Frequency modulates structural choice in Turkish suspended affixation: a latent-process account

Utku Turk[1], Eva Neu[2], Ozge Bakay[2], Brian Dillon[2], Gaja Jarosz[2]

[1]University of Maryland, College Park
[2]University of Massachusetts, Amherst
Correspondence: utkuturk@umd.edu

Abstract

Suspended affixation (SA) allows a suffix on one conjunct to scope over all coordinated elements. While inflectional SA is productive in Turkish, derivational SA is claimed to be highly restricted; yet speakers readily accept certain cases. We propose that this gradient acceptability reflects a frequency-modulated choice between two possible syntactic representations: base-generation, which licenses derivational SA, and ellipsis. To test this, we conducted a rating task on the acceptability of four derivational suffixes in SA form while manipulating the frequency of coordinations. Using a Multinomial Processing Tree model to isolate latent structural choices from surface ratings, we found that frequency modulated SA acceptability for some suffixes (i.e., *-sIz* '-less' and *-cI* '-maker'), but not others (i.e., *-lI* '-having' and *-lIk* '-for'). These findings suggest that frequency shapes syntactic parsing in morphologically complex environments.

1 Introduction

A central puzzle in the morphosyntax of agglutinative languages concerns the distribution of affixes across coordinated structures. In Turkish, a single suffix appearing on the rightmost conjunct can be semantically interpreted as scoping over multiple conjuncts—a phenomenon termed *suspended affixation* (SA). Consider (1), where the dative marker *-ye* surfaces only on the second conjunct yet receives an interpretation in which both *Ali* and *Ayşe* bear dative case.

(1) Ali ve Ayşe-ye bak-tı-m.
Ali and Ayşe-DAT look-PST-1SG
'I looked at Ali and Ayşe.'

This construction has been documented across typologically diverse languages, including Turkish (Kornfilt, 1996, 2012; Kabak, 2007; Broadwell, 2008; Akkuş, 2016; Atmaca, 2021), Mari (Guseva and Weisser, 2018), Ossetic, Iron, and Eastern Armenian (Erschler, 2012), Dagur (Gong, 2021), Japanese and Korean (Yoon and Lee, 2005), Nivkh (Gruzdeva, 1998), and Hungarian (Trommer, 2008).

While SA with *inflectional* morphology is relatively uncontroversial, the availability of *derivational* SA remains theoretically contested. While some suggest that derivational SA is not a productive process in Turkish (Atmaca, 2021; Kabak, 2007; Kornfilt, 2012), others argue that it is more productive than previously assumed (Akkuş, 2016). Drawing on these conflicting observations, the current study empirically examines whether the gradient acceptability of derivational SA in Turkish is modulated by the frequency of coordinations. By fitting a Multinomial Processing Tree model to rating data on SA forms across different suffixes and coordinations with varying frequencies, we show evidence that frequency does not simply increase surface acceptability of all suffixes, but instead influences syntactic parsing for certain suffixes.

1.1 The status of derivational SA in Turkish

Earlier observations regarding derivational SA in Turkish report that it is a rare phenomenon with considerable inter-speaker and inter-item variation. For example, Kabak (2007) argues that apparent cases of derivational SA, such as (2), are restricted to tightly bound, lexicalized collocations (see also Atmaca, 2020).

(2) a. ana ve baba-lık
mother and father-DER
'motherhood and fatherhood'

b. ay-yıldız-lı bayrak
moon-star-ADJ flag
'a flag with a moon and star' (lit. 'a moon-star-ADJ flag', i.e., the Turkish flag)

Under his view, examples such as (2) are idiomatic

chunks stored holistically in the lexicon and denoting frequently co-occurring entities (Wälchli, 2005), rather than reflecting a productive affixation process. Similarly, Kornfilt (2012) has argued that SA in Turkish is strictly limited to inflectional affixes.

However, subsequent work by Akkuş (2016) challenges this categorical restriction. Based on examples sourced online, Akkuş demonstrates that derivational SA across both nominal and verbal domains is more productive than previously thought. Crucially, he shows that the distributing affix can scope over full phrasal conjuncts—not merely adjacent, tightly-bound lexical items—e.g., '[5 lira] and [10 dollar]-DER' in (3).

(3) ... beş lira ve on dolar-lık
five lira and ten dollar-DER
banknot-lar...
banknote-PL
'5-lira and 10-dollar banknotes'

Moreover, some cases of derivational SA cited by Akkuş are acceptable regardless of the order of conjuncts, shown for the conjunct *dost ve arkadaş* 'buddy and friend' along with its version with a reversed order in (4). This flexibility is hard to reconcile with the idea that these SA constructions are simple isolated collocations stored in the lexicon.

(4) a. Bütün eğitim çalışmaları boyunca
all training sessions during
dost ve arkadaş-ça bir ortam
buddy and friend-DER an atmosphere
ol-malı-dır.
be-must-MOD
'There must be a friendly and intimate environment during the whole training session.'

b. ... buna uygun davran-mak
this-DAT appropriate behave-NMLZ
için arkadaş ve dost-ça
for friend and buddy-DER
gel-di-k.
come-1PL-PST
'... accordingly, we came in a friendly and intimate manner.'

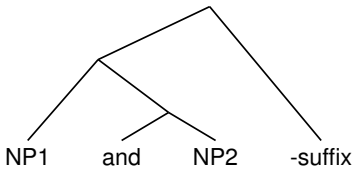
At the same time, the fact remains that derivational SA is more idiosyncratically restricted than inflectional SA. Specifically, we observe that it appears to be more readily accepted for high-frequency coordinations such as *moon and star* and *friend and buddy*, as correctly noted by Kornfilt and Kabak.

