

ARTICLE

No hesitation at the choice point: verb–particle placement is cost-free in production

Matt Hunt Gardner^{1,2}  and Xinyu Leslie Liao³

¹Linguistics, Queen Mary University of London, UK; ²Linguistics, University of Oxford, UK and ³Linguistics, University of Toronto, Canada

Corresponding author: Matt Hunt Gardner; Email: matt.gardner@qmul.ac.uk

(Received 11 November 2025; Revised 14 January 2026; Accepted 19 January 2026)

Abstract

Using 3,154 tokens from American English, we test whether optionality in verb–particle placement increases speech-planning cost, measured as pre-verbal silence. Tokens were coded for object properties, idiomaticity and verb frequency. We find that pre-verbal silence does not differ between split (*pick the book up*) and joined (*pick up the book*) orders. While idiomaticity favours the joined order, it does not raise planning cost. Verb frequency shortens pauses only in fast speech, suggesting predictability acts lexically, not structurally. Choice symmetry does not lengthen pauses. We therefore fail to reject the null hypothesis: the two orders are equally easy to plan. This null result, from tests designed to detect a theoretically predicted effect, aligns with other evidence that syntactic choice imposes no production cost. We conclude that variation in verb–particle constructions (VPCs) is cost-free; distributional differences reflect object properties and idiomaticity, not derivational markedness.

Keywords: speech planning; verb–particle constructions; syntax; English; language variation

1. Introduction

Theories of syntax and production make distinct, testable predictions about the ease of producing the two orders in English verb–particle constructions (VPCs), as in (1). If one order is derived from the other, additional structure-building or displacement should raise planning demands for the derived order. If usage-based accounts are right, frequency and idiomaticity should grant one order faster access. If information-theoretic pressures matter, greater choice uncertainty should slow utterance onset. We test these claims in *SWITCHBOARD* (Godfrey et al., 1992), a well-used and large corpus of American English telephone conversations, by asking whether pre-verbal silence differs between the orders and whether idiomaticity, frequency, and contextual uncertainty modulate any difference.

- (1) a. Joined order (**verb, particle, direct object (DO)**)
 ‘[...] you **heat up** *re-fried beans* and you **cut up** *tomatoes and lettuce and grated cheese*’.
 (Woman, b. 1958, New York City Dialect Area, SWITCHBOARD)
- b. Split order (**verb, DO, particle**)
 ‘Usually I **cut** *those little cherry tomatoes* **up** and put some colour into it [...]’
 (Woman, b. 1946, Southern Dialect Area, SWITCHBOARD)

These competing accounts yield clear empirical predictions. Derivational proposals – whether treating the joined order as basic with the split derived (e.g., Johnson, 1991) or vice versa (e.g., Kayne, 1985) – predict that the derived form should incur greater planning cost. If usage-based accounts are right, frequency and idiomaticity should grant one order faster access, as high-frequency and semantically non-compositional chunks are retrieved holistically (e.g., Bybee, 2010; Gries, 1999). If information-theoretic pressures matter, greater choice uncertainty (i.e., more balanced variant probabilities) should slow utterance onset, as predicted by theories of uniform information density (e.g., Jaeger, 2010). If any of these mechanisms are at work, we should observe differences in production ease between variants.

Throughout, we use terms like *choice* and *selection* as convenient shorthand for a variable outcome in usage and for the modelling of token-wise variant probabilities. In a variationist (Labovian) tradition, patterned distributions in production are an empirical object of study in their own right, and we do not assume that corpus conditioning profiles transparently identify an internal competence mechanism. Accordingly, our ‘choice’ models should not be read as a psychological decision tree in which speakers first decide on object type and only then decide on particle placement; rather, they summarise associations between annotated contexts and observed variant rates. The contribution of the present article is to link those distributional regularities to an independent behavioural measure that we operationalise below – pre-verbal silence – and to ask a deliberately modest question: do sites of variable realisation, or contexts in which variant probabilities are more balanced, show any reliable increase in pre-verbal planning time?

We therefore measure planning difficulty where it peaks: immediately before VPC onset (the verb head). Pauses robustly index cognitive load, and pre-verbal silences, in particular, reflect the assembly of upcoming structure (Goldman-Eisler, 1968; Schneider, 2016; Shriberg, 1996). We thus adopt pre-verbal silence as our metric of production ease (see Section 3.2) and test whether it differs significantly between VPC orders and whether idiomaticity, frequency or choice symmetry are modulating factors. If one order is harder, pauses should be longer before it. If both are equally accessible, the difference should be null. Our results support the latter.

1.1. Hypotheses

Research on syntactic variation raises two distinct questions about production cost. First, does the mere availability of choice itself impose a cost – does optionality complicate planning? Recent corpus work suggests not, finding that syntactic choice sites do not consistently attract hesitations (Gardner et al., 2021; Szmrecsanyi et al., 2026a, 2026b). Second, even if variation itself is cost-free, might individual variants

differ in production ease due to their structural properties? Experimental evidence shows that structurally complex constructions like passives take longer to produce than their active counterparts (Ferreira, 1994; Tannenbaum & Williams, 1968). VPCs provide a critical test case for the latter question: theories strongly predict structural asymmetries between orders, but whether this translates to measurable planning difficulty in spontaneous speech remains unknown.

H1 (Generative: joined underlying). Accounts that treat the joined order (*pick up the book*) as underlying analyse the split order as derived by movement of the direct object (DO). If derivation entails extra operations, the split variant should incur greater planning cost. *Prediction:* split orders should be preceded by longer pre-verbal pauses than joined orders.

H2 (Generative: split underlying). Other accounts posit the split order (*pick the book up*) as underlying and the joined order as derived by particle movement. *Prediction:* joined orders should be preceded by longer pre-verbal pauses than split orders.

H3 (Mixed base generation). Some proposals allow both orders to be underlying, often tied to idiomaticity: idiomatic VPCs pattern with joined order, compositional ones with split order. Mismatches should raise planning effort. *Prediction:* idiomatic VPCs in split order and compositional VPCs in joined order should be preceded by longer pauses.

H4 (Usage-based / choice symmetry). From a usage-based and information-theoretic view, cost reflects contextual predictability. Symmetrical choice contexts ($\approx 50/50$ split versus joined) require greater selection effort than asymmetrical contexts. *Prediction:* pauses should be longer when both variants are equally probable than when one variant is strongly favoured.

These hypotheses do not presuppose a particular syntactic theory. Our goal is to test whether their distinct assumptions – derivational direction, dual availability, predictability – entail measurable differences in production. If one order is derived from the other, we expect planning asymmetries. If both are equally available, we expect none. If predictability dominates, we expect planning effort to track choice symmetry (balanced versus skewed contexts) rather than derivational complexity.

1.1.1. Auxiliary predictions (*chunkiness*)

As a complement to the theory-driven hypotheses above, we probe a second dimension: how tightly the verb and particle cohere in memory. We refer to this cohesion as ‘chunkiness’ and define it in Section 2.3. We do not assume that either order is intrinsically harder to plan. Rather, we ask (i) whether idiomaticity shifts choice towards the joined order without lengthening pre-verbal pauses (semantic chunkiness), and (ii) whether higher head-verb frequency shortens pre-verbal pauses, especially at higher speech rates, by easing access to a familiar verb–particle template (frequency-based chunkiness). The first prediction concerns selection; the second concerns launch timing.

C1 (Semantic chunkiness / idiomaticity). Idiomatic VPCs should favour the joined order. No prediction that idiomaticity lengthens pre-verbal pauses.

C2 (Frequency-based chunkiness / entrenchment). Higher head-verb frequency should shorten pre-verbal pauses, especially at faster speech rates, by easing access to the upcoming verb–particle template.

We therefore distinguish tests of *order-cost contrast* (H1–H4) from *chunkiness effects* (C1–C2). The former target whether one order is harder to plan; the latter target which order is selected and how quickly articulation begins.

1.2. Outline of the article

Section 2 reviews prior work on VPCs and situates our study with respect to generative, usage-based and information-theoretic perspectives. It also outlines our definition of ‘chunkiness’. Section 3 describes the data, coding and our dependent measure of speech-planning time. Section 4 presents distributional patterns and tree-based regression models. Section 5 discusses implications for theories of variation and complexity. Section 6 concludes.

2. Background

2.1. Verb–particle constructions in English

VPCs are a hallmark of the Germanic languages: a verb and a particle form a unit with verbal properties (Rodríguez-Puente, 2019). In English, the particle originated as an adverbial element and has developed into a grammatical marker with varying degrees of semantic transparency. There are an estimated 3,000 productive VPCs in present-day English (McArthur & Aikins, 1974), though not all take part in the alternation studied here.

Transitive VPCs alternate in the linear order of verb, particle and DO, as in (2a). The particle can appear after the verb (*pick up something*, joined order) or after the DO (*pick something up*, split order). Intransitives (*calm down*, *come back*), as in (2b), do not alternate and are excluded in the present analysis.

- (2) a. Transitive vpc (**verb**, [DO], **particle**, [DO])
- i. ‘Just squatting down when I’m going to **pick something up** instead of bending over at the waist.’
(Woman, b. 1965, North Midland Dialect Area, SWITCHBOARD)
 - ii. ‘If I’m really bored, I’ll even **pick up something** and color it.’
(Woman, b. 1941, Western Dialect Area, SWITCHBOARD)
- b. Intransitive vpc (**verb**, **particle**)
- i. ‘That’s incredible, some of the pictures that have **come back**.’
(Man, b. 1962, North Midland Dialect Area, SWITCHBOARD)
 - ii. ‘The kids wouldn’t **calm down**. They wouldn’t go to sleep.’
(Woman, b. 1953, Northern Dialect Area, SWITCHBOARD)

A standard descriptive claim is that pronominal DOs require the split order (*pick it up*, not **pick up it*). While pronouns overwhelmingly favour the split order, they can occur in the joined order in specific contexts – for instance under indefiniteness (2a),

focus/emphasis or deictic force (3) (Bolinger, 1971; Johnson, 1991). We retain pronominal tokens because they vary in principle and because they provide a near-categorical baseline for split order in spoken discourse.

