Katz, Yuval Z. 2026. Grammatical generalizations and item-specific listing in Hebrew transitivity alternations. *Glossa: a journal of general linguistics* 11(1). pp. 1–41. DOI: https://doi.org/10.16995/glossa.27491

Grammatical generalizations and item-specific listing in Hebrew transitivity alternations

Yuval Z. Katz, Department of English Literature and Linguistics, Bar-Ilan University; Language and Brain Lab, Tel Aviv University, yuvalzkatz@gmail.com

Hebrew transitivity alternations are typically marked by a contrast in verbal template (*binyan*), raising the question of whether speakers generalize these template alternations through synchronic grammatical mechanisms or rely on stored, item-specific listing. This study presents two nonce verb experiments, production and acceptability judgment, designed to disentangle grammatical generalizations from item-specific listing in the causative alternation. Results show that speakers systematically avoid generating transitive forms in templates that do not occur as transitives in Hebrew, *niCCaC* and *hitCaCeC*, indicating that such restrictions are encoded in the grammar. Speakers also demonstrate productive use of the phonological stop-fricative alternation rule (spirantization) in the morphophonological environments tested. Additionally, speakers match the frequency of alternating transitive templates in Hebrew in their production. However, speakers' choice of specific transitive templates does not reliably follow from the intransitive template, does not align with the distribution of attested alternations, and shows no detectable sensitivity to manner-of-articulation faithfulness. Instead, speakers produce and accept a wide range of template pairings, including rare or unattested combinations. These findings challenge theories that posit systematic morphosyntactic mechanisms for template realization, specifically accounts in which verbal templates are taken to be realizations of a *Voice* head and the alternating templates to be structurally related through layering. Rather, the results support an account in which alternants are systematically related to each other at the level of argument structure and event semantics, while the morphological realization of each of them is listed independently, alongside a grammatical constraint on the transitivity of *niCCaC* and *hitCaCeC*.

1 Introduction

1.1 Hebrew verb templates

Hebrew verbs appear in one of five distinct morphological templates (and two additional passive templates), known as *binyanim* (singular *binyan*). These non-concatenative templates include vowels and consonants, and they combine with a root that is unpronounceable on its own.[1] In **Table 1**, five Hebrew verbs, one in each template, are presented together with their root and template.

Verb	Root	Template
KaTaV ('write')	K-T-V	*CaCaC*
SiXeK ('play')	S-X-K	*CiCeC*
hiTRiL ('troll')	T-R-L	*hiCCiC*
hitRaKeZ ('focus')	R-K-Z	*hitCaCeC*
niRDaM ('fall asleep')	R-D-M	*niCCaC*

Table 1: Examples of Hebrew verbs, their root, and their template. Throughout this paper, templates are represented using a notation that includes the fixed consonants and vowels of the template in lower case, with the root consonants (C) embedded between them in capital letters. Verbs are presented in the Hebrew citation form (third person masculine singular).

Theoretical accounts of the Hebrew template system face a tension. On the one hand, some theoretical accounts suggest that the selection of a verbal template for a given verb is principled and is determined by grammatical mechanisms (see Section 1.3). On the other hand, theoretical formulations of these proposed mechanisms frequently must concede that exceptions are very frequent, and sometimes they even outnumber instances in the lexicon that follow the rule. This paper contributes to resolving this tension in the domain of transitivity alternations (introduced in Section 1.2) using experimental methods. Using nonce verbs built on novel roots, I aim to tap into speakers' internalized knowledge of the system beyond the limitations of the attested lexicon.

1.2 Transitivity alternations in Hebrew

Transitivity alternations in Hebrew, and specifically the causative alternation,[2] the focus of this paper, are typically marked by a morphological difference in template between two alternants

[1] This is a descriptive characterization. The existence of an independent root morpheme in Semitic and its status is debated; see Bat-El (1994), Hever and Faust (2010), Faust (2019).

[2] I use the term *causative alternation* pre-theoretically to refer to any alternation in which one alternant denotes the causing of the other in some sense. Several researchers have noted that this phenomenon is not homogeneous

that share a consonantal root. Sentence (1) provides an example of a transitivity alternation. In (1a) the root D-L-K is in the *niCCaC* template, resulting in an intransitive verb.[3] In (1b), the same root is in the *hiCCiC* template, resulting in a transitive verb.

(1) a. ha-maxšev **niDLaK.**
the-computer turned-on$_{\text{INTR}}$

b. Dan **hiDLiK** et ha-maxšev.
Dan turned-on$_{\text{TR}}$ ACC the-computer

Whereas the contrast in morphological template is what marks the alternation, the specific pairs of templates vary. Example (2) shows an alternation in a different pair of templates.

(2) a. ha-oxel **hitXaMeM.**
the-food warmed$_{\text{INTR}}$

b. Dan **XiMeM** et ha-oxel.
Dan warmed$_{\text{TR}}$ ACC the-food

In some cases, predominantly in the *hiCCiC* template, but also in *CaCaC*, there is no morphological contrast between transitives and intransitives as shown in (3) (Borer 1991; Doron 2003; Lev 2016; Kastner 2020):

(3) a. ha-sveder **hiLBiN.**
the-sweater whitened$_{\text{INTR}}$

b. ha-xomer **hiLBiN** et ha-sveder.
the-substance whitened$_{\text{TR}}$ ACC the-sweater

1.3 Regularities and idiosyncrasies within the template system

The synchronic lexicon of Hebrew includes verbs in all five templates. For most lexical concepts in a given argument structure, there is typically a single, exclusive pairing of a root and a

---

and have identified multiple distinct alternations differing in lexical semantics, argument-structure properties (in particular the properties of the external argument), and morphological realization, including template selection (Levin & Rappaport Hovav 1995; Doron 2003; Reinhart 2003; Horvath & Siloni 2011; Kastner 2020). It is generally accepted that alternating unaccusatives differ systematically from alternating unergatives/transitives, and that the latter alternation is cross-linguistically rarer (Horvath & Siloni 2011; Laks 2011). In the stimuli of the current study, the subject of the intransitive alternant was animate and the causer (the external argument of the transitive) was inanimate. Whether particular templates are better suited to such arguments, and whether speakers consider properties of arguments such as animacy and agentivity when selecting templates in real time, are questions left to be experimentally investigated in future work. Because the items in the experiments (Sections 2.1.1.2, 2.2.1.2) could be interpreted as instances of either alternation type, I abstract away from the distinction between them.

[3] Throughout this paper I mainly deal with alternations in which one member is intransitive and the other is transitive, and therefore these are the labels I use. There are alternations in which the non-causative is transitive and the causative is ditransitive (e.g., AXaL-he'EXiL 'eat'-'feed').

template, such that speakers generally do not have multiple competing options. For example, although transitive verbs are commonly in *CiCeC* (e.g., *XiMeM*, ‘warm$_{TR}$’) or in *hiCCiC* (e.g., *hiXNiS*, ‘insert$_{TR}$’), the transitive verb *SaGaR* (‘close$_{TR}$’) is exclusively in *CaCaC*. Using the root S-G-R in *CiCeC* results in a nonword **SiGeR*. Using the root S-G-R in *hiCCiC* results in a different lexical concept, *hiSGiR* (‘extradite’, ‘hand over’).

Several regularities have been proposed to underlie the organization of the template system of Hebrew. These include semantic properties associated with templates (for example, *hiCCiC* as encoding causation), as well as argument structure properties (for example, *hitCaCeC* as intransitive). In addition, it has been proposed that certain template combinations form permissible combinations while others do not (e.g., *CaCaC* intransitives tend to alternate with *hiCCiC* and rarely with *CiCeC*). I elaborate on semantic and argument structure regularities in Section 1.4 when discussing specific theories of verbal templates.

Morphophonological regularities have also been identified in the verbal template system, such as the effect of the stop-fricative alternation on verbal template selection. In Hebrew, the consonants /p/, /b/, and /k/ surface as the fricatives [f], [v], and [x] in certain morphophonological environments (*spirantization*). Studies have shown that this alternation is productive in some environments but that there is variation in other environments, particularly in post-consonantal positions with some consonants (Adam 2002, for spontaneous speech data and analysis, Albert 2014, for elicited production, Temkin-Martínez 2010, for acceptability judgment, Temkin-Martínez 2013, and Temkin-Martínez & Müllner 2016, for elicited production of verbs and nonce-verbs).

The focus of the current discussion is not the stop-fricative alternation itself, but rather its role in template selection in transitivity alternations. Laks (2009) observed that alternating pairs tend to preserve paradigm uniformity with respect to the stop-fricative alternation. According to Laks, when the first root consonant is an underlying /p/, /b/, or /k/, the transitive-intransitive alternation is more likely to preserve the same manner of articulation between its members. For example, the intransitive verb derived from the root P-R-D appears in *niCCaC* (*niFRaD* ‘separate$_{INTR}$’), and the causative alternant appears in *hiCCiC* (*hiFRiD*, ‘separate$_{TR}$’), since in both alternants the consonant /p/ surfaces as the fricative [f]. This is in contrast to the unattested alternation for this root between *CaCaC* (*PaRaD) or *hitCaCeC* (*hitPaReD*) and *hiCCiC* (*hiFRiD*), which would result in a non-uniform realization of the consonant /p/ in the intransitive and the transitive. According to Laks (2009), uniform paradigms are more frequent than non-uniform paradigms.

Taken together, these observations reveal that Hebrew alternating verbs exhibit regularities in multiple domains: semantic, syntactic, and morphophonological. Each regularity found in the template system could in theory reflect speakers’ grammatical knowledge. However, finding regularities in the corpus does not necessarily mean that they reflect an underlying grammatical mechanism. Alternatively, regularities could also result from diachronic pressures that have shaped the language but are no longer part of the grammatical knowledge of individual speakers, from statistical tendencies that speakers may pick up but are not necessarily part of their grammar, or even from coincidence. In these cases, if there are no grammatical mechanisms that govern the selection of verbal template, we will have to concede that speakers must store the template for specific verbal instantiations even if there are some sub-regularities in the system (Pinker 1999; Jackendoff 2002).

Since Hebrew verbs constitute a relatively small and closed class, and since real verbs generally show an exclusive root-template pairing for each concept, speakers' judgments about existing words cannot straightforwardly reveal the underlying grammar. Lexical regularities may be suggestive, but given the limited lexicon, it is difficult to determine whether they reflect synchronic knowledge. For this reason, the present study uses nonce verbs.

When speakers apply a mechanism consistently to nonce verbs, this provides evidence that a grammatical mechanism (a rule or a constraint) provides them with specific ways to generate verbs and limits other options (Berko 1958; Pinker 1999). When speakers' choices on nonce verbs vary, by contrast, this indicates that no such grammatical mechanism is available to them, and, for real words, that the pairing between the root and the template must be listed in long-term memory. Listing is essential in this case, even if speakers are able to pick up some distributional patterns in the lexicon, since such generalizations are inherently partial: they hold over many lexical pairs but do not fully determine any individual one. As Jackendoff (2002:36) puts it:

> "…the outputs of semiproductive rules must be listed (at least in part) in long term memory; they cannot be a product of free combination. […] Such a rule expresses a generalization in the relationship among pairs of lexical entries. But it does not apply to all entries that satisfy its conditions, nor does it fully predict the information in one of these entries, given the other."

In the present study, using nonce verbs allows us to evaluate proposed accounts of the Hebrew verbal system and to constrain the hypothesis space for future theories. In Section 1.4, I turn to specific theories of Hebrew verbal templates and their predictions regarding abstract grammatical knowledge and item-specific listing.

1.4 Theoretical accounts for Hebrew verbal templates and their predictions

Theoretical accounts of Hebrew verbal templates aim to capture semantic, syntactic, and morphological regularities in the system. In this section, I describe theories that address how the verbs' syntactic-semantic properties relate to their morphological template, and their accounts for the causative alternation. I focus the discussion on syntactic decompositional frameworks that assume that the root composes with functional structure in the syntax to form the verb. Although none of these accounts provide explicit predictions for nonce verbs, predictions are derived based on claims for grammatical and listed information within the proposed theories.

1.4.1 Relations between templates

In approaches that decompose the verb in the syntax, transitivity alternations are commonly derived through layering (Alexiadou et al. 2015). On this view, alternating verbs are not derived from one another (that is, the transitive is not derived from the intransitive, nor vice versa). Rather, the two alternants share the same core structure and differ only in higher structural layers associated with voice and the external argument. In these theories, alternations are therefore not a grammatical primitive or a rule linking two templates, but an epiphenomenon of two independently derived structures that share a common core. I use the term *alternation* descriptively throughout, to refer to pairs of related intransitive and transitive verbs, without implying that one alternant is derived from the other or that the two are linked by a grammatical rule. The predictions regarding relations between templates are derived from claims about templates that share a core structure but minimally differ in their higher layers, the same mechanism these approaches use to explain why some template combinations alternate and others do not.

Two of the most influential approaches that assume layering to explain Hebrew verbal template alternations are Doron (2003) and Kastner (2020).

In Doron (2003) *CaCaC*, *CiCeC*, and *hiCCiC* each correspond to a syntactic configuration specified for the presence or absence of the two agency heads, the "causative" head γ, and the "intensive" head ι ([$-\gamma$, $-\iota$], [$-\gamma$, $+\iota$] and [$+\gamma$, $-\iota$] respectively). These heads Merge above the root node and constitute part of the verb's syntax. *CaCaC* ([$-\gamma$, $-\iota$]) is termed "the simple template", *CiCeC* ([$-\gamma$, $+\iota$]) is termed "the intensive template", and *hiCCiC* ([$+\gamma$, $-\iota$]) is termed "the causative template". Crucially, *niCCaC* and *hitCaCeC* are derived by layering the "middle" voice head μ on top of *CaCaC* and *CiCeC* respectively, and *hiCCiC* is derived by layering the "causative" head γ on top of *CaCaC*. Therefore, for Doron, the template pairs structurally related through layering are *niCCaC-CaCaC*, *hitCaCeC-CiCeC*, *CaCaC-hiCCiC*, and *CaCaC-CiCeC* (the intensive alternation, e.g., *ŠaLaX* 'send' *ŠiLeaX* 'send away').