While Akkuş's findings have been adopted in subsequent work on SA (e.g., Atmaca, 2021; Türk, 2025), the role of frequency in modulating the availability and parse of these constructions has not been systematically investigated. Given the speaker- and item-level variation noted in the literature, we propose that the apparent conflict between the intuitions of Kabak (2007) and the evidence provided by Akkuş (2016) may be reconciled by attending to the *probabilistic* nature of syntactic choice underlying derivational SA.

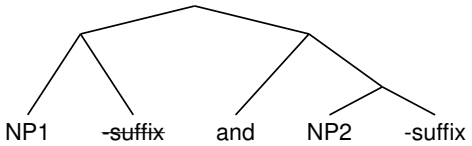
1.2 The present study

Our proposal in this study rests on the assumption that in general, SA is structurally ambiguous between two possible parses. Under a *base generation* analysis as in (5), the suspended affix combines with the coordinated noun phrase as a whole. Under an *ellipsis* analysis as in (6), the affix is underlyingly merged with each conjunct, but is deleted in the non-final conjunct(s). When encountering an SA construction, speakers are in principle free to adopt either parse.[1]

(5)



(6)



Against this background, we argue that the difference in productivity between derivational and inflectional SA can be accounted for as follows. First, we propose that derivational affixes can only participate in SA in a base generation structure and not in an ellipsis structure, whereas inflectional affixes can freely undergo ellipsis. What distinguishes derivational from inflectional affixes is that they sit closer to the root and are more tightly integrated with it semantically, phonologically and prosodically. We suggest that as a result, derivational affixes cannot be deleted at the exclusion of the root since the two form too inseparable a

[1] This holds for the items in our experimental studies. Some instances of SA are only compatible with one parse due to various syntactic, morphological and semantic factors. In Section 4, we discuss one such factor—scope interpretation—and how it can be used to further test our hypothesis that the gradient acceptability of derivational SA is linked to structural ambiguity.

unit. To our knowledge, this claim is novel. In Section 4, we discuss more in detail why it would apply; here, we focus on how it derives the gradient acceptability profile of derivational SA.

Second, we argue that the availability of the two structural paths to SA in (5) and (6) is shaped by the frequency of the coordinated structure. There is strong evidence that speakers store holistic representations of coordinated noun phrases, and that the more frequent the coordination is, the more likely speakers are to access this holistic representation in parsing (Morgan and Levy, 2015, 2016). Crucially, such a holistic representation is only compatible with a base generation analysis; in the ellipsis structure, the coordinated phrase such as *mother and father* does not form a constituent. Consequently, we propose that high-frequency structures bias the parser towards a base generation analysis.

Taken together, these two proposals account for the differences in productivity of inflectional and derivational SA. Since derivational SA is only available with a base generation analysis, it is more acceptable with high-frequency coordinations, as has been observed anecdotally in the literature and as we confirm experimentally in the following. Upon encountering a high-frequency coordinated phrase, speakers tend to access the holistic representation associated with it. This leads them to posit a base generation analysis, which is compatible with derivational SA. In contrast, inflectional SA is highly productive overall (Atmaca 2020, but see Kornfilt 1996, Türk 2025). This is as expected if inflectional affixes can freely participate in SA under both structural analyses. Even if speakers posit an ellipsis analysis of SA for low-frequency coordinations, the result is still acceptable.

Our claim that the frequency of coordinated noun phrases modulates structural choice in SA is rooted in much previous work showing evidence that frequency distributions can influence online parsing. For example, a large body of evidence supports the claim that frequency affects the processing of both single words and complex multi-word expressions (Pinker and Ullman, 2002; Arnon and Snider, 2010; Carroll and Conklin, 2020; Siyanova-Chanturia et al., 2011; Sosa and MacFarlane, 2002; Swinney and Cutler, 1979; Tremblay et al., 2011; Morgan and Levy, 2015, 2016). Frequent multi-word sequences are recognized faster and processed more fluently than matched infrequent sequences, even when individual word frequency is controlled. This suggests that speakers maintain holistic representations for recurrent phrases alongside compositional parsing.

Morgan and Levy (2015, 2016) show evidence that frequency-based storage interacts with other grammatical preferences. They show that more frequent coordinated phrases exhibit more rigid ordering preferences. For example, *ladies and gentlemen* occurs almost exclusively in this order, whereas *towels and sunglasses* alternates freely with *sunglasses and towels*. Morgan and Levy interpret this as a trade-off between compositional and holistic representations: as a coordination becomes more frequent, speakers increasingly store it as a unit rather than assembling it compositionally, resulting in frozen ordering.

A key finding that is relevant for the present study is that the effect of frequency does not have to be linear. Alternatively, the effect of frequency can be seen as a choice between competing syntactic structures, as has been suggested for the domain of argument structure alternations, most notably the English dative alternation. Bresnan and colleagues (Bresnan, 2007; Bresnan and Ford, 2010; Bresnan et al., 2004) have shown that the choice between a double object construction (*I gave Mary the book*) and a prepositional dative (*I gave the book to Mary*) is probabilistic rather than categorical, and is governed by a range of factors including animacy, givenness, and constituents' phonological weight. Crucially, verb bias—the frequency with which a verb appears in one construction over the other—shapes expectations about structural form and influences both production and comprehension, with speakers' naturalness ratings closely tracking corpus-derived construction probabilities (Bresnan, 2007).