(3) Joined order (**verb, particle, DO**)

‘Oh yeah it’s hard to **put down** those.’

(Woman, b. 1969, Western Dialect Area, SWITCHBOARD)

VPC alternation is robust across the history of English and varies across registers and varieties. Written English often favours the joined order, while spoken English shows higher split rates (Cappelle, 2009; Gries, 2003). Cross-dialectal studies report differences as well: British and Irish English show more split order than North American English, and apparent-time studies suggest possible change (Haddican & Johnson, 2012; Röthlisberger & Tagliamonte, 2020). These patterns indicate that VPC order is conditioned by grammatical, pragmatic and social factors that any predictive model must include.

2.2. Competing accounts of VPC order

Analyses differ on whether one order is derived from the other or both are independently generated. We remain agnostic about grammatical architecture; we ask what each family of accounts implies for production under the linking assumption that extra structure-building or displacement can raise planning time (see Table 1).

2.2.1. Generative accounts

One line of generative theorizing treats the split order as basic (particle projects a small clause or resultative phrase), with the joined order arising by particle raising or incorporation (e.g., den Dikken, 1995; Kayne, 1985). Another treats the joined order as basic (verb–particle complex head), with the split order derived by object movement or linearisation constraints (e.g., Dehé, 2002; Johnson, 1991). Other proposals allow both orders in the grammar, distinguished lexically or semantically – for

Table 1. Theoretical accounts of VPC variation and their predictions for production planning

Approach	Core mechanism	Prediction for planning cost	Key citations
Generative (asymmetric)	Structural derivation; one order derived from the other	Longer pauses before derived variant	Johnson (1991), den Dikken (1995), Zeller (2001)
Generative (dual-base)	Both orders base-generated; conditioned by idiomaticity	Longer pauses for mismatched orders (e.g., idiomatic split)	Ross (1967), Wurmbrand (2000), Ramchand (2008)
Usage-based/Constructional	Entrenchment, idiomaticity, construal (item-specific baselines)	Shorter pauses for frequent/entrenched items; adjacency for idioms	Gries (2003); Goldberg (2016); Bybee (2010); Luo (2019)
Information-theoretic	Higher uncertainty in balanced choice contexts	Longer pauses when variant probabilities are balanced	Jaeger (2010), Levy (2008)

instance, idiomatic VPCs as complex heads and transparent combinations as small clauses (e.g., Punske, 2013; Ross, 1967; Wurmbrand, 2000). Recent work highlights prosody and information structure, with surface realisation set at the phonological form (PF)–discourse interface (Büring, 2013; Köhn et al., 2018; Samek-Lodovici, 2017). If one order is consistently derived, a planning contrast should follow; if both are equally represented, no contrast is expected.

2.2.2. Usage-based and constructional accounts

Usage-based approaches treat both orders as entrenched patterns shaped by frequency, idiomaticity and discourse (Bybee, 2010; Cappelle, 2006; Gries, 2003). Idiomatic VPCs are more likely to be stored as holistic chunks, which favours adjacency, while compositional combinations more readily allow separation. Production ease depends on the accessibility of stored exemplars more than on derivational complexity. For Luo (2019), in a construal account, the preferred order reflects whether the event is packaged as a tight gestalt (joined order) or as a sequence that foregrounds the object (split order). Because construal attaches to particular VPC *senses*, the expectation is for item-specific baselines: some VPCs will favour adjacency across contexts, others separation. Production ease then, still, follows from entrenchment: frequent, idiomatic combinations are retrieved quickly and tend to keep verb and particle adjacent, while transparent, semantically compositional combinations more readily allow separation.

2.2.3. Information-theoretic perspectives

A complementary view links selection cost to entropy: when the distribution of variants is symmetrical, uncertainty is higher and selection may be harder; when one variant is strongly favoured, selection is easier (Jaeger, 2010). For a binary alternation, entropy is maximised at $\hat{p} = 0.5$ and decreases monotonically as $|\eta|$ increases; we therefore use $|\eta|$ as a simple proxy for choice uncertainty in the present study. This prediction is theory-neutral about representation but interacts with representational choices insofar as they shape contextual probabilities.

2.3. Chunkiness and cohesion

Across traditions, accounts of VPCs appeal to cohesion between the verb and the particle. We use *chunkiness* as a neutral cover term for this cohesion and distinguish two sources for it that matter for production.

2.3.1. Semantic chunkiness

We use ‘semantic chunkiness’ for cases in which a verb–particle pairing is stored and retrieved as a single lexicalised unit rather than assembled compositionally. Idiomaticity is graded: combinations range from transparent to fully idiomatic, with many intermediate cases arising through semantic extension (Luo, 2019). In constructional accounts, such units are multi-word constructions, that is, conventional form–meaning pairings larger than a word; the grammar licenses their internal syntax and their preferred linearisation (Culicover et al., 2017; Goldberg, 2016). Accordingly, idiomatic VPCs count as ‘chunks’, which explains their tendency towards verb–particle adjacency without any claim that they impose extra planning cost.

While idiomaticity is graded in principle, our analyses require an operational contrast; we therefore adopt a binary coding that collapses fully idiomatic and conventionalised non-literal/resultative uses into a single ‘idiomatic’ (semantically chunked) category (details in Appendix A).

2.3.2. Frequency-based chunkiness

Independent of semantic opacity, repeated experience with a verb–particle combination can make it behave as a planning unit. We use *frequency-based chunkiness* for this accessibility effect of token frequency and entrenchment in usage. Usage-based accounts predict that frequent VPCs are retrieved more readily and initiated sooner (e.g., Bybee, 2010; Gries, 2003). Frequency strengthens access to a multi-word template, which should reduce pre-verbal delay, particularly when overall speech is fast and planning time is tight.

2.3.3. Production evidence

Production studies of spontaneous speech support treating chunks as planning units. In conversational data, established expressions tend to be realised without internal hesitations, whereas frame-like, discontinuous constructions are more likely to span pause boundaries – a classic signal of increased planning work (Schneider, 2016; Xie et al., 2023). Broader reviews of chunking in spoken interaction likewise tie chunk status to prosodic phrasing and planning scope (Barth-Weingarten & Ogden, 2021). These results legitimise our use of pre-verbal silence as a window on chunk-mediated planning because pauses concentrate at constituent onsets where speakers complete selection and retrieval of the upcoming verb–object–particle material, and they shorten when that material is accessed as a single stored unit.

Taken together, the production-oriented view of chunkiness yields concrete predictions for our data. Semantic chunkiness (idiomaticity) should push speakers towards joined order but need not lengthen pre-verbal pauses. Frequency-based chunkiness should shorten pauses, especially during rapid speech, by easing access to the upcoming verb–particle template. Crucially, neither prediction entails that one *order* is intrinsically costlier to plan; the claim is that lexical cohesion affects *which* order is chosen and that frequency affects *how quickly* articulation can begin.

3. Methods

3.1. Data

Our analysis is based on a new collection of VPC tokens extracted from the SWITCHBOARD corpus of spoken American English (Godfrey et al., 1992; Linguistic Data Consortium, 1997). The corpus contains approximately 240 hours of semi-spontaneous telephone conversations between strangers who are native speakers of American English.¹ The 520 speakers are stratified by gender, age, education and dialect region. Conversations were recorded between 1989 and 1990.

¹Although SWITCHBOARD captures largely unscripted conversational speech, it is not purely spontaneous in the strict sense: calls were initiated and routed by the collection protocol and speakers were supplied with suggested topics. We therefore treat the data as *semi-spontaneous* telephone conversation, and we return in Section 5 to what this implies for generalisability beyond elicited interaction.

Table 2. Summary of speaker demographics in *SWITCHBOARD* (Godfrey et al., 1992)

Dialect	Age	Gender	Highest education
South midland (156)	15–19 (23)	Male (280)	< High School (8)
Western (82)	20–29 (185)	Female (240)	High School (40)
North midland (75)	30–39 (146)		College (298)
Northern (73)	40–49 (108)		> College (170)
Southern (55)	50–59 (54)		Unknown (4)
New York city (31)	60 + (4)		
Mixed/Unknown (27)			
New England (21)			
	Older 35 + (242)		College + (468)
	Younger \leq 35 (278)		No College (52)

We identified 3,154 tokens of transitive VPCs. Tokens were included only if they exhibited the potential for order variation; intransitives (*calm down*, *come back*), passives (*was given out*) and cases with intervening adverbs (*pay the bill right up*) were excluded. Potential preposition–verb confounds (*put the dress on the table*) were distinguished from true VPCs (*put the dress on*). Full exclusion criteria and examples are given in the Appendix.

Pronominal tokens were retained despite their near-categorical preference for split order. They occasionally occur in the joined order, which makes them variable in principle, and they form a large share of VPC tokens. Excluding them would misrepresent the distribution and remove a crucial control.

Table 2 summarises the demographic distribution of speakers. Speaker ID was retained as a random effect in statistical modelling to account for individual baselines.

3.2. Speech-planning time as a metric of cognitive effort

We operationalise cognitive effort as pre-verbal silence: the duration of unfilled pauses immediately preceding the verb head of the VPC. Silence duration is a well-established indicator of planning load in spontaneous speech (e.g., Clark & Wasow, 1998; Goldman-Eisler, 1968; Lickley, 2015). Longer silences suggest greater planning difficulty; shorter silences indicate easier retrieval or production.

We focus on the silence immediately preceding the verb because pauses at clause and constituent boundaries correlate with the complexity of upcoming material. Early work showed clustering of pauses before major constituents with lengthening as structural complexity increases (Cooper & Paccia-Cooper, 1980; Goldman-Eisler, 1968; Maclay & Osgood, 1959). Analysis of *SWITCHBOARD* – our dataset – shows hesitation peaks at constituent onsets (Schneider, 2016). Experimental results further indicate that verbs and complements are planned in advance of articulation, so pre-verbal pauses reflect the load of assembling the upcoming clause (Momma & Yoshida, 2023).

Silence duration was extracted automatically from the audio using Praat’s silence detection (*Sound: To TextGrid (silences)*; Boersma & Weenink, 2020). Silence was defined as intervals below -50 dB and longer than 10 ms (Gardner et al., 2021; Hieke et al., 1983). For each token, we measured the pause immediately before the verb. Tokens with extreme outliers (silences > 3 s) were excluded.

