


ARTICLE

Young minds' quest for regularity: Evidence from the Turkish causative

Mine NAKİPOĞLU^{a*} , Berna A. UZUNDAĞ^b, and Özge SARIGÜL^c

^aDepartment of Linguistics, Boğaziçi University, John Freely Hall 301, Bebek 34342, İstanbul, Turkey,

^bDepartment of Psychology, Kadir Has University, İstanbul, Turkey and ^cDepartment of Linguistics, Boğaziçi University, İstanbul, Turkey

*Corresponding author: Mine Nakipoğlu nakipogl@boun.edu.tr

(Received 29 October 2020; revised 5 May 2021; accepted 7 September 2021;
first published online 2 November 2021)

Abstract

Children's remarkable ability to generalize beyond the input and the resulting overregularizations/irregularizations provide a platform for a discussion of whether morphology learning uses analogy-based, rule-based, or statistical learning procedures. The present study, testing 115 children (aged 3 to 10) on an elicited production task, investigated the acquisition of the irregular distribution in the Turkish causative. Results showed that in early acquisition, to pin down the four causative suffixes, children engaged in comparisons between analogous exemplars. Thereafter to tackle the irregularity in two of the suffixes, children entertained competing hypotheses that yielded overregularizations and irregularizations. Overregularizations were instances of abstraction across the input based on type frequency; irregularizations were attempts to default to erroneous micro-generalizations. Negative correlation between errors and verb frequency suggested that recovery from errors was sensitive to token frequency. The overgeneralize-then-recover pattern that emerged in the acquisition of causative supported an integrated account of the roles of analogy, abstraction, and frequency in morphology learning.

Keywords: causative; abstraction; analogy; overregularization/irregularization errors; frequency

Introduction

Languages contain countless regularities as well as irregularities. Given that a regular system is in principle a system with less variation, the fewer the irregularities, the easier it would be to learn a language. When one adopts such a position, irregularities are like unwanted weeds in the garden. When, however, one gets intrigued by what irregularities can potentially reveal about the human mind – in particular, about a developing mind – one starts viewing irregularities in the morphology or syntax of languages as a fertile ground for an in-depth inquiry into how learning takes place. Children's erroneous uses in the form of overgeneralizations and irregularizations thus provide a way to investigate how the

© The Author(s), 2021. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

child's mind operates in discovering the patterns in her native language. In the present study, we investigated Turkish-speaking children's errors in acquiring causative morphology and what these errors imply in terms of children's learning mechanism.

Explaining children's errors: analogy-based and rule-based models

Morphological overgeneralizations, where irregular verbs are used with regular inflections, have been documented extensively in children's acquisition of language. For example in learning the English past tense, children produce morphological overregularizations like **hold-ed*, **go-ed* with irregular verbs by inducing that the majority of verbs are attached the *-ed* affix for past tense formation. In learning the argument structure of their language, children produce syntactic overgeneralizations like **I disappeared it* (= I made it disappear); *Don't giggle me!* (= Don't make me giggle!) with change-of-state verbs as most such verbs (e.g., melt, break) appear in causative alternation. Children also produce irregularization errors where they apply an irregular inflection to a regular verb (e.g., **truck* instead of *trick-ed*). In essence, any error, be it morphological or syntactic is in fact a logically possible interpretation of the input. Of interest then is how, in the time course of acquisition, children hypothesize the existence of forms unattested in the input, formulate erroneous generalizations, and then resolve them.

Inquiry into how erroneous morphological and syntactic use emerge over the course of acquisition, and what constrains the production of errors and how children recover from overgeneralizations, has stimulated a considerable amount of research both in morphology and syntax acquisition. In much of the morphological development literature to date, children's overregularizations and irregularizations have provided a platform for a discussion of whether regular and irregular forms are a product of a single, analogy-based mechanism or are products of a dual mechanism where irregulars are analogy-based but regulars are rule-based. Formulated mostly to account for the irregularity in English past tense, two theoretical positions and several gradient offshoots of these positions that vary in terms of how much they incorporate analogy and frequency to the model they advocate – suggesting a continuum from rule-based to analogy-based accounts – stand out on the issue. According to the RULE-BASED APPROACH, past tense inflection is a dual-route process involving a rule for regulars and an analogy-based mechanism for irregulars. In the earlier versions of the rule-based approach, such as the Dual-route Model (Marcus, 1995, 2001; Marcus, Pinker, Ullman, Hollander, Rosen & Xu, 1992; Pinker, 1999; Pinker & Prince, 1988; Pinker & Ullman, 2002; Prasada & Pinker, 1993), regulars are viewed as products of a default rule (e.g., add *-ed* for English) insensitive to phonology and frequency effects; whereas irregulars are viewed as products of an associative process that is sensitive to phonology and frequency. An influential version of the rule-based approach, the Multiple Rules Model of Albright and Hayes (2003), for example, posits various rules for regulars and irregulars on the grounds that they cluster around certain phonological neighbors – hence, they incorporate some sensitivity to phonology; and also holds that morphological patterns are extended on the basis of type frequency, thereby incorporating sensitivity to frequency. Thus, under the rule-based approach, overregularization errors like **hold-ed*, **drink-ed*, reflect children's tendency to default to the majority variant and the application of the regular rule *-ed* to the verb. The child must simply memorize irregulars as exceptions and, until she realizes that irregulars are unlike regulars, the

regular rule operates on any verb, yielding errors. Recovery from errors has been argued to take place via a process called BLOCKING: which is based on the idea that the learning of the correct irregular (e.g., *held*) over time would block the use of an overregularized form (**hold-ed*).

Countering the rule-based position, ANALOGY-BASED APPROACH holds that a single associative mechanism suffices to account for the behavior of regulars and irregulars. Morphology learning is thus explained not by instantiating deterministic rules but by means of probabilistic rules/schemas that implicitly encode structural similarities and effects of frequency (e.g., schema-based Model of Bybee, 1995, 2001; Bybee & Hopper, 2001; Bybee & Slobin, 1982; connectionist/single-route models of Rumelhart & McClelland, 1986; Hahn & Chater, 1998; Marchman, Plunkett & Goodman, 1997; McClelland & Patterson, 2002; Mirković, Seidenberg & Joanisse, 2011; Nakisa, Plunkett & Hahn, 2000; competition model of MacWhinney, 1987, 2004). Under this account, children learn the morphological forms by analogizing across the input. As both regulars and irregulars serve as the basis for analogical generalization, children are assumed to produce the overregularization: **hold-ed*, for example, by analogy to the regulars *fold-ed*, *scold-ed*; or **drink-ed*, by analogy to *blink-ed*, *wink-ed*, etc. This account further holds that the correct form can co-exist and compete with the overregularized form in the developing memory system of the child. Thus what governs the competition between *held* and **hold-ed* is their memory strength. The child's repeated encounters with the correct form in the input strengthen its memory representation; the incorrect form, however, does not receive support from the input and is eliminated over the course of acquisition (MacWhinney, 1987, 2004; Ramscar & Yarlett, 2007).

The Turkish causative as a window to understanding children's learning mechanisms

The present paper investigates the acquisition of causatives in Turkish, an agglutinative language with a rich set of morphological causatives that show unconditioned phonological variation. Thus the irregularity in the causative and children's regularization behavior in learning the construction offers a platform for a discussion of the implications of the rule-based vs. analogy-based models of morphological acquisition. Crosslinguistic research on the acquisition of causatives has mostly centered on the developmental stages the children go through in the productive use of causative constructions – as laid out in Berman (1993) for Hebrew; Pye (1994) for K'iche Mayan; Nomura and Shirai (1997) for Japanese; Allen (1998) for Inuktitut; Courtney (2002) for Quechua, and Lin and Tsay (2008) for Taiwanese Southern Min (TSM). Furthermore, in most of the languages studied, there is either a single morphological causative marker (e.g., *-tit-* in Inuktitut, *-chi* in Quechua, *-hoo-* in TSM, etc.) or no morphological means to express causativity (e.g., English). Unlike previous studies, here we focus on the irregularity in the causative and children's regularization behavior in learning the construction to examine whether a rule-based or an analogy-based model of morphological acquisition better accounts for the pattern in children's acquisition and use of the Turkish causative.

As the linguistic process of causativation cross-cuts morphology and syntax, it also provides an opportunity to explore whether Turkish-speaking children produce syntactic overgeneralizations in learning the causative. Though language-particular phenomena play a key role in children's syntactic overgeneralizations, a topic of inquiry in crosslinguistic literature has been whether children's overgeneralizations

are unidirectional, especially from intransitive to transitive, or bidirectional. For example, in English, as verbs do not exhibit fixed-transitivity (i.e., they are not exclusively intransitive or transitive), overgeneralizations are bidirectional. More precisely, children overgeneralize and use intransitives to express transitive meaning (e.g., **The magician disappeared the rabbit*; **She's gonna die it*); and to a much lesser extent transitives to express intransitive meaning (e.g., **But the parts might lose*) (Bowerman, 1982; Bowerman & Croft, 2007; Clark, 1993; Lord, 1979). In many other languages such as Hebrew, Inuktitut and Japanese, however, overgeneralizations are unidirectional, i.e., from intransitive to transitive.