Similarly, our proposal goes beyond the claim that frequent expressions are simply stored as unanalyzed wholes. While we draw on Morgan and Levy's idea that higher frequency coordinations are more likely to be represented holistically, our critical prediction concerns not storage *per se*, but its downstream consequences for syntactic structure assignment. In our model, the key variable modulated by frequency is the probability of adopting a base-generation parse over an ellipsis parse. Highly frequent coordinations are more likely to be accessed holistically, making the base-generation structure—in which the affix attaches to the coordination as a single constituent—more readily available. Lower-frequency coordinations remain compatible with both routes, but the ellipsis

parse—in which the affix merges with each conjunct separately—becomes relatively more available.

Thus, while we see an effect of frequency in the experimental results, we show that frequency does not uniformly boost acceptability scores across the board. This would not be an informative finding, but might simply indicate that speakers prefer expressions that they have encountered more often in their linguistic input. Instead, we argue that frequency shifts the distribution over latent structural analyses, and these structural analyses in turn determine whether a given affix type can participate in SA. More specifically, we present a formal hypothesis that captures how frequency and structural ambiguity interact to yield the gradient acceptability profile observed in Turkish derivational SA using Multinomial Processing Trees (Riefer and Batchelder, 1988; Erdfelder et al., 2009), which represent the observed acceptability responses as a process generated by a sequence of latent subprocesses rather than by a single undifferentiated acceptance probability (Riefer and Batchelder, 1988; Batchelder and Riefer, 1999; Klauer, 2010). Our model rests on the assumptions that the structural choice between a base generation and an ellipsis analysis is probabilistic and sensitive to frequency, and furthermore, that participants are not always attentive in their responses.

We proceed as follows. Section 2 presents experimental data confirming a frequency effect for SA with certain kinds of derivational affixes. Section 3 then introduces a Multinomial Processing Tree model of the results. We end on a general discussion in Section 4.

2 Experiment

We conducted a Likert-scale judgment task to test whether the acceptability of SA involving derivational suffixes is modulated by the frequency of the coordinated phrases. We predicted that if derivational SA is sensitive to frequency, the acceptability of coordinated phrases with SA would increase as a function of their frequency. Furthermore, if this sensitivity is a general property, we should observe comparable frequency effects across all tested suffixes.

2.1 Materials and methods

The stimuli were Turkish sentences including a coordinated phrase as a modifier of a noun phrase in the sentence. We employed a 2x4 within-subject design, crossing SA, i.e., coordinated phrases with SA (i.e., with the derivational suffix appearing only on the second conjunct) or in FULL form (i.e, with the derivational suffix appearing on both conjuncts), and SUFFIX TYPE, i.e., *-cI* '*-maker*', *-lI* '*-having*', *-lIk* '*-for*' or *-sIz* '*-less*'. All four suffixes are highly productive in forming modifiers in the language. Example stimuli are in Table 1. The eight experimental conditions were distributed across eight lists using a Latin square design, ensuring that each participant saw each item only once and in only one condition.

There were 32 experimental sentences in total, each with a different coordinated noun phrase. The coordinated phrases covered a wide range of frequency: their counts were taken from a Turkish Web 2012 corpus (∼3.4 billion words), and ranged from 131,657 ('mother and father') to 1 ('mouse and lion'). To ensure semantic naturalness, coordinated nouns were taken from the same conceptual realm.

The experimental sentences were interspersed with 64 filler items. 24 of the filler items were ungrammatical due to several syntactic violations such as voice, case or number mismatches. The remaining 40 were grammatical sentences, 10 of which had coordinating adjectives with no affix, eliminating any ambiguity in syntactic parse, as in *red and thick scarf*.

47 Turkish speakers participated in the task. Their task was to judge the acceptability of sentences on a 6-point Likert scale (6 = 'very good'). An even-numbered scale was chosen to encourage participants to lean towards either an acceptable or unacceptable interpretation (Schütze and Sprouse, 2013). The task was conducted online on PCIbex-Farm (Zehr and Schwarz, 2018).

2.2 Results

Figure 1 shows the observed mean ratings as a function of centered rank frequency, with higher values corresponding to more frequent coordinations. Descriptively, neither the SA penalty nor the effect of frequency seems to be uniform across suffixes.

To analyze the ordinal rating data, we fit separate Bayesian cumulative-logit models for each suffix using `brms` (Bürkner, 2017). The models included fixed effects for SA, centered rank frequency, and their interaction, as well as varying effects for subjects and items. SA was sum contrast coded, with .5 for FULL and -.5 for SA forms. We

SUFFIX	Sentence								
-cI '-maker'	Selin	tamir	için	kapı(-cı)	ve	pencere-ci	bir	marangozla	anlaşmış.
	S.	repair	for	door-MAKER	and	window-MAKER	a	carpenter	made.a.deal.with
	'Selin agreed with a door-maker and window-maker carpenter for the repair.'								
-lI '-having'	Selin	tamir	için	kapı(-lı)	ve	pencere-li	bir	atölye	hazırlamış.
	S.	repair	for	door-HAVING	and	window-HAVING	an	atelier	prepared
	'Selin prepared an atelier with a door and a window for the repair.'								
-lIk '-for'	Selin	tamir	için	kapı(-lık)	ve	pencere-lik	bir	desen	seçmiş.
	S.	repair	for	door-FOR	and	window-FOR	a	pattern	chose
	Selin chose a pattern for a door and a window for the repair.								
-sIz '-less'	Selin	tamir	için	kapı(-sız)	ve	pencere-siz	bir	atölye	hazırlamış.
	S.	repair	for	door-LESS	and	window-LESS	an	atelier	prepared
	'Selin prepared an atelier without a door or a window for the repair.'								