In the final dataset, pauses ranged from 11 ms to 1,792 ms (mean 164 ms, SD 220 ms; median 96 ms).

Because pause duration is right-skewed and bounded by zero, analyses used log-transformed values. This reduces the influence of outliers and better meets linear-model assumptions.

3.3. Coding of constraints

Each token was coded for order and for factors known to condition VPC variation. We group these as grammatical, constructional and social.

3.3.1. Grammatical conditioning

We coded the DO for type and complexity. Simple noun phrases (NPs) comprise either a pronoun or a head noun with optional determiners; intermediate NPs contain non-finite post-nominal modification or coordination; complex NPs contain finite modification. This captures the tendency for heavier objects to favour the joined order (Gries, 2003; Lohse et al., 2004) (see Figure 1). Pronominal objects were coded separately as a near-categorical baseline for split order.

Each DO was further coded for animacy (animate versus inanimate), definiteness and concreteness, given their links to particle placement (Gries, 1999, 2003). Because order choice is also sensitive to the lexical-semantic cohesion of the verb-particle pairing itself (independently of DO properties), we additionally coded each token for idiomaticity, treating it as a construction-level property that can interact with DO type and complexity in conditioning particle placement.

For transparency, (4) gives a single attested sentence that contains two VPC tokens; we list each token and its coding for DO properties and idiomaticity.

- (4) ‘When we **broke things off** I quit crocheting and I’ve never **picked up a needle** since.’
(Woman, b. 1956, Western Dialect Area, SWITCHBOARD)

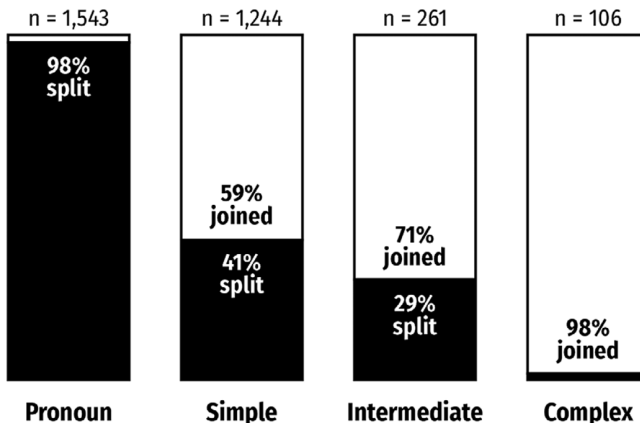


Figure 1. Distribution of joined and split VPC order by complexity of the direct object.

Token-level coding in (4):

- **Broke things off** (split order): DO = *things* (NP; simple NP; inanimate; indefinite; abstract); idiomaticity = idiomatic (conventionalised non-literal sense ‘end a relationship’).
- **Picked up a needle** (joined order): DO = *a needle* (NP; simple NP; inanimate; indefinite; concrete); idiomaticity = compositional (literal/spatial *up* ‘lift’).

Throughout, we use *idiomaticity* as a practical cover term for *semantic chunkiness*: whether the verb–particle pairing in its attested sense and context is plausibly accessed as a cohesive form–meaning unit in usage.² This operationalisation is intentionally broader than ‘idiom’ narrowly construed. In particular, it includes (i) fully idiomatic/holistic senses (e.g., *pick up guys* ‘meet/attract’) and (ii) conventionalised resultative/aspectual particle uses such as *up* in *cheer someone up* or *brighten something up*, which are productive but non-spatial and often pattern with cohesive VPC behaviour. Crucially, this does *not* claim that productive resultative/aspectual particles are strictly non-compositional; rather, because they involve a conventionalised non-spatial particle contribution and often behave as cohesive verb–particle units in usage, we group them with the ‘idiomatic’ (semantically chunked) category for the purposes of a binary operational contrast.

Coding followed an entailment-based diagnostic used in prior corpus work on VPCs (Lohse et al., 2004, 244–246; Grafmiller & Szmrecsanyi, 2018, 392) and subsequent applications (Röthlisberger & Tagliamonte, 2020). Tokens were coded as compositional only when the utterance-level meaning transparently entailed both (a) the corresponding simplex-verb event and (b) a literal/spatial paraphrase contributed by the particle; otherwise, the token was coded as idiomatic in the present operational sense. Idiomaticity was coded independently by both authors; disagreements were resolved through discussion to a single consensus coding (Table 3). The diagnostic and fully annotated examples are provided in Appendix A.

3.3.2. Constructional conditioning

We coded frequency and lexical identity. Head-verb frequency was measured as relative frequency within *SWITCHBOARD*. Verbs occurring more than 1% of the time (≥ 32 tokens) were coded as ‘frequent’ (Erker & Guy, 2012). Each VPC type (e.g., *pick up*, *get back*) was included as a random effect to account for item-specific baselines, because some combinations are near-categorical in one order.

3.3.3. Social conditioning

Tokens were linked to speaker metadata in the corpus: age, gender, education level and dialect region. These factors allow tests of social conditioning, as reported in other varieties (Röthlisberger & Tagliamonte, 2020). Speaker identity was included as a random effect to control for individual baselines.

²We retain the label *idiomaticity* for continuity with prior VPC work and for ease of reference in figures and models, but we emphasise that it is used here as an operational cover term rather than “idiom” in the narrow sense.

Table 3. Distribution of split vpc order by type of DO and idiomaticity of the vpc

do type	Idiomatic		Compositional	
	<i>n</i>	% split	<i>n</i>	% split
Pronoun	1,137	95	406	99
Simple NP	976	30	268	79
Intermediate NP	219	21	42	74
Complex NP	84	0	22	9

Analysis plan. We distinguish two analytic goals: (i) testing whether particle placement is associated with increased speech-planning cost (H1–H4) and (ii) characterising how grammatical and lexical–semantic factors condition order choice, including auxiliary chunkiness predictions (C1–C2). All primary modelling uses tree-based mixed-effects models; where a theoretically central predictor is unsupported, we quantify evidence for a negligible effect with targeted Bayesian follow-up.

Step 1: Model order choice with a GLMM tree (C1; inputs for H4). We first fit a binomial generalised linear mixed-effects model (GLMM) tree predicting split versus joined vpc order. Candidate predictors include object properties (type/complexity, animacy, definiteness, concreteness, length in words), construction-level status (idiomaticity), lexical accessibility (head-verb frequency) and speaker controls; random intercepts are included for speaker and lexical item. The GLMM tree uses model-based recursive partitioning to discover subpopulations in which effects differ, while retaining subject and item variability. Idiomaticity is treated as a carried fixed effect estimated within each terminal-node model rather than being enforced by hard partitioning, allowing us to evaluate whether semantic chunkiness systematically shifts order choice across grammatical contexts (C1).

Step 2: Derive a choice-symmetry index from fitted logits (H4). From the fitted GLMM tree we obtain, for each token, the terminal-node linear predictor (log-odds) of the split order, η_i .³ We operationalise choice symmetry as the absolute fitted logit:

$$s_i = |\eta_i|.$$

Small s_i values indicate balanced contexts with $\hat{p}_i \approx 0.5$ (high choice uncertainty); large s_i values indicate skewed contexts with \hat{p}_i near 0 or 1 (low choice uncertainty). This provides a model-based operationalisation of the usage-based/uncertainty prediction in H4.

Step 3: Model planning time with an LMM tree (H1–H4; C2). We then model log pre-verbal silence with a linear mixed-effects model (LMM) tree. Candidate predictors include vpc order (H1–H2), its interaction with idiomaticity to capture mismatch predictions (H3), the symmetry index $s_i = |\eta_i|$ (H4), speech rate (words per minute), head-verb frequency (C2), object properties (type/complexity, length in words, animacy, definiteness, concreteness) and speaker controls; random intercepts are speaker and lexical item. The LMM tree tests for parameter instability and partitions the data where relationships differ, yielding a flexible, interaction-sensitive model of planning time under established performance controls for SWITCHBOARD.

³For each token, η_i is the fitted GLMM linear predictor from its terminal-node model; the corresponding fitted probability is $\hat{p}_i = \text{logit}^{-1}(\eta_i)$.

Step 4: Quantifying evidence for negligible effects (H4 and order effects if needed). When the tree-based analyses suggest no support for a theoretically central predictor (in particular the symmetry index s_i in H4), we follow up with a Bayesian linear mixed model to quantify evidence that any remaining effect is negligible. We predefine a region of practical equivalence (ROPE) on the log-seconds scale corresponding to changes in pause duration too small to be meaningful (here, ± 0.05). We then evaluate whether the posterior for the symmetry effect lies within the ROPE and compare models with versus without the symmetry term using Bayes factors and out-of-sample predictive checks.

3.3.3.1. *Interpretive scope.* Idiomaticity and frequency are treated as sources of lexical–semantic cohesion (‘chunkiness’) that may influence (i) which order is realised (C1) and (ii) how quickly articulation begins (C2). They are not taken to imply that either order is intrinsically harder to plan. H1–H3 concern order-linked planning asymmetries; H4 concerns whether balanced choice contexts (small $|\eta|$) impose additional planning cost beyond established performance factors.

4. Results

4.1. VPC order and speech planning time

We first asked whether split and joined orders differ in pre-verbal silence. Figure 2 plots log-transformed pause duration by order, and by order crossed with DO complexity and idiomaticity. Across all comparisons, mean silence durations for split and joined tokens overlap substantially.

These results run directly against the theoretical expectation that one order should be costlier to plan – whether because it is derivationally more complex, less entrenched or less predictable. In the top-left panel of Figure 2, split and joined tokens are indistinguishable: a one-way analysis of variance (ANOVA) shows no difference in pre-verbal silence between orders, $F(1, 3149.8) = 0.41$, $p = .522$, $\eta_p^2 \approx 0$. The top-right panel, which separates Simple and Intermediate NPs (the genuinely variable cases), again shows no difference between the orders. Moreover, the near-categorical contexts (Pronouns $\approx 98\%$ split; Complex NPs $\approx 98\%$ joined) are indistinguishable from the variable Simple/Intermediate NP contexts in their pre-verbal silence; an omnibus comparison across six groups yields no differences, $F(5, 3144) = 1.33$, $p = .25$. H1 and H2, predicting longer pauses for whichever order is derived, find no support. H3, predicting a mismatch cost when idiomaticity and order ‘disagree’, is likewise not borne out: the bottom panel shows overlapping distributions with no interaction pattern. In short, where several theories would predict asymmetries, we find sameness across the board.

