JS Journal of Historical Syntax
Volume 9, Article 6: 1–38, 2025 · DOI: 10.18148/hs/2025.v9i2-10.208

VARIABILITY AND LEARNING IN LANGUAGE CHANGE: THE CASE OF V2*

Marc Meisezahl
University of Konstanz &
University of Edinburgh, Centre for Language Evolution

Simon Kirby
University of Edinburgh, Centre for Language Evolution

Jennifer Culbertson
University of Edinburgh, Centre for Language Evolution

Abstract The loss of V2 has received considerable attention in the past with some theories linking it to learning (e.g. Lightfoot 1999, Yang 2002). Here, we use artificial language learning experiments to test, in a controlled setting, what factors affect learning of V2. Specifically, we build on previous work demonstrating a general beneficial effect of input variability. We explore the role of variation in clause-initial constituents by comparing artificial languages that differ both in the kinds of grammatical functions that tend to appear in initial position, and the level of variability present. We find that these different distributions of clause-initial constituents indeed affect V2 learning outcomes. However, contrary to our predictions, a language with the highest level of variability is not the best learnt. Rather, a language containing many adjunct-initial sentences was learnt best. We discuss the possibility that a high quantity of clause-initial adjuncts is in fact important to acquiring V2 grammars in natural language. We find further support for this in corpus data indicating a high proportion of adjunct-initial sentences in stable V2 languages and a low proportion in languages that had been in

* We are indebted to Mora Maldonado for helping with the experiment design, the analysis and providing parts of the code. Furthermore, we thank Alexander Martin and the audiences of CONSOLE 30 in Nantes and DiGS 23 in New York and two anonymous reviewers for valuable input. This project received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 757643), and from ESRC grant ES/R011869/1.

the process of losing V2. We also discuss the role of variability in grammatical *categories* rather than roles, which might give languages with many clause-initial adjuncts an advantage. Taken together, our findings establish the first evidence for a causal link between the reduction of evidence and the loss of V2.

1 Introduction

Most modern Germanic languages are verb second (V2) languages. That is, the verb is obligatorily realised in the second position of a clause and no restrictions apply as to the grammatical function or category of the clause-initial constituent (Holmberg 2015). For instance in German, subjects (1a), objects (1b), adjuncts (1c) and past participles can all occupy the initial position (1d), but verbs must be in second position.

(1) a. *Die Maus* ***geniesst*** *eine Scheibe Raclettekäse* *auf der Alm.*
the mouse enjoys a slice raclette cheese on the alp
'The mouse enjoys a slice of raclette cheese on the alp.'
b. *Eine Scheibe Raclettekäse* ***geniesst*** *die Maus* *auf der Alm.*
a slice raclette cheese enjoys the mouse on the alp
c. *Auf der Alm* ***geniesst*** *die Maus* *eine Scheibe Raclettekäse.*
on the alp enjoys the mouse a slice raclette cheese
d. *Genossen* ***hat*** *die Maus* *eine Scheibe Raclettekäse* *auf der Alm.*
enjoyed has the mouse a slice raclette cheese on the alp

While on the surface, many English sentences also follow a V2 pattern, subjects must occupy the preverbal position (2a), and additional elements in the left-periphery lead to V>2 orders (2b).

(2) a. *A slice of raclette cheese **enjoys** the mouse on the alp.
b. On the meadow, the mouse **enjoys** a slice of raclette cheese.

This reflects change: earlier stages of English followed a word order pattern akin to (1) (van Kemenade 1987, Fischer, van Kemenade, Koopman & van der Wurff 2001), a pattern of change that is attested in a number of other languages (Willis 1998, Meelen 2016, Wolfe 2018). The loss of V2 has been the subject of longstanding research, with some theories tying it to learning. Like any other feature of language, evidence for V2 has to be sufficient in the input learners receive, otherwise it will not be acquired (Lightfoot 1999, Yang 2000). For V2, it has been claimed that exposure to sentences with verbs in second position does not on its own suffice. Rather, non-subject-initial sentences, like the ones in (1), form a crucial part of the evidence for V2 (Yang 2000).

This is supported by historical data showing a link between the increase in subject-initial sentences and loss of V2 (e.g. Roberts 1993: 199). Intuitively, if a large proportion of the sentences in the input involve SV orders, then the learner may acquire a grammar similar to modern English, rather than V2. In principle, exposure to a preponderance of sentences with *any* particular constituent type in first position (e.g. subject, object, etc.) could lead the learner to an analysis that favours a non-V2 grammar.

In this paper, we use a novel experimental method to explore the role of variation in the initial constituent in learning V2. Specifically, we ask whether learning V2 in a miniature artificial language is affected by the level of variability in the grammatical roles (i.e., subject, object and adjunct) of clause-initial constituents. We are specifically interested in the way participants generalise from the input. Forming abstract grammatical representations from the input constitutes a central task in language learning. In the case of V2, learners need to build representations without a fixed mapping of grammatical roles to the clause-initial position. If more variability increases the likelihood of generalisation – i.e., to novel types of constituents – and less variability decreases this likelihood, we will have the first direct evidence for a causal link between the frequency of non-subject-initial sentences and the loss of V2. If a learner fails to generalise to novel initial constituents but instead learns a more constrained grammar, where a specific grammatical role or set of roles appear in initial position, then the first steps toward the loss of V2 have essentially been taken. When considered on a larger timescale, even a weak tendency toward learning a more constrained grammar can be amplified through cultural transmission (Kirby, Cornish & Smith 2008, Kirby, Tamariz, Cornish & Smith 2015) with the result that V2 is lost at the population-level.

The results of our experiment suggest that the learning and hence generalisation of a V2 grammar is indeed affected by variation, however we found that learning is best when the language exhibits a large amount of *adjuncts* in clause-initial position and worst when objects dominate the clause-initial position. Although this does not exactly fit our initial hypothesis, which concerned variation in grammatical functions, our results still provide evidence for a fostering effect of variability on learning. Specifically, this may indicate that instead of a high variability of grammatical functions, a high variability of grammatical categories (e.g. DPs, PPs & AdvPs) in the clause-initial position is what fosters generalisation. This finding is further supported by a large-scale corpus study on the distribution of clause-initial constituents in German. In line with previous studies on Germanic and Romance, we find that adjuncts account for the most frequent constituent type in clause-initial

position after subjects.

This paper is structured as follows: We first motivate our hypothesis (§2). We will then report the results of our artificial language learning experiment (§3). In §4, we will present a large-scale, multi-corpus analysis of variation in the clause-initial constituent in a natural V2 language, German. We will conclude with a discussion of what our findings mean for research on the historical loss of V2 in §5.

2 The role of learning in the loss of V2

V2 is a cross-linguistically rare phenomenon (Holmberg 2015: 343).[1] There are some languages (including most Germanic languages other than English) which have exhibited relatively strict V2 since their earliest records, in some cases developing even stricter V2 order over time (Eythórsson 1995, Axel 2007, 2009, Axel-Tober 2018, Þorgeirsson 2012). There are other languages which have had and lost, to varying degrees, their V2 status (including English (van Kemenade 1987, Roberts 1996), the Romance languages (Benincà 1995, Wolfe 2018) and Welsh (Willis 1998, Meelen 2016)). Why is V2 rare, and what leads some languages to retain, and even strengthen V2, while others lose it?

One possibility is that the rarity and fragility of V2 is due to the kind of evidence that is needed for learners to acquire it. In a prominent account, Yang (2000, 2002) argues that the loss of V2 is tied to changes in the linguistic input of learners. For example, looking at the case of French, Yang (2000) argues that to retain a V2 grammar, unambiguous evidence to support V2 (i.e. OVS and XVSO sentences, must outnumber the evidence against V2 (i.e., SXVO and XSVO). Crucially, because Middle French was a *pro*-drop language, non-subject-initial V2 sentences could not provide unambiguous evidence for V2: a V2 analysis i.e., [X V *pro*], or an SVO analysis i.e., [X *pro* V], are both possible. Yang (2000) uses counts from Roberts (1993: 148, 155) to argue that in Middle French, unambiguous V2 sentences decreased to such an extent that a SVO grammar gained an advantage over a V2 grammar. The cue-based learning model of Lightfoot (1999, 2006) also takes a learner-centred approach to the loss of V2. In this model, cues – $_{CP}$[XP $_{C}$V…] in the case of V2 (Lightfoot 2006: 86) – need to be sufficiently expressed in the input to ensure successful learning (Lightfoot 2006: 82). Both of these accounts suggest that sufficient unambiguous evidence for V2 must be present for successful acquisition. But they also highlight the role of variability in sentence structures as a key aspect of evidence for V2. For example, it is not just the absence of V3 sentences that

1 To the best of our knowledge, 21 languages have been categorised as V2.

matters, but the presence of OVS and XVS alongside SVO sentences.

There is independent evidence from various domains suggesting that greater variability can in some cases benefit learning (see Raviv, Lupyan & Green (2022) for a recent review). For example, Gómez (2002) and Gómez & Maye (2005) show that learning of non-adjacent dependencies is successful only when there is sufficient variation in the elements that occur between two dependents. For example, learners exposed to sequences of the type 'aXc' only learn that 'a' elements must be followed by 'c' elements when there are sufficiently many different 'X' elements. Here, variability helps learners to focus on key patterns of interest, and rule out irrelevant information: variability in the intervening elements helps learners to move from a focus on transitional probabilities between adjacent elements (e.g. 'aX' or 'Xc') to the non-adjacent dependency. The elements in these experiments could in principle represent grammatical categories (e.g. pronoun + verb + agreement marker), constructions (e.g. *be* + verb + *-ing*), or grammatical functions (Gómez 2002: 431). While the current study does not focus on dependencies, the intuition remains the same: variability helps draw the learner's attention to critical constraints – here the position of verb – and provides evidence for what is unconstrained – the sentence-initial position. In line with Yang (2000) and Lightfoot (1999, 2006), we specifically pursue the idea that more variation *in the clause-initial constituent* provides more robust evidence for generalised XP-movement and a V2 grammar. In particular, we predict that V2 grammars are learnt best when variation in the types of grammatical roles (i.e., subjects, objects and adjuncts) of clause-initial constituents is highest. The degree to which participants generalise XP-movement will affect the preservation of V2 in the grammar. A weaker generalisation will eventually lead to the loss of V2 as constraints on the clause-initial position are amplified through cultural transmission.[2,3] We test this prediction using an artificial language learning experiment.

2 See also Cournane & Klævik-Pettersen (2023) for another account of the loss (and rise) of V2 that focuses on the role of learning. Their account highlights the conservative nature of learners when it comes to the acquisition of syntactic structures. Note that the account of Cournane & Klævik-Pettersen (2023) and our account are not necessarily in conflict but can complement each other.

3 It should also be noted that the positive effects of variation in this context do not contradict the potential negative effects of variation elsewhere in learning. For example, when a fixed rule must be learned, then variation in the form of exceptions can be problematic for learning. For example, according to the Tolerance Principle (Yang 2016), rule learning is sensitive to a specific threshold of exceptions. Specifically, the Tolerance Principle predicts that if the number of exceptional forms (e.g. irregular past tense forms) remains below a threshold, a rule (e.g. past tense) will be considered productive by learners. Crucially in the context of V2, the rule learners need to acquire is that the mapping between first position and grammatical role is *not* fixed. Thus in this case, more variability provides more evidence.