Table 1: Example stimuli from the acceptability rating task. Suffixes in parentheses are omitted in forms with SA.

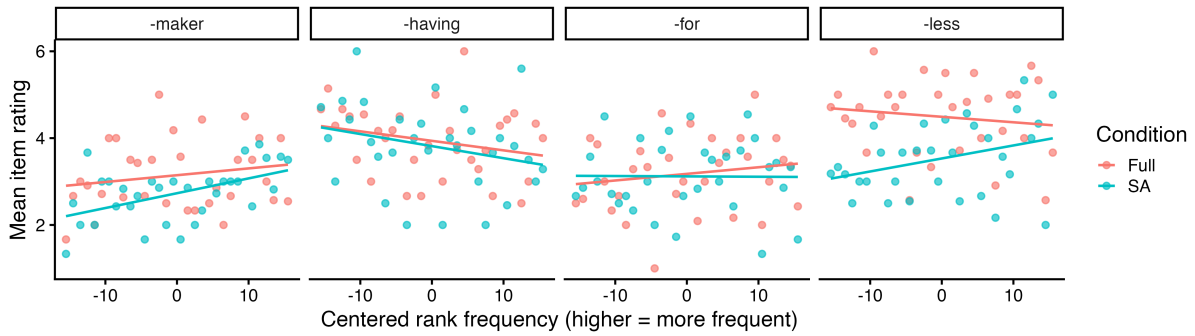


Figure 1: Observed mean item ratings by suffix and condition. Higher centered-rank values correspond to more frequent coordinations.

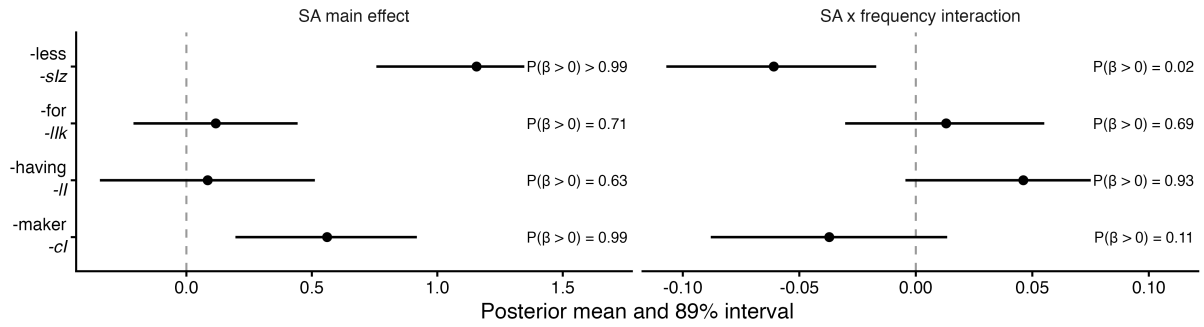


Figure 2: Posterior summaries from the suffix-specific Bayesian ordinal models. The left facet shows the SA main effect; the right facet shows the SA-by-frequency interaction. Labels at the right edge give $P(\beta > 0)$ for each coefficient.

also examined the potential role of trial order. Effects were mostly negligible for three of the four suffix-specific models (*-sIz*: $\hat{\beta} = 0.00$, 89% CrI $[-0.16, 0.17]$, $P(\beta{>}0) = 0.51$; *-cI*: $\hat{\beta} = 0.01$, 89% CrI $[-0.16, 0.17]$, $P(\beta{>}0) = 0.53$; *-lIk*: $\hat{\beta} = -0.03$, 89% CrI $[-0.19, 0.13]$, $P(\beta{>}0) = 0.38$). However, we found a negative main effect for *-lI* ($\hat{\beta} = -0.21$, 89% CrI $[-0.38, -0.05]$, $P(\beta{>}0) = 0.02$). Since it did not have a reliable effect on the ratings of all suffixes, it was excluded from the models reported here.

Figure 2 plots the posterior means, their 89% intervals and the posterior probability of a positive effect for SA and its interaction with frequency. In Figure 2, a positive sign for the SA effect suggests higher ratings for FULL relative to SA forms; a negative sign for the interaction effect suggests that frequency improves the ratings for SA forms more than it improves those for their FULL versions.

Results showed an overall decrease in the ratings

for SA relative to FULL forms with '-less' ($\hat{\beta}$=1.16, 89%CrI [0.76, 1.55]) and '-maker' ($\hat{\beta}$=0.56 [0.20, 0.92]). There was no such difference with '-having' or '-for'. The clearest evidence for an interaction between form and frequency was with '-less', evidenced by an increase in the acceptability of SA forms with higher frequency ($\hat{\beta}$=−0.06 [-0.11, -0.02]). There was no clear evidence for an interaction with '-maker' or '-for'. Lastly, '-having' showed a trend in the opposite direction, suggesting a decrease in the acceptability of SA forms with higher frequency; but this trend was not statistically reliable ($\hat{\beta}$=0.05 [0.00, 0.10]).

Overall, there was clear evidence that the four suffixes tested here differ with respect to the overall SA penalty and how this penalty is modulated by frequency. This suggests that a simple narrative where frequency uniformly boosts acceptability is insufficient. Moreover, the cumulative-logit models we report above can only describe final ratings, which entangle structural parsing preferences with baseline response biases and inattention. To systematically reveal the latent variables contributing to the ratings and determine the exact role of frequency in syntactic choice, we use a Multinomial Processing Tree process model, which we discuss next.