These results indicate that neither order carries an inherent processing cost. The top-left panel of Figure 2 compares joined versus split tokens in aggregate and shows overlapping distributions with no shift in central tendency, which directly contradicts H1 and H2 (no order-specific planning penalty). The top-right panel conditions on DO complexity and restricts attention to the most variable object types (Simple and Intermediate NPs). Within each of these classes, split and joined tokens again exhibit indistinguishable pre-verbal silences, so the null for H1/H2 persists even when we partition out near-categorical contexts.

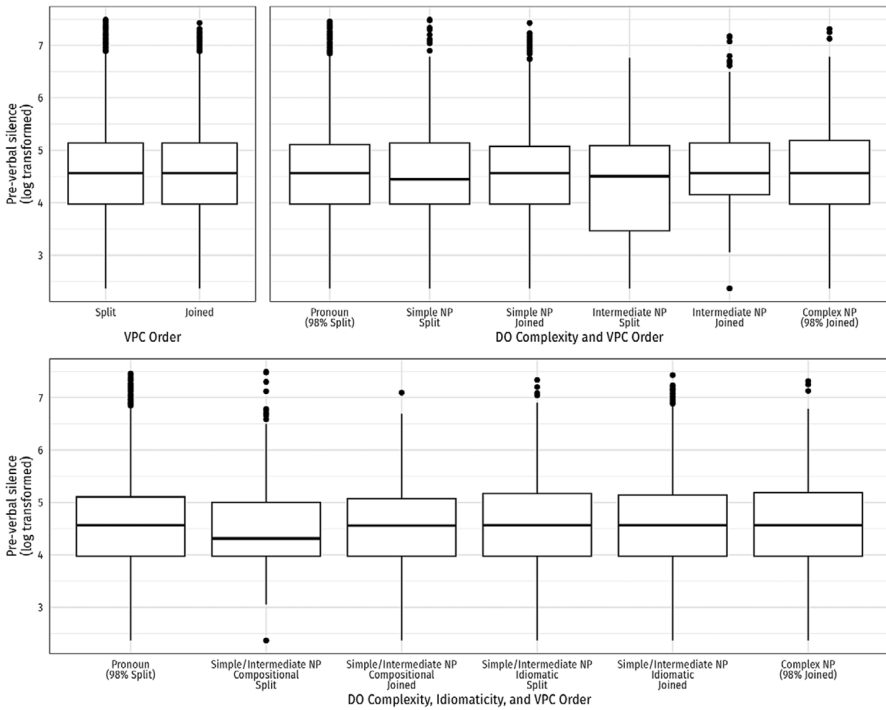


Figure 2. Distribution of pre-verbal silence (log-transformed) by VPC order (top left), by order and direct-object complexity (top right), and by order, complexity and idiomaticity (bottom). Horizontal bars mark means. Error bars show 95% CIs computed on the log scale.

Having found no order-level difference in pre-verbal silence, we now model what predicts the two orders in order to quantify the choice landscape. We fit a GLMM tree to obtain token-wise fitted probabilities of the split order. These probabilities yield a choice-symmetry index s for each token, which lets us test H4: that more balanced contexts should slow utterance initiation if selection is costly. The model also operationalises our two chunkiness claims: C1 that semantic chunkiness (idiomaticity) favours adjacency without adding cost, and C2 that frequency-based chunkiness (head-verb entrenchment) eases retrieval and could shorten pauses. We first report the distributional model of order choice and then link its fitted probabilities to pause duration, separating predictors of linearisation from predictors of planning cost.

4.2. Predictors of order choice

We model order choice as a binary outcome between the joined and split orders. Guided by Table 1 and prior sociolinguistic/corpus research, we code predictors that theory and previous evidence associate with this alternation: object properties (type, structural complexity, animacy, definiteness, concreteness), constructional status (idiomaticity), lexical pressure (head-verb frequency) and the presence of a following directional adverbial. Coding follows, e.g., Gries, 2003; Wasow, 2002; Lohse et al., 2004;

Diessel, 2005; Szmrecsanyi et al., 2016; Grafmiller & Szmrecsanyi, 2018; Röthlisberger & Tagliamonte, 2020. We also include speaker gender, age, education and dialect region as sociolinguistic controls. These linguistic and social variables constitute the predictor set; all enter as candidate predictors.

To model these effects we fit a binomial GLMM tree (Bernaisch, 2022; Fokkema et al., 2020; Gardner, 2025) with random intercepts for speaker and lexical item. The tree uses model-based recursive partitioning: it tests for parameter instability, splits the predictor space where effects differ and refits a local mixed model in each node. This retains subject and item variability, surfaces interactions and effect heterogeneity without pre-enumerating them and accommodates non-monotone, piecewise effects with minimal functional-form assumptions. We treat idiomaticity as a carried, node-wise fixed effect: it is included in the GLMM component and estimated in every node, while the partitioning set comprises the remaining candidate predictors. This lets the tree discover subpopulations via splits on pronominality, complexity, animacy, definiteness and related properties and within each subpopulation, estimate the idiomaticity effect, rather than collapsing it through a hard split.

The fitted tree echoes Figure 1; the first and strongest split is pronominality (Node 1). Definite pronouns are essentially categorically split order (Node 3); indefinite pronouns admit some joined tokens (Node 4), with a stronger joined bias when the VPC is idiomatic (node-wise idiomaticity effect). For NP objects, complexity separates a branch of complex NPs that are overwhelmingly joined (Node 11), with the residual split tokens largely compositional. Among non-complex NPs, animacy (Node 6) yields a more split-friendly branch for animate NPs (Node 7) and a more joined-friendly branch for inanimate NPs, where definiteness further divides strong joined order (definite; Node 9) from less strong joined order (indefinite; Node 10). Idiomaticity consistently biases towards the joined order; its coefficient reaches node-wise significance where data are dense (Nodes 9–10) and shows the same direction but is underpowered in (Nodes 4, 7 and 11). Other candidate predictors – speaker demographics, verb frequency, and directional adverbials – are not selected once grammatical and idiomatic cues are in the model; this is model-based non-selection rather than evidence of no effect (Figure 3).

4.2.1. How to read the tree.

Internal nodes show the variable used to partition the data (with the split criterion on the connecting edges); terminal nodes correspond to subgroups in which a local mixed model is fit. Within each terminal node, points indicate fitted probabilities of the split order (with 95% confidence intervals), and the idiomaticity effect is estimated as a fixed effect in every node rather than being enforced via a hard split.

Having fitted the model, we code each token by the node-specific log-odds of the split order η : the Node 3 intercept plus that token's terminal-node coefficient. Node 3 is the reference at $\eta = 5.267$ log-odds ($p \approx 0.995$ split). The coefficient at Node 11 is -8.316 , so $\eta_{11} = 5.267 - 8.316 = -3.0486$ ($p \approx 0.045$ split). Tokens from Nodes 3 or 11 are therefore highly skewed (one order overwhelmingly preferred). Tokens from Node 7 lie much closer to $\eta = 0$ and hence nearer 50/50; within nodes where data allow estimation (e.g., Nodes 9–10), idiomaticity shifts η toward the joined order.

Our next step asks whether highly predictive contexts are easier to plan than balanced ones: that is, whether tokens from skewed contexts (one variant strongly favoured) show shorter pre-verbal silences than tokens from balanced contexts

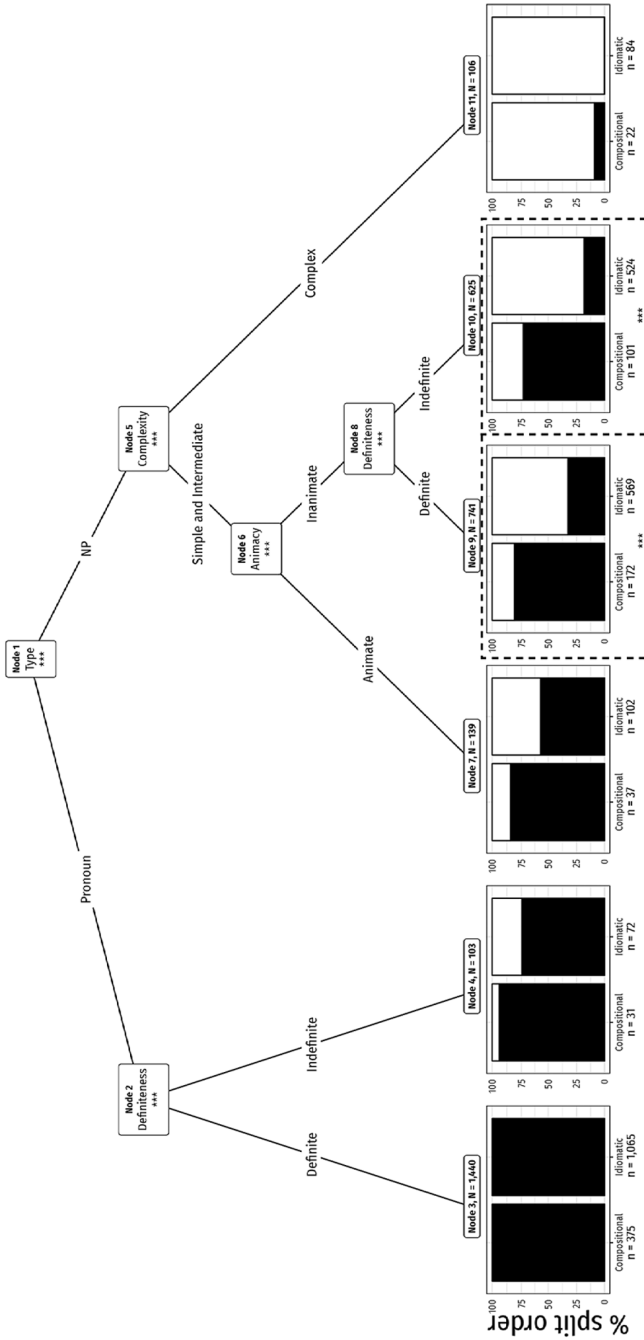


Figure 3. GLMM tree for split versus joined order in SWITCHBOARD. Predictors: object properties, idiomaticity, frequency and social factors. Random intercepts: speaker, lexical item. Subgroup partition: idiomatic versus compositional. Error ribbons indicate 95% CIs.