3 Artificial language learning experiment

To test our hypothesis experimentally, we conducted an artificial language learning (ALL) experiment. ALL studies allow researchers to create miniature linguistic systems in which variables of interest can easily be manipulated and variables not of interest controlled (for a review see Culbertson & Schuler 2019). It has been shown that artificial languages are learnt in similar ways to natural languages (Ettlinger, Morgan-Short, Faretta-Stutenberg & Wong 2016), and there is substantial evidence that learners' preferences in ALL studies align with linguistic typology (Culbertson 2012, 2017, 2023, Culbertson & Schuler 2019). Consequently, participants' behaviour in experiments provides an important source of evidence from which to draw conclusions about the link between learning and natural language structure. Previous studies have suggested that it is possible to learn V2 in an artificial language (Getz 2018, 2019, Rebuschat & Williams 2012, Tagarelli, Ruiz, Vega & Rebuschat 2016, Ruiz, Tagarelli & Rebuschat 2018). In our study, participants learn a novel miniature artificial language involving English lexical items which conform to a (non-English-like) V2 grammar. The verb always comes second, but our study design manipulates the distribution of clause-initial elements participants are exposed to. We included three conditions: a uniform condition in which subjects, direct objects and adjuncts occurred equally frequently in clause-initial position and two conditions with skewed distributions where either adjuncts or objects accounted for the majority of clause-initial constituents. We predict learners in the uniform condition will be more likely to acquire a V2 grammar compared to the skewed conditions. Importantly, we measure learning V2 in terms of the critical feature of V2 languages: generalisability of clause-initial position (equivalent to XP-fronting).

3.1 Methods

The design, the hypotheses, predictions and analyses were preregistered prior to data collection. We implemented the experiment using the JavaScript library jsPsych (de Leeuw, Gilbert & Luchterhandt 2023). All materials are available online.

3.1.1 Participants

314 participants were recruited online using Prolific. By using in-built Prolific filters, the participant pool was restricted to United States nationals, who are

monolingual speakers of English that were also raised monolingually.[4] We also used Prolific filters to exclude participants whose subject at university was English literature, English language or languages more broadly. Finally, only participants with an approval rating of 95% or higher were invited to participate. Following our pre-registered exclusion criterion, 82 participants were excluded due to low performance in the first half of the experiment (cf. §3.1.3) and two participants had to be excluded due to knowledge of a V2 language (as determined by a post-experiment questionnaire). Data analysis therefore included 74 participants in the uniform condition and 78 participants in each of the two skewed conditions.

3.1.2 Materials

Stimuli sentences were constructed using a semi-artificial language. The vocabulary of the language consisted of English lexical items, but the word order followed a non-English-like V2 pattern. All sentences were comprised of a verb, a subject, an object, an adjunct, and an additional adverb. The verb was always in second position, with either a subject, an object, or an adjunct in initial position. The adverb served as an additional cue to the non-English-like structure of the language, cf. (3). It always appeared to the right of the finite verb (unlike in English), indicating movement of the verb out of the VP/vP to a higher functional projection (Vikner 1995, Waldmann 2008, Westergaard 2009a). Using English lexical items allowed us to train participants on the grammar in a short time (given the length and complexity of the sentences), and allowed us to control for lexical novelty (as described below) in testing whether participants generalise XP-fronting to novel types of constituents in each condition. A number of other studies have demonstrated that results obtained with semi-artificial languages can be replicated with a fully-artificial language (Culbertson & Adger 2014, Martin, Ratitamkul, Abels, Adger & Culbertson 2019, Martin, Holtz, Abels, Adger & Culbertson 2020).

The distribution of elements (subject, object, adjunct) in first position depended on the condition. In the *uniform* condition, all three were equally likely to appear in initial position. In the *object-dominant* condition, objects were more likely to appear in initial position. In the *adjunct-dominant* condition, adjuncts were more likely to appear in initial position. Because we use an English lexicon, and English is SVO, we did not run a subject-dominant condition. Participants were randomly assigned to a condition.

4 The experiment received ethical approval from the ethics board of the Linguistics and English Language department at the University of Edinburgh (180-2021/2). All participants gave informed consent before their participation.

Condition	Subject-initial	Object-initial	Adjunct-initial
Uniform	30	30	30
Object-dominant	18	54	18
Adjunct-dominant	18	18	54

Table 1 Distribution of subject-initial, object-initial and adjunct-initial sentences in each of the three conditions.

We created 30 unordered sets of constituents used for training. For each set, three different sentences – subject-initial, object-initial and adjunct-initial – were created yielding 90 sentences in total. This process is illustrated in (3) (bold, italics and underline used here for illustrative purposes only). Each constituent was a phrase made up of a single word or two words. Constituents in each set were unique, and no constituent was used in more than one set.

(3) {revises, in Boston, a novel, the author}

a. *The author* **revises** eventually a novel in Boston.
b. A novel **revises** *the author* eventually in Boston.
c. In Boston **revises** *the author* eventually a novel.

Subjects were always animate and objects inanimate. Subjects were DPs or proper nouns, objects were DPs. To further facilitate identification of grammatical roles, only verbs denoting irreversible actions were included. Adjuncts were temporal or locative adverbs, PPs or few adverbially used DPs.

For the uniform condition, we used all subject-initial, object-initial and adjunct-initial sentences created from the 30 unordered sets. That is, each constituent type occurred equally frequent in clause-initial position in the training. By contrast, for the skewed conditions, the dominant element accounted for 60% of all sentences, whereas non-dominant elements accounted for 20% each (Table 1). We chose 60% as frequency for the dominant constituent type as this lies within the range of the dominant constituent in V2 languages (cf. §4). As a result of the skew, only a subset of all possible combinations could be included and some sentences were repeated. For eight of our thirty unordered sets, we replaced one of the sentence variants with a non-dominant constituent in clause-initial position with the variant featuring the dominant constituent (i.e., either adjunct or object) in clause-initial position. For example, participants in the adjunct-dominant condition could have seen (3a) once and (3c) twice during training. We repeated sentences

rather than including new ones in order to control lexical variation across conditions. Sets were randomly selected for repetition for each participant. We assigned each sentence of the same set to three different blocks. The distribution of clause-initial elements in each block adhered to the same overall pattern (i.e. uniform or skewed). For each condition, three training lists were created from the three different blocks by Latin Square. The order in each block was randomised for each participant.[5] The first training block was used for the reading trials and the second and third block for the production trials (cf. §3.1.3).

For the testing phase, two different sets of materials were constructed. The first type was used for production testing, and consisted of unordered sets of constituents that were presented to participants as buttons they could choose to create a sentence (see §3.1.3). Each set could contain the same types of constituents featured in training (i.e., subject, direct object, adjunct, adverb) or they could contain one of two novel constituent types. In *complex adjuncts* trials, adjuncts comprised of three words (i.e., containing an additional determiner or modifier as in (4b)) rather the simple adjuncts used in training (as in (4a)). These trials allow us to test whether participants are sensitive to constituent length when they make generalisations about what can be fronted. In the *indirect object* trials, an indirect object replaced the adjunct (4c). These trials allow us to test whether participants generalise XP-fronting to novel constituent types.

(4) a. {the driver, delivers, grumpily, the food, this afternoon}
b. {Jayden, sweeps, halfheartedly, the floor, in the bathroom}
c. {Charles, suggests, cheekily, a whiskey, to the friend}

Four sets of each trial type – i.e., seen in training (i.e., with a simple adjunct), complex adjunct, or indirect object – were constructed, for twelve total sets. All constituents apart from adverbs were novel in the sense that they were not words or phrases seen in the training stimuli. Adverbs could be repeated from the training stimuli.

The second set of testing items was constructed for use in the sentence judgement phase (see §3.1.3). These items were created by crossing two factors: VERB POSITION (V2 or V3) and INITIAL CONSTITUENT (simple adjunct, complex adjunct, direct object, or indirect object). (5a) and (5b) exemplify the V2 sentences with initial indirect objects and complex adjuncts, respectively. Similarly, V3 sentences with initial indirect object and complex adjunct are illustrated in (6a) and (6b). For each factor combination, four sentences were created. We applied the same construction criteria on the

5 Examples of a training set for each condition can be found here.

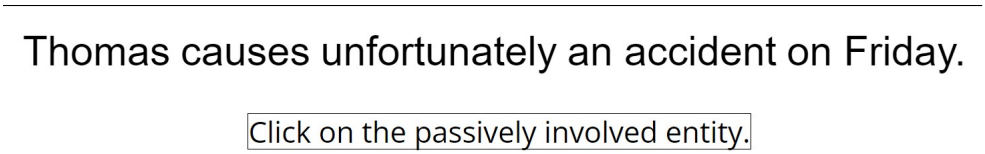

Figure 1 Example trial for reading task during training. Constituents were revealed one at a time, i.e., *Thomas*, *causes*, *unfortunately*, *an accident* and *on Friday*. Participants were then prompted to identify a particular constituent by clicking on it.

different constituents as described above. Again, only adverbs were repeated, all other constituents involved new lexical items.

(5) a. To the congregation **shows** *the priest* silently the candle.
b. In late April **regrets** *the politician* openly his misconduct.

(6) a. To the doctor *the patient* **describes** precisely the pain.
b. At the moment *the referee* **verifies** briefly the decision.

3.1.3 Procedure

Participants accessed the experiment through a web browser on their personal computer or laptop. At the start of the experiment, participants were informed that they would be learning a recently discovered dialect of English that differs from other varieties of English in greater flexibility of the word order. The experiment was divided into a training phase and a testing phase. Participants were assigned randomly to one of three training lists per condition. The training phase was comprised of two parts: sentence reading and production. Before the reading trials, participants were told they would see sentences and be asked to identify either the actor of the action (i.e., subject), the passively involved entity (i.e., object), the action (i.e., verb) or the time/location of the action (i.e., adjunct). Then, on each trial, a sentence was revealed, one constituent at a time, with a delay of 500ms between constituents in order to give participants sufficient time for reading. After the full sentence was visible, participants were asked to click on one of the constituents, as in Figure 1. Feedback was provided after each trial (shown for 450ms if correct; 850ms if incorrect). Participants' performance on this task was used as pre-registered exclusion criteria: only participants who achieved

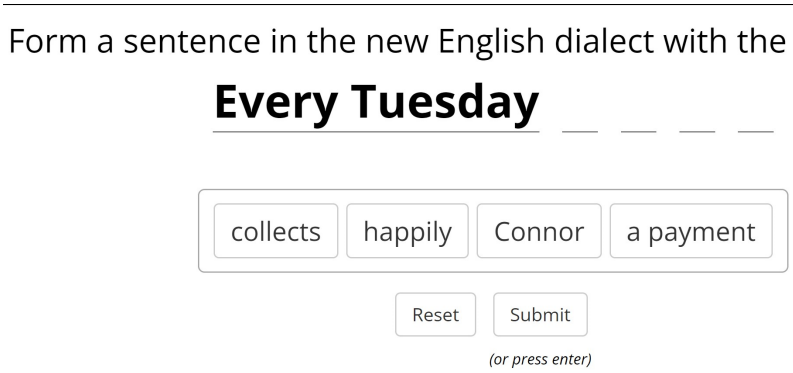

Figure 2 Example trial for production task during training. The initial constituent was always shown on the screen, the remaining constituents were provided as buttons beneath. Participants had to simply click on words to construct a sentence. The order of the constituents was randomised for each participant.

a score of 90% or higher were included for further analysis (cf. §3.1.1).[6,7] Participants completed 30 trials of this kind.