Kastner (2020) states a similar layering claim in different theoretical terms. He proposes that templates are phonological realizations of configurations that include different values of the EPP feature [D] on the *Voice* head, which determines the presence or absence of an external argument in the specifier of *Voice*. When *Voice* carries a [+D] feature, the verb is in the *hiCCiC* template; when it carries a [–D] feature, the verb is in either *niCCaC* or *hitCaCeC*; and when it is unspecified, it is in either *CaCaC* or *CiCeC*. *hitCaCeC* and *CiCeC* minimally differ from *niCCaC* and *CaCaC* since they include the modifier $\sqrt{\text{ACTION}}$ on top of *v*. For Kastner, the structurally related template pairs are therefore *niCCaC-CaCaC, hitCaCeC-CiCeC*, *CaCaC-hiCCiC* and *niCCaC-hiCCiC*. For both Doron (2003) and Kastner (2020), since layering is the grammatical mechanism for relating these patterns, it is predicted that speakers will prefer pairs of templates that share the same syntactic core but differ in their specification of *Voice*[4]: presence of the middle head μ

[4] While Kastner (2020) derives all alternations by layering of Voice, for Doron (2003) the *CaCaC-hiCCiC* alternation is derived by layering of the causative head γ, which she calls an agency head.

for Doron, and specification of the [D] features for Kastner. This yields the following structurally related pairs: *niCCaC-CaCaC*, *hitCaCeC-CiCeC*, *CaCaC-hiCCiC, niCCaC-hiCCiC* (Kastner), and *CaCaC-CiCeC* (Doron, in the intensive alternation).

Unlike Doron and Kastner, Arad (2005) puts more weight on the idiosyncrasy of the template system. For Arad, the template is a realization of the verbalizer head *v*, and in non-alternating verbs a specific template is specified under this head. She distinguishes between the syntax of alternating and non-alternating verbs, and claims that in the causative alternation, the root selects a conjugation class consisting of a pair of templates (e.g., *CaCaC-hiCCiC*). The correct template for the syntactic context is then selected based on the “flavor” of *v* (inchoative or causative). Arad proposes that there are six conjugation classes based on observed template pairs in alternating verbs: *niCCaC-CaCaC* (e.g., *niFTaX-PaTaX* – ‘open’), *CaCaC-hiCCiC* (*KaFaʔ-hiKPiʔ* – ‘freeze’), *niCCaC-hiCCiC* (*niRGaʕ-hiRGiʕa* – ‘relax’), *hitCaCeC-CiCeC* (*hitXaMeM-XiMeM* – ‘warm’), *hitCaCeC-hiCCiC* (*hitBaHeR-hiVHiR* – ‘clear’), and *hiCCiC-hiCCiC* (*hiLBiN-hiLBiN* – ‘whiten’). Additionally, four less-frequent template pairs are observed in Hebrew, not mentioned by Arad as conjugation classes: *CaCaC-CiCeC* (*SaMaX-SiMe’aX* – ‘be happy’, ‘cheer up’), *niCCaC-CiCeC* (*niRPaʔ-RiPeʔ* – ‘cure’), *CaCaC-CaCaC* (*ACaR-ACaR* – ‘stop’), and *CiCeC-hiCCiC* (*ZiNeK-hiZNiK* – ‘jump away’). Since conjugation classes lexically specify the possible template combinations for the grammar, it is predicted that speakers will prefer alternations that form a conjugation class over those that do not, i.e., that they will prefer alternations attested in the language over novel combinations.

Borer (2004) goes even further than Arad (2005) in highlighting the idiosyncrasies of the system, proposing that each root is listed with a specific template, thereby predicting an even less constrained selection of templates by speakers. She also argues that templates themselves encode transitivity, a claim I discuss in Section 1.4.2.

1.4.2 Transitivity

A recurring theme in theories of Hebrew verbal templates is the extent to which they are tied to transitivity. Borer (2004), discussing acquisition studies of Berman (1982; 1993; 1994), proposes that there are morphosyntactic rules that relate templates with the transitivity of the syntactic structure manifested in the [OM] feature denoting Object Marking, the ability to assign accusative case. In Borer’s system *hiCCiC* and *CiCeC* are realizations of [+OM], *niCCaC* and *hitCaCeC* are realizations of [–OM], and *CaCaC* is unspecified for transitivity. Since templates are listed for specific roots in Borer’s system, as discussed in Section 1.4.1, the [OM] feature does not directly determine the choice of template, but, according to Borer, may assist in lexical access by “delimiting the search” for the appropriate template for a specific root.

Similarly, Arad (2005) proposes that *niCCaC* and *hitCaCeC* are marked as [–tr] or [–ACC case], which confines their environment of insertion to intransitive verbs. This generates the correct pattern in the lexicon where all templates can be used in intransitive contexts, but normally only *CaCaC, CiCeC,* and *hiCCiC* can be used in transitive contexts.

In Kastner (2020) the templates' relation to transitivity is through *Voice*, which determines the presence of an external argument, and, through Burzio's (1986) generalization, also determines the verb's ability to assign accusative case. Therefore, the prediction is that *hiCCiC* ([+D]) will have an external argument and will be able to assign accusative case, whereas *niCCaC* and *hitCaCeC* ([–D]) will not have an external argument and will not be able to assign accusative case. *CaCaC* is unspecified for [D] and therefore also for transitivity. Similarly, for Doron (2003), the middle head μ blocks an external argument, and therefore also transitivity for the middle templates *niCCaC* and *hitCaCeC*.

Since all of these theories associate verbal templates with transitivity, the prediction is that whatever the speaker knows about this association should surface in their preference for template in nonce verbs. Specifically, speakers will rule out transitives in *niCCaC* and *hitCaCeC*. Additionally, they might prefer transitives in *CiCeC* and *hiCCiC* over transitives in *CaCaC*, since *CiCeC* and *hiCCiC* are transitive whereas *CaCaC* is unspecified in Doron (2003), Borer (2004), and Kastner (2020).

Having reviewed the main theoretical proposals, I briefly outline what the lexicon of Hebrew actually shows about the relationship between transitivity and verbal templates. Although all alternating transitives are in either *CiCeC*, *hiCCiC*, or *CaCaC*, the opposite implication is not true. That is, there are verbs in *CiCeC* and *hiCCiC* that are not transitive (e.g., *hiGiʕa* – 'arrive', *PiZeZ* – 'prance'), and at least one alternating verb in *CiCeC* that is not transitive (*ZiNeK* – 'jump away' is intransitive and alternates with hiZNiK). *hiCCiC-hiCCiC* alternations (*hiLBiN* – 'whiten') are also of this sort (although Borer 1991 claims that in these cases each alternant is independently derived from the corresponding adjective). Additionally, there are cases in which both alternants are syntactically transitive and assign accusative case, e.g., *LaVaŠ-hiLBiŠ* ('wear'-'dress'). In sum, it seems that the only generalization about the relation between transitivity and template is that *niCCaC* and *hitCaCeC* are not transitive.

The experiments presented in this paper investigate which aspects of template selection in the causative alternation are based on grammatical generalizations and which ones rely on listed idiosyncratic information about specific items. I examine this question using alternating nonce verbs in production and judgment, with the rationale that grammatical generalizations will be applied to nonce verbs, whereas listed information will be manifested as variability in template selection. Before turning to the present study, I briefly review previous work that has used nonce verbs in Hebrew.

1.5 Previous research on Semitic morphology using nonce verbs

Although questions concerning transitivity alternations in Hebrew have not previously been addressed experimentally using a nonword paradigm, the method of using nonwords to reveal speakers' knowledge (or lack thereof) about certain linguistic generalizations has been previously used for other phenomena in other languages, including Hebrew and other Semitic languages.

Bolozky conducted a series of studies over many years, using both judgment and production tasks to investigate how speakers choose verbal templates (Bolozky 1999). His experiments examined how speakers derive denominal verbs from existing nouns. The findings from his studies conducted during the 1990s show that *CiCeC* is the most common template for deriving causative denominal verbs, and that *hiCCiC* is used primarily in cases where the consonant cluster of the noun is preserved (see also Bat-El 1994).

Laks (2014b) studied passive formation in Hebrew, which comprises two types: In one type, active verbs in *CiCeC* and *hiCCiC* undergo melodic overwriting to *CuCaC* and *huCCaC* respectively (e.g., *ŠiLeV* – 'integrate', *ŠuLaV* – 'was integrated'; *hiSBiR* – 'explain', *huSBaR* – 'was explained'). This type of passive is regular in that almost every transitive in *CiCeC* and *hiCCiC* has a corresponding passive. The second type of passive, derived from transitives in *CaCaC*, is marked by the verbal template *niCCaC*, which is not unique to passives.

Laks used a written nonce verb sentence completion task and found that speakers generalize the two passivization operations differently: while they hardly show any variation when forming passives in *CiCeC* and *hiCCiC*, which are invariably generated by melodic overwriting, there was greater variability in the results when producing passives for verbs in *CaCaC*, which do not always have a predictable passive alternant. Specifically, in 59% of the cases, speakers selected *CuCaC* (which rarely alternates with *CaCaC* in Hebrew), and only in 30% of the cases did they select *niCCaC*, which is usually the passive alternant of *CaCaC*, when a passive form exists.

Laks explained this result as a difference between lexical and syntactic phenomena: *CuCaC* and *huCCaC* passives are formed by a categorical grammatical rule, and therefore the alternation is regular. In contrast, *niCCaC* passives are formed in the lexicon, and therefore their distribution in the language is less regular, leading to greater variability among speakers when generating them in the context of nonce verbs. For the present context, without committing to where in the grammar each mechanism is located, two mechanisms appear to be at work: a grammatical mechanism that applies consistently to both words and nonwords, and another mechanism that applies unpredictably to specific words and not others, and is therefore rarely extended by speakers to nonwords.

Moore-Cantwell (2013), also using nonwords, tested denominal verb formation in the *CiCeC* template in Hebrew, which shows variability in morphological structure when the nominal has only two consonants, e.g., *dam* → *dimem* ('blood'-'to bleed'), *kod* → *koded* ('code'-'to code'), *tik* → *tijek* ('file'- 'to file'), *daf* → *difdef* ('paper'-'to turn pages'). She found that speakers are sensitive to output frequency, as overall they tend to use consonant doubling, the most common strategy in the lexicon, as in *dam* → *dimem* ('blood'-'to bleed'), rather than other available strategies: vowel overwriting, coronal glide formation, labial glide formation, or reduplication. However, speakers do not learn conditional probabilities, as they do not replicate the relations between the morphological properties of nouns and denominal verbs in the corpus.

With respect to morphophonology, Temkin-Martínez (2013) and Temkin-Martínez and Müllner (2016) tested variability in speakers' adherence to the stop-fricative alternation using nonwords. They found that the morphophonological environment in which the consonant appears affects its susceptibility to underapplication of the alternation. Importantly for the current study, underapplication was more frequent in post-consonantal positions than in post-vocalic and word-initial positions, replicating findings from Temkin-Martínez (2013) on real verbs and providing experimental support for the generalization proposed by Adam (2002).

Taken together, previous studies on Hebrew show that not every regularity attested in the lexicon reflects an active synchronic grammatical mechanism, and that nonce verb experiments can reveal which regularities are due to a grammatical generalization and which are not and are therefore dependent on listed information. The finding in Laks (2014b) on passive formation illustrates this clearly: the *CuCaC/huCCaC* passives behave uniformly, as expected of the output of a productive grammatical generalization, whereas the *niCCaC* passives, whose distribution is lexically determined, elicit far more variability. Similarly, Moore-Cantwell's results demonstrate that although speakers are sensitive to overall statistical distributions, they do not acquire the conditional dependencies that characterize the denominal system in the corpus, meaning that information about the derivation strategy in this case must be individually listed for specific roots.

2 Experiments

To test speakers' generalizations about the template system I conducted a type of wug test (Berko 1958) using alternating nonce verbs in various templates, in both a production task (Experiment 1, Section 2.1) and a judgment task (Experiment 2, Section 2.2).

2.1 Experiment 1: production

2.1.1 Methods

2.1.1.1 Participants

The participants were 25 Hebrew-speaking adults, recruited via Facebook and tested online through Zoom. Participants received 20 ILS for their participation.

2.1.1.2 Procedure

Participants were presented with sentences containing nonce verbs (e.g., *PaNaG*), read by the experimenter and also presented on the screen, and were asked to complete them aloud. Sentences were designed to elicit the transitive alternant of the nonce verb. (4) is an example of a test item.

(4) Dan *PaNaG* etmol. ha-kerax garam le-ze. klomar, ha-kerax _______ et Dan.
Dan *PaNaG* yesterday. The-ice caused to-it. That is, the-ice _______ ACC Dan
'Dan *PaNaG* yesterday. The ice caused it. That is, the ice _______ Dan.'

Each participant was presented with a total of 60 sentences, consisting of 48 test items and 12 filler items. The 48 items contained nonce verbs in four binyanim, 12 each: *CaCaC*, *niCCaC*, *hitCaCeC*, and *hiCCiC*. I did not include nonce verbs in *CiCeC* in the test items, since occurrences of intransitives in *CiCeC* participating in a transitivity alternation are very rare. (There is only one occurrence in Laks' (2011; 2014a) inventory of alternating verbs: *ZiNeK-hiZNiK* – 'jump away'.) To minimize confounds from potential similarity to real verbs, I generated a randomized list of test items for each participant. For each item, I randomized a non-root from a pool of non-roots generated in advance that do not form real words in any verbal or nominal template. The nonce verbs were designed to be phonologically licit in Hebrew by avoiding roots with identical first and second radicals. Since the pool of roots was smaller than the total number of items, some roots were repeated across and within participants. Additionally, I randomized an argument for the nonce verb in the intransitive form (e.g., Dan, from a pool of 64 human proper names) and an external argument for the causative event (e.g., the ice, from a pool of 64 inanimate nouns). Finally, the order of the items in each list was also randomized for each participant to cancel ordering effects.

To test the effects of stop-fricative alternations on template selection, half of the items in each template included roots whose first consonant is /p/ or /b/, which surface as [f] and [v] in post-vocalic positions. The other half included roots whose first consonant is /g/ or /l/, which do not participate in the stop-fricative alternation.