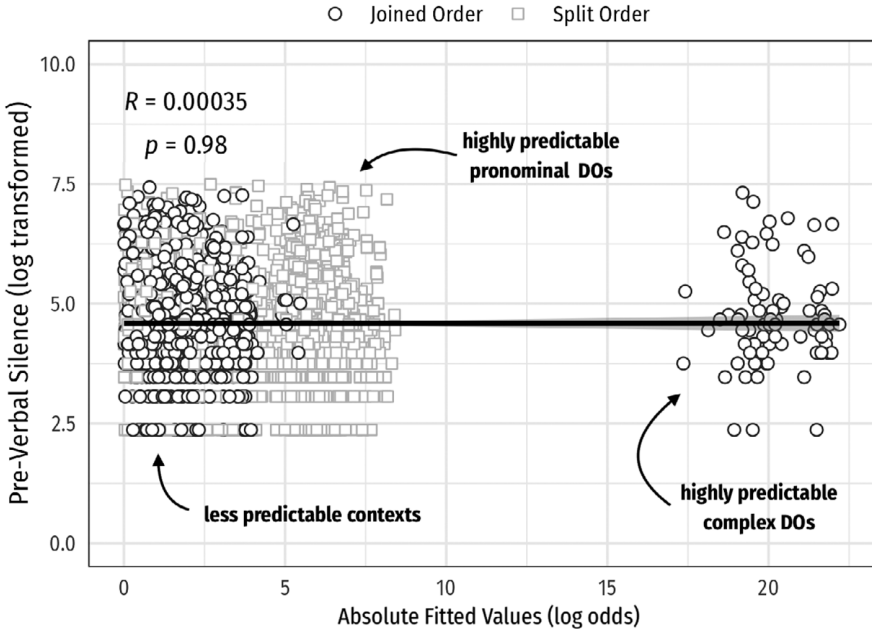


Figure 4. Log-transformed pre-verbal silence by the absolute fitted log-odds of the split order ($|\eta|$) for *vpc* tokens in *SWITCHBOARD*. Fitted values are from the GLMM tree in Section 4.2.

(variants near 50/50, predicted to be harder). We operationalise skew with the absolute value of the node-specific logit ($|\eta|$) and visualise it as the x -axis in Figure 4. Small $|\eta|$ values (on the left) mark a balanced, high-choice context; large $|\eta|$ values (on the right) mark a low-choice, skewed context. In Figure 4, black circles represent joined order tokens and grey squares are split order tokens. The y -axis represents the log-transformed amount of pre-verbal silence per token (higher values indicate longer pausing). A regression line fitted to the data is effectively flat ($R^2 = 0.00035$, $p = 0.98$), indicating no detectable association between choice symmetry and pre-verbal pausing.

We analyse pre-verbal silence directly with a LMM tree (Figure 5) under controls that are well established for *SWITCHBOARD*, most notably overall speech rate.⁴ Our predictor of interest is the symmetry index $s = |\eta|$ from the order-choice model, which operationalises how predictable the upcoming syntactic configuration is (small s = balanced; large s = skewed). To cover lexical accessibility without conflating it with order predictability, we include head-verb frequency. Accordingly, candidate predictors are s , speech rate (words per minute), idiomaticity, head-verb frequency, object properties (type, complexity, length in words, animacy, definiteness, concreteness) and speaker gender, age, education and dialect region; random intercepts are speaker and lexical item.

The fitted tree corroborates the null for the symmetry predictor. It first splits on speech rate at ≈ 176 wpm: slower turns show longer pre-verbal silences. In the faster

⁴Because s already captures predictability of the order, we do not add separate next-word predictability terms (e.g., transitional probabilities or next-word predictability measures): these are not orthogonal to s and would amount to conditioning on the construct we are testing.

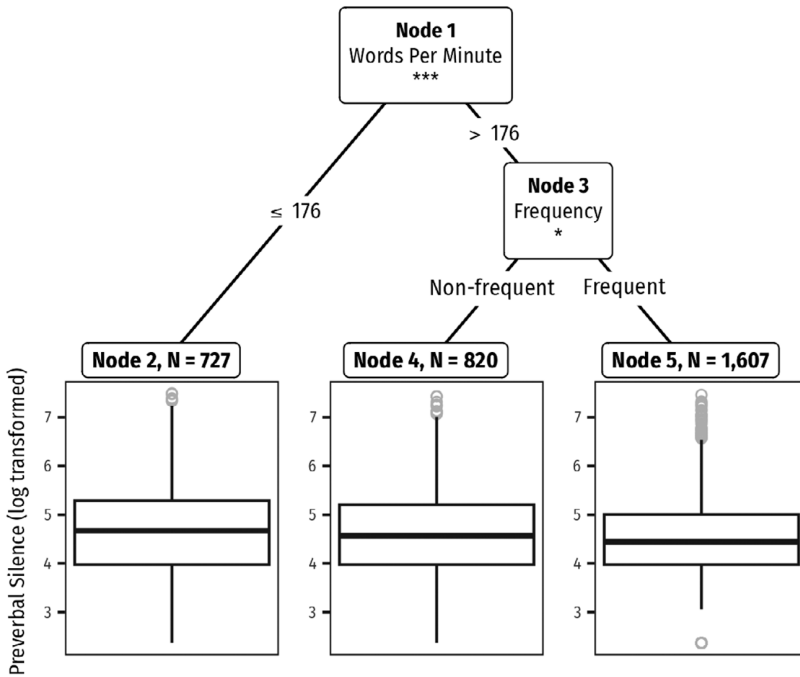


Figure 5. LMM tree of pre-verbal silence. Predictors: symmetry index $s = |\eta|$, speech rate, idiomaticity, head-verb frequency, object properties and speaker factors. Random intercepts: speaker, lexical item.

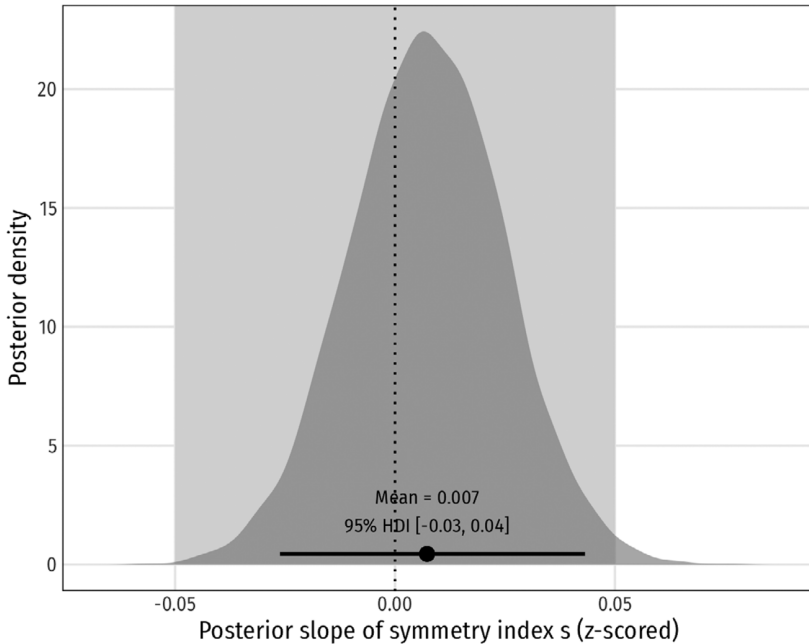
branch, head-verb frequency is selected: frequent verbs are associated with shorter pauses than infrequent verbs. The symmetry index s is not selected at any split, and idiomaticity and the social predictors likewise drop out once rate and frequency are in the model.

With respect to H4, the LMM tree does not select the symmetry index $s = |\eta|$ at any split, so there is no detectable penalty for balanced contexts (small $|\eta|$). By contrast, C2 is supported: after an initial split on speech rate (≈ 176 wpm), the fast-rate branch selects head-verb frequency, with higher-frequency verbs showing shorter pre-verbal silences – consistent with frequency-based chunking easing access to the verb-particle template. We therefore treat H4 as a null and test it with a Bayesian linear mixed model using bridge-sampled Bayes factors and a ROPE equivalence test, followed by robustness checks.

To further test H4, we use Bayesian methods to ask two simple questions: first, is any relationship between symmetry s and pause length large enough to matter; and second, does adding s to the model actually make predictions better? For the first question, we define what ‘large enough to matter’ means in advance: changes smaller than about 5% in pause duration (± 0.05 on the log-seconds scale) count as negligible. This range is called a ROPE. In the fitted model, the estimated effect of s falls entirely inside this negligible band – the 95% credible interval runs from -0.03 to 0.04 , meaning the data support at most a $\pm 4\%$ change in pause length (Table 4 and Figure 6). In other words, varying the symmetry index s , which quantifies the balance in choice between split and joined orders, has no measurable impact on pre-verbal pause duration at the scale observable in our data.

Table 4. Bayesian evaluation of the symmetry effect s (log-seconds scale). HDI = highest-density interval; ROPE = $[-0.05, 0.05]$

	Estimate	95% HDI	% in ROPE	Evidence for null
Slope of s (z)	0.007	$[-0.03, 0.04]$	100	$\log BF_{01} = 29.27$
Sensitivity (wider priors)	–	–	100	$\log BF_{01} = 2.19$

**Figure 6.** Posterior distribution of the slope for the symmetry index s (z-scored). The dark grey curve shows the posterior density; the light grey band marks the ROPE $[-0.05, 0.05]$ where effects on log pause duration are considered negligible. The black dot and horizontal bar indicate the posterior mean and 95% HDI, respectively. The x -axis gives possible effects of s on pre-verbal silence, the y -axis their relative plausibility. The entire distribution lies within the ROPE (HDI $[-0.03, 0.04]$), indicating no meaningful effect of symmetry.