3 Latent Modeling

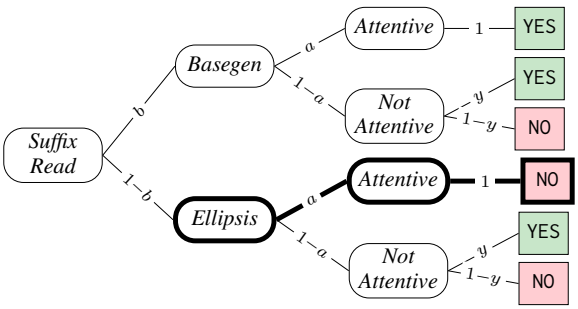
3.1 Methods

It is widely accepted that experimental measures such as acceptability judgments or reading times are not transparent reflections of a single underlying process; rather, they are generated by a mixture of latent subprocesses (Ratcliff, 1978; Riefer and Batchelder, 1988; Batchelder and Riefer, 1999; Paape and Vasishth, 2022; Nicenboim et al., 2024). Consequently, inferences based only on condition means can underrepresent some cognitive processes while amplifying others. On this view, assuming a simple, one-to-one mapping between frequency and acceptability is too coarse: frequency may influence ratings indirectly by shifting structural choice, while the observed responses are simultaneously shaped by attention and response bias. The descriptive results above align with this point, because neither frequency nor the availability of derivational SA exerts a single linear effect across suffixes, suggesting a potential role for subprocesses in acceptance ratings.

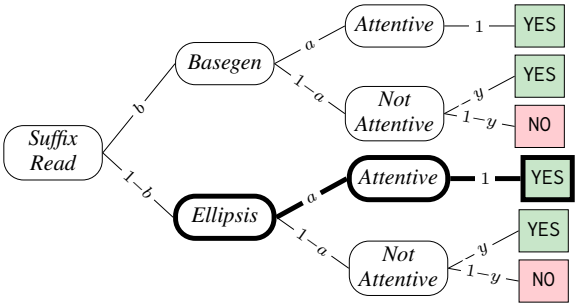
Therefore, we analyze the experimental findings using Multinomial Processing Trees (Riefer and Batchelder, 1988; Erdfelder et al., 2009; Batchelder and Riefer, 1999; Klauer, 2010). We make three main assumptions in the model. First, the structural choice between a base-generation and an ellipsis analysis is probabilistic rather than deterministic, and *probabilities for structural choices* can shift as a function of experience-based factors such as frequency (Jurafsky, 1996). The second assumption is that participants may not always be attentive in their responses. When they are inattentive, their responses are not random; instead, they are influenced by an a priori response yes/no tendency, which we model as a *subject-level response-bias*. Third, in addition to differences across participants, a single participant's attentional state may vary from trial to trial. While a participant may sometimes respond on the basis of the stimulus, they may sometimes fail to do so (Paape and Vasishth, 2022; Nicenboim et al., 2024). We model this as *trial-wise attentional variation*.

Accordingly, our model includes three latent subprocesses: structural choice (base generation vs. ellipsis), attentional state (i.e., when participants are not attending to the stimulus), and a residual yes-response tendency in an inattentive state. The trees in (7) and (8) illustrate the logic schematically:

(7) Derivational Suffix



(8) Inflectional Suffix



In the derivational case, attentive processing leads to acceptance under the base-generated parse but not under the ellipsis parse; in the inflectional case, both structural routes remain compatible with

acceptance. In both trees, inattentive trials fall back to a participant-specific yes-bias. We predict that if frequency modulates structural choice in parsing derivational SA, we should observe a higher probability for a base generation route with more frequent coordinations.

Before fitting the MPT, we converted 6-point ratings to binary accept/reject outcomes. The threshold was chosen empirically from filler items by fitting a Bayesian cumulative-logit model and evaluating each candidate threshold by its posterior balanced accuracy in separating grammatical from ungrammatical fillers. A threshold at 4 gave the best performance (balanced accuracy = 0.929, 90% interval: [0.914, 0.942]); ratings of 4–6 were treated as acceptance and 1–3 as rejection.

The MPT was then fit to these binary judgments. Let $a_{j[n]}$ denote subject j's attentiveness on trial n and $g_{j[n]}$ their filler-derived yes-bias. For SA items, the attentive state involves a further latent choice: with probability b_n the parser adopts a base-generation analysis and the SA item inherits full-form acceptance θ_n^{full}; with probability $1 - b_n$ it adopts an ellipsis analysis governed by θ_n^{ell}:

$$P(\text{yes}_n) = a_{j[n]}\,[b_n\,\theta_n^{\text{full}} + (1 - b_n)\,\theta_n^{\text{ell}}] + (1 - a_{j[n]})\,g_{j[n]} \quad (9)$$

For FULL-form items $b_n = 1$ by assumption, which simplifies Equation (9) to (10):

$$P(\text{yes}_n) = a_{j[n]}\,\theta_n^{\text{full}} + (1 - a_{j[n]})\,g_{j[n]} \quad (10)$$

The latent subprocesses of the MPT model shown in (7) and (8) are as follows.

Attention (a_j): Each participant's attentiveness probability, governing whether their response reflects the structural model or falls back to their default yes-bias, calculated by their responses to grammatical fillers.

Criterion-derived yes bias (g_j): Estimated from each participant's hits and false alarms on filler items via a signal-detection-theoretic criterion c_j, mapped to a yes-bias through $\Phi(-c_j)$. Using exclusively fillers ensures that response bias and attentiveness are identified independently of the critical SA items (Hautus et al., 2021).