The filler items also included nonce verbs, all intransitives in *CiCeC*, to balance the proportion of templates across the experiment (recall that the nonce verbs in the test items were never in *CiCeC*). In the filler items the participants were expected to change the inflection of the nonce verb without changing its template or valency. Fillers were not included in the analysis. (5) is an example of a filler item.

(5)	ha-klav-im	SiBŠ**u**,	ve-az	gam	ha-xatul-a	__.	(Target response: SiBŠ**a**)
	the-dog-PL.M	SiBŠu	and-then	also	the-cat-SG.F	__	

All responses were transcribed and coded for template. When participants had multiple responses, they were asked to choose one; if they did not, their last response was considered. Cases in which the first root consonant /b/ or /p/ was pronounced as a stop post-vocalically (e.g., *hiPNiG instead of hiFNiG), or as a fricative at the beginning of the word (e.g., *FaNaG instead of PaNaG) were marked as *template-manner mismatch*.

2.1.1.3 Analysis

Data were analyzed using mixed-effects regression models in R (R Core Team 2023). For all models, I fit the maximal random-effects structure justified by the design and simplified only when the model failed to converge following the recommendations of Barr et al. (2013).

To evaluate preferences among transitive templates, I fit a multinomial mixed-effects logistic regression model (*mclogit* package; Elff 2026) with Transitive (the transitive template) as the outcome and Intransitive (the intransitive template) as a fixed effect, with by-participant random intercepts and slopes for Intransitive. This model addresses both the overall ranking of transitive templates (averaged across intransitive template) and whether the intransitive template modulates the choice of transitive template. I did not include *niCCaC* and *hitCaCeC* as transitive templates since there were virtually no transitive responses with these templates. Marginal estimates of overall template preferences and pairwise comparisons with Tukey correction were obtained using the *emmeans* package (Lenth 2025). In addition to the model-based analyses, I report by-participant descriptive patterns to characterize variability across individuals.

To test whether speakers select potentially faithful templates when the first radical is /b/ or /p/, I fit a binomial mixed-effects logistic regression model using the lme4 package (Bates et al. 2015) predicting whether the participant selected a potentially faithful transitive template (coded as 1) or an unfaithful one (coded as 0). A response was coded as *potentially faithful* if it preserved the manner of articulation of the first radical when the radical was /b/ or /p/. Specifically, in the inflections tested, first-radical /b/ and /p/ surface as stops in *hitCaCeC*, *CaCaC*, and *CiCeC*, but as fricatives in *niCCaC* and *hiCCiC*. Consequently, targets were considered potentially faithful if they belonged to the same stop-realizing group or the same fricative-realizing group as the intransitive. The predictors were the intransitive template (Intransitive), whether the root-initial consonant participates in the stop-fricative alternation (BP), and their interaction. The model included by-participant random intercepts and slopes for BP, Intransitive, and their interaction. Model comparisons were performed using Type II Wald chi-square tests (via the car package, Fox & Weisberg 2019), and fixed effects were interpreted based on their estimated coefficients, standard errors, z-values, and associated *p*-values. Pairwise comparisons were conducted using the emmeans package with Tukey correction.

To examine whether there is a relation between the distribution of responses in the experiment and the distribution of templates in verbal alternations in the language, I compared participants' preference for transitive template (overall and across intransitive templates) with an inventory of 443 Hebrew alternating verbs collected by Laks (2011; 2014a).

2.1.2 Results

2.1.2.1 Template preference

All 25 participants produced well-formed Hebrew verbs on the majority of trials and answered the majority of filler items correctly (correct agreement and verbal template). Accordingly, none were excluded. At the item level, 27 responses were excluded since they were not in one of the five templates in Hebrew (e.g., GoCiV), the root differed from the root in the stimulus intransitive verb, or the participant did not provide a response. The percentages of transitive responses in each template for each intransitive nonce verb in each template are presented in **Table 2**.

As summarized in **Table 2**, there was variability in the selection of transitive templates across each of the conditions, and there was no case in which the selected template was consistent across participants and sentences. The most common response template for all four conditions was *CiCeC*, which accounted for 45% to 66% of responses for all four intransitive templates.

		Response (transitive)				
		CaCaC	***CiCeC***	***hiCCiC***	***hitCaCeC***	***niCCaC***
Stimulus (intransitive)	*CaCaC*	4%	64%	32%	0%	0%
	niCCaC	32%	45%	23%	0%	0%
	hitCaCeC	9%	66%	25%	0%	0%
	hiCCiC	24%	50%	26%	0%	0%

Table 2: Rate of response (transitive) templates by stimulus (intransitive) template.

Overall, *CiCeC* was the most frequent response with 56% of responses, followed by *hiCCiC* with 26%, and *CaCaC* with 17%. Among the 45% of *CiCeC* responses to intransitive *niCCaC* items, about a quarter (11% of all responses) were cases in which the *ni-* prefix was reanalyzed by participants as a root consonant, as evidenced by the vowel template characteristic of *CiCeC* with quadriliteral roots. For example, for an intransitive *niLTaG* (the triliteral root LTG in *niCCaC*), a participant might produce the transitive *NiLTeG*, apparently reanalyzing the initial /n/ of *niCCaC* as part of a quadriliteral root *NLTG* in *CiCeC*, analogous to the real verb *NiRMeL* 'normalize'.

Speakers' preferences among the transitive templates were assessed with the model described in Section 2.1.1.3. The model's output is reported in **Table 3**, and overall template preferences (averaged over Intransitive) and pairwise comparisons with Tukey correction are reported in **Table 4**. *CiCeC* was preferred over both *hiCCiC* ($p < .0001$) and *CaCaC* ($p < .0001$). The contrast between *hiCCiC* and *CaCaC* was not significant ($p = .27$). This is illustrated in **Figure 1**. At the level of individual participants, *CiCeC* was the most common response for 20 of the 25 participants. Of those, 13 exhibited a *CiCeC* > *hiCCiC* > *CaCaC* preference. That is, the dominance of *CiCeC* held for most individual participants, but the relative ranking of *hiCCiC* and *CaCaC* varied. The only invariable preference observed in the data was the virtual avoidance of *niCCaC* and *hitCaCeC* as transitive alternants, with only one response each.

	Transitive: *CaCaC* (vs. *hiCCiC*)			**Transitive: *CiCeC* (vs. *hiCCiC*)**		
Predictors	***Odds Ratios***	***CI***	***p***	***Odds Ratios***	***CI***	***p***
Intercept (Intransitive *CaCaC*)	0.13	0.05–0.30	<0.001	2.43	1.29–4.55	0.006
Intransitive *hiC*CiC	6.55	1.62–26.51	0.008	1.37	0.46–4.09	0.573
Intransitive *hit-CaCeC*	1.15	0.32–4.11	0.835	1.44	0.70–2.96	0.324
Intransitive *niCCaC*	10.50	3.89–28.34	<0.001	1.03	0.55–1.92	0.919
$N_{\text{Participant}}$	25					
Observations	1171					

Table 3: Multinomial mixed-effects logistic regression model predicting choice of transitive template (Transitive ~ Intransitive + (1+ Intransitive | Participant)). Estimates are odds ratios. The reference level is *hiCCiC* for the transitive template and *CaCaC* for the intransitive template.

Contrast	Estimate	SE	z	p
hiCCiC − *CaCaC*	0.10	0.06	1.54	.274
hiCCiC − *CiCeC*	−0.43	0.09	−4.76	<.0001
CaCaC − *CiCeC*	−0.52	0.08	−6.99	<.0001

Table 4: Pairwise comparisons among transitive templates, averaged over the intransitive templates, with Tukey correction for multiple comparisons. Estimates are on the log-odds scale and given as the first template minus the second.

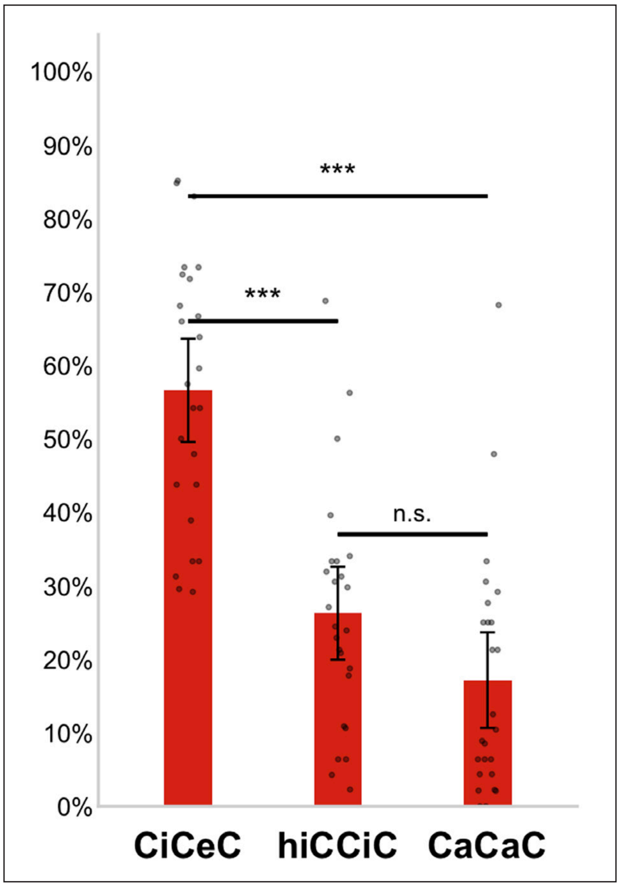

Figure 1: Mean proportion of trials on which each template was selected as the transitive form in Experiment 1. Bars show means across participants; error bars show 95% confidence intervals; dots show individual participant proportions. Significance brackets indicate pairwise differences from a multinomial mixed model (*** $p < .001$; n.s. = not significant after Tukey's correction).

The model also revealed that the intransitive template modulates the choice of transitive template. The contrast between *CiCeC* and *hiCCiC* was stable across intransitive templates, with no significant effect of Intransitive on this contrast (all *ps* $> .32$). The contrast between *hiCCiC* and *CaCaC* was modulated by Intransitive: the odds of selecting *CaCaC* over *hiCCiC* increased when the intransitive template was *niCCaC* (OR = 10.50, p < .001) or *hiCCiC* (OR = 6.55, p = .008); the shift at the *hitCaCeC* intransitive was not significant (OR = 1.15, p = .835). This is illustrated in **Figure 2**. Crucially, some significant shifts run counter to the inventory pattern, as presented in Section 2.1.2.3.

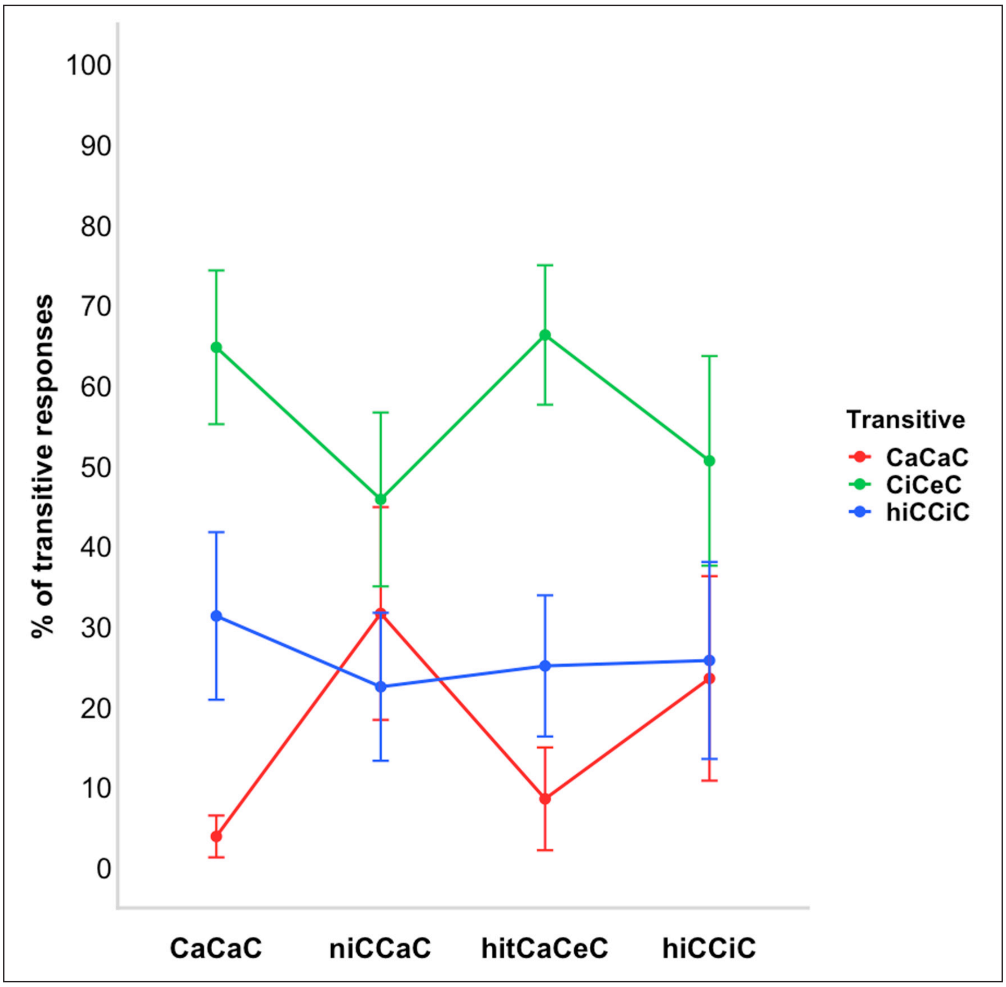

Figure 2: Mean proportion of trials on which each transitive template (*CiCeC*, *hiCCiC*, *CaCaC*) was selected in Experiment 1, plotted by intransitive template. Points show means across participants and error bars show 95% confidence intervals; lines connect the same transitive template across intransitive templates.