The majority of research on syntactic overgeneralizations – in particular, argument structure overgeneralizations and how children retreat from errors – has been conducted in English and findings point to the role of both semantics (Pinker, 1989) and frequency of constructions in the input (Braine & Brooks, 1995; Brooks & Tomasello, 1999; Brooks, Tomasello, Dodson & Lewis, 1999; Brooks & Zizak, 2002; Goldberg, 2006; Tomasello, 2003). Under the SEMANTIC VERB-CLASS HYPOTHESIS of Pinker (1989), when the child decodes that in causative-alternation verbs, a transitive verb prototypically denotes DIRECT EXTERNAL CAUSATION (e.g., *X melts Y*) and the intransitive counterpart *internal causation* (e.g., *Y melts*), she formulates a semantic rule, thereby uncovering that verbs that exhibit similar semantic properties, for example, CHANGE-OF-STATE VERBS (e.g., *break*, *open*) can alternate. The child's overextension of this observation to verbs of EXISTENCE (e.g., *disappear*, *vanish*) may give rise to overgeneralization errors like **I disappeared it*. According to the CONSTRUCTIVIST USAGE-BASED APPROACH, in learning the argument structure of their language, children track the distributional statistics of verbs, and form probabilistic inferences about attested and unattested verb constructions in the input. The repeated representation of a verb – take the verb *disappear*, for instance, in an intransitive frame (e.g., *X disappeared*) or a periphrastic construction (e.g., *X made Y disappear*) – gradually strengthens the evidence that adults do not use the verb in a transitive frame: hence, pre-empts (i.e., blocks) the child's use of the verb in that unattested frame. Fine-tuning to the semantic properties of verbs, as argued in Pinker (1989), may assist the child to retreat from errors in English; but, under the construction-based approach, errors are pruned back through the statistical mechanisms of PRE-EMPTION – which is assumed to track the token frequency of a single unattested construction (e.g., *X disappeared Y*) – and ENTRENCHMENT, which keeps track of all the attested constructions that a verb can take part in (i.e., *Disappear!*, *Why did X disappear?* etc. – Ambridge, Pine, Rowland, Chang & Bidgood, 2013).

In Turkish, non-causative predicates are productively related to causative predicates through morphological markers: hence, unlike in English, verbs are for the most part, exclusively intransitive or transitive. Though it appears that this structural property of Turkish would, on the face of it, apparently disallow any bidirectionality in syntactic overgeneralizations, a secondary goal of this paper is to investigate whether Turkish children extend the use of intransitives to transitives as observed in other languages.

Morphological causativization in Turkish

In grammatical terms, causativization is a valency changing operation where an intransitive verb is rendered transitive, or a transitive verb ditransitive, through introduction of the causee argument: in subject and indirect object positions,

respectively. In an agglutinative language like Turkish, introduction of the causee requires concatenation of a causative suffix to the verb stem. For example, the intransitive base *eri* ‘melt’ in (1a) is rendered transitive in (1b) by the introduction of the external argument ‘the girl’ (i.e., causee) and the attachment of the causative suffix -t to the verb stem. In (2b), however, the causative morpheme introduces the causee *adam-a* ‘man-DAT’ and suffixes onto the transitive base *yıka* ‘wash’ rendering it ditransitive.

- | | |
|---|--|
| (1) a. Çikolata eri-di.
chocolate melt-PAST.3g
‘The chocolate melted.’ | b. Kız çikolata-yı eri-t-ti.
girl chocolate-ACC melt-CAUS-PAST.3g
‘The girl melted the chocolate.’ |
| (2) a. Kadın araba-yı yıka-dı.
woman car-ACC wash-PAST.3g
‘The woman washed the car.’ | b. Kadın araba-yı adam-a yıka-t-tı.
womancar-ACCman-DATwash-CAUS-PAST.3g
‘The woman had the man wash the car.’ |

Unlike in English, where verbs can undergo causative alternation, in Turkish, verbs exhibit fixed-transitivity: that is, for the most part, verbs are exclusively intransitive or transitive, and causativization is expressed morphologically through affixation. There are four distinct suffixes in Turkish that convey a causative function.¹ Type-wise the most frequent causative suffix is -DİR and it is attached to the majority of monosyllabic verbs (both vowel- and consonant-final) and non-liquid-final consonant-ending multisyllabic verbs. In Turkish, the vocalic and consonantal changes that suffixes exhibit are indicated with capital letters. While D in -DİR indicates that the suffix undergoes voicing harmony and complies with the voicing features of the final consonant of the stem, I in -DİR indicates that it complies with high vowel harmony. The variants of the causative suffix -DİR are exemplified in (3):

- | | |
|--------------------------------|----------------------------------|
| (3) verb causative | verb causative |
| ye -dir ‘make someone eat’ | kes -tir ‘make someone cut’ |
| kaşın -dır ‘make someone itch’ | dolaş -tır ‘make someone walk’ |
| dol -dur ‘fill’ | konuş -tur ‘make someone talk’ |
| öl -dür ‘kill’ | dök -tür ‘have something thrown’ |

Typewise, the second most frequent causative suffix in Turkish is -t. All vowel-final and liquid-final (i.e., stems ending in /r/ and /l/) multisyllabic stems select this affix. Examples are in (4):

- | |
|--------------------------------|
| (4) verb causative |
| yakala -t ‘have someone catch’ |
| süpür -t ‘have someone sweep’ |
| düzel -t ‘fix’ |

Thus far, the distribution of the causative suffixes appears to be relatively rule-governed; however, the presence of two sets of consonant-final monosyllabic verbs that do not select -DİR obscures the picture and renders the distribution intriguing from an acquisition perspective.

¹In addition to the morphological causative forms, Turkish has periphrastic causative structures like *neden olmak* ‘to give rise to something’; *sağlamak* ‘to cause something’. The present study focuses on non-periphrastic causativization.

The first set of non-DIr selecting verbs is formed causative by appending the suffix -I/Ar. The -Ir variant, i.e., [-ır, -ir, -ür, -ur] follows the high-vowel harmony rule (5a) and the -Ar variant, i.e., [-ar, -er], the non-high vowel harmony rule (5b), whereby the suffix vowels have to agree with the stem vowel in backness and rounding.

(5) a. Verb causative

bat	-ır	'sink (tr.)/ prick (tr.)'
iç	-ir	'make someone drink'
düş	-ür	'drop'
duy	-ur	'make oneself heard'

b. Verb causative

çık	-ar	'take off'
kop	-ar	'pull off'

The second set of non-DIr selecting monosyllabic verbs is appended the causative suffix -It, which also complies with the high-vowel harmony rule. Examples are in (6).

(6) Verb causative

ak	-it	'cause to flow'
sark	-it	'dangle (tr.)'
kok	-ut	'give off a stench'
kork	-ut	'frighten; scare'

As the overall type count of the suffixes may have a defining role in the path the child follows in learning the causative, to provide the full type count of the causative affixes we consulted Nakipoğlu and Üntak (2008). According to this study, of the 4700 verbs in Turkish, 4479 are multisyllabic and 221 are monosyllabic roots. While the overwhelming majority of monosyllabic verbs are attached the causative suffix -DIr; 24 monosyllabic verbs select the suffix -I/Ar and 5 select the causative suffix -It. Thus, the two causative affixes that apply to non-DIr selecting monosyllabic verbs have a very low type frequency. Of the multisyllabic verbs, while all non-liquid-final consonant-ending verbs are attached the suffix -DIr (~3140); vowel-final- (~1220) and liquid-final- (~120) multisyllabic stems are attached the affix -t in causative formation.

In addition to the four morphological causative forms – i.e., -DIr, -t, -I/Ar, -It – there are also a handful of suppletive forms (6 in total) in Turkish in which a different lexical item is used to describe the non-causative and the causative events (see (7)).

(7) yan	'burn (int.)'	yak	'burn (tr.)'
gör	'see'	göster	'show'
kalk	'get up'	kaldır	'lift up'

What is intriguing from an acquisition perspective is the multitude of suffixes that serve as markers of causativization and the absence of unambiguous cues for dissociation – in particular, in the case of monosyllabics. Therefore, given this many-to-one mapping between form and function and the type count of the affixes, we conjecture that the acquisition of the Turkish causative cannot be rapid and error-free. Apart from a small corpus study where Ketrez (1999) investigated the causative use in the spontaneous speech samples of three girls (age range: 1;6–3;3),

acquisition of causative has not been explored in Turkish. Noticeably, our scan of this corpus revealed 33 causative types (12 -Dİr; 11 -I/Ar; 6 -t and 4 suppletives), and 89 causative tokens, where -I/Ar tokens outnumbered -Dİr tokens (33% -I/Ar tokens vs. 23% -Dİr tokens). Children's early verbs contained neither any It-taking verbs, nor morphological overregularizations. With that in mind, we now turn to the predictions of this first ever production study on the Turkish causative and set out to explore the way the child's learning mechanism imposes regularity on the causative.

The present study

In the present study, we aimed to investigate Turkish-speaking children's acquisition of the causative to examine their overregularization and irregularization errors in relation to the analogy-based and rule-based models of morphology acquisition. A secondary goal in this paper is to investigate children's syntactic overgeneralizations with respect to whether the errors reflect the use of intransitives for transitives. For these purposes, 3- to 10-year-old children participated in an elicited production task.

Given the distribution of the causative affixes in Turkish, the task facing the child can be broken down into two processes: the child has to figure out that causativization is achieved by affixation; and that there are variant affixes that take part. To untangle the distribution of the causative affixes the child has to make a set of inferences. It is through these inferences we can formulate specific hypotheses about how the causative puzzle can be solved. Here is what we conjecture with respect to the path the child may follow in learning the Turkish causative:

- i. Two of the causative affixes, -Dİr and -t, are massively frequent in Turkish and in tackling the distribution of these affixes, the child has to roughly pay attention to the consonant-final stems for the former (excluding the liquid-final multisyllabic stems); and the vowel- and liquid-final multisyllabic stems for the latter. As the distribution of these affixes is phonologically conditioned for the most part, we expect children to rapidly map these forms to causative function and use them relatively competently, i.e., Dİr-taking mono- and multi-syllabic verbs and t-taking verbs may not yield much erroneous use.
- ii. The causative affixes -I/Ar, and -It, however, are few in type and apply only to a small set of monosyllabic verbs. Thus, in learning the Turkish causative, the child has to puzzle over how -I/Ar and -It selecting verbs differ from monosyllabic verbs that are rendered causative by adding -Dİr. Evidently, the child has to tackle why some monosyllabic verbs with almost identical root-final rhymes, (i.e., nucleus and coda) as in (8) like *bat* 'sink' and *at* 'throw' surface as *bat-ır* but *at-tır* or *ak* and *tak* surface as *ak-it* but *tak-tır*.

