexts are excluded.
 - Richer agreement favors null subjects **only in genuine *pro*-drop contexts**.
- Cross-linguistically, higher VASS values are associated with lower null-subject rates.
- Strong V2 systems may result in fewer null subjects than non/weak V2 languages at comparable VASS levels.
- VASS offers a **fine-grained measure** of agreement syncretism, opening the way to quantitative studies of its relation to null subjects, verb movement, and other syntactic phenomena.

THANK YOU!

THE WORK IS FUNDED BY THE EUROPEAN UNION (ERC, CAUSALITY, GRANT NUMBER 101042427). VIEWS AND OPINIONS EXPRESSED ARE HOWEVER THOSE OF THE AUTHOR(S) ONLY AND DO NOT NECESSARILY REFLECT THOSE OF THE EUROPEAN UNION OR THE EUROPEAN RESEARCH COUNCIL. NEITHER THE EUROPEAN UNION NOR THE GRANTING AUTHORITY CAN BE HELD RESPONSIBLE FOR THEM.



YIMING LIANG, PHD
GHENT UNIVERSITY
YIMING.LIANG@UGENT.BE



REFERENCES

REFERENCES



Bobaljik, J. D. (2002). **Realizing Germanic Inflection: Why Morphology Does Not Drive Syntax.** *The Journal of Comparative Germanic Linguistics*, 6(2), 129–167.



Faarlund, J. T. (2004). **The syntax of Old Norse: With a survey of the inflection** Oxford University Press.



Kinn, K., Rusten, K. A., & Walkden, G. (2016). **Null Subjects in Early Icelandic.** *Journal of Germanic Linguistics*, 28(1), 31–78.



Koenenman, O. (2000). **The flexible nature of verb movement.** Netherlands Graduate School of Linguistics.



Koenenman, O., & Zeijlstra, H. (2014). **The Rich Agreement Hypothesis Rehabilitated.** *Linguistic Inquiry*, 45, 571–615.

REFERENCES II

-  Platzack, C., & Holmberg, A. (1989). **The role of AGR and finiteness.** Working Paper (43), 51–76.
-  Randall, B. (2010). **CorpusSearch 2..**
-  Rizzi, L. (1982). **Negation, wh-movement, and the null subject parameter.** An Annotated Syntax Reader, 169.
-  Roberts, I. (1993). **Verbs and diachronic syntax: A comparative history of English.** Kluwer.
-  Rohrbacher, B. W. (1994). **The Germanic VO languages and the full paradigm:** University of Massachusetts Amherst.
-  Rohrbacher, B. W. (1999). **Morphology-Driven Syntax.** John Benjamins Publishing Company.
-  Sigurðsson, H. Á. (1993). **Argument-drop in old Icelandic.** Lingua, 89(2), 247–280.
-  Speyer, A. (2019). **Topichood and the margins of the German clause from a historical perspective.** Architecture of Topic. Berlin/New York: Mouton de Gruyter, 337–371.



Taraldsen, K. T. (1980). **On the NIC, vacuous application and the that-trace filter.** Indiana University Linguistics Club, Bloomington.



Vikner, S. (1997). **V-to-I movement and inflection for person in all tenses.** In Elements of grammar-handbook in generative syntax (pp. 189–213). Kluwer Academic Publishers.



Walkden, G. (2021). **Do the wealthy stay healthy? Rich agreement and verb movement in early English.** Journal of Historical Syntax, 5(30), 1–28.