2.1.2.2 The effect of the stop-fricative alternation on choice of template

I compared nonce verbs with first root consonants /p/ or /b/ (which participate in the stop-fricative alternation) to nonce verbs with first root consonants /g/ or /l/ (which do not participate in the stop-fricative alternation). I examined whether speakers more frequently select "faithful" forms, templates in which the stop does not change into a fricative and vice versa, when the first root consonant is /p/ or /b/. For example, for the intransitives PaNaG and LaNaG, I investigated whether speakers prefer PiNeG over hiFNiG more often than they prefer LiNeG over hiLNiG, given that PiNeG is more faithful to PaNaG than hiFNiG, whereas there is no difference between LiNeG and hiLNiG with respect to manner of articulation.

The model revealed a significant main effect of Intransitive ($\chi^2(3) = 27.06$, $p < .001$). Fixed effect estimates indicated that "faithful" responses were significantly less likely when the intransitive template was *hiCCiC* or *niCCaC*, relative to the reference level *CaCaC*. This aligns with the observation that *hiCCiC* and *niCCaC* are the only intransitive templates in which the first radical surfaces as a fricative in relevant phonological environments, in contrast to the three other templates where the radical surfaces as a stop. Neither the main effect of BP ($\chi^2(1) = 0.95$, $p = .33$) nor the interaction between BP and Intransitive ($\chi^2(3) = 2.51$, $p = .47$) was significant. These results suggest that, within the current experimental paradigm, there is no evidence that the phonological alternation potential of the first radical (i.e., whether it participates in the stop-fricative alternation) affects the choice of transitive template, either across the board or in interaction with specific intransitive templates.[5] See **Table 5** for the full model output.

Although there is no evidence that speakers select templates based on the stop-fricative alternation, it seems that they know the morphophonological environments of the alternation: There was a very small number of responses with template-manner mismatch (9 out of 593 responses), which suggests that speakers know the grammatical mechanisms governing the distribution of stops and fricatives in the first radical, and can use them productively.

[5] A post-hoc power analysis using the *simr* package (Green & MacLeod 2016), based on a slightly simplified random-effects structure (Faithful ~ BP * Intransitive + (1 + BP | Participant)) due to numerical instability of the maximal structure on simulated datasets, indicated that the study was well-powered to detect a large BP effect (97% power for $\beta = 0.8$) but underpowered for small to moderate effects (10% power for $\beta = 0.2$; 49% for $\beta = 0.5$). Two considerations nonetheless support the interpretation of this null result. First, productive grammatical generalizations are expected to apply consistently, as, for instance, speakers' knowledge of the intransitivity of *niCCaC* and *hitCaCeC* in the current study, and speakers' use of *CuCaC* and *huCCaC* in Laks (2014b). Second, Experiment 2 tested the same prediction using acceptability judgments and a substantially larger dataset (Section 2.2), and the BP effect was again null. The convergence of null findings across two experiments provides stronger evidence against a grammatical BP effect on template selection, while leaving open the possibility of a small bias that neither experiment was designed to detect.

	Faithful		
Predictors	***Odds Ratios***	***CI***	***p***
(Intercept)	2.20	1.22–3.96	0.009
BP [1]	2.17	0.83–5.69	0.115
Intransitive [*hiCCi*C]	0.06	0.02–0.22	<0.001
Intransitive [*hitCaCeC*]	1.81	0.82–3.99	0.144
Intransitive [*niCCaC*]	0.17	0.06–0.44	<0.001
BP [1] × Intransitive [*hiCCiC*]	0.47	0.10–2.07	0.316
BP [1] × Intransitive [*hitCaCeC*]	0.63	0.18–2.21	0.467
BP [1] × Intransitive [*niCCaC*]	0.41	0.13–1.28	0.126
Random Effects			
$\tau_{00\ \text{Participant}}$	1.23		
$\tau_{11\ \text{Participant.BP1}}$	2.44		
$\tau_{11\ \text{Participant.Intransitive}\mathit{hiCCi}}$C	6.16		
$\tau_{11\ \text{Participant.Intransitive}\mathit{hitCaCeC}}$	1.59		
$\tau_{11\ \text{Participant.Intransitive}\mathit{niCCaC}}$	3.88		
$\tau_{11\ \text{Participant.BP1: Intransitive}\mathit{hiCCiC}}$	3.91		
$\tau_{11\ \text{Participant.BP1: Intransitive}\mathit{hitCaCeC}}$	3.02		
$\tau_{11\ \text{Participant.BP1: Intransitive}\mathit{niCCaC}}$	1.91		
$N_{\text{Participant}}$	25		
Observations	1171		

Table 5: Logistic mixed-effects model of faithful responses in Experiment 1. Model: Faithful ~ BP * Intransitive + (1 + BP * Intransitive | Participant). Reference levels: BP = 0, Intransitive = *CaCaC*. Estimates shown as odds ratios with 95% Wald confidence intervals.

2.1.2.3 Comparing the results to template distribution in an inventory of Hebrew alternating verbs

To examine whether there is a relation between the distribution of responses in the experiment and the distribution of templates in verbal alternations in the language, I compared the results with an inventory of 443 Hebrew alternating verbs collected by Laks (2011; 2014a).[6] Comparing the rate of transitive templates in the causative alternation overall, regardless of the intransitive

[6] Laks categorized causative alternations into two types: the decausative alternation, which forms unaccusatives and subject experiencers from transitives, and the causative alternation, which forms causatives from unergatives (Horvath & Siloni 2011). To avoid assuming any theoretical differences between these alternations, I combined them and calculated the percentage of transitives for each intransitive template.

template, the experiment's results align closely with the inventory (frequency matching, Ernestus & Baayen 2003; Hayes et al. 2009, among others). As **Figure 3** illustrates, *CiCeC* is the most prevalent transitive template in both the inventory and the experiment, followed by *hiCCiC*, with *CaCaC* being the least common. Neither *niCCaC* nor *hitCaCeC* is ever a transitive alternant in the inventory and are also virtually absent as transitive responses in the experiment.

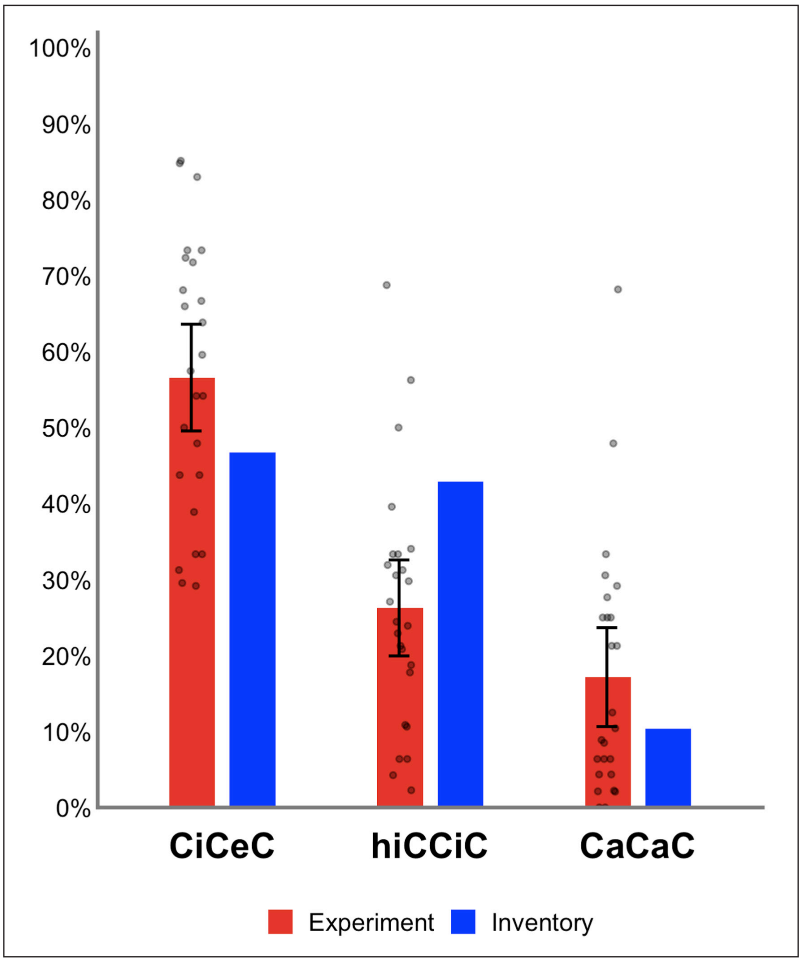

Figure 3: Mean proportion of trials on which each transitive template was selected in Experiment 1 (Experiment), compared to the distribution of transitive templates in the Hebrew verbal inventory (Inventory). Bars show means across participants; error bars show 95% confidence intervals; dots show individual participant proportions.

However, when comparing the conditional distribution of transitive templates given an intransitive template, a different picture emerges. As demonstrated in **Figure 4**, speakers do not seem to know which template pairs are frequent in Hebrew, and do not systematically consider the intransitive template when selecting a transitive template. As presented in Section 2.1.2.1, the intransitive form does not modulate speakers' preference for *CiCeC* over *hiCCiC*, even though in the inventory *CiCeC* is the most frequent transitive only for *hitCaCeC* intransitives. For instance, intransitives in *CaCaC* very infrequently alternate with *CiCeC* in the inventory, yet *CiCeC* was the predominant response in the experiment. Particularly striking is the case of intransitive *hiCCiC*, where the inventory shows that 100% of its transitive alternants are in *hiCCiC*, while the majority of responses in the experiment were in *CiCeC*.

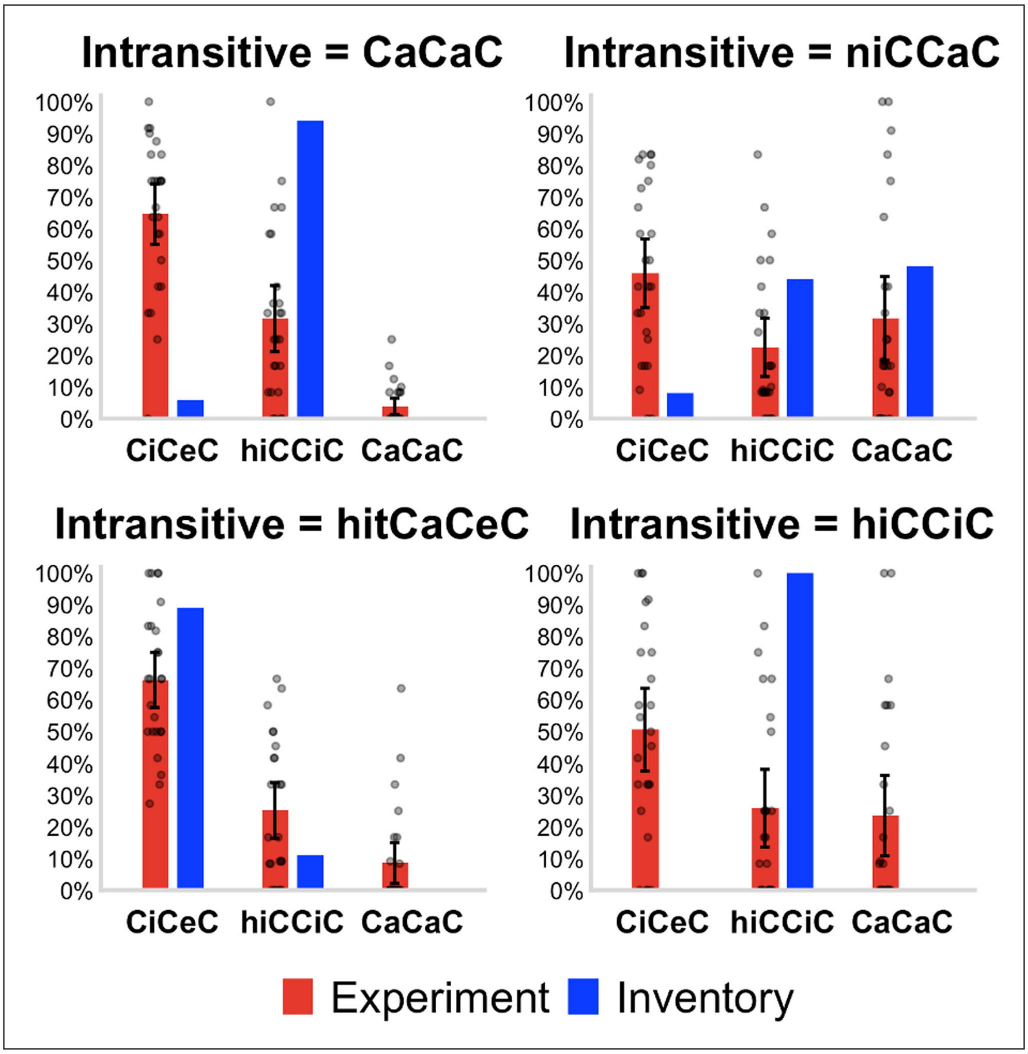

Figure 4: Mean proportion of trials on which each transitive template was selected in Experiment 1 (Experiment), broken down by intransitive template, compared to the distribution of transitive templates in the Hebrew verbal inventory (Inventory). Experiment bars show means across participants; error bars show 95% confidence intervals; dots show individual participant proportions. Inventory bars show corpus proportions from Laks (2011; 2014a).

The preference for *CaCaC* over *hiCCiC* was modulated by intransitive in the experiment, but not consistently following the pattern of the inventory. The shift toward *CaCaC* in *niCCaC* intransitives is compatible with the inventory, where *CaCaC* and *hiCCiC* alternate with *niCCaC* at roughly equal rates. The shift in *hiCCiC* intransitives, however, runs counter to the inventory: *hiCCiC* intransitives always alternate with *hiCCiC* transitives, yet speakers shifted away from *hiCCiC* toward *CaCaC*.

Together, these comparisons indicate that while speakers' aggregate template preferences resemble the inventory, their behavior in choosing a transitive template for a given intransitive template does not reflect the conditional distributions attested in the language.

2.2 Experiment 2: acceptability judgment

To further test the preferences revealed in production, I conducted an acceptability judgment task. Using acceptability judgment, where speakers evaluate the fit between two alternating nonce verbs rather than deriving one alternant from the other (as in Experiment 1), makes it possible to abstract away from assumptions about directional derivation (whether the transitive is derived from the intransitive or vice versa) and to focus instead on how well template combinations align with speakers' grammatical knowledge.