Attentive acceptance (θ_n^{full}): Suffix-specific acceptance probability for FULL forms, with varying effects for subjects and items.

Structural route (b_n): For SA forms, the probability of adopting a base-generation analysis (Equation (9)). Under base generation, the SA item inherits θ_n^{full}; under ellipsis $(1 - b_n)$, acceptance is governed by a separate suffix-specific parameter θ_n^{ell}.

Frequency effect on b_n: The base-generation probability is modeled on the logit scale with a suffix-specific intercept and slope for centered rank frequency, plus varying effects for subjects and items. A positive slope means higher frequency favors base generation—the structural route that licenses derivational SA.

The MPT model was implemented directly in Stan (Carpenter et al., 2017).

3.2 Results

The MPT outperformed a maximal `brms` baseline on PSIS-LOO (ΔELPD = 7.47; Sivula et al. 2025; Vehtari et al. 2017), and is therefore used as the primary model. Nuisance parameters were well-identified: mean attentiveness was 0.81 ($SD = 0.01$), and the filler-derived criterion was mildly liberal ($M = -0.24$, $SD = 0.41$), corresponding to an inattentive yes-bias of 0.59 ($SD = 0.15$), ensuring that structural parameters do not absorb response bias.

Table 2 reports the structural parameters. SA-form acceptability (row 2) is not a free parameter: it is the weighted average $b \cdot \theta_{\text{full}} + (1 - b) \cdot \theta_{\text{ellip}}$, so the gap between rows 1 and 2 reflects both how penalized the ellipsis route is and how often it is taken.

Two patterns stand out. First, the SA penalty varies substantially across suffixes: largest for '-less' and '-maker', modest for '-having', and negligible for '-for'. The '-for' case shows why frequency cannot act as a direct acceptability booster: when both routes already yield poor acceptance, shifting route probability has little observable effect on ratings.

Second—and this is the central contribution of the MPT—frequency shifts $P(\text{base-gen})$ in the predicted direction for the two suffixes with the clearest SA penalty ($P(\beta{>}0) = 0.83$ for '-less'; 0.80 for '-maker'). The '-maker' effect had no surface signature in the ordinal model, because the low full-form acceptability (0.34) limits the observable gain from route-shifting. For '-having', the frequency effect runs negative ($P(\beta{>}0) = 0.12$), and for '-for', it is uninformative ($P(\beta{>}0) = 0.62$).

Overall, frequency is better understood as modulating structural route availability than as adding a uniform boost to surface acceptability.

Parameter	Suffix	M [Lower, Upper]	Interpretation
Full-form acceptability	*-sIz* '-less'	0.86 [0.79, 0.92]	Attentive-state $P(\text{accept})$ for full forms
	-cI '-maker'	0.34 [0.24, 0.44]	
	-lI '-having'	0.77 [0.69, 0.85]	
	-lIk '-for'	0.36 [0.26, 0.46]	
SA-form acceptability	*-sIz*	0.49 [0.37, 0.62]	Marginal $P(\text{accept})$ at mean frequency
	-cI	0.19 [0.11, 0.29]	
	-lI	0.65 [0.53, 0.76]	
	-lIk	0.32 [0.23, 0.42]	
Ellipsis acceptability	*-sIz*	0.33 [0.17, 0.52]	Attentive-state $P(\text{accept})$ via ellipsis route
	-cI	0.11 [0.04, 0.21]	
	-lI	0.52 [0.32, 0.70]	
	-lIk	0.24 [0.11, 0.38]	
$P(\text{base-gen})$ at mean freq.	*-sIz* ($P(\beta{>}0.5){=}0.08$)	0.30 [0.11, 0.54]	Ellipsis strongly dominates
	-cI ($P(\beta{>}0.5){=}0.26$)	0.38 [0.12, 0.71]	Ellipsis dominates
	-lI ($P(\beta{>}0.5){=}0.50$)	0.50 [0.19, 0.81]	Completely uncertain
	-lIk ($P(\beta{>}0.5){=}0.69$)	0.60 [0.23, 0.89]	Base-gen. modestly dominant
Effect of freq. on $P(\text{base-gen.})$	*-sIz* ($P(\beta{>}0){=}0.83$)	$+0.08$ [-0.06, $+0.26$]	suggestive positive effect
	-cI ($P(\beta{>}0){=}0.80$)	$+0.15$ [-0.16, $+0.45$]	suggestive positive effect
	-lI ($P(\beta{>}0){=}0.12$)	-0.20 [-0.49, $+0.08$]	suggestive negative effect
	-lIk ($P(\beta{>}0){=}0.62$)	$+0.06$ [-0.33, $+0.43$]	uncertain

Table 2: Posterior means and 89% credible intervals for key MPT parameters. In the lower section, parenthetical values next to each suffix are $P(\beta{>}\text{chance})$: the posterior probability that base generation is the dominant parse at mean frequency. In the frequency effect section, parenthetical values are $P(\beta{>}0)$: the posterior probability that the effect of frequency on base-generation probability is positive. The frequency effect is on the logit scale.

4 General Discussion

This study investigated how lexical frequency modulates the acceptability of derivational suspended affixation and modeled the underlying mechanism using a Multinomial Processing Tree. While our ordinal regression models demonstrate that suffixes do not behave uniformly with respect to frequency, the MPT's central contribution was to explicitly map out the underlying mechanism driving these differences. Specifically, the MPT model decomposes what might look like a direct frequency-acceptability link into a frequency effect on structural route probability: higher-frequency items are more likely to receive a base-generation analysis, and acceptability then follows from this structural choice together with nuisance processes. Critically, this decomposition revealed a frequency effect on structural choice for '-maker' that was entirely absent from the surface ordinal ratings—demonstrating that condition means alone are insufficient and that the computational model is necessary to isolate latent structural preferences.