- (8) a. *bat* 'sink (int.)' *bat-ır* 'sink (tr.)'
 at 'throw' *at-tır* 'make someone throw'
 b. *ak* 'flow' *ak-it* 'make flow'
 tak 'attach' *tak-tır* 'have someone insert'

Adopting the key assumptions of the theory of structural alignment invoked to explain similarity in various cognitive domains (Gentner, 1983; and, in particular, for word similarity, Hahn & Bailey, 2005), we posit that in learning the causative, to untangle the causative use, the child has to engage in similarity comparisons between verbs. As similarity between verbs would depend on comparisons between aligned

parts of the roots, the rhyme would stand out as the aligned part. Thus, following a similarity-driven path in early acquisition, the child would notice that verbs sharing the same root-final rhyme like *bat* 'sink' and *at* 'throw' are causativized with the suffixes, -Ir and -Dİr respectively, as in (8a); whereas verbs like *ak* 'flow' and *tak* 'attach' are causativized with the suffixes -It and -Dİr as in (8b). Crucially, through similarity comparisons the child would realize that a monosyllabic verb can be suffixed three variant causative morphemes, and the phonological features of the root-final consonants do not provide a clear-cut cue for a dissociation among the suffixes. Thus by engaging in similarity comparisons between verbs, the child can construct a hypothesis space on the basis of the syllable-count of the verb and the type-count of the causative affixes. As a result she may deem -Dİr the default causative affix as it attaches to the majority of roots both monosyllabic and multisyllabic. Therefore, in learning the Turkish causative the child may use non-Dİr-selecting monosyllabic verbs (9a) and (9b), and suppletives (9c) with the affix -Dİr yielding overregularization errors.

(9) a. I/Ar > *Dİr	<i>iç-ir</i>	'make drink'	>	* <i>iç</i>	-tir
	<i>duy-ur</i>	'announce'	>	* <i>duy</i>	-dur
b. It > *Dİr	<i>ak-ıt</i>	'make flow'	>	* <i>ak</i>	-tır
	<i>kork-ut</i>	'scare'	>	* <i>kork</i>	-tur
c. <i>yan</i> 'burn (intr.)/ <i>yak</i> 'burn (tr.)'			>	* <i>yan</i>	-dır
<i>kalk</i> 'get up' (intr.)/ <i>kaldır</i> 'lift up (tr.)'			>	* <i>kalk</i>	-tır

iii. Variant causative forms with an ambiguous phonological distribution within monosyllabic verbs may give rise to competition among affixes and yield erroneous productions with Dİr-selecting verbs as well. Any non-Dİr suffix appearing on a Dİr-selecting verb can be considered an irregularization error. Thus, if the child draws an erroneous generalization on the basis of the syllable count of the verbs and hypothesizes that monosyllabic verbs may select -It or -I/Ar rather than -Dİr when causativized, errors on Dİr-selecting verbs like (10a) or (10b) would be conceivable.

(10) a. Dİr > *It (-It irregularization)	<i>öp-tür</i>	'make kiss'	>	* <i>öp</i>	-üt
b. Dİr > *I/Ar (-I/Ar irregularization)	<i>aç-tır</i>	'make open'	>	* <i>aç</i>	-ır

iv. Similarly the presence of three causative morphemes (i.e., -Dİr, -I/Ar and -It) that can be suffixed to a monosyllabic verb and the absence of any phonological cues that can assist the child in dissociating among the suffixes may lead the child to erroneously conclude that -I/Ar-taking verbs may be formed causative by attaching the -It affix, or -It verbs by attaching the -I/Ar affix, yielding other irregularization errors as in (11):

(11) a. I/Ar > *It	<i>uç-ur</i>	'fly'	>	* <i>uç</i>	-ut
	<i>iç-ir</i>	'drink'	>	* <i>iç</i>	-it
	<i>doğ-ur</i>	'bear'	>	* <i>doğ</i>	-ut
b. It > *I/Ar	<i>sark-ıt</i>	'dangle'	>	* <i>sark</i>	-ır
	<i>ak-ıt</i>	'flow'	>	* <i>ak</i>	-ır

v. There may be a negative correlation across verbs between the overregularization/irregularization rates and the token counts. High-frequency verbs may be resistant to errors, whereas low-frequency verbs may be more error prone.

vi. To unpack whether morphological errors are analogy-driven or rule-driven, one needs to investigate the overregularization/irregularization errors in terms of whether they appear to be produced by analogizing across the input; or abstracting across the input based on type frequency. The learning system may opt for statistically tracking how often a particular rhyme – for instance, the rhyme [at] as in *bat-ır* vs. *at-tır* in (8) above – occurs in the environment of -ır vs. -Dir. Many analogous -Dir exemplars like *k[at]-tır* ‘make add’, *s[at]-tır* ‘make sell’, *t[at]-tır* ‘make taste’, etc., may exert analogic pressure on the root *bat-*, yielding the overregularization, **bat-tır*. Thus, if overregularization errors are mostly observed in instances where a particular root-final rhyme has many -Dir neighbors (or irregularization errors are observed in instances where root-final rhymes have -I/Ar or -It neighbors), then the generalization mechanism the child implements and the resulting errors are likely to be analogy-driven. Another possible line of analysis for the learning system may be to seek ways of reducing variation in the data. A comparison of the pairs *b[at]-ır* / *[at]-tır* vs. *[ak]-ıt* / *t[ak]-tır* may suggest that -Dir is the majority variant and the system may thus infer that adhering to the type count of -Dir would provide the simplest encoding of the data. This would then suggest that the generalization mechanism is rule-driven based on type and the overregularization errors are products of the application of the rule *add -Dir* on roots that do not take -Dir.

vii. Finally, we conjecture that in learning the argument structure of Turkish – as verbs are for the most part exclusively intransitive or transitive (hence, do not exhibit causative alternation) – children would be expected not to overgeneralize syntactically, i.e., use intransitives to express transitive meanings (or transitives for intransitives). That said, sorting out the intricate relation between transitivity and causativity – more precisely, the difference between verbs that are inherently transitive (e.g., *yıka* ‘wash’), those that are transitive-causative (e.g., *eri-t* ‘melt (tr.)’), and those that are rendered ditransitive (hence, causative by suffixing a causative morpheme on a transitive base e.g., *yıka-t* ‘make wash’) – must be challenging for Turkish-speaking children. In particular, in early acquisition, if children assume that a transitive base is already a causative, they may refrain from attaching a causative morpheme onto the base and hence produce affix omission errors. For example, to describe a picture where a girl is having her hair cut, if the child opts for the use of the transitive base *kes* ‘cut’ instead of the causative attached verb *kes-tir* ‘have cut’ to express a causative meaning as in (12), then the child would be producing an affix omission error. In a sense the child would be erroneously overgeneralizing a transitive to express a causative meaning in a ditransitive construction.

- (12) Kız saç-ın-ı berber-e *kes-iyor / kes-tir-iyor.
 girl hair-POSS-ACC hairdresser-DAT cut- CAUS-PROG.3sg
 ‘The girl is having the hairdresser cut her hair.’

Method

Participants

115 Turkish-speaking typically developing monolingual children (58 females) ranging in age from 3;1 to 10;1 and five adult controls were included in the study. Six children, in terms of the items for which they did not provide any response, were outliers (i.e., > mean + 2.5*SD) and were eliminated from the analyses. Participants were recruited from two daycare centers and one elementary school in upper-middle class neighborhoods of İstanbul. The control group consisted of undergraduate students studying at a university in İstanbul.

Materials and procedure

We tested a total of 50 verbs using a picture-cued sentence completion task (see Appendix for the test items). The test materials of the study consisted of 68 pictures presented to the children on a laptop in random order. 18 (18 × 2 pictures) of the verbs we tested were presented first in an intransitive/ non-causative frame, followed by a transitive/ causative frame. For example, in (13a) in Figure 1 the child was first shown a picture of a baby holding a baby bottle and drinking her milk and was asked what the picture depicts, requiring the use of the transitive/ non-causative *iç* ‘drink’. Then, a picture of the causative counterpart of the same event was shown to the child where this time the mother was feeding the baby, where the context required the use of the causative *iç-ir* ‘make drink’ as in (13b).



Figure 1. Sample test item.

To elicit the remaining set of verbs (32 verbs in total), we used a single causative frame, as the intransitive counterparts of these verbs were difficult to elicit under the same semantic context. For example, we elicited the causative verb *ak-ıt* ‘make run/ flow’ by presenting the child with the picture of a girl brushing her teeth while leaving the tap water running. The causative *sus-tur* ‘silence’ was elicited by showing a picture of a teacher silencing a class of students, etc. In the experimental items, in all the events the causative verbs depicted, an animate agent made an (in)animate patient undergo an external or internal change of state. Overall, 19 items were asked to elicit -DİR, 16 items for -I/Ar, 8 items for -t, 3 items for -It, and 4 items for suppletive forms. The children were tested individually in a quiet room in the daycare centers and the elementary school. The sessions were audio-recorded for later transcription and coding.