Participants then moved on to production training (see Hopman & MacDonald 2018 for evidence of the benefit of production during language learning). On each trial, an initial constituent was provided together with four blank lines, as in Figure 2. The remaining constituents appeared underneath, each in a separate box. Participants were instructed to fill in the blanks by clicking to insert each constituent into the sentence. All words had to be used. The order of the buttons was randomised for each trial and participant. Participants received feedback after each trial (shown for 1500ms if correct; 3000ms if incorrect). Any incorrectly placed constituents in the sentence produced were highlighted in red. Participants completed 60 trials of this type.

6 Note that this task is mainly to check that participants are attending to the training. It does not require participants to actually learn anything about the language.

7 A reviewer asks whether the high number of exclusions (cf. §3.1.1) is caused by the difficulty of the attention task, citing the technical terminology (e.g. *passively involved entity* for objects) as a potential reason. We do not think that this is the case. First, participants were given a thorough explanation of this task. Second, conditions were affected differently. In the uniform and adjunct-dominant condition, we had to exclude 23 and 19 participants, respectively. The attrition rate for the object-dominant condition was much higher with 40 participants. This suggests that the issue lies with the difficulty participants have in learning the object-dominant pattern, discussed further below.

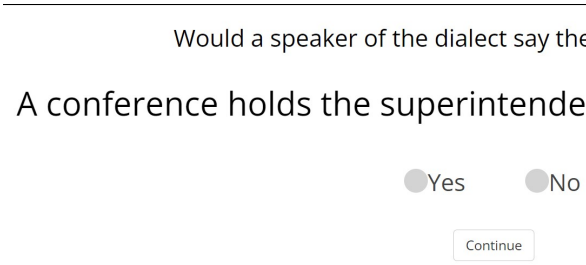

Figure 3 Example trial for judgement task during testing. Participants had to decide whether the sentence could be uttered by a speaker of the language.

Part	Task type	*n* trials	Feedback
Training	Reading	30	yes
	Production	60	yes
Testing	Production	12	no
	Judgement	32	no

Table 2 Summary of experimental design.

The testing phase was comprised of two parts: sentence production and judgement. Sentence production in testing was identical to production in training, except no initial constituent was provided, and no feedback was given. Participants completed 12 trials of this kind. In the judgement task, participants were asked to judge whether a speaker of the dialect would say a given sentence. On each of the 32 trials, a sentence appeared, and participants chose 'yes' or 'no' as in Figure 3. Again no feedback was given. The experiment finished with a questionnaire inquiring participants' language background and strategies employed during the experiment. Table 2 provides a summary of the outlined design.

3.2 Predictions

Recall that we operationalised learning of a V2 language as generalising, or extrapolating XP-fronting to novel constituent types: namely indirect objects and complex adjuncts. Our hypothesis was that greater variability in types of clause-initial constituents should aid learning of V2. Accordingly, we made

two specific predictions, both measuring the degree to which participants generalise XP-fronting. First, we predicted that participants in the uniform condition would be more likely to produce or accept V2 sentences with initial novel constituent types than participants in the skewed conditions. Second, we predicted that the difference in acceptability between ungrammatical V3 sentences and grammatical V2 sentences with novel constituent types would be greater in the uniform condition compared to the skewed conditions. Recall that neither of these sentence types will have been seen by participants, thus in principle they could both be treated as ungrammatical in the language. However, if participants learn XP-fronting as a generalisable feature, then they should judge fronted novel constituent types as grammatical but V3 sentences as ungrammatical. We predicted learners in the uniform condition should be more likely to do this than learners in the skewed conditions. We made no predictions regarding any differences between the two skewed conditions.

3.3 Results

Our analyses were conducted in R (R Core Team 2020) using the packages lme4 for the statistical analysis (Bates, Mächler, Bolker & Walker 2015) and ggplot2 for plotting (Wickham 2016). The results of the hypothesis-confirming and exploratory analyses are summarised in Table 3.

3.3.1 Hypothesis-confirming analysis

Figure 4 shows the proportion of different constituent types in V2 sentences produced by learners in each condition. Recall that indirect objects and complex adjuncts are *novel* constituent types, and all remaining constituents are *familiar* (i.e. seen in initial position during training). Regardless of the type, all constituents use novel lexical items. We first assess whether learners in the uniform condition produced more V2 sentences with the novel constituent types in clause-initial position. We fitted a mixed-effects logistic regression model to all V2 sentences that include a novel constituent type in the sentence production test. The dependent variable was production of a clause-initial novel constituent (=1) or an familiar constituent type (=0). The model included CONDITION (object-dominant, adjunct-dominant or uniform) as a fixed effect, and by-participant and by-item random intercepts. CONDITION was treatment coded with object-dominant as baseline. The model revealed that participants in the adjunct-dominant condition used more novel constituents clause-initially than the participants in the object-dominant condition ($\beta = 3.46$, SE $= .45$, $p = 2.16{\times}10^{-14}$). Learners in

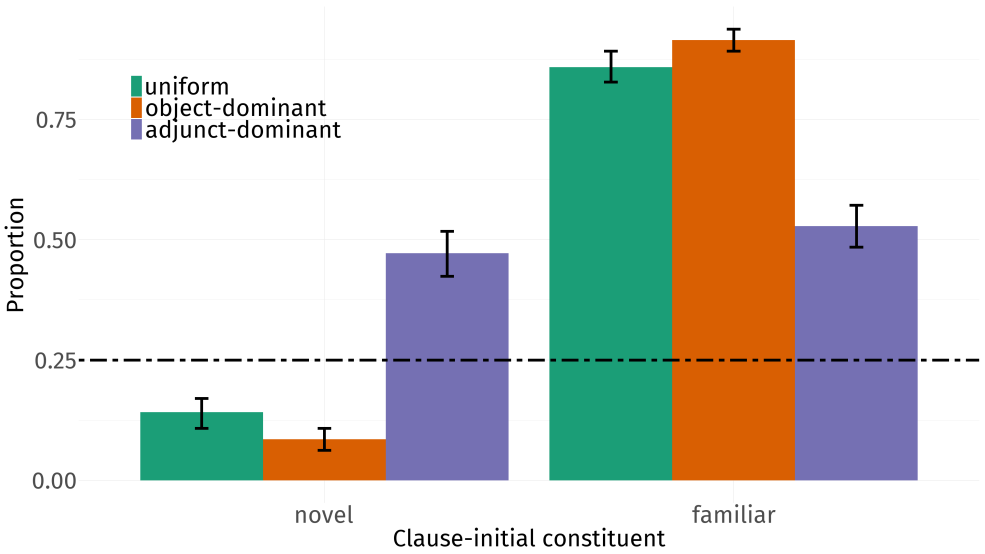

Figure 4 Proportion of novel and familiar constituent types in V2 sentences by condition in the sentence production test. The latter are the complement of the former, hence the values of each condition add up to 1. Error bars represent bootstrapped 95% confidence intervals around by-participant means; the dotted line indicates chance level. Participants in the adjunct-dominant condition produce significantly more V2 sentences with novel constituents in initial position than the other two conditions. Participants in the uniform condition in turn produce significantly more novel constituent types in the clause-initial position.

the uniform condition also produced significantly more novel constituents clause-initially than learners in the object-dominant condition (β = .91, SE = .45, p = .04). To directly compare the adjunct-dominant and uniform condition, we fitted a model with identical effect structure to the same data with the adjunct-dominant condition as baseline. This model revealed that participants in the adjunct-dominant condition used more novel constituents clause-initially than the participants in the uniform condition (β = −2.55, SE = .43, p = 2.75×10^{-9}). These findings do not straightforwardly match our predictions. First, the skewed conditions unexpectedly differ from one another. Second, while learners in the uniform condition indeed fronted

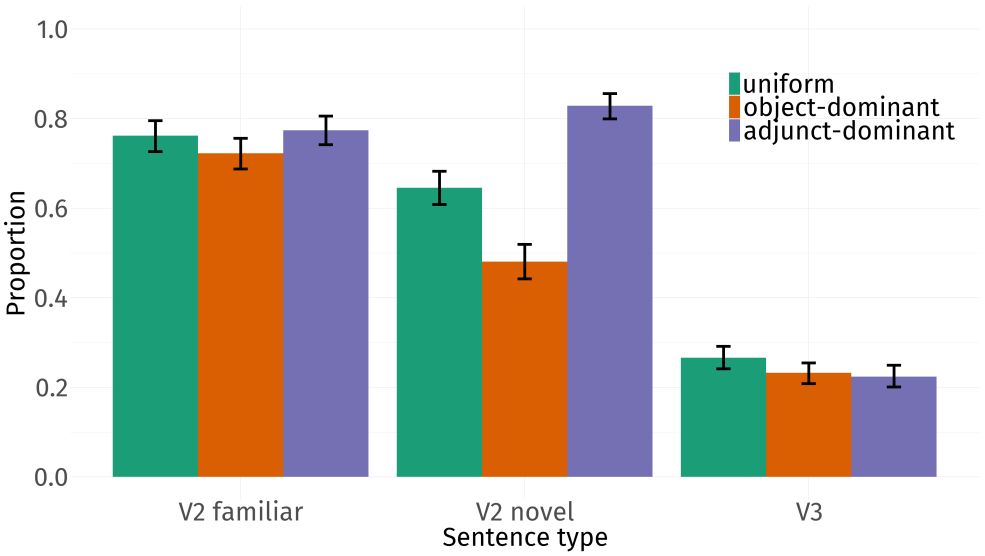

Figure 5 Acceptance rates of V2 sentences in the sentence judgement test across three sentence types: V2-familiar, i.e. with clause-initial constituent types seen during training, V2-novel, i.e., with novel clause-initial constituent types, V3, i.e., ungrammatical sentences. Error bars indicate bootstrapped 95% confidence intervals of the mean. V2-novel sentences were accepted at a higher rate in the adjunct-dominant condition compared to other two conditions. V2-novel sentences were also accepted at a higher rate in the uniform condition compared to the object-dominant condition.

more novel constituents than those in the object-dominant condition, they fronted *fewer* novel constituents than learners in the adjunct-dominant condition.

