Journal of Child Language (2026), **53**, 950–971
doi:10.1017/S0305000925000200

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Children Learn Causatives Despite Pervasive Ellipsis: Evidence from Turkish

Guanghao You[1], Ebru Ger[2,3], Moritz M. Daum[2,4] and Sabine Stoll[1]

[1]Institute for the Interdisciplinary Study of Language Evolution (ISLE), University of Zurich, Zürich, Switzerland; [2]Department of Psychology, University of Zurich, Zurich, Switzerland; [3]Institute of Psychology, University of Bern, Bern, Switzerland and [4]Jacobs Center for Productive Youth Development, University of Zurich, Zurich, Switzerland
Corresponding author: Guanghao You; Email: guanghao.you@uzh.ch

(Received 22 March 2023; revised 09 December 2024; accepted 21 February 2025; first published online 21 May 2025)

Abstract

In this study, we explore how children learn causatives from the language they are exposed to in their everyday lives. Previous research has argued that argument structure is a crucial facilitator for learning causatives. Here, we examine the role of argument structure in the acquisition of morphological and lexical causatives. We use Turkish as a test case which allows argument ellipsis and ask whether reliable argument expression is a key factor in learning causatives. We use longitudinal data of 7 children spanning from 8 to 36 months of age and their surrounding environment. We show that ellipsis in child-directed speech is pervasive for both lexical and morphological causatives. However, such pervasive ellipsis does not seem to influence the acquisition process. This suggests that argument structure may not play a major role in the learning of causatives in a natural learning environment.

Keywords: Turkish; ellipsis; causative; argument structure

Özet

Bu çalışmada çocukların dildeki neden-sonuç yapılarını günlük yaşamlarında duyduklari dil üzerinden nasıl öğrendiklerini inceledik. Önceki araştırmalar, üye yapısının neden-sonuç yapılarını öğrenmede önemli bir rol oynadığını öne sürmüştür. Bu çalışmada, morfolojik ve sözcüksel neden-sonuç yapılarını öğrenmede üye yapısının rolünü inceledik. Türkçe'yi bir test vakası olarak kullandık ve üye eksiltiminin neden-sonuç yapılarını öğrenmede önemli bir faktör olup olmadığını araştırdık. 7 çocuğun 8 ila 36 ay yaş aralığında ürettiği ve duyduğu boylamsal dil verilerini kullandık. Bulgular çocuklara yöneltilen konuşmada hem sözcüksel hem de morfolojik neden-sonuç yapıları için üye eksiltiminin yaygın olduğunu gösterdi. Ancak, bu yaygın üye eksiltiminin öğrenme sürecini etkilemediği

M.M.D. and S.S. share the senior authorship.

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial licence (http://creativecommons.org/licenses/by-nc/4.0), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original article is properly cited. The written permission of Cambridge University Press must be obtained prior to any commercial use.

https://doi.org/10.1017/S0305000925000200 Published online by Cambridge University Press

gözlendi. Bu, üye yapısının doğal bir öğrenme ortamında neden-sonuç yapılarını öğrenmede önceki araştırmaların ileri sürdüğü kadar önemli bir rol oynamadığını düşündürmektedir.

1. Introduction

Causality is a key category in human cognition and language (Baillargeon et al., 1995; Corrigan & Denton, 1996; Piaget, 1927). By inferring causality, humans manage to understand the interrelations between entities (Pearl, 2009; Trabasso & Van Den Broek, 1985). The ability of causal inference is attested in human infants as early as 6 months (Leslie & Keeble, 1987). Later on, causal expressions become a crucial component of children's language. Learning how causality is expressed in language is an early challenge for first-language learners. Verbs take center stage in this process (Bartsch & Wellman, 1995; Dunn, 2010). The goal of the present study is to test which cues impact the learning of causatives in a naturalistic setting.

Languages vary in the types of verbal constructions they use to express causality. Two main types of causatives are prevalent in the languages of the world (Dixon, 2000; Shibatani, 2002). First, there are verbs, which only express causality via their inherent semantics. These verbs are known as lexical causatives. Lexical causatives commonly exist across languages. They have no specific formal marker expressing causality (e.g., "raise" in English, denoting "cause to rise"; "yak" in Turkish, denoting "set on fire," namely "cause to be on fire"). Lexical words pertaining to this category, such as "open," "close," and "raise," convey an event that causes a change of state in the patient (Shibatani, 2002). For instance, the verb "raise" in "I raise the chair" entails the event "the chair rises," that is, a change of state occurs in the patient, therefore making "raise" a lexical causative. Most studies, when defining lexical causatives, emphasise the involvement of two events and the presence of both the agent and the patient of the verb (Comrie & Polinsky, 1993; Haspelmath, 2001). To learn about the common meaning of lexical causatives, children have to rely on the immediate linguistic context and the extra-linguistic context.

Hence, one would expect that the processing of lexical causatives, by definition, fundamentally relies on the recognition of the two arguments in the utterance, especially for the patient who experiences the change. Acquiring lexical causatives can thus be a difficult process, where children have to disentangle the agent–patient information in the speech they hear to achieve causal understanding. Second, some verbs formally mark causation, usually by affixation. Turkish, for instance, renders non-causative verbs into causatives by adding a suffix (e.g., the suffix *-Ar* changes *çık* "go off" to *çıkar* "take off"). Such causativised verbs are called morphological causatives, which exist in many morphologically complex languages, such as Japanese, Korean, Inuktitut, and Turkish. Children acquiring these languages thus have access to the additional formal marking to infer the causal meaning.