(13) a. Non-causative frame:

Bebek süt iç-iyor.
baby milk drink-PROG.3sg

‘The baby is drinking milk.’

b. Causative frame:

Anne bebeğ-e süt
mother baby-DAT milk
iç-ir-iyor.
drink-CAUS-PROG.3sg

‘The mother is having the baby drink milk.’

The study was designed to elicit the targeted causative attached verb rather than to make children produce the entire causative construction with intact argument structure, i.e., with the accusative-marked/ null-marked direct object and the dative-marked indirect object as in (12) and (13) above. For the purposes of this study we aimed at eliciting the verb only, as we were mainly interested in finding out how children master the correct use of the causative affix.

Data coding and analysis

Two linguistics graduate students who were also experimenters in the test transcribed the data. The authors of the study later coded the transcribed data. A response was coded as correct if the child produced the correct causative marker in her response. A response was coded as incorrect if the child produced

- i. no causative marker for verbs that require an affix (omission);
- ii. -DİR for -I/Ar-taking and -It-taking monosyllabic verbs (overregularization errors);
- iii. -I/Ar for -DİR-taking monosyllabic verbs (-I/Ar irregularizations);
- iv. -It for DİR-taking monosyllabic verbs (-It irregularizations);
- v. -I/Ar for -It-taking or -It for -I/Ar-taking monosyllabic verbs (other irregularizations);
- vi. a causative affix for suppletive forms.

Sometimes children failed to respond to a question or used another verb than the target verb to respond. 402 such instances were eliminated from the data before the analyses. In case of the verb *sevin-dir* 'please' for example, 15 of the children preferred another verb (i.e., *mutlu et* 'make happy') which did not require a causative affix. Therefore, children's responses on the item eliciting this verb were eliminated from the analyses.

On some occasions, in particular with t-taking stems, children produced multiple causative affixes (e.g., *temizle-t-tir* instead of *temizle-t* 'make clean' in a sentence like *The boy made the genie clean the house* or *yıka-t-tır* instead of *yıka-t* 'have someone wash' in the sentence *The woman had someone wash the dishes*). These instances were deemed as correct responses as adults also often produce multiple causatives in such contexts where in fact a single causative affix would suffice. While multiple causative use was rarely observed in stems that are attached affixes other than -t, somewhat higher rates of multiple causative use in t-selecting stems suggests that in learning the causative – perhaps by analogy with frequent roots like *kapa/ kapa-t*, both meaning 'close (tr.)', or *çık-ar/ çık-ar-t* meaning 'take off (tr)' in Turkish – children erroneously assume that the affix -t is part of the verb root: hence, an extra causative morpheme (i.e., -DİR has to be suffixed onto the verbs for a causative reading).

In the following section, we present both descriptive analyses and mixed effects logistic regression models that enable the inclusion of both participant- and item-related factors (Baayen, 2008). Mixed effects models were built with the *lme4* package (Bates, Maechler, Bolker & Walker, 2015) in R (R Core Team, 2013). All models included by-item and by-participant random intercepts. For each analysis, a model having only random effects was compared to a model having random effects and child age as a fixed effect to test whether age improved model fit. In each model, age improved model fit and was included as a fixed effect. Post-hoc pairwise comparisons corrected with Tukey's multiple comparisons test were run with the

emmeans package (Lenth, 2020). Significance values of categorical predictors were obtained with the *simr* package (Green & MacLeod, 2016). Finally, influential data points in terms of participants and items were detected by using the *influence.ME* package (Nieuwenhuis, te Grotenhuis & Pelzer, 2012) and eliminated from the models. Influential data negatively affect the generalizability and statistical fit of a model. To decide whether a data point is influential, Cook's distance was used. Following Van der Meer, Te Grotenhuis and Pelzer (2010), cases were regarded as too influential if they had a larger Cook's distance than $4/n$, n being either the number of children in the sample or the number of verbs in a particular analysis.

Sometimes children's responses lacked a causative morpheme. These cases of omission were not entered into the mixed effects models and were analyzed in more detail at the end of the next section.

Results

Our findings yielded no evidence for an error-free acquisition path for the Turkish causative. Of the 5048 valid responses the children gave, 4578 were correct and 470 were incorrect (i.e., overregularizations, irregularizations, and omissions). A Spearman's rank correlation showed a positive association between children's age and percentage of correct responses, $r_s = .55$, $p < .001$ (Figure 2).

Errors on DIr-taking (2%), t-taking verbs (3%) and suppletive forms (3%) were relatively infrequent followed by a considerably higher number of errors on I/Ar- (13%) and It-taking verbs (21%). A mixed effects logistic regression was conducted with response (i.e., correct/ incorrect) as the outcome variable and age and verb category (i.e., DIr, I/Ar, It, t, suppletive) as predictor variables. Both age ($estimate = 0.67$, $SE = 0.11$, $p < .001$) and verb category ($p = .002$) were significant predictors.² Follow-up analyses for verb category showed that children made considerably fewer errors on DIr-taking verbs compared to I/Ar-taking verbs ($estimate = 2.66$, $SE = 0.87$, $p = .019$) and It-taking verbs ($estimate = 5.37$, $SE = 1.71$, $p = .015$). The difference in correct responses between t-taking and It-taking verbs was also significant where children gave more erroneous responses on It-taking verbs ($estimate = 6.02$, $SE = 1.95$, $p = .017$).

Overregularization errors

Children causativized non-DIr-taking verbs with the affix -DIr in various proportions. Overall, there were 248 overregularization errors. On the average, the proportion of overregularization errors was 11% for I/Ar-taking verbs, 19% for It-taking verbs, and 3% for suppletive forms. Verbs that take the -t affix were overregularized in only three instances where the verb was first intransitivized and then suffixed the causative marker -DIr as in *yıka* 'wash (tr.)' > *yıka-n* 'wash-oneself' (intr.) >

²A 9-year-old participant with a low correct response rate of .80 and the items eliciting the verbs *korkut* 'frighten' and *süpürt* 'have s.o. sweep' were found to be influential data points and removed from the data set for this analysis. The verb *süpürt* behaved differently from other verbs as the correct response rate decreased instead of increasing with age. The verb *korkut* differed from the other verbs in its category (i.e. *akıt* 'cause to flow' and *sarkıt* 'dangle (tr.)') in that children gave no incorrect responses for *korkut*. The model with an interaction term between age and verb category did not improve model fit.

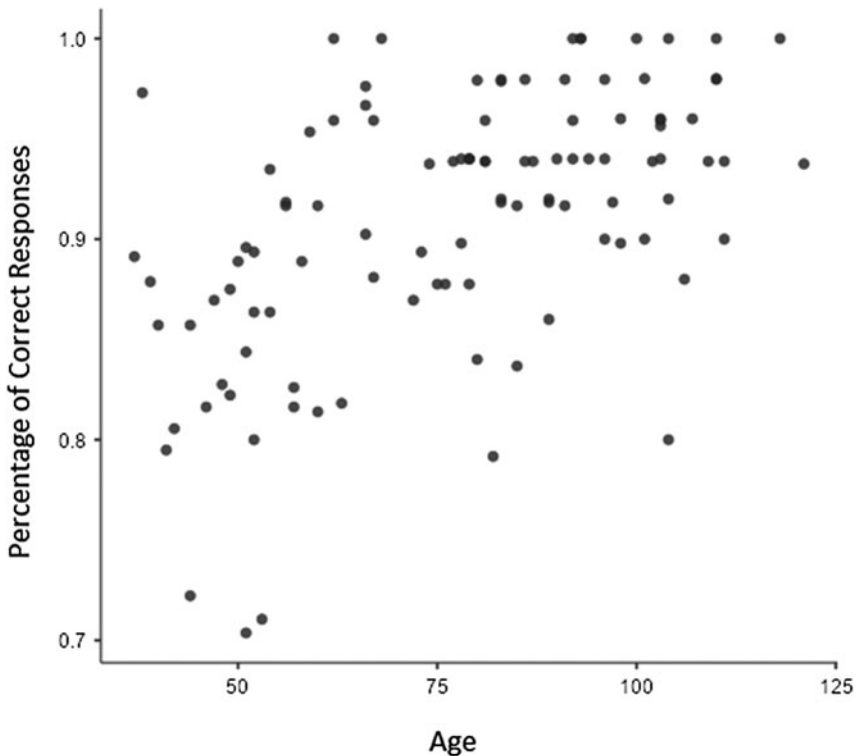


Figure 2. Scatterplot showing the relation between children's age (in months) and percentage of correct responses

**yıka-n-dır* instead of *yıka-t* 'make wash'; as there were only three such instances, these verbs were eliminated from further analyses regarding overregularizations.

Mixed effects logistic regression analyses were conducted with the presence/absence of overregularization as the outcome variable, and age and verb category (i.e., I/Ar, It, suppletive) as predictor variables. The model with age as the only predictor ($AIC = 1060$, $BIC = 1083$) was not significantly different from the model which additionally included the verb category as a predictor variable ($AIC = 1063$, $BIC = 1098$), $p = .57$. In other words, overregularization errors were similar across verb categories. However, as age increased, the rate of overregularization decreased ($estimate = -0.63$, $SE = 0.11$, $p < .001$).³

Irregularization errors

In contrast to overregularization errors, irregularization errors were rare where only 82 out of 5048 responses contained these errors. In 36 of these responses, DIr-taking verbs were produced with non-DIr causativizing suffixes (14a). In 25 cases, children made irregularization errors on I/Ar-taking verbs (14b). The remaining errors mostly

³Two 3- and 9-year-old participants with a high rate of overregularization errors (31% and 23%) were found to be influential data points and removed from the data set for this model.

included cases where t-taking verbs were irregularized (e.g., **taşı-r-t* for *taşı-t* ‘have carry’; **süpü-t* for *süpür-t* ‘make sweep’).