Turning to the judgement data, we created a new factor SENTENCE TYPE by grouping V2 sentences with initial simple adjuncts and direct objects together as *V2-familiar*.[8] V2 sentences with an initial complex adjunct or an indirect object were grouped together as *V2-novel*. All remaining combinations were grouped together as *V3*. The ratings for all three sentence types are shown in Figure 5. We then tested whether learners in the uniform condition were (i)

8 We did not include subject-initial *V2-familiar* sentences as those would be too similar to participants' L1.

more likely to accept V2-novel sentences compared to the skewed conditions and (ii) less likely to accept V3 sentences compared to learners in the skewed conditions. We fitted a mixed-effect logistic regression model to the V2-novel and V3 data. The model included CONDITION and SENTENCE TYPE as fixed effects as well as an interaction term for both. The model also included by-participant and by-item random intercepts and by-participant random slopes for SENTENCE TYPE. Both fixed effects were treatment coded with the object-dominant condition and V2-novel as reference level. To assess the first prediction, we investigated the simple effect of condition. The model indicated that participants in the uniform condition were more likely to accept V2-novel sentences than those in the object-dominant condition, as predicted ($\beta = .99$, SE $= .31$, $p = .001$). In fact, the latter group were not significantly more likely than chance to accept such sentences ($\beta = -0.096$, SE $= .25$, $p = .70$). The model further showed that V2-novel sentences were significantly more likely to be accepted by learners in the adjunct-dominant condition compared to the object-dominant condition ($\beta = 2.44$, SE $= .32$, $p = 3.16\times10^{-14}$). To compare the adjunct-dominant and uniform conditions, a further model was fitted to the data with identical effect structure but the adjunct-dominant condition as baseline. This model revealed that learners in the uniform condition were significantly less likely to accept V2-novel sentences compared to participants in the adjunct-dominant condition, contrary to our prediction ($\beta = -1.45$, SE $= .32$, $p = 7.01\times10^{-6}$).

To assess the second prediction, i.e., the discrimination of grammatical V2-novel sentences and ungrammatical V3 sentences, we looked at the interaction between CONDITION and SENTENCE TYPE. As Figure 5 suggests, V3 clauses were generally less likely to be accepted than V2-novel sentences by learners in the object-dominant condition ($\beta = -1.58$, SE $= .36$, $p = 1.04\times10^{-5}$). We did not find a significant interaction between V3 and the uniform condition ($\beta = -0.82$, SE $= .46$, $p = .07$) suggesting that the discrimination was not greater in the uniform condition, contrary to our prediction. At the same time, participants in the adjunct-dominant condition were better than those in the object-dominant condition at discriminating V2-novel and V3, indicated by a significant interaction of V3 and the adjunct-dominant condition ($\beta = -2.79$, SE $= .48$, $p = 4.63\times10^{-9}$). When learners in the uniform condition were directly compared to those in the adjunct-dominant condition, we again found greater discrimination of V2-novel and V3 for learners in the adjunct-dominant condition ($\beta = 1.95$, SE $= .47$, $p = 2.87\times10^{-5}$). Taken together, our second prediction was therefore not borne out: we only found a learning advantage for participants in the adjunct-dominant condition.

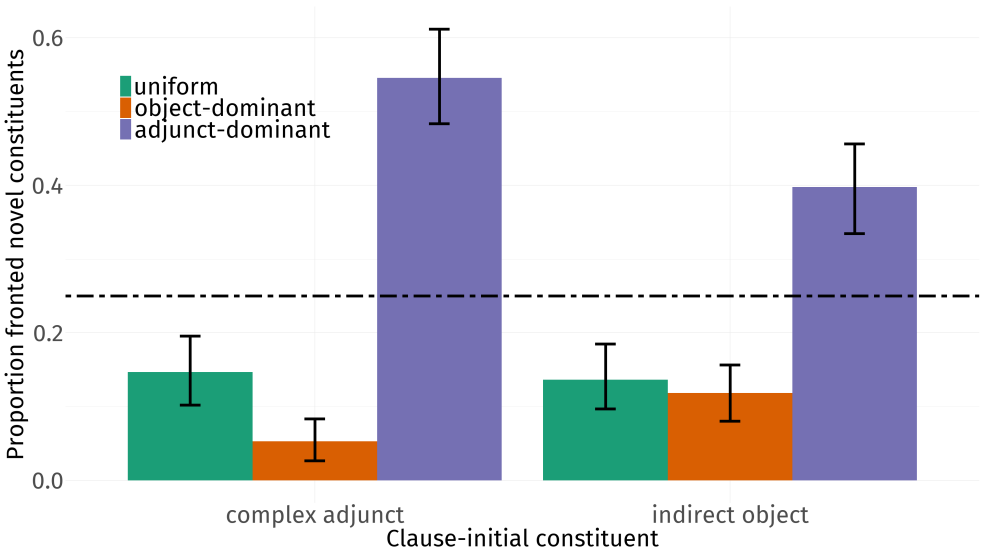

Figure 6 Proportion of fronted complex adjuncts and indirect objects in V2 sentences by condition in the sentence production test. Error bars indicate bootstrapped 95% confidence intervals, while the dotted line represents the chance level. Learners in the adjunct-dominant condition use both constituents types with significantly higher proportion in initial position than learners in the other two conditions.

3.3.2 Exploratory analysis

Recall that we used indirect objects and complex adjuncts to measure participants' generalisation of V2 in both testing tasks. Arguably, these two types of novel constituents are different from one another. Complex adjuncts are the same type of constituent as simple adjuncts, but longer than any of the constituents encountered during training. Indirect objects are a completely novel type of constituent, which participants have not encountered in the language at all. It might be that learners in the adjunct-dominant condition exhibited a clear advantage over learners in the other two conditions on the grounds of their familiarity with clause-initial adjuncts in general. If this were the case, we should find a difference between complex adjuncts and indirect objects. As Figure 6 shows, participants in the adjunct-dominant condition produced both complex adjuncts and indirect objects more frequently in initial

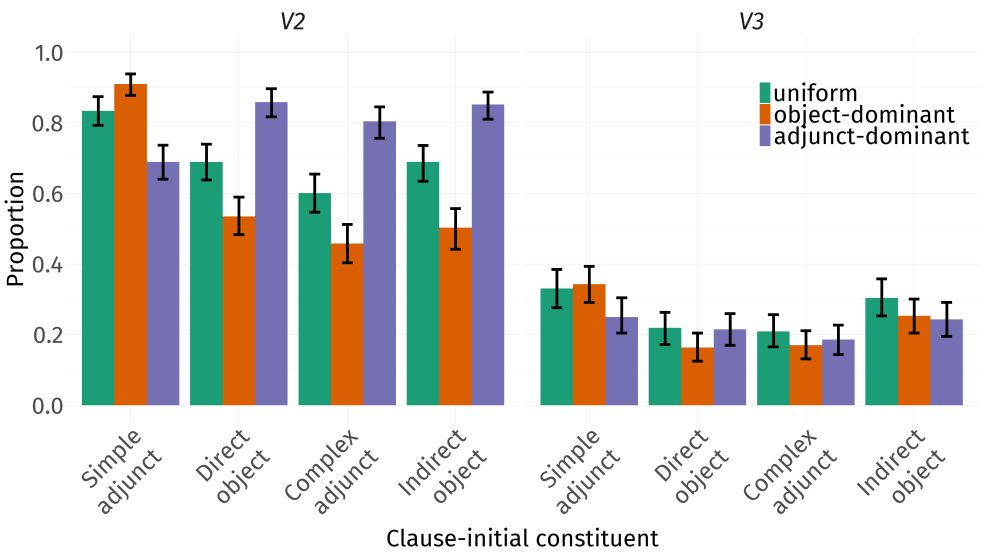

Figure 7 Acceptance rate of V2 and V3 sentences with different clause-initial elements by condition in the sentence judgement task. Error bars indicate bootstrapped 95% confidence intervals. Participants in the adjunct-dominant condition are more likely to produce both simple and complex adjuncts in clause-initial position compared to the other two conditions.

position than participants in the other conditions. This is confirmed in two mixed-effects logistic regression models fitted to V2 sentences that include a complex adjuncts and indirect objects, respectively. The dependent variable was initial constituent type (either complex adjunct or indirect object=1, other constituents=0). The model included CONDITION as fixed effect and by-participant and by-item random intercepts. CONDITION was treatment coded with the adjunct-dominant condition as baseline. Participants in the adjunct-dominant condition placed significantly more complex adjuncts and indirect objects in clause-initial position compared to participants in the object-dominant condition (complex adjuncts: $\beta = -5.96$, SE $= .95$, $p = 2.84\times10^{-10}$; indirect objects: $\beta = -3.57$, SE $= .82$, $p = 1.41\times10^{-5}$) and the uniform condition (complex adjuncts: $\beta = -4.43$, SE $= .92$, $p = 1.72\times10^{-6}$; indirect objects: $\beta = -3.18$, SE $= .85$, $p = 1.75\times10^{-4}$).

Figure 7 depicts the acceptance rate of V2 and V3 sentences with different initial constituents. Similar to the production data, V2 sentences with initial complex adjuncts and indirect objects were both more likely to be accepted in the adjunct-dominant condition compared to the other two. This was confirmed with two mixed-effect logistic regression models fitted to participants' judgements of sentences with clause-initial complex adjuncts and indirect objects, respectively. The model included CONDITION as fixed effect and by-participant and by-item random intercepts. CONDITION was treatment coded with adjunct-dominant as baseline (complex adjuncts: uniform: $\beta = -0.55$, SE = .21, $p = .01$; object-dominant: $\beta = -1.17$, SE = .22, $p = 6.52\times10^{-8}$; indirect objects: uniform: $\beta = -0.31$, SE = .22, $p = .02$; object-dominant: $\beta = -1.02$, SE = .22, $p = 3.15\times10^{-6}$). To summarise, neither the production nor judgement data support the idea that the learning advantage observed for participants in the adjunct-dominant condition can be attributed to the similarity between simple and complex adjuncts.

Our hypothesis was specifically about the generalisability of XP-fronting, the placement of verbs within sentences is obviously one of the defining features of V2 too. We thus conducted an exploratory analysis of the verb placement in sentence production. We fitted a mixed-effect logistic regression model to participants' production data. The dependent variable was the position of the verb (second position=1, not in second position=0). The model included CONDITION and ADDITIONAL CONSTITUENT (i.e., simple adjunct, complex adjunct and indirect object), as fixed effects as well as their interaction. The latter was included to be sure that verb order was not conditioned on which other constituent was present apart from subject, verb and direct objects. The model also included by-participant random slopes for ADDITIONAL CONSTITUENT. Both fixed effects were treatment coded with object-dominant and simple adjunct as baseline. The model revealed no differences in the likelihood of producing the verb in second position across conditions or additional constituents ($\beta_{min} = -0.62$, $\beta_{max} = 0.14$, $p_{min} = .12$, $p_{max} = .91$).