To check whether including s improves prediction, we compare two otherwise identical models – one with s , one without it – using a Bayes factor, which expresses how much more likely the data are under one model than the other. The ratio overwhelmingly favours the model without s ($\log BF_{01} = 29.27$, roughly $5 \times 10^{12} : 1$ against an effect). When we deliberately relax our prior assumptions about the size of the effect, the conclusion is the same ($\log BF_{01} = 2.19$, about 9 : 1 for no effect). Out-of-sample checks tell the same story: when predicting unseen data, the model with s does not perform better than the simpler model ($\Delta\text{ELPD} = -1.0$), and information-weighting methods put essentially all weight on the model without s . A speaker-wise permutation test shows that the observed slope for s is entirely consistent with chance ($p = 0.63$), and allowing a non-linear effect of s or using projection-based variable selection likewise offers no support for including it. Taken together, these converging results indicate that choice symmetry (the $|\eta|$ index) does not measurably affect pre-

verbal silence, with any possible influence smaller than our 5% threshold. We therefore regard H4 as unsupported (effectively rejected) and turn to the broader implications of this finding in the discussion.

5. Discussion

Our findings provide a focused test of whether syntactic variation incurs measurable production cost. Across more than 3,000 VPCs in *SWITCHBOARD*, we find no evidence that one order (split versus joined) is harder to produce than the other. Pre-verbal silences do not differ systematically between orders, even when controlling for object type, complexity, animacy and idiomaticity. This result directly bears on H1 and H2: the split and joined orders are produced with comparable ease, and there is no indication that either ordering imposes extra planning load. This is consistent with the fact that analyses trade off different notions of complexity (e.g., syntactic configuration versus morphological cohesion), and it provides no production-based support for accounts that predict a general timing penalty from additional operations in one ordering. This undermines the assumption, common to many derivational accounts, that formally ‘marked’ variants require greater production effort. Pauses reliably lengthen before syntactically complex clauses, infrequent words and unexpected constituents (e.g., Clark & Wasow, 1998; Ferreira, 1991; Goldman-Eisler, 1968; Momma & Yoshida, 2023; Shriberg, 1994); the absence of any such pattern here suggests that both orders are equally available to speakers in real time despite their formal differences.

H3 and H4 addressed whether optionality itself, as reflected in more balanced choice contexts (small $|\eta|$), adds to planning difficulty. If symmetrical contexts (where two variants are equally likely) impose additional cognitive load, tokens with fitted probabilities near 0.5 should show longer pre-verbal silences. No such pattern emerged. Pauses are no longer in balanced contexts than in skewed contexts. In information-theoretic terms, choice-set uncertainty, operationalised here as $|\eta|$, does not predict production cost in our production measure. Bayesian model comparison, equivalence testing and permutation analysis all converge on the same conclusion: any effect of symmetry, if present at all, is smaller than a few percentage points on the log-seconds scale.

The chunkiness hypotheses (C1–C2) were partially borne out. Speech rate and head-verb frequency together predict planning time: slower turns show longer pre-verbal pauses, and within fast speech, frequent verbs shorten them. This confirms that frequency-based chunking eases retrieval, not that any specific order is inherently costlier. Idiomaticity, as a form of semantic chunkiness, strongly predicts order choice but not planning duration. Thus both frequency-based and semantic chunkiness shape *which* variant is chosen, but neither increases the time needed to produce it.

These findings align with prior corpus work on production fluency. Gardner et al. (2021) found that optionality in itself does not attract disfluency across more than 20 alternations in *SWITCHBOARD*, and subsequent studies report no correlation between uncertainty/balance in syntactic choice and filled pauses, repairs or hesitations (Szmrecsanyi et al., 2026a, 2026b). By isolating a single alternation and modelling choice symmetry directly (via $|\eta|$), our results confirm this pattern with greater precision: across alternations and measures, syntactic variation is not harder

to produce. Absolute complexity in the grammar does not entail relative complexity in use.

5.1. *Easy, chunky, all the time*

Evidence from related systems that particles can be syntactically and prosodically *phrasal* rather than clitic-like (e.g., Swedish; Toivonen, 2003; Svenonius, 1996) weakens any expectation that a single head-movement derivation should govern English verb–particle order. If surface order is resolved at the PF–discourse interface – via information structure, prosody and weight – then neither order is inherently ‘derived’ in a way that should increase planning cost. Our null difference in pre-verbal silence is consistent with this view: idiomaticity influences adjacency preferences, but there is no general production penalty for either order.

The distinction between semantic and frequency-based chunkiness clarifies why both matter yet remain independent of processing cost. Semantic chunkiness (idiomaticity) guides adjacency and lexical cohesion; frequency-based chunkiness (entrenchment) facilitates retrieval, particularly under temporal pressure. Neither drives measurable changes in pausing. This dissociation helps explain why production is sensitive to lexical accessibility but buffered against the structural ‘difficulty’ often assumed in derivational models.

Theoretical implications follow on both sides. For generative accounts that posit one order as derived, there is no processing evidence of a derivational direction: if additional operations imposed cost, silences would lengthen before the derived form, but they do not. For usage-based approaches, idiomaticity effects are real but bounded: idiomaticity chiefly shapes selection (order choice), while frequency shows a limited fluency benefit under time pressure (shorter pre-verbal silences in faster speech). Processing ease depends more on lexical entrenchment than on structural markedness, while semantic cohesion mainly constrains linearisation.

Finally, these results help clarify how production and comprehension can diverge in what counts as ‘difficulty’ in the presence of variation. In comprehension, processing cost is often linked to the unexpectedness of upcoming material, typically operationalised as surprisal (negative log probability of a specific ‘next thing’, e.g., word/structure/sound, etc.), whereas entropy quantifies uncertainty over a set of alternatives prior to observing a particular outcome (approximated in our production analyses by $|\eta|$ for this binary alternation) (Frank, 2013; Hale, 2001; Levy, 2008). These are related but distinct notions: a context can have high uncertainty (several plausible ‘next things’) without any single one being especially surprising, and a context can have low entropy (i.e., few possible ‘next things’) while still yielding a highly surprising outcome. Our production measure tests a different question: whether the availability of two surface realisations, or balanced variant probabilities, is associated with longer pre-verbal planning time. We find no evidence that it is: pre-verbal silences are insensitive to choice symmetry in this alternation. This does not imply that vPC variation would be cost-free in comprehension. A comprehension-focused test would instead ask whether listeners incur greater cost when the realised order is less expected in its context (high surprisal), and whether any such effects track item-specific entrenchment (idiomaticity, frequency) and discourse cues. The broader implication is that distributional predictability can matter for comprehension and production in different ways, and that the absence of an entropy effect on production timing here sets a concrete boundary condition for theories that would equate variant uncertainty with planning difficulty.

5.2. Limitations and robustness checks

The evidence derives from semi-spontaneous American English telephone speech and may not generalise to other registers or languages. Because the conversations are semi-spontaneous and topic-prompted, some planning behaviour may differ from fully unelicited interaction; however, the design still captures natural turn-taking, lexical retrieval and incremental utterance launch under real-time constraints. Pre-verbal silence is an indirect index of planning effort and can misclassify short low-amplitude segments. We mitigated this with conservative thresholds (-50 dB, ≥ 10 ms), log-transformation and exclusion of outliers (> 3 s). The key results are robust to alternative preprocessing choices (e.g., different thresholds or winsorising) and to modelling on the raw millisecond scale. Beyond the GLMM and LMM trees reported, the null effect of symmetry was verified through a Bayesian mixed model with ROPE and Bayes factor tests, sensitivity analyses with wider priors, out-of-sample model comparison and a speaker-wise permutation test. All converge on the same outcome: symmetry does not measurably influence pre-verbal silence, and any residual effect is smaller than the predefined 5% equivalence margin.

6. Conclusion

The study asked whether syntactic variation in VPCs adds to production complexity. It does not. Across multiple models and validation methods, pre-verbal silence is unaffected by order, idiomaticity or choice symmetry. Planning effort reflects only general performance factors – speech rate and under time pressure, head-verb frequency. Within this alternation, syntactic variation carries no measurable cost.

This result does not imply that derivational complexity never affects timing: clause-initial pauses and repair sites remain sensitive to structural load. But at the verb–object juncture, where lexical access and incremental planning occur, no cost appears. At the point immediately before VPC onset (the verb), we find no timing penalty associated with the eventual ordering: pre-verbal silences are indistinguishable across split and joined tokens. Variation in VPCs is structured by grammatical and semantic constraints, stable across speakers and efficient in production.

In short, syntactic alternations can be both probabilistic and fluent. Variation is not a symptom of inefficiency but a property of a production system that is flexible, resilient and low-cost. Variation is sometimes treated as a challenge for strongly economy-driven or efficiency-shaped views of grammar and processing (e.g., Chomsky, 1995; Fukui, 1993; Gibson et al., 2019; Hawkins, 2004; Henry, 1998), but our results suggest that (at least for VPC placement) probabilistic alternation can be fully fluent in production. Future work should test whether this pattern generalises to other alternations and languages, but within this domain the evidence is clear: VPC variation does not slow speakers down.

Data availability statement. An Open Science Framework (OSF) repository hosting R scripts, derived (non-identifying) data and session info (exact package versions) can be found here: osf.io/pr2bw. Raw transcripts and audio files are not available as they are licensed from the Linguistic Data Consortium; however, interested readers may seek out the data there.

Acknowledgements. The authors thank Benedikt Szmrecsanyi, Thomas Van Hoey, and the broader Complexity and Change research project at KU Leuven, Belgium. The authors also thank early reviewers for their valuable feedback. This manuscript was prepared in L^AT_EX using the CUP-JNL-LCO class. We used an AI-assisted editor to copy-edit wording and refactor L^AT_EX (preamble, citation commands, and cross-

references). The model did not generate results, annotations, analyses or claims; all analyses and interpretations are by the authors.