- | | | | | | |
|---------|---------------------|------------|----|------------------------|------------|
| (14) a. | öp-tür ‘make kiss’ | > *öp -üt | b. | iç-ir ‘drink’ | > *iç -it |
| | çiz-dir ‘make draw’ | > *çiz -it | | doğ-ur ‘give birth to’ | > *doğ -ut |
| | sus-tur ‘silence’ | > *sus -ut | | uç-ur ‘fly’ | > *uç -ut |

Furthermore, in a few instances, the causative marker -I/Ar was suffixed onto a DIr-taking verb or an It-taking verb, rendering further irregularizations. So, children produced **ölç-ür-t* for *ölç-tür* ‘have someone measure’ or **ak-ır-t* for *ak-it* ‘make flow’ where in both cases the causative marker -t was also suffixed onto the stems, resulting in multiple causative use.

Since there were no irregularization errors on suppletive forms, mixed effects logistic regression analyses were conducted on the other categories of verbs with the presence of irregularization error as the outcome variable. The model with age as the only predictor ($AIC=688$, $BIC=713$) was not significantly different from the model which additionally included verb category (i.e., DIr-, I/Ar-, t-, and It-taking) as a predictor variable ($AIC=694$, $BIC=739$), $p=.99$, indicating that irregularization errors were similar across verb categories. With increasing age, irregularization errors decreased ($estimate = -0.50$, $SE = .18$, $p = .006$).⁴

Errors in relation to verb frequency

We expected children to produce more overregularization errors with less frequent verbs. To test this hypothesis, we gathered frequency information about verbs in child-directed speech via examining the Koç University Longitudinal Language Development Database (KULLDD) (Küntay, Koçbaş & Taşçı, 2015). KULLDD is a longitudinal corpus consisting of video recordings of eight children interacting with their caregivers (e.g., mother, grandmother) in their home environment. Between 8 and 36 months of age, families were visited twice a month for 1-hour video-recordings of daily activities (e.g., playing, eating). We searched the corpus for all I/Ar- and It-taking verbs and suppletive forms, and computed the frequency of these verbs in child-directed speech. Supporting our hypothesis, a Spearman’s rank correlation analysis showed that children in our study made more overregularization errors on less frequent verbs in child-directed speech, $r_s = -.62$, $p = .002$. Figure 3 plots the 23 verbs we tested (19 non-DIr verbs and 4 suppletives) and demonstrates the relationship between verb frequency and the rate of overregularization errors.

Overregularizations: are errors analogy-based or rule-based?

The non-DIr-taking verbs we tested in the present study varied in how often they were overregularized. Of the 19 non-DIr-taking verbs, 5 -I/Ar and one -It verbs (i.e., *bit-ir* ‘finish’, *çık-ar* ‘take off’, *piş-ir* ‘cook’, *şiş-ir* ‘blow’, *uç-ur* ‘fly’, and *kork-ut* ‘frighten’) did not yield any overregularizations, and another five verbs rendered erroneous use in very small proportions (i.e., *düşür* ‘drop’ (1%), *bat-ır* ‘prick’ and *kop-ar* ‘pull off’

⁴Two 8- and 9-year-old participants with irregularization rates of 7% and 4% were found to be influential data points and removed from the data set for this model.

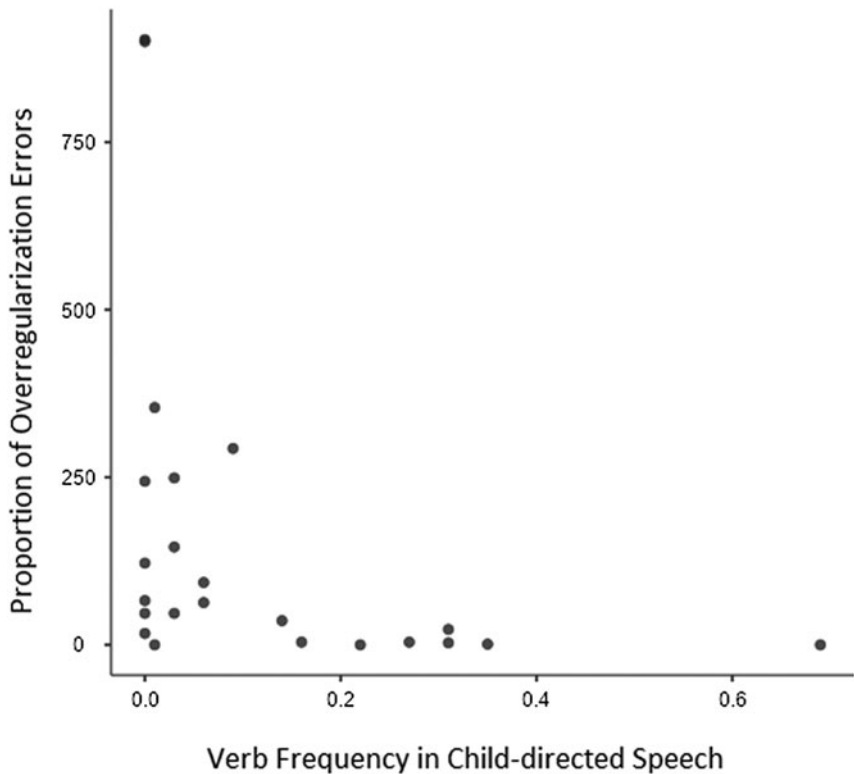


Figure 3. The relationship between verb frequency and children's overregularization errors

(3% each) and *yat-ır* 'put to bed' and *geç-ır* 'get through' (6% each)). The remaining 8 verbs, however, yielded high rates of overregularization errors (see Table 1).

To bear on the question of whether the generalization at issue is analogy-driven, we explored the type and token counts of the non-DIr vs. -DIr exemplars of the root final VC(C) sequences for the eight verbs that yielded high rates of overregularizations. Two of these verbs, i.e., *s[ark]-ıt* and *t[aş]-ır*, overregularized with rates of 35% and 27% respectively, were observed not to have any -DIr neighbors; a finding that casts doubt, even rules out any possible effect of analogy on the resulting overregularization errors. The rhymes, which the remaining six verbs fall into, are listed in Table 2. The frequency counts for the -I/Ar and -DIr exemplars are obtained from *TS Corpus V2* (Sak, Güngör & Saraçlar, 2008), a corpus of adult language with more than 491 million tokens.

As Table 2 illustrates, the rhymes (1), (5) and (6), i.e., *d[uy]*, *d[oğ]* and *[iç]* with single -I/Ar and -DIr types and more frequent -I/Ar tokens, (for example, 70,762 vs. 15,768 in the case of *duy-ur* vs. *uy-dur*) render any potential role of analogy dubious. Of the remaining three verbs, while type-wise -DIr has the upperhand in each case, token-wise -I/Ar appears to be more prevalent in *d[oy]-ur* (rhyme (2)) and *k[aç]-ır* (rhyme (4)) and the token counts of -It and -DIr exemplars are quite close (i.e., 4264 vs. 4422) in *[ak]-ıt* rhyme (3). To rule out any further possible role of analogy, we expanded the search domain to multisyllabic verbs. Quite remarkably,

Table 1. -Dir overregularizations on non-Dir-taking verbs.

Verb	Overregularization	Error rate
duy -ur	*duy-dur	69%
sark-it	*sark-tir	35%
doy-ur	*doy-dur	31%
ak-it	*ak-tir	31%
taş-ır	*taş-tır	27%
kaç-ır	*kaç-tır	22%
doğ-ur	*doğ-dur	16%
iç-ır	*iç-tır	14%

apart from the rhyme [aş] in the verb *taş-ır* ‘make overflow’ that lacked a -Dir neighbor on monosyllabic verbs, our search yielded no -Dir exemplars on multisyllabic bases in Turkish sharing the root-final rhymes listed in Table 2⁵ (see Appendix B for a list of the rhyme sets that I/Ar-taking verbs fall into and the frequency of neighboring monosyllabic and multisyllabic -Dir types). These results lend further credence to the conclusion that overregularization errors are not analogy-driven, rather they are rule-driven and are products of the child’s gradual abstraction across the input.

Omission of the causative marker

We observed a total of 140 affix omissions in the entire data. A Spearman’s rank correlation showed that age was negatively associated with the rate of omissions ($r_s = -0.40$, $p < .001$) indicating that mostly younger children committed omission errors. Next, we examined whether children tended to omit the causative affix for transitive verbs to a greater extent compared to intransitive ones. Thirteen out of 19 Dir-taking, 3 out of 16 I/Ar-taking, 1 out of 3 It-taking, and all t-taking verbs had transitive bases. Suppletive forms were not entered into this analysis as they do not have an overt causative marker. In the analysis model, age, verb base (i.e., transitive vs. intransitive) and the interactions between these two variables were taken as predictor variables and whether the causative affix was omitted was taken as the outcome variable. A likelihood ratio test showed that the model including the interaction term between age and verb base ($AIC = 829$, $BIC = 867$) provided better fit than a simpler model without the interaction term ($AIC = 834$, $BIC = 866$), $\chi^2(2) = 7.00$, $p = .008$. However, according to the preferred model, only age ($estimate = -2.12$, $SE = 0.53$, $p < .001$) and verb base ($estimate = -1.64$, $SE = 0.78$, $p < .036$) were significant predictors. These results showed that children tended to omit the causative affix to a greater extent if the base to which the affix is to be attached was transitive.