We conducted a further exploratory analysis, to investigate the acceptance rate of V2-familiar sentences in the sentence judgement test. We did this to check whether the learning advantage for the adjunct-dominant condition is also visible for sentence types that are familiar to participants. Figure 5 indicates a generally high acceptance rate for V2-familiar sentences across all three conditions, although the adjunct-dominant condition does show the highest ratings. We fitted a mixed-effect logistic regression model to all V2-familiar sentences with CONDITION as fixed effect and by-participant and by-item random intercepts. CONDITION was again treatment coded with

the object-dominant condition as the baseline. The model did not indicate a significant difference between the object-dominant condition and the uniform condition ($\beta = .376$, SE $= .23$, $p = .11$) and the adjunct-dominant condition ($\beta = .31$, SE $= .23$, $p = .19$). This finding suggests that participants in all three conditions learnt the language during training equally well. However, it is worth noting that acceptance ratings do differ to some degree depending on the type of initial constituent, as shown in Figure 7. Surprisingly, learners in the object-dominant condition are most likely to accept sentences with clause-initial simple adjuncts, while learners in the adjunct-dominant condition are most likely to accept sentences with clause-initial direct objects. To test this statistically, we fitted a mixed-effects logistic regression model to all V2 sentences with direct objects and simple adjuncts in clause-initial position. As we were interested in the difference between the skewed conditions, the uniform condition were not included in this analysis. The model included fixed effects for CONDITION and INITIAL CONSTITUENT as well as their interaction and by-participant random slopes for INITIAL CONSTITUENT. Both fixed effects were sum-coded with the object-dominant condition and direct objects as reference levels. The model revealed a main effect for INITIAL CONSTITUENT ($\beta = .36$, SE $= .16$, $p = .02$) but not for CONDITION ($\beta = .12$, SE $= .15$, $p = .43$). This suggests the skewed conditions do not differ with respect to their grand means, and that adjuncts were overall rated a bit higher than direct objects. The model further showed a significant interaction between INITIAL CONSTITUENT and CONDITION ($\beta = -1.12$, SE $= .14$, $p = 9.84{\times}10^{-16}$). This confirms our observation that the acceptance rates for these two constituent types differ across conditions. It is unclear why we see this unexpected pattern of results, and it is contradicted by the production data, where, for example, participants in the object-dominant condition were highly likely to front direct objects (see Figure 9).

3.4 Discussion

This experiment investigated whether the distribution of initial constituents in the input impacts learning of V2. Following Yang (2000), we identified non-subject-initial V2 sentences as a crucial type of evidence for V2. Our hypothesis, inspired by evidence for the benefit of variability in other domains (cf. Raviv et al. 2022), was that high variability in initial grammatical functions would aid learners in identifying a key feature of V2, the generalisability of XP-fronting. To test this, we compared three distributions of clause-initial elements: a uniform distribution, an object-dominant skewed distribution, and an adjunct-dominant skewed distribution. We taught participants a semi-artificial language, with English vocabulary but V2 word order. We then

	Comparison	Finding
Hypothesis-conf. analysis	Produced novel const. types in initial position in V2 sentences	A-dom > O-dom Uni > O-dom A-dom > Uni
	Acceptance rate V2-novel	Uni > O-dom A-dom > O-dom A-dom > Uni
	Discrimination V2-novel & V3 in judgements	Uni = O-dom A-dom > O-dom A-dom > Uni
Exploratory analysis	Produced complex adjuncts in initial position in V2 sentences	A-dom > O-dom A-dom > Uni
	Produced indirect objects in initial position in V2 sentences	A-dom > O-dom A-dom > Uni
	Acceptance rate V2 sentences with initial complex adjuncts	A-dom > O-dom A-dom > Uni
	Acceptance rate V2 sentences with initial indirect objects	A-dom > O-dom A-dom > Uni
	Produced V2 sentences	no differences
	Acceptance rate V2-familiar sentences	O-dom = A-dom O-dom = Uni

Table 3 Summary of the main findings of the hypothesis-confirming and exploratory analyses for the production and judgement data comparing the uniform (uni), adjunct-dominant (A-dom) and object-dominant (O-dom) conditions. '>' indicates a statistically significant contrast, '=' a non-significant contrast.

asked them to produce and judge sentences with novel constituents in the clause-initial position. These could be either complex adjuncts – longer than any initial constituents seen during training – or indirect objects – a grammatical role not seen in initial position during training. We analysed participants' extrapolation of the clause-initial position to these novel constituent types.

First, it is worth noting that participants in all three conditions were able to learn the requirement that the verb be in second position. This suggests that at least this aspect of V2 is readily learnable in an artificial language. However, the results regarding the generalisability of XP-fronting were mixed. Participants in the uniform condition fronted more novel constituent types in pro-

duction and were more likely to accept sentences with novel constituent types in clause-initial position compared to participants in the object-dominant condition. However, the apparent advantage of the uniform condition over the object-dominant condition did not extend to the adjunct-dominant condition. Instead, participants in the adjunct-dominant condition were more likely to produce and accept sentences with novel clause-initial constituents compared to both other conditions. Participants in the adjunct-dominant condition were also better at discriminating between grammatical and ungrammatical (V3) sentences. Our exploratory analysis suggested that the learning advantage for the adjunct-dominant condition is not attributable to the similarity between simple and complex adjuncts: participants in the adjunct-dominant condition were also more likely to produce and accept fronted indirect objects. Further exploratory analyses also indicated no overall differences between the three conditions for ratings of V2 sentences with initial familiar constituent types and the frequency of V2 productions. This suggests that a skewed distribution with adjuncts as the dominant element mainly affected generalisation to novel structures.

To summarise, the uniform language gave learners an advantage over the object-dominant condition, but contrary to our prediction, the opposite was the case for the adjunct-dominant condition. Why would there be such a substantial difference between the two skewed conditions? One possibility is that this difference is due to the influence of English on participants' perceptions of sentences in the novel language. This is particularly relevant here as we used English vocabulary. For example, it may be that for English speakers, object-initial sentences are particularly unexpected. Recall that all objects in the language were inanimate NPs. If participants generally assume that whichever NP is first will be the subject, then object-initial sentences will result in a garden-path effect, or at least a semantic clash, since the inanimate NPs cannot be coerced into subjects. By contrast, the violation induced by adjunct-initial sentences may not be perceived as equally serious; it is a syntactic violation, since the subject is not in the expected position, but it is less likely to induce a garden-path. Further, in English, although displacement of both arguments (7a) and adjuncts (7b) to the clause-initial position is possible, displaced adjuncts are preferred over displaced arguments (Doherty 2005).[9]

(7) a. A block of Emmentaler *the mouse* **found** in the pantry.
b. In the morning, *the mouse* **devours** Appenzeller cheese.

9 Overall, structures like these are a minority. According to Yang (2000: 242), non-subject-initial sentences account for less than 10% of all cases in the Penn Treebank.

Thus, compared to the adjunct-dominant condition, the object-dominant condition may be more different from participants' native language. This might lead participants in the object-dominant condition to learn the V2 grammar less well, particularly relative to the adjunct-dominant condition, but also to the uniform condition (which still has fewer object-initial sentences). However, it is worth noting that our results suggest that familiar V2 sentences were actually learnt equally well across conditions. It is not entirely clear why a difference in similarity to English would specifically affect generalisation.

It is perhaps worth noting here that English has been described as residual V2 language (e.g Rizzi 1996, Holmberg 2015, Sailor 2020). That is, certain structures still require V2 orders in present-day English such as *wh*-questions (8a) and locative inversions (8b). One might thus expect that these vestiges of a former fully-fledged V2 grammar could constitute another way in which participants' native language affect their performance in the experiment.

(8) a. What kind of cheese **does** the *cheese monger* recommend?
b. Here **is** *your loaf of cheese*.

However, there are good reasons to believe that V2 residues in the English grammar played little to no role in the experiment. Work in L1 acquisition has demonstrated the conservative manner with which learners approach the learning task (Westergaard 2009a). For instance, the Norwegian dialect of Tromsø does not require a strict V2 order in *wh*-questions in that V2 is contingent on the length of the *wh*-word and information-structural aspects (e.g. Westergaard & Vangsnes 2005). Crucially, Westergaard (2009a) showed that learners exhibit target-like structures from early on while no erroneous generalisations to other contexts (such as declaratives) are being made. Similarly for L2/L3/L*n* acquisition, Westergaard (2021) argues for a property-by-property transfer from previously learnt language(s) during the acquisition of a new language. Under such a model, any influence from the structure of English *wh*-questions on declaratives (as used in the present experiment) is not predicted. Accordingly, locative inversion in English often occurs with unaccusative verbs (e.g, *come*, *sit*) and *be*, that is a well-defined class of verbs (Westergaard 2009b: 68). Generalising this pattern is thus also not expected. Moreover, structures like (8b) are relatively rare (Anderssen & Bentzen 2018: 15) such that any influence would presumably be minimal. Besides, the same question we noted in the previous paragraph arises: why would the residual V2 grammar affect generalisation in particular?

A second possibility is that there is a genuine learning advantage of having a high proportion of adjunct-initial sentences. In our stimuli, subjects and objects were always DPs, and were not distinguished based on any formal

criteria (i.e., case marking). If learners were sensitive to variability in initial constituents defined in terms of grammatical category ([XP-V]) instead of grammatical function ([S|O|A-V]), as we assumed, a high proportion of adjuncts would mean more variability in the input. While Yang (2000, 2002) assumes the latter, the view that grammatical categories are the relevant cue to V2 is indeed taken by Lightfoot (1999, 2006). When the conditions are reconsidered from this perspective, learners in the uniform condition saw 33.3% non-DPs and 66.6% DPs clause-initially, learners in the object-dominant condition saw 20% non-DPs and 80% DPs, but learners in the adjunct-dominant condition saw 60% of the initial constituents were non-DPs and 40% DPs. The object-dominant condition thus exhibits the greatest skew and hence the least evidence for generalised XP-fronting, while the adjunct-dominant condition which exhibits the most uniform distribution and thus the most evidence for generalised XP-fronting. This aligns perfectly with our results: generalisation was best in the adjunct-dominant condition followed by the uniform condition, followed by the object-dominant condition. Learning in the adjunct-dominant condition may have been further facilitated by our use of both PPs and AdvPs as adjuncts; this would increase the variability of grammatical categories even more.

It is also possible that both of these explanations for our results are at play: the adjunct-initial condition might facilitate learning the most due to its high level of variability in the grammatical category of initial constituents, while the object-dominant condition might lead to particular poorly learning because it has both low category variability *and* is least similar to English. Future experimental work could tease these apart.[10] Below, we further examine the distribution of clause-initial elements in natural V2 languages. If adjuncts

10 We thank an anonymous reviewer for pointing out a third possible explanation for the advantage of the adjunct-dominant condition over the object-dominant condition: learners may tend to search for the base word order of a new language using the position of verbal arguments. This is conceivable given the early sensitivity of children to the argument structure of verbs (Naigles 1990, Perkins & Lidz 2021). If such a bias was in place, participants in the object-dominant condition would face a significantly more difficult task in that they need to overcome the native English SVO order in favour of an OVS word order. This is potentially supported by the fact that the majority of all productions exhibit an OV order and hardly any SV orders (cf. Figure 9). At the same time, there are two issues with this analysis. On the one hand, the results illustrated in Figure 7 are not compatible with this explanation. If learners posit an underlying OVS order, sentences with clause-initial direct objects should exhibit the highest acceptance rate, not sentences with clause-initial adjuncts. On the other hand, it remains unclear what base word order learners in the adjunct-dominant condition would assume as both subjects and objects occur with the same frequency in the clause-initial position. That is the input contains inconclusive evidence as to the base order. One would need to stipulate that learners transfer the SVO order of their native language in order to explain the observed learning advantage of the adjunct-dominant condition.

are indeed beneficial for the acquisition of V2, one may expect to find them over-represented in initial position.