The two types of causatives provide different learning challenges. Lexical causatives can be challenging because the causal meaning is implicit and the meaning needs to be inferred from verb external cues such as the extralinguistic context or the linguistic meaning of the utterance as a whole. Morphological causatives might be challenging because causative markers are usually affixes that are not necessarily salient and, hence, tend to escape notice. Thus, the question remains as to how and when children learn about causatives or, more broadly, the expression of causality. As Piaget (1955) first pointed out decades ago, understanding the causal link between events may not mature

https://doi.org/10.1017/S0305000925000200 Published online by Cambridge University Press

even for children up to 8 years of age, and such understanding is often achieved with careful guidance by adults. But how do they extract this information? It is well known that language learning is an implicit process, and also, in learning about causatives, children have to rely on their surrounding language to develop (Bonawitz et al., 2010; Callaghan et al., 2011; Ger et al., 2021; Legare et al., 2017). Here, we focus on the role of linguistic cues in the learning process. Linguistic cues in adult–child interaction, such as the surrounding nominals and the respective argument structure have been claimed crucial for inferring the semantics of causatives (e.g., Fisher et al., 2010; Gleitman, 1990). This cue-based acquisition, however, becomes problematic for languages that allow pervasive omission of linguistic units, that is, ellipsis. So far, it has been unclear how the reliability of these cues impacts the acquisition of causatives. In the present study, we test (i) whether the input in the two types of causatives differs with respect to the distributions of object ellipsis and (ii) whether the acquisition of lexical and morphological causatives differs. We choose Turkish because it constitutes an ideal natural laboratory to test these questions, with both types of causatives extensively used to convey causality.

Previous studies have focused on the structural information that accompanies lexical causatives in their acquisition (Gleitman, 1990; Naigles, 1996). Naigles (1996) tested children's comprehension of causal events upon hearing structural frames that contained different numbers of arguments in speech. They found that children capitalised on such structures to discriminate between causal and non-causal events. Others have followed this track to suggest the important role of syntactic frames in verb learning (Lidz et al., 2003; Messenger et al., 2015; Yuan et al., 2012). In Turkish, particularly, Göksun et al. (2008) tested the role of the number of noun phrases (NPs) for Turkish preschool children (up to age 5) in the causal enactment upon linguistic prompts. The results showed that the number of NPs in an utterance was a predictor for causal enactment, although the effect was weaker than that in English (Naigles et al., 1993). These results suggest that children's lexical learning relies on at least some syntactic structures, either granted as innate capability or developed in an item-specific manner (Bates & Goodman, 1997; Fisher, 2002; Tomasello, 2003; Yang, 2004). Either way, children are believed to infer, or activate, the structural rules from the input they hear to support further verb learning (e.g., Tomasello, 2003). A critical issue arises that syntactic structures can lack consistency and reliability when used as a cue for linguistic interpretation (Bates et al., 1989). For instance, the word order of arguments in German is not a reliable cue for determining the subject and the object of a verbal construction (e.g., "ich" [1SG.NOM] as the subject in both "Ich mag das" [1SG.NOM like.PRS.1SG 3NSG.ACC] and "Das mag ich" [3NSG.ACC like.PRS.1SG 1SG.NOM]). Such arguments can even be elided in many languages, resulting in unreliable elliptical structures for linguistic interpretation.

In Turkish, for example, both the agent and the patient of a causal event can be omitted in an utterance without rendering the sentence ungrammatical. While the information about the omitted agent may be retrieved from the subject agreement in the verb, the discourse, or the extra-linguistic context, there is a lack of morphological marking to indicate the omitted patient. Such missing patients can only be referenced in discourse and extra-linguistic contexts (Şener & Takahashi, 2010), making the processing rather implicit.

(1) (sen o-nu) Aç-acak mı-sın?
(2SG 3G-ACC) open-FUT Q–2SG
"Will (you) open (it)?"

https://doi.org/10.1017/S0305000925000200 Published online by Cambridge University Press

For instance, in example (1), the object "it" that bears the caused event is missing, similar to an intransitive construction. Thus, the argument structure here is implicit, and no formal marking can be capitalised on to make further causal inferences. The causal inference might be possible from the discourse, such as a preceding utterance "Bak, burada bir kutu var (Look, there is a box here)," where the object "it" can be recovered as "box" in the discourse context. When an object of a transitive verb is evident in context, object ellipsis can extensively occur, even in child-directed speech (CDS) in Turkish as evidenced in corpus studies (Küntay & Slobin, 1996; Sofu & Diser, 2018). Nevertheless, the ellipsis of objects may still pose a great challenge for processing the causal meaning in lexical causatives as recovering objects from the surrounding context in discourse requires attention beyond the sentence frame. Consequently, even if argument structure is facilitative for causal inference, it is unlikely to be the primary source children capitalise on for inferring causal meaning. It has been claimed that cues lacking reliability might severely compromise the learning of related linguistic aspects (Bates et al., 1989; Bates & MacWhinney, 1987; Dittmar et al., 2008). The question, therefore, arises whether ellipsis affects children's acquisition of lexical causatives.

In addition to argument structure, previous studies have explored other morpho-syntactic and pragmatic cues that may facilitate children's verb learning (Allen, 2000; Clancy et al., 2003; Göksun et al., 2008; Huang, 2012; Narasimhan et al., 2005). For instance, in the same study where the number of NPs was tested, Göksun et al. (2008) also examined the morphological cues. They found a significant effect of accusative marking for acting out causal actions. Also, Narasimhan et al. (2005) revealed that despite substantial argument ellipsis in Hindi CDS, 3–4 year olds managed to produce error-free transitive constructions thanks to their exploitation of multiple cues such as the context in discourse, verbal morphology, and extra-linguistic cues. Similar discourse cues have been found to play a role in other typologically diverse languages as well, such as Inuktitut (Allen, 2000), Korean (Clancy et al., 2003), and Mandarin (Huang, 2012). With various cues, it is plausible that children can remedy the issue of unreliable syntactic frames for their causal inference. Yet, little evidence has been supplied regarding the potential hindrance by ellipsis for the development of lexical causatives for Turkish children.