Funding statement. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Competing interest. The authors declare none.

Ethics statement. This study involves the secondary use of data exclusively. No identifying information about speakers is presented.

References

- Allaire, J., Xie, Y., Dervieux, C., McPherson, J., Luraschi, J., Ushey, K., Atkins, A., Wickham, H., Cheng, J., Chang, W., & Iannone, R. (2025). *rmarkdown: Dynamic Documents for R*. R package version 2.30.
- Barth-Weingarten, D., & Ogden, R. (2021). “Chunking” spoken language. *Open Linguistics*, 7(1), 531–548.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- Bernaish, T. (2022). Comparing generalised linear mixed-effects models, generalised linear mixed-effects model trees and random forests: Filled and unfilled pauses in varieties of English. In Schützler, O. & Schlüter, J. (Eds.), *Data and methods in corpus linguistics: Comparative approaches* (pp. 163–193). Cambridge University Press.
- Boersma, P., & Weenink, D. (2020). *Praat: Doing phonetics by computer*. Computer Program.
- Bolinger, D. (1971). *The phrasal verb in English*. Harvard University Press.
- Borkovec, M., & Madin, N. (2025). *ggparty: 'ggplot' Visualizations for the 'partykit' Package*. R package version 1.0.0.1.
- Büring, D. (2013). Syntax, information, structure, and prosody. In den Dikken, M. (Ed.), *The Cambridge handbook of generative syntax* (pp. 861–895). Cambridge University Press.
- Bürkner, P.-C. (2025). *brms: Bayesian Regression Models using Stan*. R package version 2.23.0.
- Bybee, J. (2010). *Language usage and cognition*. Cambridge University Press.
- Cappelle, B. (2006). Particle placement and the case for “allostructions”. *Constructions online*, 1–7, 1–28.
- Cappelle, B. (2009). Contextual cues for particle placement: Multiplicity, motivation, modeling. In Bergs, A. T. & Diewald, G. (Eds.), *Contexts and constructions* (pp. 145–192). John Benjamins.
- Chomsky, N. (1995). *The minimalist program*. MIT Press.
- Clark, H. H., & Wasow, T. (1998). Repeating words in spontaneous speech. *Cognitive Psychology*, 37(3), 201–242.
- Cooper, W. E., & Paccia-Cooper, J. (1980). *Syntax and speech*. Harvard University Press.
- Culicover, P. W., Jackendoff, R., & Audring, J. (2017). Multiword constructions in the grammar. *Topics in Cognitive Science*, 9(3), 552–568.
- Dehé, N. (2002). *Particle verbs in English: Syntax, information structure, and intonation*. John Benjamins.
- den Dikken, M. (1995). *Particles: On the syntax of verb particle, triadic, and Causitive constructions*. Oxford University Press.
- Diessel, H. (2005). Particle placement in early child language: A multifactorial constraint-based analysis. *Corpus Linguistics and Linguistic Theory*, 1(1), 89–112.
- Erker, D., & Guy, G. R. (2012). The role of lexical frequency in syntactic variability: Variable subject personal pronoun expression in Spanish. *Language*, 88(3), 526–557.
- Ferreira, F. (1991). Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language*, 30(2), 210–233.
- Ferreira, F. (1994). Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language*, 33(6), 715–736.
- Fokkema, M., Edbrooke-Childs, J., & Wolpert, M. (2020). Generalized linear mixed-model (GLMM) trees: A flexible decision-tree method for multilevel and longitudinal data. *Psychotherapy Research*, 31(3), 329–341.

- Fokkema, M., Smits, N., Zeileis, A., Hothorn, T., & Kelderman, H. (2018). Detecting treatment-subgroup interactions in clustered data with generalized linear mixed-effects model trees. *Behavior Research Methods*, 50(5), 2016–2034.
- Fokkema, M., & Zeileis, A. (2024). Subgroup detection in linear growth curve models with generalized linear mixed model (GLMM) trees. *Behavior Research Methods*, 56, 6759–6780.
- Frank, S. L. (2013). Uncertainty reduction as a measure of cognitive load in sentence comprehension. *Topics in Cognitive Science*, 5(3), 475–494.
- Fukui, N. (1993). Parameters and optionality. *Linguistic Inquiry*, 24(3), 399–420.
- Gardner, M. H. (2025). A guide to build (ING) Generalized Linear Mixed Model trees in Canadian Maritime English; Part 1, social factors. *Language and Linguistic Compass*, 19(6), e70020.
- Gardner, M. H., Uffing, E., Van Vaeck, N., & Szmrecsanyi, B. (2021). Variation isn't that hard: Morpho-syntactic choice does not predict production difficulty. *PLoS One*, 16(6), e0252602.
- Gibson, E., Futrell, R., Piantadosi, S. T., Dautriche, I., Mahowald, K., Bergen, L., & Levy, R. (2019). How efficiency shapes human language. *Proceedings of the National Academy of Sciences*, 116(52), 26064–26076.
- Godfrey, J. J., Holliman, E. C., & McDaniel, J. (1992). Switchboard: telephone speech corpus for research and development. In *Proceedings of the 1992 IEEE International Conference on Acoustics, Speech and Signal Processing - Volume 1* (pp. 517–520). IEEE Computer Society.
- Goldberg, A. E. (2016). Tuning into the verb-particle construction in English. In L. Nash & Samvelian, P. (Eds.), *Approaches to complex predicates* (pp. 110–141). Brill.
- Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. Academic Press.
- Grafmiller, J., & Szmrecsanyi, B. (2018). Mapping out particle placement in Englishes around the world: A study in comparative sociolinguistic analysis. *Language Variation and Change*, 30(3), 385–412.
- Gries, S. T. (1999). Particle movement: A cognitive and functional approach. *Cognitive Linguistics*, 10(2), 105–145.
- Gries, S. T. (2003). *Multifactorial analysis in corpus linguistics: A study of particle placement*. Continuum.
- Gronau, Q. F., & Singmann, H. (2021). *bridgesampling: Bridge Sampling for Marginal Likelihoods and Bayes Factors*. R package version 1.1–2.
- Haddican, B., & Johnson, D. E. (2012). Effects on the particle verb alternation across English dialects. *U. Penn Working Papers in Linguistics*, 18(2), 31–40.
- Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In *Second Meeting of the North American Chapter of the Association for Computational Linguistics*, (pp. 1–8). Association for Computational Linguistics.
- Hawkins, J. A. (2004). *Efficiency and complexity in grammars*. Oxford University Press.
- Henry, A. (1998). Dialect variation, optionality and the learnability guarantee. *Linguistica Atlantica*, 20, 51–71.
- Hieke, A. E., Kowal, S., & O'Connell, D. C. (1983). The trouble with “articulatory” pauses. *Language and Speech*, 26(3), 203–214.
- Hothorn, T., Hornik, K., & Zeileis, A. (2006). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics*, 15(3), 651–674.
- Hothorn, T., & Zeileis, A. (2015). Partykit: A modular toolkit for recursive partytioning in R. *Journal of Machine Learning Research*, 16, 3905–3909.
- Jaeger, T. F. (2010). Redundancy and reduction. *Cognitive Psychology*, 61(1), 23–62.
- Johnson, K. (1991). Object positions. *Natural Language and Linguistic Theory*, 9(4), 577–636.
- Kassambara, A. (2025). *ggpubr: 'ggplot2' Based Publication Ready Plots*. R package version 0.6.2.
- Kayne, R. (1985). Principles of particle constructions. In Guéron, J., Obenauer, H., & Pollock, J. (Eds.), *Grammatical representations*. De Gruyter.
- Köhn, A., Baumann, T., & Dörfler, I. (2018). An empirical analysis of correlation in syntax and prosody. In *Proceedings of Interspeech 2018* (pp. 2157–2161). International Speech Communication Association.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26.
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106(3), 1126–1177.
- Lickley, R. J. (2015). Fluency and disfluency. In Redford, M. A. (Ed.), *The handbook of speech production* (pp. 445–469). Wiley-Blackwell.
- Linguistic Data Consortium (1997). Switchboard-1 Release 2. LDC Catalog No. LDC97S62.