4 The distribution of clause-initial constituents in Germanic

In the previous section, we tested whether greater variability in the distribution of grammatical roles in clause-initial position in the input would lead to better learning of V2. Contrary to our predictions, participants in the adjunct-dominant condition exhibited the best learning outcomes. As discussed above, one possible explanation is that learners benefit from a high frequency of clause-initial adjuncts in the input for the acquisition of V2. Under such an analysis, although there will be many factors that determine the distribution of different elements in initial position, one may expect to find adjuncts over-represented clause-initially in V2 languages precisely because they support learning of V2. The present section will summarise findings in the literature suggesting that adjuncts indeed occur frequently in initial position (§4.1). We will corroborate these earlier findings with the results of a large-scale corpus study of German (§4.2).

4.1 Previous evidence

A number of previous studies have explored the distribution of clause-initial constituents in contemporary Germanic languages.[11] In this section, we will discuss some of these studies, summarised in Table 4. An early study on German by Winter (1961), examined 63,000 sentences from diverse sources (theatre, fiction, non-fictional prose, scientific texts) and reported a heavily skewed distribution for written German: subjects dominate the clause-initial position (66.7%) followed by adverbs and PPs (28.1%). Clause-initial direct objects, on the other hand, are very infrequent, occurring in only 2.9% of all clauses. Crucially, such a skewed distribution appears not to be confined to formal registers: Bohnacker & Rosén (2008) observe a similar distribution in elicited informal texts. Similarly, Engel (1974: 212) provides evidence for a skewed distribution in spoken language, with subjects occurring most in clause-initial position (51.31%), again followed by adjuncts (35.35%) and objects (9.25%).

A skewed distribution of initial constituents is not unique to German and has been noted for other Germanic languages as well. Bohnacker & Rosén (2008) and Bohnacker & Lindgren (2014) show for spoken and written

11 Note that there is not a consistent definition of adjuncts in the studies reported here and in §4.3 and results have been reported with different levels of detail.

Author	Lang	Modality	*n*	S	DO	A	Other
Win(61)	Ger	written	63,000	66.7	2.9	28.1	2.3
Eng(74)	Ger	spoken	5,000	54.31	9.25	35.35	4.09
B&R(08)	Ger	written	1,173	50	7	42	1
B&R(08)	Sw	written	545	71.2	4.4	22.6	1.8
B&L(14)	Sw	written	680	67.0	1.3	32.4	0.3
B&L(14)	Sw	spoken	755	64.4	1.6	33.8	0.3
B&L(14)	Dt	written	646	53.0	2.3	44.7	9.0
B&L(14)	Dt	spoken	711	60.9	1.0	38.1	0.0
Yan(00)	Dt	spoken	n/a	66.8	1.2	23	7
Pug(19)	Dn	spoken	500	62.0	9.4	24.4	4.2

Table 4 Proportion of clause-initial subjects (S), direct objects (DO), adjuncts (A) and other elements in German (Ger), Swedish (Sw), Danish (Da) and Dutch (Dt). Authors are abbreviated as follows: Win(61) = Winter (1961), Eng(74) = Engel (1974), B&R(08) = Bohnacker & Rosén (2008), B&L(14) = Bohnacker & Lindgren (2014), Yan(00) = Yang (2000), Pug(19) = Puggaard (2019). Despite differences between but also within languages, all of the listed studies found a skewed distribution of clause-initial elements in different V2 languages. Adjuncts are the second-most frequent element in initial position after subjects. However, note the small sample size of many of these studies.

Swedish that, as in German, subjects are the most frequent clause-initial element while direct objects are the least frequent type. Similar patterns have also been observed for spoken Danish (Puggaard 2019)[12] and for spoken and written Dutch (Bohnacker & Lindgren 2014). Finally, Yang (2000: 242) found a similar skew even in child-directed speech in Dutch. Despite these similarities, the actual frequencies of different constituent types varies between (but also within) languages. Two observations from Table 4 are worth pointing out here: first, adjuncts occur relatively frequently in clause-initial position even though preposed adverbs and PPs occur more frequently in Dutch and German than in other Germanic languages. Second, other frontable constituents such as indirect objects are exceedingly rare in first position (Winter 1961: 201, Puggaard 2019: 298).

12 The authors thank Rasmus Puggaard for bringing this paper to our attention.

Corpus	*n* sentences
wiki	45.5 M
europarl	2.2 M
speeches	619,152

Table 5 Number of sentences in the three examined subcorpora of TüBa-D/DP, i.e. Wikipedia (wiki), proceedings of the European Parliament (europarl) and speeches of German officials (speeches) (de Kok & Pütz 2019: 1).

Although frequent clause-initial adjuncts appear to be a cross-Germanic feature, a cautious interpretation of the previous studies is still warranted. All studies (with the notable exception of Winter (1961)) relied on a relatively small sample. Here, we conducted a large-scale corpus study on the distribution of clause-initial elements in present-day German aimed at replicating the general findings reported in previous smaller studies.

4.2 Evidence from a large-scale corpus study on German

To test the validity of previous findings, we conducted a corpus study using a significantly larger sample.[13] We chose the dependency treebank TüBa-D/DP (de Kok & Pütz 2019) as all sentences are annotated for their position in the topological field (cf. Drach 1937, Wöllstein 2010). This annotation enables automatic identification of V2 and non-V2 clauses as well as clause-initial constituents. TüBa-D/DP consists of four different subcorpora: Wikipedia articles (wiki), proceedings of the European Parliament (europarl), speeches of German officials (speeches) and newspaper articles from *taz*. The first three are publicly available and were used as the basis for the present study. While the three corpora differ in size, as illustrated in Table 5, even the smallest of the three (speeches) is still significantly larger than any of the corpora in §4.1.

We built a custom Python script to count where and with what frequency each syntactic function (as annotated in the corpora) occurs relative to the finite verb. Only sentences ending in a full stop or colon were taken into consideration as the word order in interrogative as well as imperative and exclamative sentences may differ. Furthermore, sentences with fewer than two words were excluded. Crucially, each clause in a sentence was considered

13 The experimental code as well as the analysis is available online.

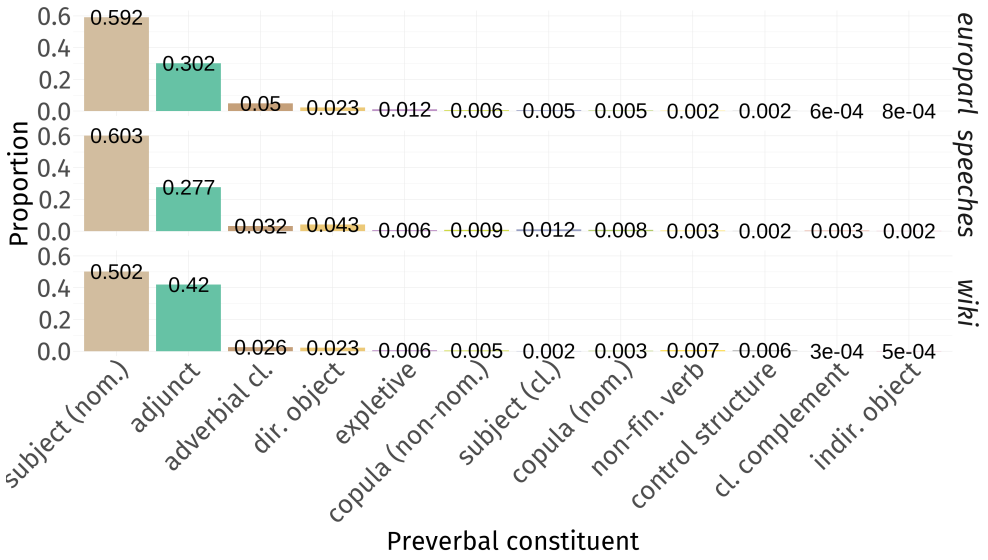

Figure 8 Proportion of different constituents in clause-initial position in V2 clauses in the Proceedings of the European Parliament (europarl), speeches of German officials (speeches) and Wikipedia (wiki). All three corpora show a skewed distribution similar to the ones previously observed in the literature (cf. §4.1). Note the parallels between the two spoken corpora europarl and speeches in terms of the frequencies of subjects, adjuncts and direct objects.

separately given that multiple V2 clauses are grammatical.

Due to the different sizes of the three corpora, distributions were plotted separately for each corpus in Figure 8.[14] However, all three corpora illustrate a similar pattern: non-clausal subjects are the dominating clause-initial element followed by adjuncts, while direct objects occur only very infrequently in initial position. Adverbial clauses (which were considered distinct from adjuncts due to their clausal status) are more frequently attested in initial position in europarl and wiki corpus than direct objects.[15] All other types of

14 Attentive readers may have noticed that the proportions do not add up to 1 but rather to a value between approximately 0.96 and 0.98. A likely explanation for this divergence lies in errors in the part-of-speech tagging.

15 Note that this is a somewhat artificial distinction as adverbial clauses are formally adjunct CPs. However, we wanted to underscore the fact that they are comparable in frequency with direct objects.

Language	Subj	Obj	Adj
Old French	46.32	12.84	40.84
Old Occitan	23.78	25.00	51.22
Old Sicilian	47.63	12.62	39.75
Old Venetian	74.34	9.26	16.40
Old Sardinian	62.26	18.24	19.50
Old Spanish	35.32	6.42	58.26

Table 6 Distribution of different clause-initial constituents in V2 sentences in several Old Romance languages (Wolfe 2018: 25). Note that Wolfe (2018) makes a more fine-grained distinction for adjuncts that are subsumed under the label adjunct here. All of the listed languages show a substantial amount of adjunct-initial sentences. In Old Venetian and Old Sardinian the proportion of initial adjuncts is lower but still larger than the proportion of initial objects.

constituents that the grammar permits in initial position are very rare. This is particularly striking in the case of indirect objects, which are perfectly grammatical in initial position, and yet almost never occur there in these corpora.

To summarise, our corpus study confirms earlier findings of a skewed distribution of clause-initial constituents in V2 sentences.[16] To summarise, there is robust evidence that adjuncts appear very frequently in initial position in the Germanic languages.

4.3 Comparison with Old Romance

The way in which different constituent types are distributed in the clause-initial position has also been systematically studied for the Medieval Romance languages (e.g. Labelle & Hirschbühler 2018: 281, Wolfe 2018: 25). Here, there is also evidence that adjuncts were strikingly frequent. Consider the proportion of different types of constituents in clause-initial position according to Wolfe (2018: 25) summarised in Table 6. The frequency of adjuncts in

16 We also examined whether the skew persists when frequencies of clause-initial elements proportional to their base rates are considered. Intuitively, one could imagine that the base rate of subjects occurring in a sentence is greater than the base rate of adjuncts, and thus the former have a greater probability of appearing in initial position simply due to this. However, Monte Carlo simulations revealed that constituents are still skewed, with subjects and adverbial clauses more likely to be fronted given their base rate. In the europarl and wiki corpus, adjuncts also appear more frequently in clause-initial position than expected.

these languages was comparable, if not higher, to the ones found for modern Germanic languages (cf. Table 4). Importantly, in later stages, at least for French, when V2 is being lost, a different picture emerges. Steiner (2014: 129), for instance, notes an increase of SV structures from the 13th (47.11%) to the 16th century (62.88%) in V2 sentences.[17] Simultaneously, the frequency of null subjects remained mostly constant (approx. 30%) in the same contexts. These data support the possibility that the decline of V2 grammar is connected to the decline in adjunct-initial sentences.