Ellipsis concerns the acquisition of morphological causatives as well. In addition to lexical causatives, morphologically complex languages such as Turkish employ morphological constructions to express causal meaning. For instance, the suffixes "-Ir/Ar" and "-DIr"[1] can directly causativise a non-causative verb. As in "düşürmek" (to drop, to make something fall off), "-Ir" marks the causal meaning as opposed to "düşmek" which denotes the non-causal meaning "to fall off." Examples (2) and (3) show how such morphological causativisation is concretely realised.

(2) ayakkabı-m ayağ-ım-dan düş-tü
shoe-POSS.1SG foot-GEN-ABL fall.off-PST.3SG
"My shoe fell off my foot."

(3) ayakkabı-m-ı düş-ür-dü-m
shoe-POSS.1SG-ACC fall.off-CAUS-PST.1SG
"I dropped my shoe."

[1]Variant forms due to consonant assimilation and vowel harmony in Turkish (e.g., for -DIr: -dir, -dir, -dür, -dur, -tir, -tir, -tür, -tur). Children have proven to show sensitivity to vowel harmony and the ability to employ it in linguistic tasks in their first year (e.g., Altan et al., 2016; Hohenberger et al., 2017).

https://doi.org/10.1017/S0305000925000200 Published online by Cambridge University Press

The non-causative verb *düştü* in example (2) has only one argument *ayakkabım* (*my shoe* in English). With the causative marker *-Ir*, which surfaces as *-ür*, in example (3), the verbal construction now can take *my shoe* as the patient and express that someone is causing *my shoe* to fall off, namely *I* as the agent. In this example, a valency change (i.e., a change in the number of verbal arguments) can be seen.

To showcase the productive use of causative markers in Turkish, example (4) displays a sentence involving secondary causation rendered by *-t*. That is, causativisation with morphological marking is permissive even for verbs that already involve causal meaning, such as *düşür* (drop).

(4) (ban-a) ayakkabı-m-ı düş-ür-t-tü
(1SG-DAT) shoe-POSS.1SG-ACC fall.off-CAUS-CAUS-PST.3SG
"S/he made (me) drop my shoe."

Eliding the arguments, especially the patients, in the morphological causative constructions above can jeopardise causal inference in the same way as in lexical causative constructions if inferring causality of both causatives relies on the recognition of the agent and the patient. For instance, if *ayakkabımı* in example (3) is dropped, the causal relation becomes opaque without the presence of a patient. Concretely, if *ayakkabımı* is dropped in example (3), namely if given only the verb form *düşürdüm*, without knowing the causative marker, no semantic difference can be established between this causativised verb and the non-causative form *düştü* for a language learner. Both *düşür* (with an elided patient) and *düş* can stand alone in an utterance without linguistic context, and the causal meaning of *düşür* is difficult to be distinguished from the non-causal meaning of *düş*, if the causative marker *-ür* is not known. Children may recognise the same verb stem in both forms, but with the patient ellipsis, the discrimination of causal meaning lacks syntactic support. Teasing apart the meaning of *düş* and *düşür* here thus becomes almost as challenging as for lexical pairs in distinct forms such as *düş* (non-causative) and *aç* (causative; translated as "open" in English).

Even in the situation with secondary causation in example (4), ellipsis poses a similar, if not more severe, issue for processing the causal meaning. Here, the morphological causative *düşür* (drop, make it fall off) is attached with a second causative marker *-t*, and this introduces an extra dative argument that bridges two agent-patient relations, namely *s/he* & *me* and *me* & *my shoe*. If the intermediary argument *bana* is elided in this construction, as shown in example (4), there is a lack of cues for differentiating the resulting sentence *ayakkabımı düşürttü* (he/she made me drop my shoe) from *ayakkabımı düşürdü* (he/she dropped my shoe).

Without reliable arguments in the context, the acquisition of the causal meaning of morphological causative markers in Turkish might be delayed (e.g., Bates et al., 1989). In fact, previous studies have supplied mixed results regarding the development of morphological causative, perhaps also indicating the different stages of morphological development. For instance, Aksu-Koç and Ketrez (2003) and Ketrez (1999) both found that children's productivity of causative marking emerges at around 2 years of age using a corpus of Turkish child speech (CS). Ger et al. (2022a) found that 3- to 4-year-old Turkish-learning children inferred causality from sentences containing pseudo-verbs marked with a morphological causative marker, even when the object argument was elided. Yet, Ger et al. (2022b) found the abstract understanding of morphological causatives to emerge in experimental settings using pseudo-verbs only

https://doi.org/10.1017/S0305000925000200 Published online by Cambridge University Press

at 5 years. Moreover, Göksun et al. (2008) revealed no facilitation of causative marking up to 5 years of age for the task of enactment of causal actions. Such a delay of acquisition has been shown for essential morphemes such as the evidential marker as well, whose comprehension was not attested in experimental settings until about 6–7 years of age (Ozturk & Papafragou, 2016). Another interesting observation by Ketrez (1999) was that children produced wrong argument structures for morphological causatives, even for those without complex secondary causation, whereas they had no difficulty uttering simplex lexical words with appropriate argument structures. There is also recent evidence that Turkish-speaking parents use more lexical causatives than morphological causatives with their 1.5-year-old children (Aktan-Erciyes & Göksun, 2023), suggesting a potential explanation as to why children might be erring more in the argument structure of morphological causatives and abstracting them rather late. However, we do not know so far if the argument structure in which parents use these causatives differs or not, which would shed more light on these possibilities. If argument structure is indeed a factor for causal inference, ellipsis should exert an equally large influence on the learning of both lexical and morphological causatives. Moreover, as Bates et al. (1989) would predict in the competition model, the more reliable cue of explicit morphological marking, with its productive use in the Turkish language, should aid the acquisition of morphological causatives. Hence, it is worth asking to what extent ellipsis occurs in both causatives so that it could hint at the development of causatives.