2.2.1 Methods

2.2.1.1 Participants

The participants were 75 Hebrew-speaking adults, recruited through Facebook. Participants received 10 ILS for completing the experiment.

2.2.1.2 Procedure

The experiment was conducted online using PCIbex (Zehr & Schwarz 2018). Participants were presented with a written sentence featuring a nonce verb and its causative alternant (in boldface) and were asked to rate how well the alternant "fits the sentence" on a scale from 1 to 6. The nonce verbs were designed to be phonologically licit in Hebrew by avoiding roots with identical first and second radicals.[7] An even-numbered scale was chosen to encourage participants to commit to a preference rather than defaulting to a neutral midpoint. An example of a test item

---

[7] Beyond avoiding identical first and second radicals, the stimuli were not otherwise controlled for OCP violations: 9 sets contained roots whose first two radicals share place of articulation. Although Yeverechyahu (2014) found that Hebrew speakers rate such nonwords lower, OCP-violating and non-violating sets had virtually identical means in the current experiment (M = 3.62 vs. 3.63), and a cumulative link mixed-effects model predicting rating from OCP status and its interactions with Intransitive and Transitive (same random-effects structure as the main model) showed no OCP effect ($\chi^2(1) = 0.95$, $p = .33$; all interactions ps > .40). This may reflect a task difference: unlike Yeverechyahu's word-likelihood task, here OCP status was constant across the intransitive-transitive pair.

is provided in (6). I constructed 48 sets, each consisting of 16 pairings of intransitive templates (*CaCaC*, *hiCCiC*, *hitCaCeC*, *niCCaC*) and transitive templates (*CaCaC*, *CiCeC*, *hiCCiC*, *hitCaCeC*). As with Experiment 1, *CiCeC* was not included as an intransitive template. Additionally, *niCCaC* was excluded as a causative alternant to avoid it being analyzed as a verb in *CiCeC* with quadriliteral roots, given that this was a prominent strategy in Experiment 1. Each participant was randomly assigned to one of 16 lists and viewed one item from each set, totaling 48 experimental items. The experiment also included 24 fillers, all with intransitives in *CiCeC*. Half of the fillers contained correct agreement morphology, while the other half contained incorrect agreement morphology.

(6)	Dan	*PaNaG*	etmol.	ha-kerax	garam	le-ze.	klomar,	ha-kerax	***hiFNiG***	et	Dan.
	Dan	*PaNaG*	yesterday.	The-ice	caused	to-it.	That is,	the-ice	***hiFNiG***	ACC	Dan.

‘Dan *PaNaG* yesterday. The ice caused it. That is, the ice ***hiFNiG*** Dan.’

2.2.1.3 Analysis

Ratings were treated as ordered categorical responses. Mixed-effects ordinal logistic regression was used to model participants’ acceptability ratings using the clmm function from the ordinal package (Christensen 2023). For all models I fit the maximal random-effects structure justified by the design and simplified only when the model failed to converge.

Type II likelihood ratio tests (via the Anova function from the car package; Fox & Weisberg 2019) were used to assess the overall significance of predictors, and pairwise comparisons between transitive templates were computed using the emmeans package (Lenth 2025) with Tukey correction.

To assess whether identity between the intransitive and transitive templates predicts acceptability ratings, I fit an ordinal logistic regression model with SameIT as a fixed effect (a binary variable coded 1 when Intransitive $=$ Transitive and 0 when Intransitive $\neq$ Transitive). The maximal random-effects structure included by-participant and by-Set random intercepts and slopes for SameIT.

To examine participants’ preferences among transitive templates, I fit an ordinal logistic regression model predicting Rating from Transitive (the transitive template), using only trials in which the intransitive and transitive templates differed. The maximal random-effects structure included by-participant and by-Set random intercepts and slopes for Transitive. In addition to the model-based analyses, I report by-participant descriptive patterns: for each pairwise template contrast, participants are classified as showing a clear preference (mean rating difference > 0.5) or no preference (≤ 0.5).

Finally, to test whether manner faithfulness affects template selection, I fit a cumulative link mixed-effects model predicting Rating from BP, Faithfulness, and Intransitive (the intransitive template: *CaCaC*, *hiCCiC*, *hitCaCeC*, or *niCCaC*), including all two-way and three-way interactions.

The fully maximal structure failed to converge, so the structure was simplified to include by-participant random intercepts and slopes for BP, Faithfulness, and their interaction, and a by-Set random intercept.

2.2.2 Results

Six participants were excluded from the analysis: two assigned the same score to all sentences including fillers, three did not show a difference in rating between fillers with correct agreement and fillers with incorrect agreement (difference in average scores < 1), and one selected an incorrect answer for a comprehension question in the instructions. I therefore analyzed results from 69 participants, with each list containing a minimum of four participants (one list had six participants, one had seven). The average ratings for each template pair are presented in **Table 6**, and **Figure 5** presents the distribution of ratings across combinations of transitive and intransitive templates.

		Transitive			
		CaCaC	***CiCeC***	***hiCCiC***	***hitCaCeC***
Intransitive	*CaCaC*	2.6	4.9	4.8	1.9
	niCCaC	4.3	4.5	4.8	2.1
	hitCaCeC	4.1	4.8	4.5	1.6
	hiCCiC	3.7	4.3	3.0	2.0

Table 6: Acceptability ratings for intransitive and transitive template pairings in Experiment 2 on a scale from 1 (unacceptable) to 6 (excellent).

Most template combinations, including non-existent and rare combinations, were judged as acceptable with a rating between 4 and 5. *hitCaCeC* was consistently judged as an unacceptable transitive regardless of the intransitive template. In cases where the intransitive and transitive templates were identical (e.g., *CaCaC-CaCaC*, *hiCCiC-hiCCiC*, and *hitCaCeC-hitCaCeC*), the ratings were the lowest compared to all other transitives for a given intransitive (apart from *hitCaCeC* transitives). Consistent with this pattern, the model testing template identity yielded a negative and significant coefficient for SameIT ($p < .001$). In other words, identity between the intransitive and the transitive template had a reliably negative effect on acceptability ratings, with some variability across participants and items for non-identical pairs (SameIT $= 0$), as illustrated in **Figure 6** and in the analyses that follow. This result supports the conclusion that participants systematically dispreferred combinations where the transitive template preserved the intransitive template. The full model output is presented in **Table 7**.

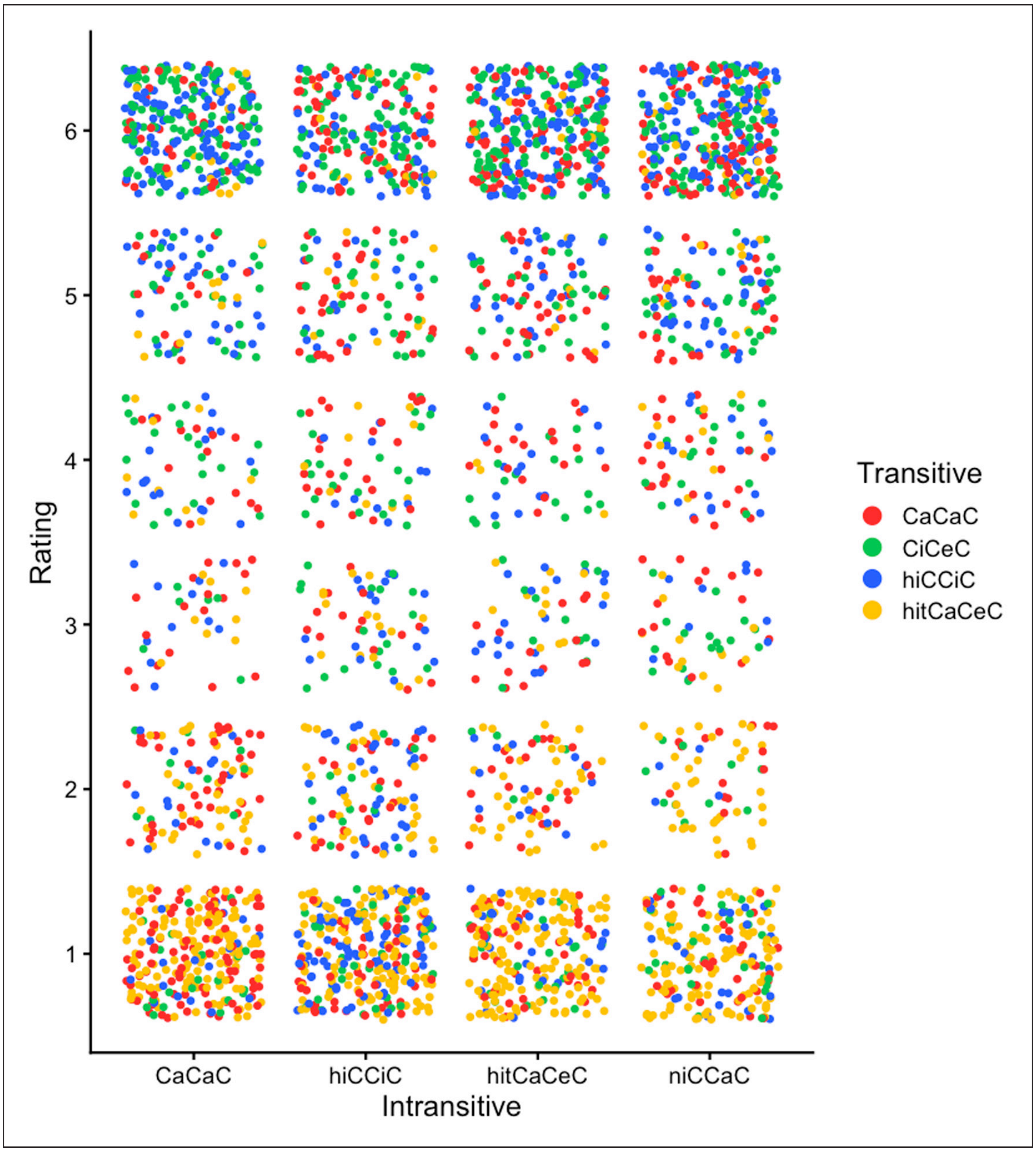

Figure 5: Acceptability ratings by intransitive and transitive template. Each point represents a single response. Colors indicate the transitive template presented for each intransitive template.

Given the observed dispreference for identity pairs, I excluded these cases from the analysis about overall template preference. An additional advantage for the exclusion is that the experiment did not include items with a *CiCeC* intransitive, and therefore there were no *CiCeC-CiCeC* pairs tested. This exclusion helps prevent potential confounding by eliminating the possibility that the absence of *CiCeC* intransitive-transitive pairs could artificially inflate the average score for *CiCeC* transitives.

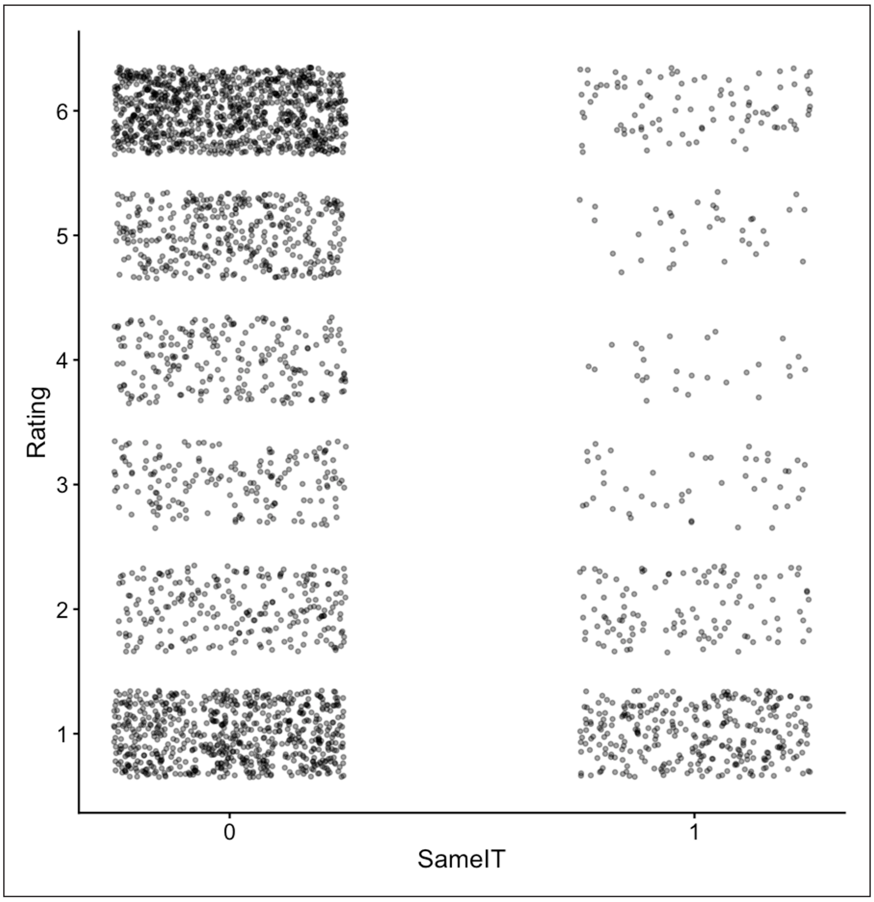

Figure 6: Distribution of acceptability ratings by identity between intransitive and transitive template (SameIT).

	Rating		
Predictors	***Odds Ratios***	***CI***	***p***
SameIT	**0.22**	**0.16–0.31**	**<0.001**
Random Effects			
$\tau_{00\ \text{Participant\_ID}}$	1.17		
$\tau_{00\ \text{Set}}$	0.00		
$\tau_{11\ \text{Participant\_ID.SameIT}}$	1.31		
$\tau_{11\ \text{Set.SameIT}}$	0.06		
ICC	0.28		
$N_{\text{Participant\_ID}}$	69		
N_{Set}	48		
observations	3312		

Table 7: Ordinal logistic mixed-effects model of acceptability ratings in Experiment 2. Model: Rating ~ SameIT + (1 + SameIT | Participant_ID) + (1 + SameIT | Set). Reference level: SameIT = 0. Estimates shown as odds ratios with 95% Wald confidence intervals.