In English, which has also lost V2, non-subject-initial constructions like (9) are possible.

(9) Here **is** *the platter with gruyère cheese*.

However as noted above, there is evidence that these are extremely low frequency (Anderssen & Bentzen 2018: 15).[18] More generally, in English, it is very uncommon for any element other than the subject to be in initial position (less than 10% of sentences according to Yang 2000: 242).[19] This is in line with the idea that there is something special about having adjuncts frequently in initial position for the maintenance of V2.

5 General discussion

V2 is a striking feature of almost all Germanic languages, and yet it is cross-linguistically rare, and its loss has been well-documented, including in English, Welsh and almost all Romance language. The loss of V2 has been tied to a lack of sufficient evidence in learners' input, crucially a lack of non-subject-initial V2 sentences (Lightfoot 1999, 2006, Yang 2000, 2002, Willis 1998, Meelen 2016). Combined with domain-general evidence for the benefit of variability on learning (cf. Raviv et al. 2022), we hypothesised that a V2 language will be learnt best if the clause-initial position is occupied by a maximally diverse

17 Adjunct-initial and object-initial sentences are not separately considered by Steiner (2014). However, assuming a lower proportion of object-initial sentences is justified given the prevalence of this pattern across V2 language.

18 Roeper (1999: 175) however reports incidental evidence that children generalise this pattern to some extent for a very brief period:

(i) what **calls** *that*
'What is it called'
(Roeper 1999: 175)

19 See also Westergaard, Lohndal & Lundquist (2023) for evidence of the relationship between the production of non-subject-initial declaratives and V2 errors in heritage speakers of Norwegian living in the USA.

set of grammatical functions, i.e. subjects, direct objects and adjuncts. We conducted an ALL experiment to test this hypothesis. We compared three conditions, a uniform condition in which subjects, direct objects and adjuncts occurred with identical frequency in the clause-initial position. Additionally, we ran two skewed conditions – one skewed towards direct objects and one towards adjuncts. Contrary to our prediction, learning was best in the adjunct-dominant condition followed by the uniform condition. Two possible explanations were identified. First, it could be that object-initial sentences are particular odd for our native English-speaking participants, and the condition which has the least such sentences – the adjunct-dominant condition – is the best learnt. Second, there could be a genuine advantage of having adjuncts in clause-initial position with a high frequency. Although further work is necessary to distinguish between the two explanations, the second is supported by the fact that our adjunct-initial condition actually had the most evidence for variability in terms of grammatical categories (rather than roles or functions). Our results therefore provide some degree of support for the view held by Lightfoot (1999, 2006) who points to changes in the distribution of clause-initial categories as an explanation for the loss of V2. In our case, the adjunct-initial condition featured a more balanced mixed of DPs, PPs, and AdvPs, while the object-initial condition featured an over-representation of DPs in initial position. The former led to better generalisation of XP-fronting, and the latter to diminished generalisation of XP-fronting. Interestingly, in this condition, participants' productions were even more skewed than their input, as can be seen in Figure 9. Such a heavily skewed distribution in productions was not observed for any of the other conditions. In fact, participants in the adjunct-dominant condition produced the least skewed distribution. If the output of learners in the object-dominant condition served as the input for a next generation of learners, then the evidence for V2 in this generation would be even more reduced (Kirby et al. 2008). This constitutes a plausible trajectory for the loss of V2, paralleling the situation e.g. in French where the evidence for V2 was more and more reduced (Roberts 1993, Steiner 2014).

If adjuncts are important more generally in the learning of V2, and not just in our experiment, one might expect to see a significant number of clause-initial adjuncts in exactly those Germanic languages that have retained a V2 grammar. To test this claim, we reviewed work on the distribution of clause-initial constituents in Germanic languages. We also reported a large-scale corpus study on German. This analysis confirmed that while subjects occur most frequently in the initial position, the next most common constituent type is adjuncts. By contrast, in languages which are in the process of losing V2, there is evidence for a particular reduction in frequency of adjuncts in initial

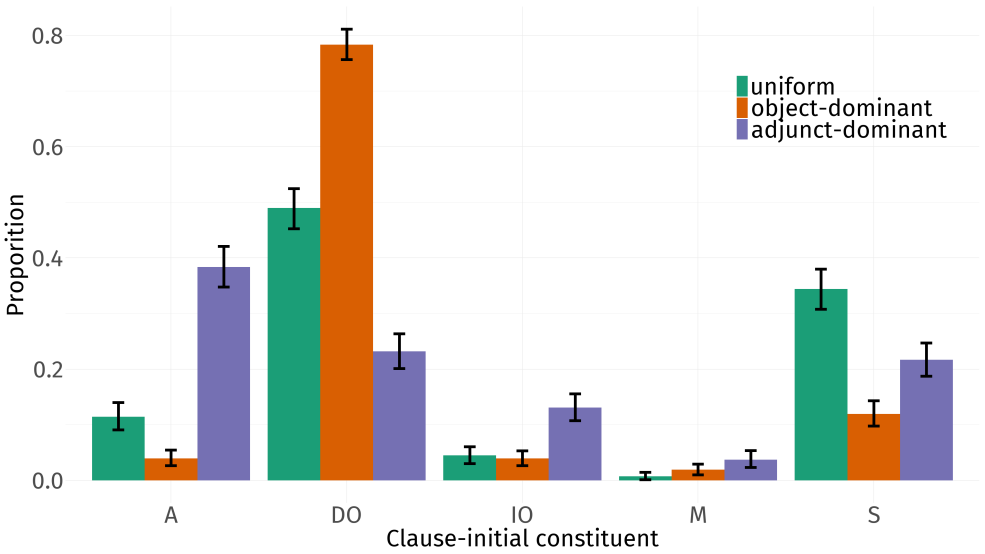

Figure 9 Proportion of different clause-initial constituent types across different trial types in V2 sentences by condition in the sentence production test of the experiment (cf. §3). A = adjuncts, DO = direct objects, IO = indirect objects, M = adverbials to mark verbal movement, S = subjects. Learners in the object-dominant condition produced a high number of object-initial sentences. This exceeds their input significantly, which contained only 60% of clause-initial objects.

position. In languages, like English, which no longer have V2, adjuncts are very unlikely to occur in initial position. These distributional differences between natural languages support the idea that adjuncts may play a special role in the learning and maintenance of V2.

6 Conclusion

In this paper, we investigated the hypothesis that learning of a V2 language is fostered by a maximally variable distribution of clause-initial constituents in the input. This hypothesis was derived from work suggesting that a V2 grammar will be lost if learners do not receive sufficient evidence for the grammar (Lightfoot 1999, 2006, Yang 2000, 2002), and that variability in general is good

for learning (Gómez 2002). We tested this by comparing learning of an artificial V2 language with subjects, objects and adjuncts in clause-initial position with equal frequency, with learning of languages with skewed frequencies – either object-dominant or adjunct-dominant. While our results suggest that different distributions of clause-initial elements do indeed affect learning outcomes, learners were best able to generalise XP-fronting to novel constituent types when the distribution of initial elements was skewed towards adjuncts. They were least able to generalise when the distribution was skewed towards objects. We propose that a high frequency of adjuncts in initial position is in fact likely to be an important feature of V2. It may lead to higher variability in the grammatical categories of elements in first position, which could be more important than the variation in grammatical roles (Lightfoot 1999, 2006). Further, there is robust evidence, including from our large-scale corpus study of German, showing that adjuncts are highly frequent in initial-position in current (or historical) V2 languages, but not in a non-V2 language like English. Our results therefore support the idea that diminished evidence in the input can lead to the loss of V2. Our study also adds to the body of literature now demonstrating the utility of artificial language learning studies in understanding language typology and change.

References

Anderssen, Merete & Kristine Bentzen. 2018. Different outcomes in the acquisition of residual V2 and do-support in three Norwegian-English bilinguals: Cross-linguistic influence, dominance and structural ambiguity. *Frontiers in Psychology* 9. 1–18. doi:10.3389/fpsyg.2018.02130.

Axel, Katrin. 2007. *Studies on Old High German syntax: Left sentence periphery, verb placement and verb-second*. Amsterdam, Philadelphia: John Benjamins. doi:10.1075/la.112.

Axel, Katrin. 2009. The consolidation of verb-second in Old High German: What role did subject pronouns play? In Paola Crisma & Giuseppe Longobardi (eds.), *Historical syntax and linguistic theory*, 128–143. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199560547.003.0008.

Axel-Tober, Katrin. 2018. Origins of verb-second in Old High German. In Agnes Jäger, Gisella Ferraresi & Helmut Weiß (eds.), *Clause structure and word order in the history of German*, 22–47. Oxford: Oxford University Press. doi:10.1093/oso/9780198813545.003.0003.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. doi:10.18637/jss.v067.i01.

Benincà, Paola. 1995. Complement clitics in medieval Romance: the Tobler-

Mussafia law. In Adrian Battye & Ian Roberts (eds.), *Clause structure and language change*, 325–344. New York, Oxford: Oxford University Press. doi:10.1093/oso/9780195086324.003.0012.

Bohnacker, Ute & Josefin Lindgren. 2014. Fundament, formellt subjekt och frekvens: Ordföljdsmönster i svenska, nederländska och hos vuxna inlärare av svenska. *Språk & Stil* 24. 33–71.

Bohnacker, Ute & Christina Rosén. 2008. The clause-initial position in L2 German declaratives: Transfer of information structure. *Studies in Second Language Acquisition* 30(4). 511–538. doi:10.1017/S0272263108080741.

Cournane, Ailís & Espen Klævik-Pettersen. 2023. The role of the conservative learner in the rise and fall of verb-second. *Journal of Historical Syntax* 7(9). 1–48. doi:10.18148/hs/2023.v7i6-19.160.

Culbertson, Jennifer. 2012. Typological universals as reflections of biased learning: Evidence from artificial language learning. *Language and Linguistics Compass* 6(5). 310–329. doi:10.1002/lnc3.338.

Culbertson, Jennifer. 2017. New approaches to Greenbergian word order dependencies. In N. J. Enfield (ed.), *Dependencies in language: On the causal ontology of linguistic systems*, 23–38. Berlin: Language Science Press. doi:10.5281/zenodo.573778.

Culbertson, Jennifer. 2023. Artificial language learning. In Jon Sprouse (ed.), *The Oxford handbook of experimental syntax*, 271–300. Oxford: Oxford University Press. doi:10.1093/oxfordhb/9780198797722.013.9.

Culbertson, Jennifer & David Adger. 2014. Language learners privilege structured meaning over surface frequency. *Proceedings of the National Academy of Sciences* 111(16). 5842–5847. doi:10.1073/pnas.1320525111.