⁵Though at first blush, the overregularization errors on the rhyme *t[aş]-ır*, with 6 -Dir exemplars on multisyllabic bases like *dol[aş]-tır* ‘take for walk’, *uğr[aş]-tır* ‘cause extra work’, etc. and a total token frequency of 43,300 – seven times more than -I/Ar tokens (5934) – can be argued to stem from analogous -Dir neighbors, that another -I/Ar verb, i.e., *d[üş]-ür* with 6 -Dir exemplars again and an overall 81,588 -Dir tokens yielded an error rate of 1% renders this possibility contentious.

Table 2. Non-Dlr and Dlr types/ tokens for rhyme sets that yielded the most overregularization errors.

Rhyme	Error rate (%)	I/Ar	Token	Dlr	Token	Type ratio I/Ar vs. Dlr
1. [uy]	69	duy-ur	70,762	uy-dur	15,768	1/1
2. [oy]	31	doy-ur	6626	koy-dur	1867	1/3
				oy-dur	17	
				soy-dur	197	
					2081	
3. [ak]	31	ak-it	4264	bak-tır	971	1/5
				çak-tır	1759	
				kak-tır	53	
				tak-tır	1356	
				yak-tır	283	
					4422	
4. [aç]	22	kaç-ır	68,758	aç-tır	3437	1/2
				saç-tır	2	
					3439	
5. [oğ]	16	doğ-ur	27,157	boğ-dur	148	1/1
6. [iç]	14	iç-ir	1897	biç-tır	40	1/1

Note. Shaded areas indicate higher token frequency.

Discussion

Research into the learning puzzle the Turkish causative presents and how children constrain their grammar to recover from errors has been deeply instructive in many ways. First and foremost, the rare case of irregularity the causative exhibits, especially in a language that behaves regularly for the most part, revealed that when there is many-to-one mapping between form and function, acquisition is demanding and is nowhere close to being error-free, countering what has in general been suggested for the learning of Turkish (Aksu-Koç & Slobin, 1985). Overall, with a large sample of children ranging from three to ten years of age, we found that children's overregularization errors were frequent, declined with age, were not analogy-driven, and were negatively associated with the verb frequency in child-directed speech.

Our results show that the acquisition of the Turkish causative with four variant affixes and no clear-cut phonological cues for dissociation (in particular, for monosyllabic verbs) proves to be challenging for Turkish-speaking children. Adult-like causative use was not observed even in the oldest children tested, and children in all ages entertained competing hypotheses regarding which causative morpheme goes with which verb, yielding erroneous productions.

Supporting our first hypothesis, the causative findings suggest that at the earliest age we were able to elicit the causative attached verbs, the child was already dissociating

between consonant-final and vowel-/liquid-final multisyllabics, successfully suffixing the former with -Dİr and the latter with -t: hence, the noticeably few errors on these verbs. Sensitivity to phonology from early on thus appears to render a sizable portion of Turkish verbs within the reach of children in terms of the licit causative use.

The causative use on the monosyllabic bases, however, proved to be puzzling for an extended period of acquisition. In line with our second hypothesis, the findings clearly suggest that to detect the irregularity in the distribution of the suffixes the child initially engages in similarity comparisons between monosyllabic verbs. The contrast between *bat-ır* vs. *at-tır* and *ak-it* vs. *tak-tır*, for example, not only directs the child's attention to the variant causative suffixes but also helps her to restrict the hypotheses domain regarding the irregularity with monosyllabics. Thus in early acquisition the learning system appears to be driven by similarity to pin down the variant suffixes that take part and to restrict the domain in which they occur. Thereafter, the child is helpless as there are no phonological cues that can guide the learning path for the correct use of causative on the monosyllabic verbs. As a result, the child entertains all the logically possible interpretations of the data; that is, she tries on suffixing -Dİr on non-Dİr-taking verbs, yielding overregularizations; and on rare occasions, -It on non-It taking verbs and -I/Ar on non-I/Ar taking verbs, yielding irregularizations.

As we have conjectured, the affix -Dİr with its high type frequency was deemed to be the default causative affix by the children. This inductive step resulted in licit use for Dİr-taking monosyllabics but overregularizations for non-Dİr-taking ones. The proportions of Dİr-overregularizations on It-taking (19%) and I/Ar-taking verbs (11%) clearly suggest that to reduce the variation in the input, children overextended the use of -Dİr to non-Dİr selecting verbs and produced Dİr-overregularization errors like **doğ-dur* for *doğ-ur* 'give birth to' and **ak-tır* for *ak-it* 'make flow'. Furthermore, in line with our hypotheses regarding irregularization errors (Predictions (iii) & (iv)) children also produced irregularizations like **öp-üt* for *öp-tür* 'make kiss' and **ölç-ür-t* for *ölç-tür* 'have measure' on Dİr-taking verbs. Although compared to overregularizations the rate of irregularizations was low, erroneous use clearly suggests that in early acquisition (just as in the case of overregularization errors, to tackle the irregularity in the causative), children narrow down the hypothesis space with monosyllabic verbs – and, until they deem non-Dİr taking verbs as exceptions to be memorized, they entertain competing hypotheses that naturally yield erroneous use.

The findings of the present study further showed that children's overregularizations were negatively correlated with token frequency. The two verbs that yielded the most erroneous use i.e., *duyur* (error rate 69%) and *sarkıt* 'dangle' (35%) did not show up in the speech directed to children (age range: 8–36 months) in the KULLDD corpus that we examined. Erroneous use that could have been impossible to capture by studying naturalistic corpora alone underscores the importance of running elicitation tasks on children. Thus when especially tested on low-frequency verbs, children productively applied the *add -Dİr* rule and produced **duy-dur* for *duy-ur* or **sark-tır* for *sark-it*, etc. Though we do not have corpus evidence that shows the input Turkish children receive in later years, the persistence of errors in older children suggests that these verbs continue to occur rarely in the input the child receives; consequently it takes longer for children to retreat from overregularization errors.

Any overregularization error implicates a generalization mechanism underlying children's productivity and, in principle, the mechanism can be rule-driven or analogy-driven. To unpack to what extent children's overregularization errors reflect

the use of a symbolic morphological rule (i.e., *add -Dir*) or analogy across stored -Dir exemplars, as spelled out in prediction (vi), we explored whether errors were mostly observed in verbs that share the same root-final rhyme with stored -Dir exemplars or not. Crucially, we found no evidence that would suggest that the overregularization errors are analogy-driven; thus, the errors cannot be viewed as stemming from the intrusion of analogous -Dir exemplars on monosyllabic bases. The errors rather suggested that in learning the Turkish causative, children appear to impose regularity on the input by defaulting to the -Dir type in the production of non-Dir-taking verbs.

In line with the findings in the existing literature, learning of the Turkish causative shows that children productively abstract across the input based on the type count of the affix -Dir. Thus, high proportions of overregularizations on non-Dir taking verbs (as opposed to few errors on -Dir-taking verbs) suggest that, over the course of acquisition, irregularly behaving verbs (i.e., non-Dir-taking verbs) yield to an emerging rule. The causative findings are thus consistent with the observations that morphological patterns are extended on the basis of type frequency (Albright & Hayes, 2003; Bybee, 1995, 2001; Plunkett & Marchman, 1993).

Taken together, the causative findings support a model of morphology acquisition that combines various aspects of the analogy-based and rule-based models. We thus suggest that an integrated model of analogy and rules should invoke manipulation of analogy through similarity comparisons in the earlier stages of acquisition so that the system moves beyond rote-learning and that it should invoke rule-induction on the basis of type-frequency in later stages. While application of the rule strengthens the memory representation of the licit forms in the input, the resulting overregularization errors, like **duy-dur*, can only be pruned by tracking the token frequency of the unattested **duy-dur* against the attested *duy-ur* so that the learning system pre-empts **duy-dur*. Therefore, the overgeneralize-then-recover pattern that emerges in the acquisition of causative, perhaps in the acquisition of morphology in general, suggests adherence to an inductive learning model that uses analogy-based, rule-based and statistical learning procedures.

In a nutshell, the key finding of our study is that in tackling the unconditioned phonological variation the causative exhibits, children generalize beyond what they witness in the input. How do children generalize? They set up hypotheses that are possible but unattested interpretations of the data. The hypotheses the children entertain yield erroneous use in the form of morphological overregularizations and irregularizations. In a sense, the child uses an implicit learning system: that first constructs morphological forms that never appear in the input, and then tracks the frequency of these unattested forms to see whether the input would support them. While licit forms are supported by the input, illicit forms are not: hence, they are pruned over the course of acquisition. Therefore errors are essential inductive leaps in the learning path of the causative and provide landmarks so that the child can statistically track the distribution of the attested and the unattested forms in the input.

As causativization also lies at the boundary of syntax, the findings of this study also provided the opportunity to sift through how children tackle causative use in syntactic terms. That said, the present study attempted to elicit the causativized verb *per se* – rather than the entire causativized event where the child would be expected to describe how the event comes about, with the dative-attached causee argument and the accusative-attached affected argument. Thus, when the child was presented with the causativized event and expected to complete the sentence with the causative-attached verb, younger children tended to omit the causative suffix in the

description of the causative event. Furthermore, the omission of the affix was correlated with whether the verb base to which the child suffixed the causative marker was transitive or intransitive. As the omission errors were observed mostly on transitive rather than intransitive bases, the Turkish child in the early stages of acquisition appears to overgeneralize transitive verbs to causative constructions – deeming that a transitive verb is already a causative verb: hence, it does not have to be attached a causative marker. Affix omissions on transitives may of course be correlated with children's developing knowledge of causativity and failure to pay attention to the causee argument which renders the event causative. As discussed in Ketrez (1999), where she reports lack of causative morphology in a handful of verbs, they may also reflect children's inability to transfer the agent role from the self to others: hence, occasionally yielding the use of an intransitive in a transitive context or a transitive in a causative context.