So far, we know very little about the extent of the ellipsis of arguments, especially the ellipsis of objects, and its potential effect on the learning of causal meaning. It is also largely unknown how object ellipsis with its effect varies between morphological and lexical causatives. Whereas ellipsis has not been clearly quantified as of yet, several corpus studies have examined the development of causatives, albeit without concretely depicting the developmental trajectories. For example, Ketrez (1999) evaluated children's productivity of morphological causatives, with a focus on the causativisation upon verbs within the same recording session of spontaneous speech (i.e., both the non-derived verb form and the form with a causative marker should exist in the same session, such as *düş* and *düş-ür*). However, this scenario rarely occurs in that it requires particular contexts and a substantial amount of speech. In a different vein, Aksu-Koç and Slobin (1985) and Nakipoğlu et al. (2022) looked at overregularisation as an indication of children's productivity of causative marking. Both approaches limit the scope to specific lexical items only and thus do not reflect the general development of causative use. Hence, we know little about the generalisation of causative marking.

Here, we rely on entropy estimation to depict the developmental trajectory of both lexical and morphological causatives in naturalistic data. The entropy measure helps to quantify the variability of the use of linguistic units, as pioneered by Prado Martín et al. (2004) and employed in many previous studies, especially on morphological paradigms (e.g., Ackerman & Malouf, 2013; Stoll et al., 2012). Concretely, the entropy measure does not look at the units in isolation but rather quantifies their probabilities in a distribution and gauges the uncertainty of units from such a distribution. This takes into account both types and tokens of units, which avoids the bias of specific items in traditional measures (Mazara & Stoll, 2019). Thus, the entropy measure is useful for examining how flexibly children use lexical and morphological causative constructions. The flexibility in verb use denotes the use of verbs in diverse forms and contexts (Tomasello, 2003). The flexibility measured from naturalistic data can indicate the

https://doi.org/10.1017/S0305000925000200 Published online by Cambridge University Press

proficiency of causative use (Stoll et al., 2012), namely the overall generalised use of causatives, even in unseen scenarios and contexts. This is particularly helpful for comparing the extent of children's causative acquisition and adults' fully proficient use of causatives. The widely used empirical Shannon entropy (Shannon, 1948), however, could result in severe underestimation when the observed data are under-sampled (Chao & Shen, 2003). As an example, longitudinal acquisition corpora, including the corpus we use in this study (Küntay et al., unpublished), are often collected with a small sample size. Without compensation for the unseen data, such sampled data cannot fully reflect the situation of the entire population. Thus, we employ a new measure, namely the Nemenman–Shafee–Bialek (henceforth referred to as NSB) entropy estimator (Nemenman et al., 2001), which tackles the under-sampling issue in our naturalistic data. The NSB estimator, by employing a Bayesian framework in its estimation, remedies this issue and has shown desirable results (Nemenman et al., 2001). With this estimator, we seek to maximally capture the extent of the generalisation of causative use despite the challenge posed by the small data size.

We quantify the extent of object ellipsis in both causative constructions to look into the impact of ellipsis on the development of causatives. To investigate the role of ellipsis in Turkish children's development of both causatives, we aim to answer the following two questions: (1) How often is the patient argument in lexical and morphological causatives elided in CDS? and (2) How do Turkish children acquire these two types of causatives over time?

2. Methods

2.1. Data

Our data come from a longitudinal corpus of Turkish child language (Küntay et al., n.d.; Moran et al., 2016), which includes the naturalistic speech of 8 target children and their respective surrounding speakers, predominantly caregivers. The recordings cover an age span from 8 to 36 months (see corpus details in Table 1). Here, we included the data of 7 children and left out the data of one child whose recordings stopped before 2;0. The recordings were made in children's home setting for 1 hour every 2 weeks over a period of

Table 1. Basic information of the Turkish longitudinal corpus

			CDS		CS	
Child	Sex	Hours of recordings	Utterances	Words	Utterances	Words
1	F	57	15,121	53,073	12,061	23,224
2	M	40.5	25,272	81,484	17,118	32,526
3	F	51	27,610	105,653	17,472	32,601
4	F	56	34,241	105,966	19,561	30,473
5	M	41.75	20,600	67,259	15,547	25,755
6	F	46	30,192	103,689	15,495	32,342
7	F	51	31,371	102,033	16,107	23,725

https://doi.org/10.1017/S0305000925000200 Published online by Cambridge University Press

29 months (8–36 months of age). Four of the children came from a low socioeconomic background (parents with 5 or 8 years of education), and three came from a high socioeconomic background (parents with 11 or 15 years of education). At the beginning of the data collection, mothers had a mean age of 25.5 (*SD* = 4.7, range = 21–34) and fathers had a mean age of 29 (*SD* = 3.5, range = 26–35)[2]. None of the children had siblings.

2.2. Automatic and manual annotation

The corpus was transcribed by native speakers. With the transcribed texts, we employed an automatic parser, the ITU Turkish Natural Language Processing Pipeline (Eryiğit, 2014), to add the annotation part-of-speech (POS), morphological analysis, and dependency parsing, thus obtaining the information of causative marking and argument structure of verbs.

We focused only on the ellipsis of objects and not the intermediary agents because, conceptually, the presence of an object is equally essential in all cases of morphological causatives to infer at least one causal event. In addition, because our focus was on object ellipsis, we checked the performance of the parser with respect to the annotation of objects in the dependency parsing. We randomly extracted a subset of 100 utterances with at least one verb in each utterance and manually examined the object ellipsis of all verbal constructions. A total of 133 verb tokens were correctly POS-tagged in these 100 sample utterances, and 32 out of the 133 tokens came with an object argument. All these 133 tokens were used in the evaluation. The evaluation was only for the parsing of objects, so other aspects of the parse were ignored. The parsing results reached a micro F1-score of 0.82 (see more details in Supplementary Table S1), ensuring the reliability of the ellipsis analyses.