Culbertson, Jennifer & Kathryn Schuler. 2019. Artificial language learning in children. *Annual Review of Linguistics* 5. 353–373. doi:10.1146/annurev-linguistics-011718-012329.

de Kok, Daniël & Sebastian Pütz. 2019. TüBa-D/DP stylebook.

de Leeuw, Joshua R., Rebecca A. Gilbert & Björn Luchterhandt. 2023. jsPsych: Enabling an open-source collaborative ecosystem of behavioral experiments. *Journal of Open Source Software* 8(85). 1–4. doi:10.21105/joss.05351.

Doherty, Monika. 2005. Topic-worthiness in German and English. *Linguistics* 43(1). 181–206. doi:10.1515/ling.2005.43.1.181.

Drach, Erich. 1937. *Grundgedanken der deutschen Satzlehre*. Frankfurt am Main: Diesterweg.

Engel, Ulrich. 1974. Syntaktische Besonderheiten der deutschen Alltagssprache. In Hugo Moser (ed.), *Gesprochene Sprache. Jahrbuch 1972*, 199–228. Düsseldorf: Schwann.

Ettlinger, Marc, Kara Morgan-Short, Mandy Faretta-Stutenberg &

Patrick C.M. Wong. 2016. The Relationship between artificial and second language learning. *Cognitive Science* 40(4). 822–847. doi:10.1111/cogs.12257.

Eythórsson, Thórhallur. 1995. *Verbal syntax in the early Germanic languages*. Ithaca, NY: Cornell University dissertation.

Fischer, Olga, Ans van Kemenade, Willem Koopman & Wim van der Wurff. 2001. *The syntax of Early English*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511612312.

Getz, Heidi. 2018. *Sentence first, arguments after: Mechanisms of morphosyntax acquisition*. Washington, DC: Georgetown University dissertation.

Getz, Heidi. 2019. Acquisition of morphosyntax: A pattern learning approach. In David W. Lightfoot & Jonathan Havenhill (eds.), *Variable properties in language. Their nature and acquisition*, 103–114. Washington, DC: Georgetown University Press.

Gómez, Rebecca & Jessica Maye. 2005. The developmental trajectory of nonadjacent dependency learning. *Infancy* 7(2). 183–206. doi:10.1207/s15327078in0702_4.

Gómez, Rebecca L. 2002. Variability and detection of invariant structure. *Psychological Science* 13(5). 431–436. doi:10.1111/1467-9280.00476.

Holmberg, Anders. 2015. Verb second. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – Theory and analysis. An international handbook*, 342–383. Berlin, München, Boston: De Gruyter Mouton. doi:10.1515/9783110377408.342.

Hopman, Elise W.M. & Maryellen C. MacDonald. 2018. Production practice during language learning improves comprehension. *Psychological Science* 29(6). 961–971. doi:10.1177/0956797618754486.

Kirby, Simon, Hannah Cornish & Kenny Smith. 2008. Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences* 105(31). 10681–10686. doi:10.1073/pnas.0707835105.

Kirby, Simon, Monica Tamariz, Hannah Cornish & Kenny Smith. 2015. Compression and communication in the cultural evolution of linguistic structure. *Cognition* 141. 87–102. doi:10.1016/j.cognition.2015.03.016.

Labelle, Marie & Paul Hirschbühler. 2018. Topic and focus in Old French V1 and V2 structures. *Canadian Journal of Linguistics* 63(2). 264–287. doi:10.1017/cnj.2017.52.

Lightfoot, David. 1999. *The development of language: Acquisition, change, and evolution*. Malden, MA: Blackwell.

Lightfoot, David. 2006. *How new languages emerge*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511616204.

Martin, Alexander, Annie Holtz, Klaus Abels, David Adger & Jennifer Culbertson. 2020. Experimental evidence for the influence of structure and meaning on linear order in the noun phrase. *Glossa: a journal of general linguistics* 5. 1–21. doi:10.5334/gjgl.1085.

Martin, Alexander, Theeraporn Ratitamkul, Klaus Abels, David Adger & Jennifer Culbertson. 2019. Cross-linguistic evidence for cognitive universals in the noun phrase. *Linguistics Vanguard* 5(1). 1–13. doi:10.1515/lingvan-2018-0072.

Meelen, Marieke. 2016. *Why Jesus and Job spoke bad Welsh: The origin and distribution of V2 orders in Middle Welsh*. Leiden: Universiteit Leiden Proefschrift.

Naigles, Letitia. 1990. Children use syntax to learn verb meanings. *Journal of Child Language* 17(2). 357–374. doi:10.1017/S0305000900013817.

Þorgeirsson, Haukur. 2012. Late placement of the finite verb in Old Norse Fornyrðislag meter. *Journal of Germanic Linguistics* 24(3). 233–269. doi:10.1017/S1470542712000037.

Perkins, Laurel & Jeffrey Lidz. 2021. Eighteen-month-old infants represent nonlocal syntactic dependencies. *Proceedings of the National Academy of Sciences* 118(41). 1–7. doi:10.1073/pnas.2026469118.

Puggaard, Rasmus. 2019. Flexibility of frequent clause openers in talk-in-interaction: Det 'it, that' and så 'then' in the prefield in Danish. *Nordic Journal of Linguistics* 42(3). 291–327. doi:10.1017/S0332586519000088.

R Core Team. 2020. R: A language and environment for statistical computing. Tech. rep. R Foundation for Statistical Computing Vienna.

Raviv, Limor, Gary Lupyan & Shawn C. Green. 2022. How variability shapes learning and generalization. *Trends in Cognitive Sciences* 26(6). 462–483. doi:10.1016/j.tics.2022.03.007.

Rebuschat, Patrick & John N. Williams. 2012. Implicit and explicit knowledge in second language acquisition. *Applied Psycholinguistics* 33(4). 829–856. doi:10.1017/S0142716411000580.

Rizzi, Luigi. 1996. Residual verb second and the wh-criterion. In Adriana Belletti & Luigi Rizzi (eds.), *Parameters and functional heads: Essays in comparative syntax*, 63–90. New York, Oxford: Oxford University Press. doi:10.1093/oso/9780195087932.003.0003.

Roberts, Ian. 1993. *Verbs and diachronic syntax: A comparative history of English and French*. Dordrecht, Boston, London: Kluwer. doi:10.1007/978-94-011-2910-7.

Roberts, Ian. 1996. Remarks on the Old English C-system and the diachrony of V2. In Ellen Brandner & Gisella Ferraresi (eds.), *Language change and generative grammar*, 154–167. Opladen: Westdeutscher Verlag. doi:10.1007/978-

3-322-90776-9_6.
Roeper, Thomas. 1999. Universal bilingualism. *Bilingualism: Language and Cognition* 2(3). 169–186. doi:10.1017/S1366728999000310.
Ruiz, Simón, Kaitlyn M. Tagarelli & Patrick Rebuschat. 2018. Simultaneous acquisition of words and syntax: Effects of exposure condition and declarative memory. *Frontiers in Psychology* 9. 1–11. doi:10.3389/fpsyg.2018.01168.
Sailor, Craig. 2020. Rethinking 'residual' verb second. In Rebecca Woods & Sam Wolfe (eds.), *Rethinking verb second*, 126–149. Oxford: Oxford University Press. doi:10.1093/oso/9780198844303.003.0006.
Steiner, Brittany Devan Jelm. 2014. *The evolution of information structure and verb second in the history of French*. Bloomington: Indiana University dissertation.
Tagarelli, Kaitlyn M., Simón Ruiz, José Luis Moreno Vega & Patrick Rebuschat. 2016. Variability in second language learning. *Studies in Second Language Acquisition* 38(2). 293–316. doi:10.1017/S0272263116000036.
van Kemenade, Ans. 1987. *Syntactic case and morphological case in the history of English*. Dordrecht: Foris. doi:10.1515/9783110882308.
Vikner, Sten. 1995. *Verb movement and expletive subjects in the Germanic languages*. New York, Oxford: Oxford University Press. doi:10.1093/oso/9780195083934.001.0001.
Waldmann, Christian. 2008. *Input och output. Ordföljd i svenska barns huvudsatser och bisatser*. Lund: Lunds Universitet dissertation.
Westergaard, Marit. 2009a. *The acquisition of word order. Micro-cues, information structure, and economy*. Amsterdam, Philadelphia: John Benjamins. doi:10.1075/la.145.
Westergaard, Marit. 2009b. Word order in Old and Middle English: The role of information structure and first language acquisition. *Diachronica* 26(1). 65–102. doi:10.1075/dia.26.1.03wes.
Westergaard, Marit. 2021. Microvariation in multilingual situations: The importance of property-by-property acquisition. *Second Language Research* 37(3). 379–407. doi:10.1177/0267658319884116.
Westergaard, Marit, Terje Lohndal & Björn Lundquist. 2023. Variable V2 in Norwegian heritage language: An effect of crosslinguistic influence? *Linguistic Approaches to Bilingualism* 13(2). 133–162. doi:10.1075/lab.20076.wes.
Westergaard, Marit R. & Øystein A. Vangsnes. 2005. Wh-questions, V2, and the left periphery of three Norwegian dialect types. *Journal of Comparative Germanic Linguistics* 8(1-2). 117–158. doi:10.1007/s10828-004-0292-1.
Wickham, Hadley. 2016. *ggplot2. Elegant graphics for data analysis*. Cham:

Springer. doi:10.1007/978-3-319-24277-4.
Willis, David. 1998. *Syntactic change in Welsh: A study on the loss of verb-second*. Oxford: Clarendon. doi:10.1093/oso/9780198237594.001.0001.
Winter, Werner. 1961. Relative Häufigkeit syntaktischer Erscheinungen als Mittel zur Abgrenzung von Stilarten. *Phonetica* 7(4). 193–216. doi:10.1159/000258110.
Wolfe, Sam. 2018. *Verb second in Medieval Romance*. Oxford: Oxford University Press. doi:10.1093/oso/9780198804673.001.0001.
Wöllstein, Angelika. 2010. *Topologisches Satzmodell*. Heidelberg: Winter.
Yang, Charles. 2000. Internal and external forces in language change. *Language Variation and Change* 12(3). 231–250. doi:10.1017/S0954394500123014.
Yang, Charles. 2002. *Knowledge and learning in natural language*. New York: Oxford University Press. doi:10.1093/oso/9780199254149.001.0001.
Yang, Charles. 2016. *The price of linguistic productivity*. Cambridge, MA: MIT Press. doi:10.7551/mitpress/9780262035323.001.0001.

Marc Meisezahl
University of Konstanz
PO Box 175
78457 Konstanz
Germany
marc.meisezahl@uni-konstanz.de
www.meisezahl.net
ORCID: 0000-0002-0153-9396

Simon Kirby
University of Edinburgh
3 Charles Street
Edinburgh
EH8 9AD
United Kingdom
ORCID: 0000-0002-6496-1340

Jennifer Culbertson
University of Edinburgh
3 Charles Street
Edinburgh
EH8 9AD
United Kingdom
ORCID: 0000-0002-1737-6296