In Experiment 1, *CiCeC* was the preferred transitive template overall, *hiCCiC* was second, *CaCaC* the third, and *hitCaCeC* and *niCCaC* were essentially absent as transitives. Therefore, I tested whether the same preferences would also emerge in Experiment 2.

For preferences among the transitive templates (non-identical trials only), a likelihood-ratio test against a null model with the same random-effects structure but no fixed effect of Transitive revealed a significant effect of transitive template ($\chi^2(3) = 83.29$, $p < .001$). Fixed-effects estimates showed that all three alternative templates were rated significantly higher than the reference level *hitCaCeC* (see **Table 8** for the full model output). Pairwise comparisons with Tukey correction (**Table 9**) confirmed that *hiCCiC* and *CiCeC* were both rated significantly higher than *CaCaC* (both $p < .001$) but did not differ significantly from each other ($p = .96$). This suggests a ranking of preferences: *hiCCiC* ≈ *CiCeC* > *CaCaC* > *hitCaCeC*. **Figure 7** presents the average rating for each transitive template.

	Rating		
Predictors	***Odds Ratios***	***CI***	***p***
Transitive [*CaCa*C]	20.32	11.44–36.09	<0.001
Transitive [*CiCeC*]	49.54	25.53–96.12	<0.001
Transitive [*hiCCiC*]	55.14	27.79–109.43	<0.001
Random Effects			
$\tau_{00\ \mathrm{Participant\_ID}}$	5.14		
$\tau_{00\ \mathrm{Set}}$	0.01		
$\tau_{11\ \mathrm{Participant\_ID.\ Transitive}\mathit{CaCaC}}$	3.92		
$\tau_{11\ \mathrm{Participant\_ID.\ Transitive}\mathit{CiCeC}}$	5.56		
$\tau_{11\ \mathrm{Participant\_ID.\ Transitive}\mathit{hiCCiC}}$	6.04		
$\tau_{11\ \mathrm{Set.\ Transitive}\mathit{CaCaC}}$	0.02		
$\tau_{11\ \mathrm{Set.\ Transitive}\mathit{CiCeC}}$	0.14		
$\tau_{11\ \mathrm{Set.\ Transitive}\mathit{hiCCiC}}$	0.02		
$N_{\mathrm{Participant\_ID}}$	69		
N_{Set}	48		
Observations	2691		

Table 8: Ordinal logistic mixed-effects model of acceptability ratings in Experiment 2. Model: Rating ~ Transitive + (1 + Transitive | Participant_ID) + (1 + Transitive | Set). Reference level: Transitive = *hitCaCeC*. Estimates are odds ratios with 95% CIs.

Contrast	Estimate	SE	z.ratio	p.value
hitCaCeC − CaCaC	−3.01	0.29	−10.27	<.001
hitCaCeC − CiCeC	−3.90	0.34	−11.54	<.001
hitCaCeC − hiCCiC	−4.01	0.35	−11.47	<.001
CaCaC − CiCeC	−0.87	0.18	−4.96	<.001
CaCaC − hiCCiC	−1.00	0.22	−4.50	<.001
CiCeC − hiCCiC	−0.11	0.21	−0.50	.96

Table 9: Pairwise comparisons between transitive templates with Tukey adjustment for multiple comparisons.

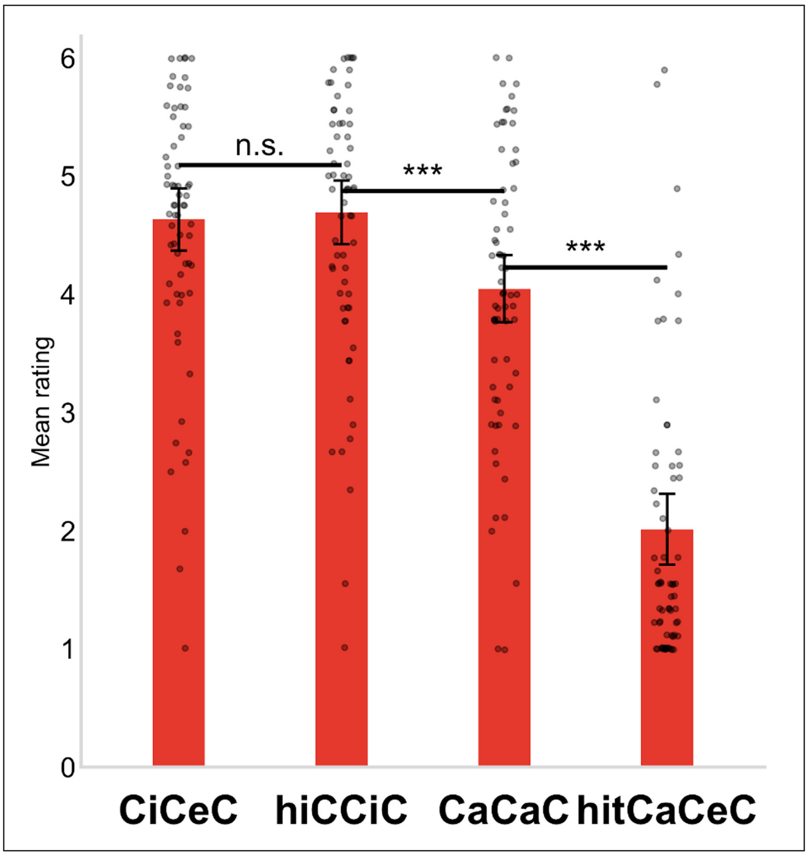

Figure 7: Mean acceptability ratings for each transitive template in Experiment 2 (SameIT = 0 trials only). Bars show means across participants; error bars show 95% confidence intervals; dots show individual participant mean ratings. Significance brackets are shown for adjacent pairs (*** $p < .001$; n.s.). All non-adjacent comparisons were significant at $p < .001$.

At the individual participant level, *CiCeC* and *hiCCiC* were rated similarly (a mean difference of 0.5 points or less) by 38 of the 69 participants, consistent with the group-level non-significant difference between the two. Of the remaining participants, 16 gave higher ratings to *CiCeC* and 15 gave higher ratings to *hiCCiC*, showing a roughly even split. The preference for *CiCeC* over *CaCaC* was present in 33 of the 69 participants, and the preference for *hiCCiC* over *CaCaC* in 36. The strongest and most consistent contrast between adjacent templates in the ranking was between *CaCaC* and *hitCaCeC*: 60 of the 69 participants rated *CaCaC* higher than *hitCaCeC*, mirroring the clear group-level avoidance of *hitCaCeC* as a transitive template.

The full group-level ranking (*CiCeC* ≈ *hiCCiC* > *CaCaC* > *hitCaCeC*) was observed in 16 participants. As with Experiment 1, the preferences held at the group level but with considerable variability across individuals.

Finally, the manner-faithfulness model revealed a significant main effect of Faithfulness and significant effects of two of the intransitive templates compared to the reference *CaCaC*: *hiCCiC* and *niCCaC*. Crucially, the main effect of BP was not significant, nor were the two-way interactions between BP and Faithfulness or BP and Intransitive, indicating that there is no evidence that BP uniformly modulates the preference for faithful template responses. A marginally significant three-way interaction between BP, Faithfulness, and Intransitive was observed for the intransitive *niCCaC* ($p = .04$), indicating that in this intransitive template the effect of BP on Faithfulness was negative, in the opposite direction to the hypothesis that participation in the stop-fricative alternation increases manner faithfulness. See **Table 10** for the full model output.

3 Discussion

3.1 Grammatical generalizations and item-specific listing

The goal of this study was to investigate which aspects of Hebrew transitivity alternations involve abstract grammatical generalizations and which ones involve item-specific listing. I used a nonce verb experiment with the rationale that grammatical mechanisms will surface as clear generalizations in nonce verbs, whereas listed information will manifest itself as variability in template selection.

3.1.1 Grammatical generalizations

I found that the intransitivity of *niCCaC* and *hitCaCeC* and the application of the stop-fricative alternation in the contexts tested in this study are known synchronically by speakers of Hebrew based on the finding that they can apply them to nonce verbs with minimal variability.

In Experiment 1, there were virtually no instances of transitive verbs in these templates. In Experiment 2 (where *niCCaC* was not tested), *hitCaCeC* was rated the lowest compared to all other transitive templates. The dispreference for *niCCaC* and *hitCaCeC* as transitives mirrors their absence as alternating transitives in the language. This finding is consistent with several theoretical accounts, as summarized in Section 3.3.2.

	Rating		
Predictors	***Odds Ratios***	***CI***	***p***
BP [1]	0.71	0.41–1.22	0.218
Faithful	0.18	0.11–0.28	$<$0.001
Intransitive [*hiCCi*C]	0.22	0.14–0.35	$<$0.001
Intransitive [*hitCaCeC*]	0.81	0.47–1.39	0.450
Intransitive [*niCCaC*]	0.29	0.19–0.45	$<$0.001
BP [1] × Faithful	1.32	0.71–2.44	0.380
BP [1] × Intransitive [*hiCCiC*]	1.21	0.65–2.23	0.551
BP [1] × Intransitive [*hitCaCeC*]	1.00	0.47–2.10	0.994
BP [1] × Intransitive [*niCCaC*]	1.31	0.71–2.42	0.395
Faithful × Intransitive [*hiCCiC*]	4.81	2.61–8.84	$<$0.001
Faithful × Intransitive [*hitCaCeC*]	1.69	0.91–3.13	0.094
Faithful × Intransitive [*niCCaC*]	23.26	12.30–44.00	$<$0.001
(BP [1] × Faithful) × Intransitive [*hiCCiC*]	0.66	0.28–1.56	0.347
(BP [1] × Faithful) × Intransitive [*hit-CaCeC*]	1.12	0.48–2.64	0.793
(BP [1] × Faithful) × Intransitive [*niCCaC*]	0.40	0.17–0.97	0.042
Random Effects			
$\tau_{00\ \mathrm{Participant\_ID}}$	1.38		
$\tau_{00\ \mathrm{Set}}$	0.00		
$\tau_{11\ \mathrm{Participant\_ID.Faithful}}$	0.12		
$\tau_{11\ \mathrm{Participant\_ID.BP1}}$	0.13		
$\tau_{11\ \mathrm{Participant\_ID.Faithful:BP1}}$	0.01		
$N_{\mathrm{Participant\_ID}}$	69		
N_{Set}	48		
Observations	3312		

Table 10: Ordinal logistic mixed-effects model testing the effect of manner faithfulness. Model: Rating ~ BP * Faithful * Intransitive + (1 + Faithful * BP | Participant_ID) + (1 | Set). Reference levels: BP = 0, Faithful = 0, Intransitive = *CaCaC*. Estimates are odds ratios with 95% CIs.

Similarly, speakers know how to apply the stop-fricative alternation to novel items at the first root consonant, the morphophonological environment tested in Experiment 1, as can be seen by the scarcity of misapplication errors. As an anonymous reviewer pointed out, previous studies have found variation in the application of the stop-fricative alternation, but crucially, the locus of variation has been mostly in post-consonantal positions, following certain consonants (e.g., *yiTFoR* 'sew.3SG.M.FUT'; Adam 2002, for spontaneous speech data and analysis, Albert 2014, for elicited production, Temkin-Martínez 2010, for acceptability judgment, Temkin-Martínez 2013, and Temkin-Martínez & Müllner 2016, for elicited production of verbs and nonce-verbs). In Experiment 1 of the current study, the potential locus of variation was the first root consonant, target transitive verbs were all in third person singular past tense, and *hitCaCeC* forms were not produced by participants; consequently, all relevant morphophonological environments were word-initial or post-vocalic. The lack of variation in the application of the stop-fricative alternation is therefore consistent with previous literature.

Speakers' consistent and reliable application of these generalizations to nonwords in these domains suggests that the intransitivity of *niCCaC* and *hitCaCeC* (or at least their inability to be transitive alternants) and knowledge about the stop-fricative alternation in the tested environments are both encoded in the synchronic grammar and considered in real-time production and judgment.

3.1.2 Item-specific listing

Conversely, speakers are inconsistent in their choice of transitive template, and are permissive when asked to rate template combinations. They are not constrained by specific templates for the expression of causation, nor by strict pairings between templates, nor even by the pairings attested in the lexicon. I also found no evidence that they adhere to paradigm uniformity with respect to the stop-fricative alternation. This implies that speakers cannot reliably generate the appropriate template based on grammatical generalizations, and therefore the template must be listed for each specific root and for both alternants.

Two caveats are worth highlighting. First, the finding that template selection in alternations is mostly not based on grammatical generalizations does not imply that the relation between alternants is not grammatically determined. While there is a grammatical relation between abstract alternants at the thematic, syntactic, and semantic levels, their morphological realization is idiosyncratic and listed, except for the grammatical constraint on *niCCaC* and *hitCaCeC* transitives. Second, the present work focused specifically on template selection in transitivity alternations, and the conclusions drawn here are limited to this domain. The Hebrew verbal system is morphologically rich and encompasses a range of systematic relationships that were not investigated in the current experiments. Whether similar patterns emerge in nonce verb experiments targeting other types of verbs and alternations remains an open question for future experimental work.

With these clarifications in mind, the consistency of these findings across tasks and modalities (including an acceptability judgment task in Experiment 2 which minimizes the effect of performance factors) strengthens the reliability of the result as a reflection of the grammatical system underlying verbal template selection in Hebrew.

3.2 Frequency-based generalizations and their limitations

In the aggregate, speakers match the overall distribution of transitive verbal templates in the lexicon when they produce novel forms, as seen in the overall preference in Experiment 1. This is a case of frequency matching, a phenomenon observed across languages and domains (Ernestus & Baayen 2003; Hayes et al. 2009, among others). Although this is a generalization in the sense that it reflects speakers' abstract knowledge about the domain of Hebrew templates, it is not a grammatical generalization like the ones outlined in Section 3.1.1, since it does not provide speakers with specific ways to generate verbs and does not limit the alternatives. To illustrate this point, speakers choose *CiCeC* for transitive verbs in about 56% of the cases, and it is also the most frequent transitive template in the inventory. However, they also produce *hiCCiC* in 26% of the cases, meaning that the preference for *CiCeC* is not absolute (compare this to the virtual absence of *niCCaC* and *hitCaCeC* transitives). As Experiment 2 shows, speakers rate *CiCeC* and *hiCCiC* transitives roughly the same, meaning that their grammar does not rule out the less frequent form, and it is not clear whether one form is the default.