Not all suffixes showed the predicted frequency-driven increase in base-generation probability, and the MPT explains why. For '-for', base generation is already the modestly dominant parse at mean frequency ($P(b{>}0.5) = 0.69$), and both full-form and SA-form acceptability are low—a configuration in which frequency-driven route-shifting yields little observable leverage on ratings. For '-having', frequency moved base-generation probability in the negative direction ($P(\beta{>}0) = 0.12$), possibly reflecting this suffix's relatively higher ellipsis-route acceptability compared to the other suffixes. These patterns are not unexplained residuals: they follow naturally from the model, because the observable consequence of route-shifting depends on the suffix-specific acceptability. It is also possible that these null or unexpected trends reflect a lack of statistical power to detect smaller effects due to our limited number of participants. Given the robust effects observed with the other suffixes, we consider this less likely, but leave further testing with a larger sample size to future work.

A crucial question raised by our proposal is why certain suffixes would resist an ellipsis analysis in the first place—which we have assumed rather than motivated—and what role is played by the difference between derivational and inflectional morphemes. Theoretical frameworks such as Distributed Morphology (Halle and Marantz, 1993, 1994) have argued that the traditional derivational/inflectional divide is better conceptualized as a continuum determined by syntactic attachment height and the position of functional heads

in the syntactic spine. Suffixes that attach low in the structure—close to the root—are more lexically integrated and therefore a poorer target for deletion. This predicts that suffixes like '-less' and '-maker', whose base-generation probability is most strongly modulated by frequency, are precisely those whose attachment height makes ellipsis less viable. Making this link precise by independently characterizing the attachment heights of Turkish derivational suffixes is an important direction for future work. More generally, connecting our proposal more closely to the syntactic literature on ellipsis analyses of SA (e.g., Despić, 2017; Erschler, 2018; Guseva and Weisser, 2018; Türk, 2025) would sharpen the predictions considerably.

We see an additional future direction for testing and extending the account. The base generation and ellipsis parses make different scopal predictions: under base generation, the derivational suffix takes scope over the full coordination, applying to both conjuncts jointly, while under ellipsis, each conjunct contains the suffix independently. These parses are therefore distinguishable by their interpretation. For example, *kitap ve dergi-siz* under a base-generation parse means 'lacking books-and-magazines (jointly)', whereas under ellipsis, the meaning is 'lacking books and lacking magazines (separately).' If frequency biases readers towards base generation as the MPT suggests, higher-frequency items should preferentially support the joint, wide-scope reading in a forced-choice interpretation task. This task would provide direct, independent evidence that frequency operates on parse selection rather than on acceptability directly–the core theoretical claim of this paper.

References

Faruk Akkuş. 2016. Suspended affixation with derivational suffixes and lexical integrity. In Proceedings of IMOG 2016, Patras.

Inbal Arnon and Neal Snider. 2010. More than words: Frequency effects for multi-word phrases. Journal of Memory and Language, 62:67–82.

Furkan Atmaca. 2020. Suspended Affixation in Turkish. MA Thesis, Boğaziçi University.

Furkan Atmaca. 2021. Suspended affixation needs no morphological word: The suffix -(y)Ip. In Proceedings of the Workshop on Turkic and Languages in Contact with Turkic, volume 6, pages 5035–5035.

William H. Batchelder and David M. Riefer. 1999. Theoretical and empirical review of multinomial process tree modeling. Psychonomic Bulletin & Review, 6:57–86.

Joan Bresnan. 2007. Is syntactic knowledge probabilistic? Experiments with the English dative alternation. In Sam Featherston and Wolfgang Sternefeld, editors, Roots: Linguistics in Search of Its Evidential Base, pages 77–96. Mouton de Gruyter, Berlin.

Joan Bresnan, Anna Cueni, Tatiana Nikitina, and R. Harald Baayen. 2004. Predicting the dative alternation. In Gerlof Bouma, Irene Kraemer, and Joost Zwarts, editors, Cognitive Foundations of Interpretation, pages 69–94. Royal Netherlands Academy of Arts and Sciences, Amsterdam.

Joan Bresnan and Marilyn Ford. 2010. Predicting syntax: Processing dative constructions in American and Australian varieties of English. Language, 86:186–213.

George Aaron Broadwell. 2008. Turkish suspended affixation is lexical sharing. In Miriam Butt and Tracy Holloway King, editors, Proceedings of the LFG08 Conference, pages 198–213. CSLI Publications, Stanford.

Paul-Christian Bürkner. 2017. brms: An R package for Bayesian multilevel models using Stan. Journal of Statistical Software, 80:1–28.

Bob Carpenter, Andrew Gelman, Matthew D. Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus A. Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2017. Stan: A probabilistic programming language. Journal of Statistical Software, 76:1–32.

Gareth Carroll and Kathy Conklin. 2020. Is all formulaic language created equal? Unpacking the processing advantage for different types of formulaic sequences. Language and Speech, 63:95–122.

Miloje Despić. 2017. Suspended morphology in Serbian: Clitics vs. affixes. Glossa: a journal of general linguistics, 2:12.

Edgar Erdfelder, Tina-Sarah Auer, Benjamin E. Hilbig, Andreas Aßmann, and Morten Moshagen. 2009. Multinomial processing tree models: A review of the literature. Zeitschrift für Psychologie / Journal of Psychology, 217:108–124.