Crosslinguistically, overgeneralizations of causatives show an intransitive-to-causative bias, i.e., children tend to use intransitive verbs to express causative meanings in Hebrew, Inuktitut, Japanese, Quechua and TSM. In Turkish, affix omissions on intransitive bases were rare: hence, our findings do not support an intransitive-to-causative bias. However, as children extended the use of transitives to ditransitives and omitted the causative affix, in line with the most crosslinguistic findings, the pattern of extension is asymmetrical. That said, the complete picture behind the acquisition of the syntactic aspects of causativization awaits future research.

We believe the present study underscores how important it is to run elicited production experiments on children to get at the heart of the inductive learning mechanism that operates in acquisition. In naturalistic corpora studies, errors may go unnoticed. Therefore a more insightful account of what gets grammar off the ground should be pursued by running production studies with a large sample size from a wide age, as we have attempted to do in this study.

Differences in the presentation of different verbs can be considered as a limitation of the study as presenting some verbs in a non-contrastive context (e.g. *ak-ıt* 'make flow', *sus-tur* 'silence') may have led to more omission errors, particularly for younger children. However, this explanation was ruled out by an additional analysis showing no effect of the type of presentation (i.e., elicitation via one picture vs. two pictures) on children's omissions when the type of presentation and age were entered as predictor variables ($p = .064$).

Although children's regularization behavior in learning their native languages has been at the center of attention since Berko (1958), a growing body of research has started inquiring into the regularization bias in cognition in general (Ferdinand, Kirby & Smith, 2019; Perfors, 2016; Realı & Griffiths, 2009; among many others). The present study uncovering the regularization behavior of Turkish children in acquiring the causative makes a modest contribution to this inspiring literature.

Acknowledgements. This work was supported by Boğaziçi University Research Fund Grant to Mine Nakipoğlu (Grant No: 7949). Part of the present study was presented at BUCLD 35. We thank Aylin C. Küntay for sharing the Koç University Longitudinal Language Development Database.

References

- Aksu-Koç, A., & Slobin, D. (1985). The acquisition of Turkish. In D. Slobin (Ed.) *The cross-linguistic study of language acquisition* (Vol. I, pp. 839–880). Hillsdale, NJ: Lawrence Erlbaum.

- Albright, A., & Hayes, B.** (2003). Rules vs. analogy in English past tenses: a computational/experimental study. *Cognition*, 90(2), 119–161. doi: 10.1016/s0010-0277(03)00146-x
- Allen, S. E.** (1998). Categories within the verb category: learning the causative in Inuktitut. *Linguistics*, 36, 637–669.
- Ambridge, B., Pine, J. M., Rowland, C. F., Chang, F., & Bidgood, A.** (2013). The retreat from overgeneralization in child language acquisition: word learning, morphology, and verb argument structure. *Cognitive Science*, 4, 47–62.
- Baayen, R. H.** (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge, UK: Cambridge University Press.
- Bates, D., Maechler, M., Bolker, B., & Walker, S.** (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48. doi:10.18637/jss.v067.i01.
- Berko, J.** (1958). *The child's learning of English morphology*. Ph. D. Thesis. Radcliffe College.
- Berman, R.** (1993). Marking of verb transitivity by Hebrew-speaking children. *Journal of Child Language* 20, 641–669.
- Bowerman, M.** (1982). Evaluating competing linguistic models with language acquisition data: Implications of developmental errors with causative verbs. *Quanderni di Semantica* 3, 5–66.
- Bowerman, M., & Croft, A.** (2007). The acquisition of English causative alternation. In M. Bowerman & P. Brown (Eds.), *Crosslinguistic perspectives on argument structure: Implications for learnability*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Braine, M. D. S., & Brooks, P. J.** (1995). Verb argument structure and the problem of avoiding an overgeneral grammar. In M. Tomasello & W. E. Merriman (Eds.) *Beyond names for things: Children's acquisition of verbs* (pp. 353–376). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Brooks, P. J., & Tomasello, M.** (1999). How children constrain their argument structure constructions. *Language*, 75 (4), 720–738.
- Brooks, P. J., & Tomasello, M., Dodson, K., & Lewis, L. B.** (1999). Young children's overgeneralizations with fixed transitivity verbs. *Child Development*. 70(6), 1325–1337.
- Brooks, P. J., & Zizak, O.** (2002). Does preemption help children learn verb transitivity? *Journal of Child Language*, 29, 759–781.
- Bybee, J.** (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, 10(5), 425–455. doi: 10.1080/01690969508407111
- Bybee, J.** (2001). *Phonology and language use*. (Cambridge Studies in Linguistics 94) Cambridge: Cambridge University Press.
- Bybee, J., & Slobin, D.** (1982). Rules and schemas in the development and use of the English past tense. *Language* 58, 265–289. doi: 10.2307/414099
- Bybee, J., & Hopper, P. J.** (2001). *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins.
- Clark, E.** (1993). *The lexicon in acquisition*. Cambridge: Cambridge University Press.
- Courtney, E.** (2002). Child acquisition of Quechua causatives. *First Language*, 22, 29–71.
- Ferdinand, V., Kirby, S., & Smith, K.** (2019). The cognitive roots of regularization in language. *Cognition* 184: 53–68.
- Gentner, D.** (1983). Structure Mapping: A theoretical framework for analogy. *Cognitive Science*, 7, 155–170. doi: 10.1207/s15516709cog0702_3
- Goldberg, A.** (2006). *Constructions at work: the nature of generalization in language*. Oxford: Oxford University Press.
- Green, P., & MacLeod, C. J.** (2016). simr: an R package for power analysis of generalised linear mixed models by simulation. *Methods in Ecology and Evolution*, 7(4), 493–498. doi: 10.1111/2041-210X.12504
- Hahn, U., & Bailey, T. M.** (2005). What makes words sound similar? *Cognition* 97, 227–267. doi: 10.1016/j.cognition.2004.09.006
- Hahn, U., & Chater, N.** (1998). Similarity and rules: distinct? exhaustive? Empirically distinguishable? *Cognition*, 65, 197–230. doi: 10.1016/s0010-0277(97)00044-9
- Ketrez, F. N.** (1999). *Early verbs and the acquisition of argument structure*. Unpublished MA thesis in Linguistics. Boğaziçi University.
- Küntay, A. C., Koçbaş, D., & Taşçı, S. S.** (2015). *Koç University Longitudinal Language Development Database as a part of the 'ACQDIV Corpus'*. retrieved from <<http://www.acqdiv.uzh.ch/en/resources.html>>.

- Lenth, R.** (2020). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.4.6. <https://CRAN.R-project.org/package=emmeans>
- Lin, H., & Tsay, J. S.** (2008). Acquiring causatives in Taiwan Southern Min. *Journal of Child Language*, 35, 467–487.
- Lord, C.** (1979). Don't fall me down': children's generalizations regarding cause and transitivity. In *Papers and Report on Child Language Development* [Stanford University], 17, 81–89.
- MacWhinney, B.** (1987). The competition model. In B. MacWhinney (ed.), *Mechanisms of language acquisition*. 249–308. Hillsdale, NJ: Erlbaum.
- MacWhinney, B.** (2004). A multiple process solution to the logical problem of language acquisition. *Journal of Child Language*, 31, 883–914.
- Marchman, V. A., Plunkett, K., & Goodman, J.** (1997). Overregularization in English plural and past tense inflectional morphology: A response to Marcus (1995). *Journal of Child Language*, 24, 767–779.
- Marcus, G. F.** (1995). Children's overregularization of English plurals: a quantitative analysis. *Journal of Child Language*: 22 (2): 447–459. doi: 10.1017/S0305000900009879
- Marcus, G. F.** (2001). *The Algebraic Mind*. The MIT Press, Cambridge, Massachusetts. doi: 10.7551/mitpress/1187.001.0001
- Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., & Xu, F.** (1992). Overregularization in language acquisition. *Monographs of the Society of Research in Child Development*, 57(4), 34–69. doi: 10.2307/1166115
- McClelland, J. L., & Patterson, K. E.** (2002). Rules and Connections in past tense inflections: What does the evidence rule out? *Trends in Cognitive Sciences* 6: 465–472. doi: 10.1016/s1364-6613(02)01993-9
- Mirković, J., Seidenberg, M. S., & Joanisse, M. F.** (2011). Rules vs. Statistics: Insights from a Highly Inflected Language. *Cognitive Science* 35(4): 638–681. doi: 10.1111/j.1551-6709.2011.01174.x
- Nakipoğlu, M., & Üntak, A.** (2008). A complete verb lexicon of Turkish based on morphemic analysis. *Turkic Languages* 12, 221–280.
- Nakisa, R. C., Plunkett, K., & Hahn, U.** (2000). A cross-linguistic comparison of single and dual-route models of inflectional morphology. In P. Broeder & J. Murre, (Eds.), *Cognitive models of language acquisition*. Cambridge, MA: MIT Press.
- Nieuwenhuis, R., te Grotenhuis, M., & Pelzer, B.** (2012). influence.ME: Tools for Detecting Influential Data in Mixed Effects Models. *R Journal*, 4(2), 38–47.
- Nomura, M., & Shirai, Y.** (1997). Overextension of intransitive verbs in the acquisition of Japanese. *Proceedings of the 28th Annual Child Language Research Forum*, ed. E.V. Clark, 233–42. Stanford: CSLI.
- Perfors, A.** (2016). Adult regularization of inconsistent input depends on pragmatic factors. *Language Learning and Development*, 12: 138–155.
- Pinker, S.** (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.
- Pinker, S.** (1999). *Words and Rules: The Ingredients of Language*. New York, NY: Harper Perennial.
- Pinker, S., & Prince, A.** (1988). On language and connectionism: analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28 (1–2): 73–193. doi: 10.1016/0010-0277(88)90032-7
- Pinker, S., & Ullman, M. T.** (2002). The past and future of the past tense. *Trends in Cognitive Sciences*, 6 (11): 456–63. doi: 10.1016/s1364-6613(02)01990-3
- Plunkett, K., & Marchman, V.** (1993). From rote learning to system building: acquiring verb morphology in children and connectionist nets. *Cognition*, 48, 21–69. doi: 10.1016/00100277(93)90057-3
- Prasada, S., & Pinker, S.** (1993). Generalization of regular and irregular morphological patterns. *Language and Cognitive Processes*, 8, 1–56. doi: 10.1080/01690969308406948
- Pye, C.** (1994). A cross-linguistic approach to the causative alternation. In Y. Levy (ed.), *Other children, other languages*. Hillsdale, NJ: Erlbaum.
- Ramscar, M., & Yarlett, D.** (2007). Linguistic self-correction in the absence of feedback: a new approach to the logical problem of language acquisition. *Cognitive Science*, 31: 927–60. doi: 10.1080/03640210701703576
- R Core Team** (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Real, F., & Griffiths, T. L.** (2009). The evolution of frequency distributions: relating regularization to inductive biases through iterated learning. *Cognition* 111, 317–328.