Therefore, even though speakers follow some generalization about the statistical distribution of templates when they are asked to derive nonce verbs, there is no reliable grammatical mechanism for choosing a transitive template, and therefore the template must be listed. Frequency matching is not very useful outside the context of a wug test. When a speaker needs to encode a newly learned word or to use a known one, the distribution of verbal templates in the lexicon is not sufficient, since the choice is not a statistical one: there is only one correct template, which must be learned in the case of a new word or already listed in the case of a known one.

Although speakers are able to track the overall distribution of verbal templates, their statistical sensitivity is limited, as the intransitive template of a verb does not appear to influence their choice of a specific transitive template the same way it does in the lexicon. Examining each intransitive template individually shows that while the intransitive template in the stimulus sometimes affected speakers' selection of a transitive template, this effect was inconsistent and did not mirror the distribution of template pairs in Hebrew.

In Experiment 1, *CiCeC* was the most common transitive template for all intransitive templates. However, in the lexicon, while it is very common for *hitCaCeC* intransitives to alternate with *CiCeC*, it is very uncommon for *CaCaC* intransitives, and non-existent for *hiCCiC* intransitives. It appears that in both production and judgment, speakers do not attribute a considerable role to the template of the intransitive alternant when selecting a transitive alternant, at least not systematically and uniformly.

The finding that speakers imitate the distribution of output forms but not the conditional distribution of output forms given properties of an input is similar to the finding reported by Moore-Cantwell (2013) regarding denominals in *CiCeC*, where speakers generalized over output forms but not over the vowels of the input nominal. Overall, speakers tend to produce denominal verbs using consonant doubling, the most common strategy in the lexicon, as in *dam* → *dimem* ('blood' → 'to bleed'), rather than other available strategies: vowel overwriting, coronal glide formation, labial glide formation, or reduplication. In the lexicon, the choice of strategy is tied to morphological properties of the noun, but speakers do not systematically adjust their choice based on these properties.

Moore-Cantwell suggested that one reason speakers did not show knowledge of conditional frequencies in her experiment was the sparse evidence, with only 52 pairs in the lexicon. The results of the present study show that this can occur even with much more evidence (443 pairs of alternating verbs in the inventory, and more in the lexicon). The failure is not confined to gradient tendencies: even categorical dependencies in the inventory are not reproduced. This is not a general inability to learn morphological co-occurrences, since speakers can, in principle, extend such dependencies to novel forms (Tabachnick 2024, for Hungarian possessives predictable from the plural).

Speakers' overall preference for *CiCeC* transitives in production can be attributed to several factors beyond its high frequency as a transitive alternant. The limited occurrences of consonant clusters in the *CiCeC* paradigm (for triliteral roots), and the lack of clusters in the inflections tested in Experiment 1, may make it easier to produce and inflect than *hiCCiC*, which contains a consonant cluster between the first and second radicals (Schwarzwald 1996). Additionally, *CiCeC* is highly productive: it accounts for 66% of verb innovations in Laks's (2018) inventory, it is the most common causative template in Bolozky's studies from the 1990s (Section 1.5), and, more generally, it is prominent among verbs coined in Modern Hebrew rather than in earlier stages of the language (Agranovsky 2025).

The productivity of *CiCeC*, which may itself stem in part from its morphophonological simplicity, is likely to affect speakers' preferences in production.

These considerations help explain the difference between the two experiments: speakers preferred *CiCeC* over *hiCCiC* in the production experiment but did not show this preference in the judgment experiment. The production task is more similar to verb innovation than the judgment task, since in the production task participants must generate a new verb themselves, rather than evaluate a form provided to them. As a result, generalizations about verb innovation may influence speakers' choices in production in addition to generalizations about the lexical distribution of templates. The production task may also be more susceptible to morphophonological pressures toward simplicity, especially given that the target transitive forms were in the third-person singular, where *CiCeC* has no consonant cluster and *hiCCiC* does. As an anonymous reviewer pointed out, the fillers in Experiment 1, which all required producing an inflected *CiCeC* form, may also have played a role in the increased preference for *CiCeC*, although they involved *CiCeC* as an intransitive rather than a transitive.

3.3 Implications for the theory of Hebrew verbal templates

3.3.1 Relations between templates

The findings indicate that participants produce and accept a broader range of template combinations than those that are frequent or even attested in the Hebrew lexicon. Thus, any theory that restricts the set of possible grammatical template combinations or assumes strict grammatical relationships between members of certain template combinations and not others is too strong in light of the current data. Specifically, the layering approaches of Doron (2003) and Kastner (2020) treat the following template pairs as structurally related (sharing a core structure and differing only in higher layers): *niCCaC-CaCaC*, *hitCaCeC-CiCeC*, and *CaCaC-hiCCiC*. Kastner (2020) additionally takes *niCCaC-hiCCiC* to be structurally related, and Doron (2003) additionally takes *CaCaC-CiCeC* to be structurally related (in the intensive alternation). See Section 1.4.1 for the details.

This is too restrictive given that speakers produce and accept other combinations as well. In fact, speakers prefer *niCCaC-CiCeC* alternations over *niCCaC-CaCaC* or *niCCaC-hiCCiC*, and *CaCaC-CiCeC* alternations over *CaCaC-hiCCiC* in the causative alternation.

Even Arad's (2005) far less restrictive conjugation classes are too strong. Speakers' preferences are not limited to the six conjugation classes proposed by Arad, nor to pairs attested in Hebrew. For example, the speakers in the current study produce and give high ratings to *CaCaC-CiCeC* pairs (rare in Hebrew and not a conjugation class in Arad's theory), and even *hiCCiC-CiCeC* pairs (unattested in Hebrew). The implication of these findings is that abstract relations between templates are not specified in the grammar, and therefore the two templates must be listed for each root, as proposed by Borer (2004).

3.3.2 Transitivity

Speakers know that *niCCaC* and *hitCaCeC* cannot be transitive. This aligns with Borer's (2004) and Arad's (2005) proposal that a verb's ability to assign object marking is marked by the template, and that transitive verbs must be in *CaCaC*, *CiCeC*, or *hiCCiC*. This finding is also consistent with accounts in which *niCCaC* and *hitCaCeC* cannot project an external argument, since by Burzio's (1986) generalization they would also be unable to assign accusative case. However, this is a stronger claim than is necessary to account for the current findings. In Kastner's (2020) analysis, *niCCaC* and *hitCaCeC* are spellouts of [-D], indicating the absence of an external argument; in Doron (2003), the middle head μ blocks the external argument. In both accounts, the inability to assign accusative case follows from Burzio's generalization. While these predictions are consistent with the findings of this study, only the claim regarding accusative case is necessary. The stronger claim about the absence of an external argument is not required, and indeed is not fully supported by the empirical evidence from the Hebrew lexicon.

There are verbs in *niCCaC* and *hitCaCeC* with an external argument: non-alternating verbs with PP objects, alternating reflexives, and simple agentive verbs in *hitCaCeC*. This claim is based on their failure to pass unaccusativity diagnostics, even when tested experimentally against unaccusatives in the same template (Plotnik et al. 2024). To reconcile this apparent contradiction, Kastner proposes that PP-verbs in *niCCaC*, which do not pass unaccusativity diagnostics, are figure reflexive. For this case, following Wood (2015), Kastner proposes that it is the *p* head rather than the *v* head that possesses the [–D] feature. This prohibits an external argument for the preposition while allowing an external argument for the verb, thereby enabling an agent to appear in the subject position. Both *p*[–D] and *Voice*[–D] are spelled out as *niCCaC*, and therefore formally there is no contradiction with the claim that Voice[–D] is unaccusative. For reflexives in *hitCaCeC*, which also do not pass unaccusativity diagnostics, Kastner argues against the diagnostics themselves, and therefore he claims that they cannot be used to determine the presence of an external argument in the syntax. However, experimental evidence suggests that unaccusativity diagnostics in Hebrew are relevant even when controlling for semantic and morphological factors (Plotnik et al. 2024). In addition, Ahdout (2021) notes that there are some clear instances of agentive verbs in *hitCaCeC* such as *hitLONeN* ('complain'), *hiŠtoLeL* ('act wildly'). These verbs are not reflexive, fail unaccusativity diagnostics, and pass agentivity diagnostics. Additionally, there are PP-verbs in *niCCaC* that Kastner categorizes as figure reflexives such as *niLXaM* ('fight'), but it is not clear that semantically they have a self-directed motion component that defines figure reflexives. To sum up this argument, verbs in *hitCaCeC* and *niCCaC* indeed cannot have a direct object, and this generalization is part of speakers' synchronic knowledge, as Borer (2004) proposed. The same prediction is derived by Doron (2003) and Kastner (2020), but in a version that is too strong: they suggest that verbal templates are a morphological realization of a "Burzio head" that marks both lack of external argument and inability to take a direct object, but *hitCaCeC* and *niCCaC* only mark the lack of a direct object, without prohibiting an external argument.

Since speakers generalize the intransitivity of *niCCaC* and *hitCaCeC* to nonce verbs in both production and judgment, it seems that their intransitivity is encoded in the grammar. This grammatical generalization seems to apply to a syntactic notion of objecthood. Verbs in *niCCaC* and *hitCaCeC* can select PP objects, i.e., they can be semantically transitive. They can also appear with a cognate object, even with the accusative case marker *et*, as demonstrated in (7)–(8):

(7)

Dan	kvar	hitPaLeL	et	ha-tfila	ha-zot.
Dan	already	prayed	ACC	the-prayer	the-this

‘Dan prayed this prayer.’

(8)

Dan	niXŠaL	et	ha-kišalon	ha-gadol	beyoter	be-xay-av.
Dan	failed	ACC	the-failure	the-big	most	in-life-his

‘Dan failed the biggest failure of his life.’

The fact that verbs in *hitCaCeC* and *niCCaC* can be semantically transitive suggests that the generalization speakers make is not based on a simple ban against an internal argument. Moreover, the verbs’ ability to appear adjacent to an accusative-marked argument indicates that surface frequency alone is insufficient for learning the behavior of these templates. Rather, speakers must generalize over true instances of direct objects to acquire this rule.

3.3.3 Toward a formal analysis

The experiments were designed to adjudicate between claims about the transitivity of templates and claims about the relations between verbs in different templates based on existing theoretical accounts of the Hebrew verbal system. The findings constrain the hypothesis space for existing and future theories in both respects: theories must include a component that makes *hitCaCeC* and *niCCaC* verbs intransitive, and at the same time a list with specific root-template combinations. Multiple frameworks can be made compatible with these requirements, and committing to any single one would require taking a stance on questions that the current experiments are not designed to address; I therefore leave this choice for future research. In this section, I demonstrate how these findings can be derived in principle, focusing for concreteness on the theoretical accounts of the type presented in Section 1.4, which assume syntactic decomposition of the root and the functional elements that are realized as templates.[8] I show that the model in Borer (2004) is largely compatible with the findings of the current paper, and propose a modification that makes it fully compatible. I demonstrate how the same idea can be implemented in the Distributed Morphology framework (Halle & Marantz 1993).

[8] This is of course not meant to imply that lexicalist models cannot capture the same generalizations.

Borer's (2004) model directly predicts the results of the current experiment. She proposes two components in the grammar: a syntactic feature responsible for (in)transitivity whose realization constrains template selection, and a post-syntactic list of root-template combinations. She explicitly argues that both are needed:

> "While the syntactic structure would exclude some possible morphophonological binyanim from consideration, […] the presence or the absence of the [+OM] feature still leaves the final morphophonology rather underdetermined and subject to root-specific listed information." (Borer 2004: 34)

This is largely in line with the present findings. The intransitivity of *niCCaC* and *hitCaCeC* is categorical and productive: speakers treat it as a grammatical constraint, consistent with the role of the [+OM] feature, which can be realized as *CiCeC*, *hiCCiC*, or *CaCaC* (the latter unspecified for [OM]). Within each set (transitive and intransitive), the choice of template in nonwords shows preferences rather than categorical behavior and does not track pair-specific generalizations, consistent with Borer's claim that, for real words, the specific template is a matter of root-specific listed information. The one area where the findings diverge from Borer's proposal is her claim that [–OM] is realized by *niCCaC*, *hitCaCeC*, or unspecified *CaCaC*, whereas in the lexicon (and by speakers' judgment in Experiment 2) there are also *CiCeC* and *hiCCiC* intransitives. This could be resolved by removing the morphophonological restrictions on [–OM], making the realization of intransitive verbs unrestricted by template, and leaving only the restriction on the [+OM] feature.

Arad's (2005) proposal that *niCCaC* and *hitCaCeC* are marked as [–tr] or [–ACC case], which confines their environment of insertion to intransitive verbs, is similar to Borer's modified account, and is consistent with the findings about (in)transitivity. However, as discussed in Section 3.3.1, Arad's theory about the relation between alternating templates, namely that alternations are constrained to six template combinations, is too restrictive to account for the experimental findings.

The general idea of a restriction on transitivity together with full listing of templates can also be implemented in the Distributed Morphology framework. For the syntactic representation of alternating verbs I assume, following Arad (2005), that the relation between alternants is attributed to the flavor of the verbalizing head *v* above the shared root (v_{TR} and v_{INTR}), which mark the verb's transitivity, replacing Borer's [OM] feature. As Arad (2005) notes (footnote 42, p. 277), this kind of syntactic account, presented here for simplicity, does not treat transitivity alternants as derivationally related on the syntactic level, but as independently derived verbs with a shared root. Arad points out that the same morphological system could in principle be recast in configurational terms. However, recasting the findings in configurational terms is non-trivial: if the template is not the realization of *Voice*, it is unclear what the syntactic relation between the alternants is, which heads are relevant for spelling out the template, and whether they are local enough to the root for it to define the environment for template selection. I leave these questions open.

The realization of the verbalizer is determined by lexical insertion rules in List 2 (the Vocabulary), where there is conditional allomorphy between possible templates, conditioned by the root. Crucially, *hitCaCeC* and *niCCaC* are not in the list of allomorphs for v_{TR}, which directly captures their inability to act as transitive alternants without further stipulations. A sketch of the lexical insertion rules is presented in (9).