- Lohse, B., Hawkins, J. A., & Wasow, T. (2004). Domain minimization in English verb-particle constructions. *Language*, 80(2), 238–261.
- Lüdecke, D. (2025). *sjPlot: Data Visualization for Statistics in Social Science*. R package version 2.9.0.
- Luo, H. (2019). *Particle verbs in English*. Springer.
- Maclay, H., & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word*, 15(1), 19–44.
- Makowski, D., Lüdecke, D., Ben-Shachar, M. S., Patil, I., Wilson, M. K., & Wiernik, B. M. (2025). *bayestestR: Understand and Describe Bayesian Models and Posterior Distributions*. R package version 0.17.0.
- McArthur, T., & Aikins, B. T. (1974). *A dictionary of phrasal verbs and their idioms*. Collins.
- Momma, S., & Yoshida, M. (2023). Planning multiple dependencies in sentence production. *Language, Cognition and Neuroscience*, 38(9), 1183–1213.
- Müller, K. (2025). *here: A Simpler Way to Find Your Files*. R package version 1.0.2.
- Pedersen, T. L. (2025). *patchwork: The Composer of Plots*. R package version 1.3.2.
- Punske, J. (2013). Three forms of English verb particle constructions. *Lingua*, 135, 155–170.
- R Core Team. (2025). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Ramchand, G. (2008). *Verb meaning and the lexicon: A first phase syntax*. Cambridge University Press.
- Rodríguez-Puente, P. (2019). *The English phrasal verb, 1650–present: History, stylistic drifts, and lexicalisation*. Cambridge University Press.
- Ross, J. R. (1967). *Constraints on variables in syntax*. Phd dissertation. Massachusetts Institute of Technology.
- Röthlisberger, M., & Tagliamonte, S. A. (2020). The social embedding of a syntactic alternation: Variable particle placement in Ontario English. *Language Variation and Change*, 32(3), 317–348.
- Samek-Lodovici, V. (2017). The syntax of information structure and the PF interface. *Glossa*, 2(1), 1–46.
- Schneider, U. (2016). Hesitation placement as evidence for chunking: A corpus-based study of spoken English. In *Frequency effects in language* (pp. 61–89). De Gruyter.
- Shriberg, E. (1994). *Preliminaries to a Theory of Speech Disfluencies*. University of California, Berkeley.
- Shriberg, E. (1996). *Disfluencies in Switchboard*. In *International Conference on Spoken Language Processing*, 11–14.
- Svenonius, P. (1996). The optionality of particle shift. *Working Papers in Scandinavian Syntax*, 57, 47–75.
- Szmrecsanyi, B., Grafmiller, J., Heller, B., & Röthlisberger, M. (2016). Around the world in three alternations. *English World-Wide*, 37(2), 109–137.
- Szmrecsanyi, B., Gardner, M. H., Ma, R., & Van Hoey, T. (2026a). Empirical accountability meets theorizing about language variation: Optionality does not attract dysfluencies. In Cukor-Avila, P., Tagliamonte, S. A., & Bailey, G. (Eds.), *Empirical accountability in variationist linguistics: Taking the next step*. Cambridge University Press.
- Szmrecsanyi, B., Gardner, M. H., & Van Hoey, T. (2026b). Choice and complexity: In naturally occurring data, absolute complexity does not necessarily trigger relative complexity. *Linguistic Typology at the Crossroads*, 5, 323–351.
- Tannenbaum, P. H., & Williams, F. (1968). Generation of active and passive sentences as a function of subject or object focus. *Journal of Verbal Learning and Verbal Behavior*, 7(1), 246–250.
- Toivonen, I. (2003). Non-projecting words: A case study of Swedish particles. *Natural Language & Linguistic Theory*, 21, 347–389.
- Vehtari, A., Gabry, J., Magnusson, M., Yao, Y., Bürkner, P.-C., Paananen, T., & Gelman, A. (2024). *loo: Efficient Leave-One-Out Cross-Validation and WAIC for Bayesian Models*. R package version 2.8.0.
- Wasow, T. (2002). *Postverbal behavior*. CSLI Publications.
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., ... Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43), 1686.
- Wickham, H., Chang, W., Flight, R., Müller, K., & Hester, J. (2025a). *sessioninfo: R Session Information*. R package version 1.2.3.
- Wickham, H., Pedersen, T. L., & Seidel, D. (2025b). *scales: Scale Functions for Visualization*. R package version 1.4.0.
- Wurmbrand, S. (2000). *The structure(s) of particle verbs*. McGill University. <https://ucloud.univie.ac.at/index.php/s/dhZtAocQg3qTHvi> (accessed 2026-05-01).

- Xie, D., Chen, H., & Li, B. (2023). Chunks, pauses, and holistic processing in mandarin spontaneous speech. *Frontiers in Psychology*, 14, 1071729.
- Xie, Y. (2014). Knitr: A comprehensive tool for reproducible research in R. In Stodden, V., Leisch, F., & Peng, R. D. (Eds.), *Implementing reproducible computational research*. Chapman and Hall/CRC.
- Xie, Y. (2015). *Dynamic documents with R and knitr* (2nd ed.). Chapman and Hall/CRC.
- Xie, Y. (2025a). *knitr: A General-Purpose Package for Dynamic Report Generation in R*. R package version 1.50.
- Xie, Y. (2025b). *xfun: Supporting Functions for Packages Maintained by 'Yihui Xie'*. R package version 0.54.
- Xie, Y., Allaire, J., & Grolemund, G. (2018). *R markdown: The definitive guide*. Chapman and Hall/CRC.
- Xie, Y., Dervieux, C., & Riederer, E. (2020). *R markdown cookbook*. Chapman and Hall/CRC.
- Zeileis, A., Hothorn, T., & Hornik, K. (2008). Model-based recursive partitioning. *Journal of Computational and Graphical Statistics*, 17(2), 492–514.
- Zeller, J. (2001). *Particle verbs and local domains*. John Benjamins Publishing Company.

Appendices

A. Inclusion and exclusion criteria: annotated examples

The following examples (all from SWITCHBOARD; Godfrey et al., 1992) illustrate the data conditions under which tokens were *included* in, or *excluded* from, analysis. Speaker metadata are given in parentheses.

A.1. Included: transitive VPCs with variable particle placement

- (5) **Joined order (verb, particle, DO)**
 '[...] you **heat up** re-fried beans and you **cut up** tomatoes and lettuce and grated cheese'.
 (Woman, b. 1958, New York City Dialect Area, SWITCHBOARD)
- (6) **Split order (verb, DO, particle)**
 'Usually I **cut those little cherry tomatoes up** and put some colour into it [...]'
 (Woman, b. 1946, Southern Dialect Area, SWITCHBOARD)

Pronominal objects and complex NPs were *retained* (as controls and because they exhibit limited variability):

- (7) **Pronoun, split (near-categorical but retained)**
 'A slug is actually kind of slimy and ...he **picks them up** and rolls them across the patio.'
 (Man, b. 1950, Northern Dialect Area, SWITCHBOARD)
- (8) **Complex NP, joined (near-categorical but retained)**
 'It's not like these other programs where they just **make up anything they want**.'
 (Man, b. 1962, Mixed Regional Dialect, SWITCHBOARD)

A.2. Excluded contexts

A.2.1. Intransitives (no DO)

- (9) 'I was just about to **freak out**, you know.'
 (Woman, b. 1964, Mixed Regional Dialect, SWITCHBOARD)

A.2.2. Passives (no active particle-placement alternation)

- (10) 'Uh, you know, just like, software is only **given out** to customers ...'
 (Man, b. 1961, Western Dialect Area, SWITCHBOARD)

A.2.3. Intervening adverbs (e.g., right) between DO and particle

- (11) ‘...and **pay the bill right up**’.
(Woman, b. 1960, New England Dialect Area, SWITCHBOARD)

A.2.4. Extracted DO (object outside the VP domain of the VPC)

- (12) ‘What have you **got on**?’
(Woman, b. 1946, North Midland Dialect Area, SWITCHBOARD)

A.2.5. Preposition–verb constructions (non-alternating) versus particle verbs

- (13) **Preposition verb (excluded):** ‘**put it on** the steak’ (**put on it the steak*)
(Man, b. 1933, South Midland Dialect Area, SWITCHBOARD)
- (14) **Particle verb (included):** ‘**Mix all that up** together ...’ (mix up all that together)
(Man, b. 1933, South Midland Dialect Area, SWITCHBOARD)

B. Idiomaticity coding: diagnostic and examples

Idiomaticity was coded following an entailment-based diagnostic (Lohse et al., 2004, 244–246) and subsequent applications (Grafmiller & Szmrecsanyi, 2018; Röthlisberger & Tagliamonte, 2020, 392). Throughout, we use *idiomaticity* as an operational cover term for *semantic chunkiness*. A token was coded as *compositional* only if the utterance-level meaning transparently entailed both (i) the corresponding simplex-verb event and (ii) a literal/spatial paraphrase contributed by the particle. Otherwise, it was coded as *idiomatic* in the present operational sense, which includes both fully idiomatic/holistic senses and conventionalised non-spatial resultative/aspectual particle uses (e.g., *up* in *cheer someone up*, *brighten something up*). The following tokens illustrate the distinction; coded values for other variables are provided for transparency.

- (15) a. **Compositional** (*pick them up* ‘lift’):
‘A slug is actually kind of slimy and- and- these are just- he plays with them all the time, he **picks them up** and rolls them across the patio.’
(Male, b. 1950, Northern Dialect Area, SWITCHBOARD)
- b. **Idiomatic** (*pick up guys* ‘meet/attract’):
‘We do social things, like you were saying, you know, going out and **picking up guys** or, or whatever.’
(Woman, b. 1965, New England Dialect Area, SWITCHBOARD)

The following examples illustrate the broader operationalisation used here, where conventionalised non-spatial resultative/aspectual particle uses are also coded as *idiomatic* (i.e., semantically chunked) for the purposes of a binary contrast. As noted above, this grouping is an *operational* decision for a binary contrast and should not be read as a claim that all resultative/aspectual uses are categorically non-compositional:

- (16) a. **Idiomatic (resultative/aspectual up)** (*beat them up* ‘whisk thoroughly’):
‘I’ll put the eggs in the same dish and **beat them up**.’
(Woman, b. 1960, New England Dialect Area, SWITCHBOARD)
- b. **Idiomatic (resultative/augmentative up)** (*brighten up their lives* ‘improve/cheer’):
‘The kids just **brighten up their lives**.’
(Woman, b. 1957, South Midland Dialect Area, SWITCHBOARD)

C. Colophon

Software and reproducibility

We used R v. 4.5.2 (R Core Team, 2025) and the following R packages: ggparty v. 1.0.0.1 (Borkovec & Madin, 2025), ggpubr v. 0.6.2 (Kassambara, 2025), glmertree v. 0.2.6 (Fokkema et al., 2018; Fokkema & Zeileis, 2024),

here v. 1.0.2 (Müller, 2025), knitr v. 1.50 (Xie, 2014, 2015, 2025a), lme4 v. 1.1.37 (Bates et al., 2015), lmerTest v. 3.1.3 (Kuznetsova et al., 2017), partykit v. 1.2.24 (Hothorn et al., 2006; Hothorn & Zeileis, 2015; Zeileis et al., 2008), patchwork v. 1.3.2 (Pedersen, 2025), rmarkdown v. 2.30 (Allaire et al., 2025; Xie et al., 2018, 2020), scales v. 1.4.0 (Wickham et al., 2025b), sessioninfo v. 1.2.3 (Wickham et al., 2025a), sjPlot v. 2.9.0 (Lüdtke, 2025), tidyverse v. 2.0.0 (Wickham et al., 2019), xfun v. 0.54 (Xie, 2025b), brms v. 2.23 (Bürkner, 2025), bayestestR v. 0.17 (Makowski et al., 2025), bridgesampling v. 1.x1–2 (Gronau & Singmann, 2021), and loo v. 2.8 (Vehtari et al., 2024). Audio segmentation for silence durations used Praat v. 6.1.19 (Boersma & Weenink, 2020).