Our analyses of entropy estimation were based on the verb stems. Hence, after the POS tagging by the ITU NLP pipeline, we extracted all the tokens tagged as verbs (864 stem types) that appeared in both CS and CDS. Subsequently, the second author (a native Turkish speaker) checked all the verb stems and removed 51 wrongly tagged stems (non-verbs). There were also 53 stems that were, in fact, only partially parsed for morphological markers. These were further manually parsed, especially for causative and passive marking. These manual checks resulted in a total of 763 unique verb stems, and the corrected parsed data were included in our analyses.

To identify lexical causatives among the verbs, we used Shibatani's criterion (Shibatani, 2002) that a lexical causative has to entail a change of state in its object (including a location state, a form state, and a psychological state in our categorisation). Relying on this criterion, the second author (a Turkish native speaker) and another native speaker manually coded the verbs that occurred in the corpus for causatives or non-causatives. The coding was done for each verb type, independent of the utterances it was situated in. The agreement was 88.9% (Kappa = .76).

Table 2 shows the numbers of tokens for different verb groups in the corpus summarised from the annotations. It should be noted that the utterances transcribed in the corpus were intonational units that did not necessarily contain verbs (e.g., an utterance can be made up of just one interjection such as "hey" in English). Hence, the number of verb tokens shown in Table 2 is generally much lower than the number of utterances reported in Table 1,

[2]Note that the data for the age of parents of one child were missing.

https://doi.org/10.1017/S0305000925000200 Published online by Cambridge University Press

Table 2. Summary of the Turkish longitudinal corpus: Number of tokens for different verb groups

	CDS			CS		
Child	Verbs	Lexical causatives	Morphological causatives	Verbs	Lexical causatives	Morphological causatives
1	13,619	3,965	550	4,598	1,739	222
2	19,886	5,540	813	6,073	1,533	160
3	27,392	7,133	953	5,588	1,841	263
4	29,665	7,088	931	4,626	1,497	262
5	17,082	4,269	622	3,922	1,114	124
6	24,875	6,517	1,084	7,126	1,875	280
7	27,768	7,032	931	3,680	1,302	102
Total	160,287	41,544	5,884	35,613	10,901	1,413

especially in child speech. The development of the numbers of tokens and types over age for different verb categories is reported in Supplementary Figures S1 and S2.

2.3. Entropy estimation

Shannon (1948) originally proposed the entropy measure for quantifying the variability of a population. The entropy can be calculated as below:

$$H = -\sum_{i=1}^{N} p_i \log p_i, \tag{1}$$

where N is the number of types in the population and p_i is the probability of the i-th type.

In practice, the estimation is done in a plug-in manner, namely the observed probability of each type, and the observed total number of types are plugged into the formula 1 to determine the estimate. In our case of the present study, take a simple sample of just two verb stem types (i.e., $N = 2$) with two tokens each, such as ("çık," "çık," "gül," "gül"), both "*çık*" and "*gül*" have an observed probability of 0.5 (i.e., $p_1 = p_2 = 0.5$), respectively. The resulting plug-in entropy estimation, based on the formula 1, would thus be calculated as $-\log 0.5$. It has been shown that the negative bias of this plug-in estimator is inevitable, and it can be particularly severe if there are unobserved types (e.g., Chao & Shen, 2003). The NSB entropy estimator (Nemenman et al., 2001) was therefore proposed to remedy the under-sampling issue and showed desirable results. NSB is a Bayesian estimator that employs a mixture of Dirichlet distributions as its priors, which are later updated by the observed counts in the sample to yield the posterior estimate. Simply put, an NSB entropy estimator could be seen to smooth the estimate of the entropy of the population by performing integration over continuous parameter space instead of a discrete one. It is particularly helpful when there is a large number of unseen types in the sample, since the cardinality can be set for the priors to enable estimation of the unseen types, thus compensating for the overall entropy estimation. It has been shown by Archer et al. (2014) that NSB can

https://doi.org/10.1017/S0305000925000200 Published online by Cambridge University Press

stably estimate the entropy of a population with a sample size as low as 10. You et al. (2024) further supplied evidence that NSB best estimates the true entropy of a population even under the extremely under-sampled regime with a sample size of less than 10.

We hence used this estimator to measure the flexibility of causative use in this article. As mentioned above, NSB requires information on the cardinality of the prior distribution. This cardinality means the total number of types in the population for which we estimate the entropy. In the context of the present study, cardinality means the total number of verb stem types in the population, including the types not seen in the sample. We thus further employed the Smoothed Good–Toulmin estimator (SGT) (Orlitsky et al., 2016) to help estimate this cardinality, which is important, especially for data with low variability in the severely under-sampled regime.

For lexical and morphological constructions, the cardinality estimated from SGT was based on different observations. For lexical causatives, the potential number of verb stem types was estimated with the observed lexical causatives. By contrast, the cardinality for morphological causatives should be the number of verb stem types that can theoretically take a morphological marker. Thus, the second author examined the verb stems in the corpus and ruled out a small number of stems that are ineligible to take a causative marker (26 stems, including equipollent or suppletive pairs and some reflexive verbs). The tokens of the remaining eligible stems were used to estimate the cardinality for the entropy estimation for morphological causatives.

2.4. Analyses

We conducted two analyses to investigate the object ellipsis of both causative constructions in CDS and children's individual developmental trajectories of both constructions.

Analysis 1: Ellipsis in CDS. We first analysed the object ellipsis of both causative constructions in CDS. The idea was to quantify the proportion of object ellipsis in causative constructions that children hear, thus examining the reliability of the argument structure for causal inference. We excluded verb tokens used in passive voice, as the passive construction raises the patient to the subject position and might exhibit a different pattern of ellipsis (only 1.89% of the tokens were excluded; 3,697 out of 195,900). For each session, the causative constructions were bootstrapped (with replacement) for 1,000 iterations, with 1,000 samples in each iteration. We examined the change in ellipsis level over age and took the bootstrapped mean of ellipsis level in each session for the statistical analysis.