- Rumelhart, D. E., & McClelland, J. L.** (1986). On learning the past tense of English verbs: implicit rules or parallel distributed processing?: In McClelland J.L., Rumelhart D.E., The PDP Research Group (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 2: pp. 216–271. MIT Press. doi: 10.7551/mitpress/5236.003.0008
- Sak, H., Güngör, T., & Saraçlar, M.** (2008). Turkish language resources: Morphological parser, morphological disambiguator and web corpus. In *Advances in Natural Language Processing* (pp. 417–427). Springer Berlin Heidelberg. doi: 10.1007/978-3-540-85287-2_40 TS Corpus - <http://tscorpus.com>
- Tomasello, M.** (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Van der Meer, T., Te Grotenhuis, M., & Pelzer, B.** (2010). Influential cases in multilevel modeling. A methodological comment. *American Sociological Review*, 75, 173–178.

Appendix A

Test Items

a. Dir-taking verbs			
Items	Verb-CAUS		Gloss
1. The boy had the genie hang the clothes.	as	-tır	'have hang'
2. The mum is mounting her daughter on the horse.	bin	-dir	'mount'
3. The girl had ice-cream smudged on her dress.	bulaş	-tır	'smear'
4. The girl is having her father pull herself.	çek	-tir	'have pull'
5. He made the genie draw a dog.	çiz	-dir	'make draw'
6. The girl is having her feet touch her head.	değ	-dir	'touch'
7. Let X fill the tub with water.	dol	-dur	'fill'
8. The girl is walking her dog.	dolaş	-tır	'walk'
9. The mum is dressing the baby.	giy	-dir	'dress'
10. The clowns make the children laugh.	gül	-dür	'make laugh'
11. The girl is having her hair cut.	kes	-tir	'have cut'
12. The boy is having his father set up the tent.	kur	-dur	'have set up'
13. The woman is making the puppet talk.	konuş	-tur	'make talk'
14. The girl is having her height measured.	ölç	-tür	'have measure'
15. The woman is having her hand kissed.	öp	-tür	'have kiss'
16. The gift pleased the child.	sevin	-dir	'please'
17. The boy blew the candles out.	sön	-dür	'blow out'
18. The teacher silenced the children.	sus	-tur	'silence'
19. We stick the stickers	yapış	-tır	'stick'

b. I/Ar-taking verbs			
Items	Verb-CAUS		Gloss
1. The boy pricked the balloon.	bat	-ır	'prick'
2. The girl finished her ice-cream.	bit	-ir	'finish'
3. The girl is taking her socks off.	çık	-ar	'take off'
4. This woman has given birth to a baby.	doğ	-ur	'give birth to'
5. The boy is upset as he dropped his icecream.	düş	-ür	'drop'
6. The girl is feeding the cat.	doy	ur	'feed'
7. The girl is trying to make her mum hear her.	duy	-ur	'shout'
8. The woman is inserting the thread into the needle eye.	geç	-ir	'get through'

(Continued)

(Continued.)

b. I/Ar-taking verbs			
Items	Verb-CAUS		Gloss
9. Mom is helping the baby/having the baby drink his milk.	ıç	-ir	'make drink'
10. The scarecrows are for scaring the crows away.	kaç	-ir	'scare away'
11. The girl is picking fruits.	kop	-ar	'pick'
12. Let X cook the meal.	piş	-ir	'cook'
13. The father has blown the balloon.	şış	-ir	'blow'
14. The man caused the meal to overflow.	taş	-ir	'cause to overflow'
15. The girl is flying a kite.	uç	-ur	'fly'
16. The dad puts the child to bed.	yat	-ir	'put to bed'

c. -t taking verbs			
Items	Verb-CAUS		Gloss
1. The girl is having her mum read her a book.	oku	-t	'make read'
2. He made the genie sweep the floor.	süpür	-t	'make sweep'
3. He had the genie decorate the house with balloons.	süsle	-t	'have decorate'
4. The girl is having her hair combed.	tara	-t	'have comb'
5. The man is having his luggage carried.	taşı	-t	'have carry'
6. He had the genie clean the house.	temizle	-t	'have clean'
7. He had the Genie gather the toys.	topla	-t	'have gather'
8. The woman had her husband wash the dishes.	yıka	-t	'have wash'

d. -It taking verbs			
Items	Verb-CAUS		Gloss
1. The girl leaves the tap water running while brushing her teeth.	ak	-it	'make run/ flow'
2. The dog frightened the girl.	kork	-ut	'frighten'
3. Rapunzel let her hair dangle.	sark	-it	'dangle'

e. Suppletives			
Items	Verb-CAUS		Gloss
1. The father is showing something (to the girl).	göster		'show'

(Continued)

(Continued.)

e. Suppletives		
Items	Verb-CAUS	Gloss
2. The woman is lifting the baby up.	kaldır	'lift/raise'
3. The boy inserted his finger into the toy.	sok	'insert'
4. Let X light/set the fire.	yak	'light fire'

Sample Protocol

Let us look at the pictures. We have cartoon characters X and Y here, and there are certain chores to be done. For example, the fire has to be set; meal has to be cooked and watertub has to be filled. Let us choose who should do what.

- (i) Experimenter: Ateşin yanması gerekiyor ve yemeğin pişmesi gerekiyor.
'The fire has to be set and the meal has to be cook'

Experimenter: X ne yapsın?
'What shall X do?'

Child: Yemeğ-i **piş-ir-sin.**
meal-ACC cook-CAUSATIVE-2sg
'Let him cook the meal.'

Experimenter: Y ne yapsın?
'What shall Y do?'

Child: Ateş-i **yak-sın.**
Fire-ACC burn(suppletive)-2sg.
'Let her set the fire.'

Appendix B

The 16 I/Ar-taking monosyllabic verbs we tested in the study fall into 10 rhyme sets. The chart below lays out the type and token counts of -I/Ar and -Dir neighbors on mono- and multisyllabic verbs for each root-final rhyme. Frequency information is obtained from *TS Corpus V2* (Sak, Güngör & Saraçlar, 2008).

	Rhyme	I/Ar	Type	Token	Dir (monosyllabic)	Type	Token	Dir (multisyllabic)	Type	Token
1.	[uy]	duy-ur		70762	uy-dur	'make fit'	15768	none		
	total		1	70762			1	15768	0	0
2.	[oy]	doy-ur		6626	koy-dur	'make put'	1867	none		
					oy-dur	'make carve'	17			
					soy-dur	'make peel'	197			
	total		1	6626			3	2081	0	0
3.	[aş]	taş-ır		5934	none			kamaş-tır 'dazzle'		2294
								dolaş-tır 'take for a walk'		2310
								kaynaş-tır 'integrate'		1239
								uğraş-tır 'cause extra work'		877
								ulaş-tır 'transport'		36536
								anlaş-tır 'make accord'		44
	total		1	5934			0	0	6	43,300
4.	[aç]	kaç-ır		68758	aç-tır	'make open'	3437	none		
					saç-tır	'make spread'	2			
	total		1	68758			2	3439	0	0
5.	[oğ]	doğ-ur		27157	boğ-dur	'make strangle'	148	none		
	total		1	27157			1	148	0	0

6.	[iç]	iç-ir	1897			40	none			
	total		1	1897	biç-tir	'make cut'	1	40	0	0
7.	[eç]	geç-ir	241,123			930	none			
	total		1	241,123	seç-tir	'make choose'	1	930	0	0
8.	[at]	yat-ır	242,645			4476	none			
		bat-ır	6399		çat-tır	'make get into trouble'		2		
					kat-tır	'make add'		7		
					sat-tır	'make sell'		766		
					tat-tır	'make taste'		1246		
	total		2	249,044			5	6497	0	0
9.	[op]	kop-ar	17532		none			none		
	total		1	17532			0	0	0	0
10.	[üş]	düş-ür	81,588		none			dönüş-tür 'transform'		2640
								bölüş-tür 'divide among a group'		26
								büzüş-tür 'cause to wrinkle'		1
								dövüş-tür 'cause to fight'		54
								görüş-tür 'make people meet'		3
								sürtüş-tür 'rub things'	0	26
	total		1	81,588			0	0	6	2724