(9) v_{TR}
↔ *hiCCiC* / K-P-C, R-G-Z, R-D-M, Z-N-K, L-B-N …
↔ *CiCeC* / X-M-M, S-M-X …
↔ *CaCaC* / S-G-R, P-T-X …

v_{INTR}
↔ *niCCaC* / S-G-R, P-T-X, R-D-M …
↔ *hitCaCeC* / X-M-M, R-G-Z …
↔ *CaCaC* / K-P-C, S-M-X …
↔ *CiCeC* / Z-N-K …
↔ *hiCCiC* / L-B-N …

I emphasize that this sketch is intended as a proof of concept rather than a full analysis. The details of the syntactic representation and of the Vocabulary Items remain to be worked out in future research. For instance, roots that share consonants but are related only abstractly in meaning (Arad's 2005 Multiple Contextual Meaning roots) are not distinguished by (9) when the two verbs share a transitivity value (e.g. P-N-H in *PiNaH* 'clear' and *hiFNaH* 'direct'). This could be resolved by treating them as distinct roots in the syntax, or by a difference in their functional context other than transitivity (such as properties of the other arguments). Such factors may influence template selection more generally but were not tested in the present experiments, and whether they affect it synchronically remains a question for future work.

4 Conclusion

In two nonce verb experiments I found that some aspects of the causative alternation involve abstract grammatical generalizations, while others are based on listed information. I showed that speakers do not select *niCCaC* and *hitCaCeC* as transitive alternants and know how to apply the stop-fricative alternation (within the limited contexts tested in the experiment). This suggests that these generalizations are governed by grammatical mechanisms, which can be applied by speakers to novel verbs. I propose that *niCCaC* and *hitCaCeC* grammatically mark intransitivity but are not a realization of a “Burzio head” which also marks the absence of an external argument.

Conversely, while speakers imitate the overall distribution of transitive alternants in the language, they exhibit great variability in their choice of transitive templates based on a given intransitive template, and their choices do not mirror the distribution of pairs in the language, nor did I find evidence that they consider the stop-fricative alternation in doing so. This indicates that templates are listed for specific roots, and that the set of possible pairs is not categorically restricted by the grammar, apart from the ban on the transitivity of *niCCaC* and *hitCaCeC*.

Abbreviations

3 = 3rd person, ACC = accusative, F = feminine, FUT = future, INTR = intransitive, M = masculine, PL = plural, SG = singular, TR = transitive.

Data availability

The data and analysis code used in this study are available in a public repository: https://doi.org/10.17605/OSF.IO/7QHN8.

Ethics and consent

The study was approved by Tel Aviv University's Ethics Committee. Participants provided informed consent to participate in the study.

Funding information

This research was supported by the Argov Center for the Study of Israel and the Jewish People at Bar-Ilan University, MiLa: The Cukier-Goldstein-Goren Center for Mind and Language, and Israel Science Foundation Grant 1781/22 to Malka Rappaport Hovav.

Acknowledgements

I thank Vera Agranovsky, Mandy Cartner, Kyle Gorman, Zara Harmon, Masoud Jasbi, Iris Kamil, Roni Katzir, Inbal Kuperwasser, Lior Laks, Alec Marantz, Aya Meltzer-Asscher, Yaron Mor, Malka Rappaport Hovav, Ezer Rasin, Yuval Schieber, and Tal Siloni for helpful discussions of this study. I am especially grateful to Naama Friedmann, who encouraged me to investigate transitivity alternations using nonce verbs and provided feedback on several versions of the paper. The editor, Sol Lago, and three anonymous reviewers provided invaluable comments that significantly improved this paper. I presented this study at the *Abstract and Item-Specific Knowledge Across Domains and Frameworks* conference at the 2025 LSA Summer Institute at the University of Oregon, which I attended on a scholarship from the organizers. I am grateful to Canaan Breiss, Gaja Jarosz, Volya Kapatsinski, and Emily Morgan for this opportunity, which has substantially shaped this paper.

Competing interests

The author has no competing interests to declare.

Author's contributions

The author is responsible for all aspects of the study and manuscript preparation.

References

Adam, Galit. 2002. *From variable to optimal grammar*. Tel Aviv: Tel Aviv University dissertation.

Agranovsky, Vera. 2025. *The productive mechanisms of derivation in the Modern Hebrew verbal system*. Jerusalem: The Hebrew University of Jerusalem dissertation.

Ahdout, Odelia. 2021. *Deverbal nouns in Modern Hebrew: Between grammar and competition*. Berlin: Humboldt University of Berlin dissertation.

Albert, Aviad. 2014. *Phonotactic universals in Modern Hebrew: Evidence for prosodic alignment of stops*. Tel Aviv: Tel Aviv University MA thesis.

Alexiadou, Artemis & Anagnostopoulou, Elena & Schäfer, Florian. 2015. *External arguments in transitivity alternations: A layering approach* (Oxford Studies in Theoretical Linguistics 55). Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199571949.001.0001

Arad, Maya. 2005. *Roots and patterns: Hebrew morpho-syntax*. Dordrecht: Springer.

Barr, Dale J. & Levy, Roger & Scheepers, Christoph & Tily, Harry J. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3). 255–278. DOI: https://doi.org/10.1016/j.jml.2012.11.001

Bat-El, Outi. 1994. Stem modification and cluster transfer in Modern Hebrew. *Natural Language & Linguistic Theory* 12(4). 571–596. DOI: https://doi.org/10.1007/BF00992928

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

Berko, Jean. 1958. The child's learning of English morphology. *Word* 14(2–3). 150–177. DOI: https://doi.org/10.1080/00437956.1958.11659661

Berman, Ruth A. 1982. Verb-pattern alternation: The interface of morphology, syntax, and semantics in Hebrew child language. *Journal of Child Language* 9(1). 169–191. DOI: https://doi.org/10.1017/S030500090000369X

Berman, Ruth A. 1993. Marking of verb transitivity by Hebrew-speaking children. *Journal of Child Language* 20(3). 641–669. DOI: https://doi.org/10.1017/S0305000900008527

Berman, Ruth A. 1994. Developmental perspectives on transitivity: A confluence of cues. In Levy, Yonata (ed.), *Other children, other languages*, 189–243. Hillsdale, NJ: Lawrence Erlbaum.

Bolozky, Shmuel. 1999. *Measuring productivity in word formation: The case of Israeli Hebrew*. Leiden: Brill. DOI: https://doi.org/10.1163/9789004348431

Borer, Hagit. 1991. The causative-inchoative alternation: A case study in parallel morphology. *The Linguistic Review* 8(2–4). 119–158. DOI: https://doi.org/10.1515/tlir.1991.8.2-4.119

Borer, Hagit. 2004. The grammar machine. In Alexiadou, Artemis & Anagnostopoulou, Elena & Everaert, Martin (eds.), *The unaccusativity puzzle: Explorations in the syntax-lexicon interface*, 288–331. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199257652.003.0012

Burzio, Luigi. 1986. *Italian syntax: A government-binding approach*. Dordrecht: Reidel. DOI: https://doi.org/10.1007/978-94-009-4522-7

Christensen, Rune Haubo Bojesen. 2023. ordinal: Regression models for ordinal data (R package version 2023.12-4.1). Comprehensive R Archive Network (CRAN). https://CRAN.R-project.org/package=ordinal

Doron, Edit. 2003. Agency and voice: The semantics of the Semitic templates. *Natural Language Semantics* 11(1). 1–67. DOI: https://doi.org/10.1023/A:1023021423453

Elff, Martin. 2026. mclogit: Multinomial logit models for categorical responses and discrete choices (R package). http://melff.github.io/mclogit/

Ernestus, Mirjam & Baayen, Harald. 2003. Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language* 79(1). 5–38. DOI: https://doi.org/10.1353/lan.2003.0076

Faust, Noam. 2019. New reasons to root for the Semitic root from Mehri and Neo-Aramaic. *The Linguistic Review* 36(3). 575–599. DOI: https://doi.org/10.1515/tlr-2019-2030

Fox, John & Weisberg, Sanford. 2019. *An R companion to applied regression*. Thousand Oaks, CA: Sage.

Green, Peter & MacLeod, Catriona J. 2016. SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution* 7(4). 493–498. DOI: https://doi.org/10.1111/2041-210X.12504

Halle, Morris & Marantz, Alec. 1993. Distributed morphology and the pieces of inflection. In Hale, Kenneth & Keyser, Samuel Jay (eds.), *The view from Building 20*, 111–176. Cambridge, MA: MIT Press.

Hayes, Bruce & Zuraw, Kie & Siptár, Péter & Londe, Zsuzsa. 2009. Natural and unnatural constraints in Hungarian vowel harmony. *Language* 85(4). 822–863. DOI: https://doi.org/10.1353/lan.0.0169

Hever, Ya'ar & Faust, Noam. 2010. Empirical and theoretical arguments in favor of the discontinuous root in Semitic languages. *Brill's Journal of Afroasiatic Languages and Linguistics* 2(1). 80–118. DOI: https://doi.org/10.1163/187666310X12688137960704

Horvath, Julia & Siloni, Tal. 2011. Causatives across components. *Natural Language & Linguistic Theory* 29(3). 657–704. DOI: https://doi.org/10.1007/s11049-011-9138-0

Jackendoff, Ray. 2002. What's in the lexicon? In Nooteboom, Sieb & Weerman, Fred (eds.), *Storage and computation in the language faculty*, 23–58. Dordrecht: Kluwer Academic Publishers. DOI: https://doi.org/10.1007/978-94-010-0355-1_2

Kastner, Itamar. 2020. *Voice at the interfaces: The syntax, semantics, and morphology of the Hebrew verb*. Berlin: Language Science Press. DOI: https://doi.org/10.5281/zenodo.3865067

Laks, Lior. 2009. How to choose a prosodic template: On verb innovation and variation in Hebrew. *Language Typology and Universals* 62(4). 307–324. DOI: https://doi.org/10.1524/stuf.2009.0023

Laks, Lior. 2011. *Morpho-phonological and morpho-thematic relations in Hebrew and Arabic verb formation*. Tel Aviv: Tel Aviv University dissertation.

Laks, Lior. 2014a. Morpho-thematic mismatches in Hebrew: What happens when morphological and thematic criteria collide? *Lingua* 138. 107–127. DOI: https://doi.org/10.1016/j.lingua.2013.10.011

Laks, Lior. 2014b. The Hebrew nif'al template: Two patterns of verbal passivization. *Anuario del Seminario de Filología Vasca «Julio de Urquijo»* 46(1). 183–193.

Laks, Lior. 2018. Verb innovation in Hebrew and Palestinian Arabic: The interaction of morpho-phonological and thematic-semantic criteria. *Brill's Journal of Afroasiatic Languages and Linguistics* 10(2). 238–284. DOI: https://doi.org/10.1163/18776930-01002003

Lenth, Russell. 2025. emmeans: Estimated marginal means, aka least-squares means (R package). https://CRAN.R-project.org/package=emmeans

Lev, Shaul. 2016. *Hebrew labile alternation*. Tel Aviv: Tel Aviv University MA thesis.

Levin, Beth & Rappaport Hovav, Malka. 1995. *Unaccusativity: At the syntax-lexical semantics interface*. Cambridge, MA: MIT Press.

Moore-Cantwell, Claire. 2013. Over- and under-generalization in learning derivational morphology. In Keine, Stefan & Sloggett, Shayne (eds.), *Proceedings of the 42nd meeting of the North East Linguistic Society (NELS)*, 41–54. Amherst, MA: GLSA.

Pinker, Steven. 1999. *Words and rules: The ingredients of language*. New York: Basic Books.

Plotnik, Ziv & Meltzer-Asscher, Aya & Siloni, Tal. 2024. The relevance of unaccusativity to possessive datives. *Linguistic Inquiry*. 1–44. DOI: https://doi.org/10.1162/ling_a_00532

R Core Team. 2023. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. https://www.R-project.org/

Reinhart, Tanya. 2003. The theta system: An overview. *Theoretical Linguistics* 28(3). 229–290. DOI: https://doi.org/10.1515/thli.28.3.229

Schwarzwald, Ora. 1996. Syllable structure, alternations, and verb complexity: Modern Hebrew verb templates reexamined. In Izre'er, Shlomo & Raz, Shlomo (eds.), *Studies in Modern Semitic Languages*, 95–112. Leiden: Brill. DOI: https://doi.org/10.1163/9789004659384_007

Tabachnick, Guy. 2024. Hungarian speakers use morphological dependencies in inflecting novel forms. *Glossa: a journal of general linguistics* 9(1). 1–38. DOI: https://doi.org/10.16995/glossa.10260

Temkin-Martínez, Michal. 2010. *Sources of non-conformity in phonology: Variation and exceptionality in Modern Hebrew spirantization*. Los Angeles, CA: University of Southern California dissertation.

Temkin-Martínez, Michal. 2013. Variation and preferences in Modern Hebrew nonce verbs. In *LSA Annual Meeting Extended Abstracts* 4. 44:1–5. DOI: https://doi.org/10.3765/exabs.v0i0.793

Temkin-Martínez, Michal & Müllner, Ivana. 2016. Specific exceptions driving variation: The role of orthography in Modern Hebrew spirantization. In Hansson, Gunnar Ólafur & Farris-Trimble, Ashley & McMullin, Kevin & Pulleyblank, Douglas (eds.), *Supplemental proceedings of the 2015 Annual Meeting on Phonology (AMP)*. Washington, DC: Linguistic Society of America. DOI: https://doi.org/10.3765/amp.v3i0.3695

Wood, Jim. 2015. *Icelandic morphosyntax and argument structure*. Dordrecht: Springer. DOI: https://doi.org/10.1007/978-3-319-09138-9

Yeverechyahu, Hadas. 2014. *The role of similarity in co-occurrence restrictions: Evidence from the Hebrew verbal system*. Tel Aviv: Tel Aviv University dissertation.

Zehr, Jeremy & Schwarz, Florian. 2018. *PennController for internet based experiments (IBEX)*. DOI: https://doi.org/10.17605/OSF.IO/MD832