A multi-level Bayesian model of zero/one inflated beta regression (with children as the group effect) was then built to test the difference between the ellipsis rate in lexical causative constructions and that in morphological causative constructions. More specifically, the ellipsis rate was the dependent variable of the regression model, and the type of causative constructions was the only independent variable, for which dummy coding was applied with the lexical causatives as the reference level. Both the random intercept and the random slope were included in the model.

Analysis 2: Development of causative constructions in children's speech. We applied the entropy estimation in both CDS and CS, and later calculated the ratio of the entropy estimate in CS to that in CDS (henceforth referred to as "developmental ratio") (Mazara & Stoll, 2019; Stoll et al., 2012), to examine to which extent children's production has reached the respective adults' level. The adults' entropy level, as the denominator, could

https://doi.org/10.1017/S0305000925000200 Published online by Cambridge University Press

also indicate how variable the use is for a measured verb group generally in the Turkish language. We first measured this developmental ratio for the causative use. Then, to better position the development of causatives in the general picture of verb learning, we also measured the ratio for two other verb groups, namely non-causatives (e.g., bak (look), çık (come out), düş (fall off), gül (laugh)), and all verbs. More concretely, we gathered all verb tokens, which were first put in a full category called "all verbs," and then grouped them into categories of lexical causatives, morphological causatives, and a third category, non-causatives (which included both intransitives and non-causative transitives). We stripped the inflections in the verb tokens afterwards to obtain the verb stems, based on which we measured the entropy for each group. The raw entropy measures in CDS and CS before the ratio calculation are reported in Supplementary Figures S3 (CDS) and S4 (CS).

Preliminary data analysis of local regression (see Supplementary Figures S5–S8) revealed a stage-like development of verbs for many individuals, where the child rapidly approaches the adults' level after an initial flat growing stage before s/he again stabilises the development. We thus employed segmented analysis, using the libraries *segmented* in R (Muggeo, 2020; Muggeo et al., 2008), to identify the breakpoints for such developmental trajectories. Concretely, we treated the developmental ratio as the dependent variable and the age as the independent variable, while searching for the existence of breakpoints with hypothesis testing implemented by the *segmented* library. We primarily modelled the trajectories with two breakpoints, with one breakpoint as the alternative when a second change of slope did not occur. We refrained from searching for more than two break points due to the short time span of our data. These breakpoints formed piece-wise regression lines, where we examined when major changes of development occurred in different verb groups.

3. Results

3.1. Analysis 1

Ellipsis in CDS is shown to be pervasive in both constructions (lexical: $M = .65$, $SD = .08$; morphological: $M = .54$, $SD = .20$). Figures 1 and 2 show the ellipsis levels for each child at each age for lexical and morphological causatives, respectively. As shown in Figure 3, with the bootstrapped means of ellipsis levels, the multi-level Bayesian model of zero/one inflated beta regression exhibits a substantially lower ellipsis rate for morphological causatives than lexical causatives. The difference is confirmed by the 95% credible interval of $(-0.63, -0.27)$ (i.e., a negative interval excluding the null effect) for the posterior distribution of the coefficient for the effect of morphological causatives, compared with the lexical causatives as the reference level.

3.2. Analysis 2

First, we compared lexical causatives with two other groups: non-causatives and all verbs. The development of lexical causatives shows little difference from the other two groups in terms of the breakpoints and developmental rate in each stage. All three groups generally experience a 3-stage development for each child. For some children, nonetheless, the development of lexical causatives can exhibit a slight time lag or a slower developmental rate (see Child 4, 5, and 6 in Figure 4: the development slope of lexical causatives in the

https://doi.org/10.1017/S0305000925000200 Published online by Cambridge University Press

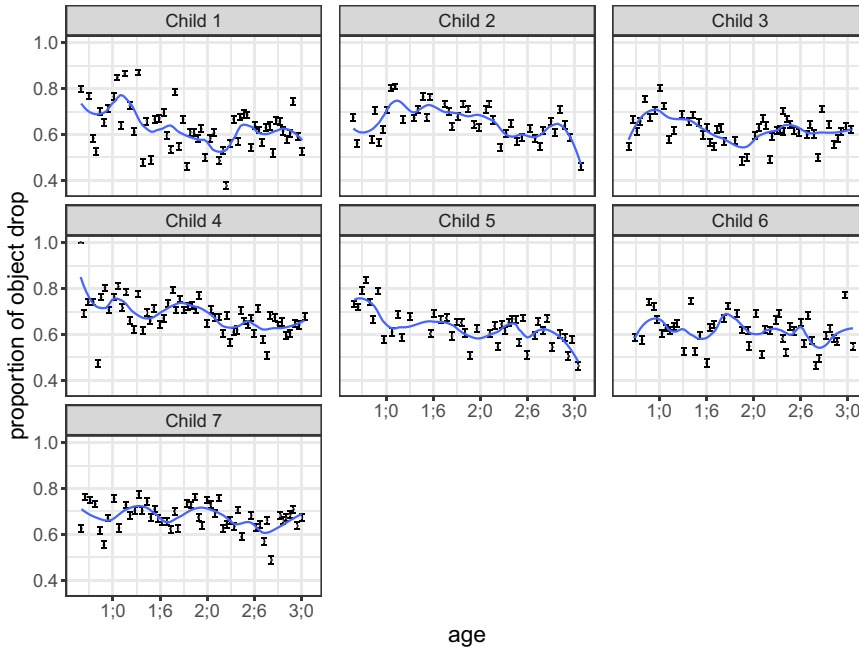


Figure 1. Object ellipsis of lexical causatives in child-directed speech.
Note. The line represents the local regression of the ellipsis level, with the span set to 0.3. The points are the means of the bootstrapped ellipsis levels, with the error bars showing the bootstrapped standard deviations.

first phase for Child 4 is slightly flatter; Child 5 starts the rapid development of lexical causatives a bit later than the other two groups; Child 6's second phase of development is slower than the other two groups). It should also be noted that the development of all three groups stabilises at a ratio close to 1.0 from the age of around 2;2, which suggests an adult-like level of generalisation of lexical causative semantics.