David Erschler. 2012. Suspended affixation in Ossetic and the structure of the syntax-morphology interface. Acta Linguistica Hungarica, 59:153–175.

David Erschler. 2018. Suspended affixation as morpheme ellipsis: Evidence from Ossetic alternative questions. Glossa: a journal of general linguistics, 3:12.

Zhiyu Mia Gong. 2021. Postsyntactic lowering and linear relations in Dagur noun phrases. Glossa: a journal of general linguistics, 6:42.

Ekaterina Gruzdeva. 1998. Nivkh. Languages of the World/Materials. Lincom Europa, Munich.

Elina Guseva and Philipp Weisser. 2018. Postsyntactic reordering in the Mari nominal domain: Evidence from suspended affixation. Natural Language & Linguistic Theory, 36:1087–1120.

Morris Halle and Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In Kenneth Hale and S. Jay Keyser, editors, The View from Building 20, pages 111–176. MIT Press, Cambridge, MA.

Morris Halle and Alec Marantz. 1994. Some key features of Distributed Morphology. In MIT Working Papers in Linguistics, pages 275–288.

Michael J Hautus, Neil A Macmillan, and C Douglas Creelman. 2021. Detection theory: A user's guide. Routledge.

Daniel Jurafsky. 1996. A probabilistic model of lexical and syntactic access and disambiguation. Cognitive Science, 20:137–194.

Barış Kabak. 2007. Turkish suspended affixation. Linguistics, 45:311–347.

Karl Christoph Klauer. 2010. Hierarchical multinomial processing tree models: A latent-trait approach. Psychometrika, 75:70–98.

Jaklin Kornfilt. 1996. On some copular clitics in Turkish. In ZAS Papers in Linguistics, volume 6, pages 96–114. Zentrum für Allgemeine Sprachwissenschaft, Berlin.

Jaklin Kornfilt. 2012. Revisiting "suspended affixation" and other coordinate mysteries. In Laura Brugè, Anna Ferro, Carla Ferraro, Anna Ferro, and Lisa Ferraro, editors, Functional Heads: The Cartography of Syntactic Structures, pages 181–196. Oxford UP, Oxford.

Emily Morgan and Roger Levy. 2015. Modeling idiosyncratic preferences: How generative knowledge and expression frequency jointly determine language structure. Proceedings of the 37th Annual Conference of the Cognitive Science Society, pages 1649–1654.

Emily Morgan and Roger Levy. 2016. Abstract knowledge versus direct experience in processing of binomial expressions. Cognition, 157:384–402.

Bruno Nicenboim, Daniel Schad, and Shravan Vasishth. 2024. An Introduction to Bayesian Data Analysis for Cognitive Science. Cambridge UP. Online version: https://bruno.nicenboim.me/bayescogsci/.

Dario Paape and Shravan Vasishth. 2022. Estimating the true cost of garden-pathing: A computational model of latent cognitive processes. Cognitive Science, 46:e13186.

Steven Pinker and Michael T. Ullman. 2002. The past and future of the past tense. Trends in Cognitive Sciences, 6:456–463.

Roger Ratcliff. 1978. A theory of memory retrieval. Psychological Review, 85:59.

David M. Riefer and William H. Batchelder. 1988. Multinomial modeling and the measurement of cognitive processes. Psychological Review, 95:318–339.

Carson T. Schütze and Jon Sprouse. 2013. Judgment data, pages 27–50. Cambridge University Press.

Tuomas Sivula, Måns Magnusson, Alejandro Andres Matamoros, and Aki Vehtari. 2025. Uncertainty in Bayesian leave-one-out cross-validation based model comparison. Bayesian Analysis.

Anna Siyanova-Chanturia, Kathy Conklin, and Walter J. B. van Heuven. 2011. Seeing a phrase "time and again" matters: The role of phrasal frequency in the processing of multiword sequences. Journal of Experimental Psychology: Learning, Memory, and Cognition, 37:776–784.

Anna Vogel Sosa and James MacFarlane. 2002. Evidence for frequency-based constituents in the mental lexicon: Collocations involving the word *of*. Brain and Language, 83:227–236.

David A. Swinney and Anne Cutler. 1979. The access and processing of idiomatic expressions. Journal of Verbal Learning and Verbal Behavior, 18:523–534.

Antoine Tremblay, Bruce Derwing, Gary Libben, and Chris Westbury. 2011. Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks. Language Learning, 61:569–613.

Jochen Trommer. 2008. Hungarian has no portmanteau agreement. In Proceedings of the 26th West Coast Conference on Formal Linguistics, pages 478–486, Somerville, MA. Cascadilla Proceedings Project.

Utku Türk. 2025. Controlling morphosyntactic competition through phonology. In Pavel Caha, Karen De Clercq, and Guido Vanden Wyngaerd, editors, Nanosyntax and the Lexicalization Algorithm. Oxford UP, Oxford.

Aki Vehtari, Andrew Gelman, and Jonah Gabry. 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Statistics and Computing, 27:1413–1432.

Bernhard Wälchli. 2005. Co-Compounds and Natural Coordination. Oxford UP, New York, NY.

James Hye-Suk Yoon and Wooseung Lee. 2005. Conjunction reduction and its consequences for noun phrase morphosyntax in Korean. In Proceedings of the 24th West Coast Conference on Formal Linguistics, pages 379–387, Somerville, MA. Cascadilla Proceedings Project.

J. Zehr and F. Schwarz. 2018. PennController for Internet Based Experiments (IBEX).