Second, we compared morphological causatives with all verbs. The development of morphological causatives displays a clear difference compared with the overall development of the verb lexicon (see Figure 5). First, in the corpus, the observed onset of flexibility in the use of morphological causatives (i.e., morphological causatives occurring with two different verb stem types) is usually much later (ranging from 1;10 to 2;3) than that of general verb use (ranging from 0;8 to 1;2). Secondly, two major patterns of developmental trajectories are shown among different individuals. For Children 2 and 5, there is a clear stabilisation phase of development in the end, but the developmental ratio for morphological causatives stays lower than that for all verbs, and this gap remains in the final phase. For the other children, there is not yet a stabilisation phase for the development of morphological causatives (e.g., for Children 1, 6, and 7, the ratio still fluctuates at around 3;0). Instead, there seems to be steady development, eventually approaching the level of the development of all verbs. Additionally, an alternative segmented analysis with BIC (Bayesian Information criterion), which tends to select a more parsimonious model, also corroborated the distinct developmental trajectory of morphological causatives compared with other verb groups (see Supplementary Section S4 and Supplementary Figures S9 and S10).

https://doi.org/10.1017/S0305000925000200 Published online by Cambridge University Press

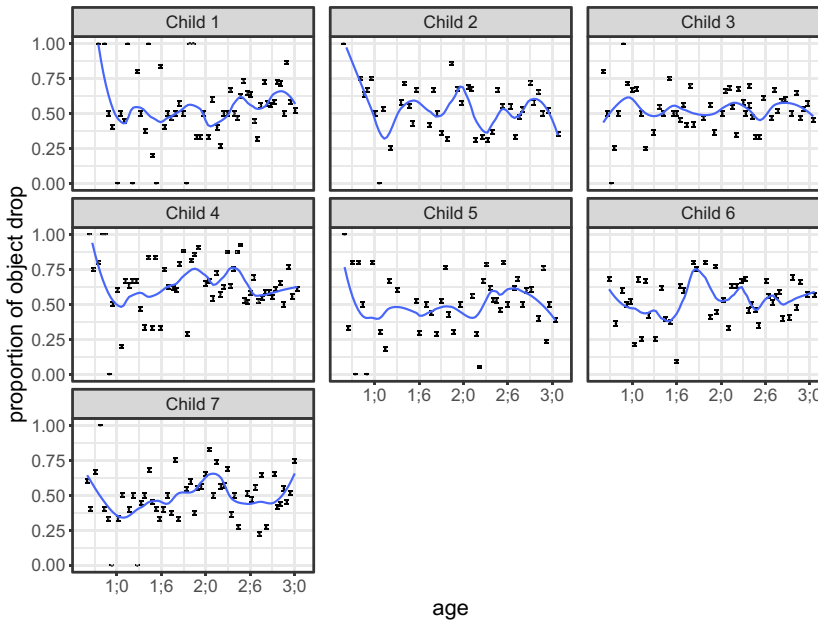


Figure 2. Object ellipsis of morphological causatives in child-directed speech.
Note. The line represents the local regression of the ellipsis level, with the span set to 0.3. The points are the means of the bootstrapped ellipsis levels, with the error bars showing the bootstrapped standard deviations.

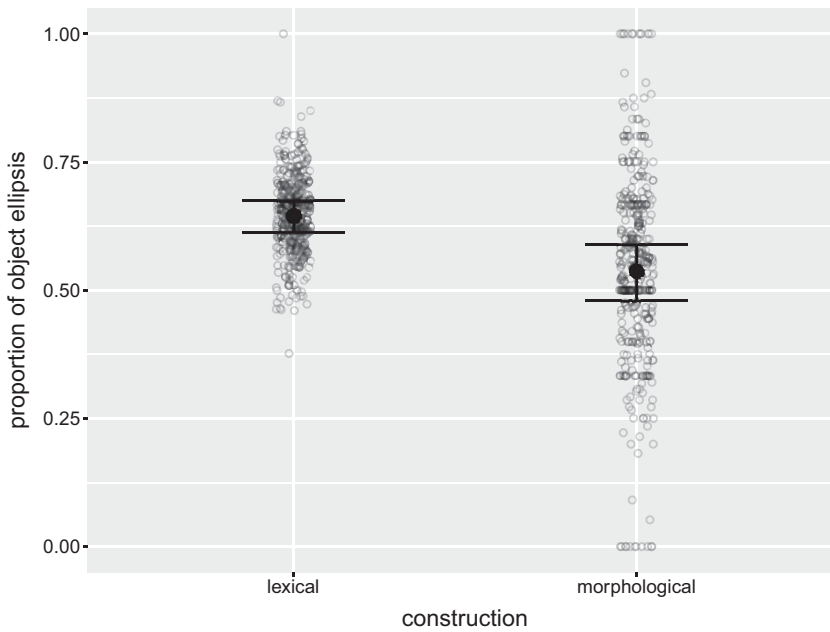


Figure 3. Comparison of object ellipsis between causative constructions.
Note. 95% credible interval of the posterior draws of the mean of the posterior predictive distribution is displayed for each construction. Lexical: (0.61, 0.68); morphological: (0.47, 0.59). The plotted points stand for the original data points.

https://doi.org/10.1017/S0305000925000200 Published online by Cambridge University Press

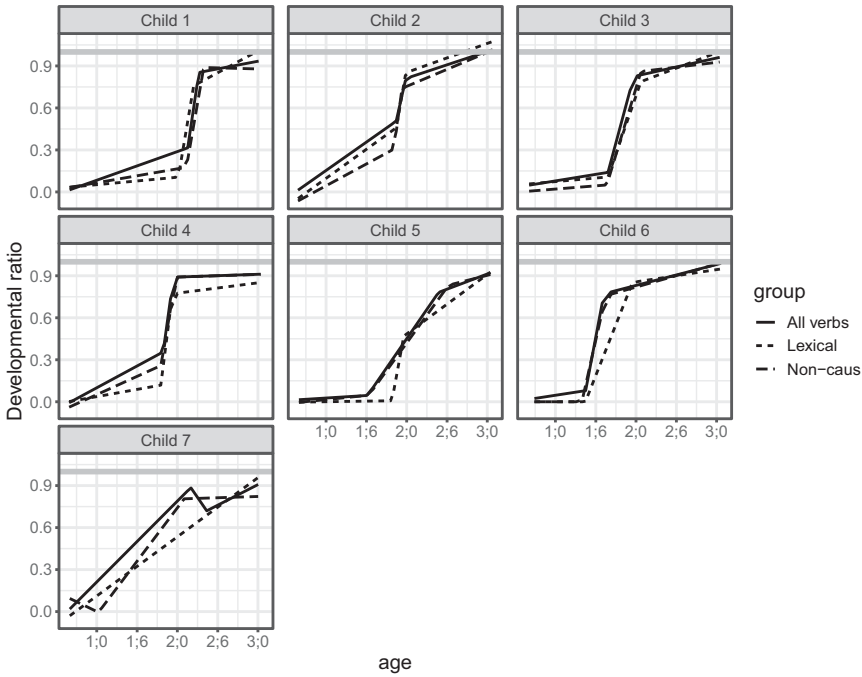


Figure 4. Development of entropy ratio for all 7 children: lexical causatives, non-causatives, and all verbs.

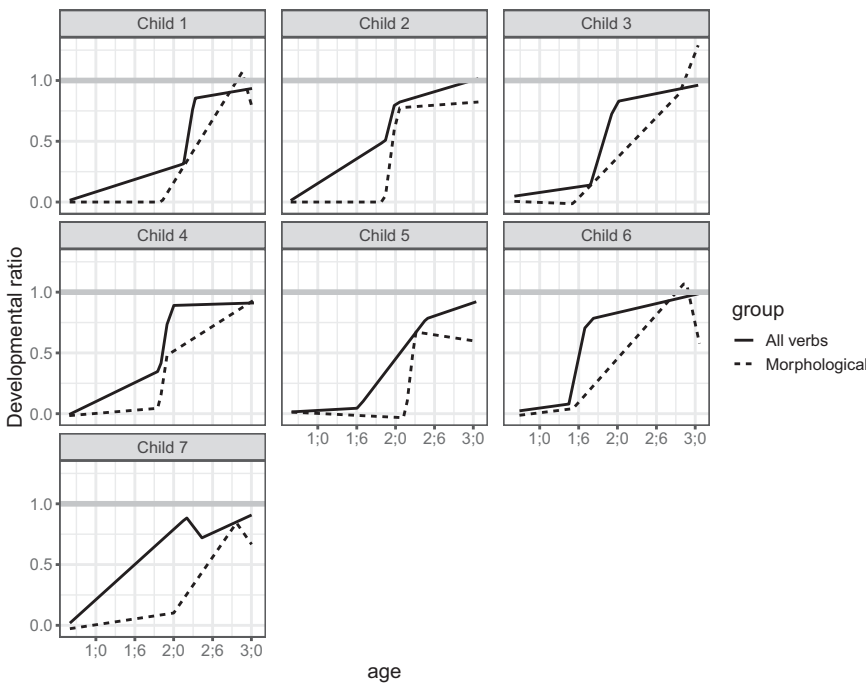


Figure 5. Development of entropy ratio for all 7 children: Morphological causatives versus all verbs.

https://doi.org/10.1017/S0305000925000200 Published online by Cambridge University Press

4. Discussion

Overall, our results suggest that children learning Turkish develop their causative use in their first 3 years of life, despite the pervasive ellipsis in the speech they hear. In the first analysis of object ellipsis, we found a large proportion of such ellipsis for both lexical and morphological causative constructions in CDS. More than half of the causative constructions, in either category, elide their object. Such ellipsis of arguments poses a great challenge for children to rely on argument structure as a cue to infer the causativity (Bates et al., 1989; Bates & MacWhinney, 1987; Kempe & MacWhinney, 1998). However, no substantial delay in learning was observed for either lexical or morphological causatives, although the pattern of development differs. It is, therefore, unclear whether argument structure is an essential cue in children's early acquisition of causatives.

The unhindered development is especially prominent for lexical causatives. Our Analysis 2 shows no major difference between the developmental trajectories of lexical causatives, non-causatives, and all verbs. Children's use of lexical causatives by the age of 3 years becomes nearly as flexible as observed in adult use. As for morphological causatives, we observed deviated development from the growth of overall verb use. Overall, there was a lag in the development of morphological causatives, and they thus did not seem to benefit from the explicit causative marking. It should be noted that the entropy of morphological causatives in CDS was not low at all compared with other verb groups (see Figure S3). Hence, morphological causatives are commonly and flexibly used by adults in the Turkish language, and the pattern of the developmental ratio as seen in Figure 5 is likely a result of lagged development of morphological causative use in child speech. In other words, morphological causatives, as a common device for expressing causality, have not consistently become a substantial part of children's linguistic causal repertoire by the age of 3 years. Interestingly, as shown in Figure 3, morphological causative constructions, in fact, involve less object ellipsis, which theoretically provides a more consistent and reliable cue for acquiring causal meaning (e.g., Bates & MacWhinney, 1987). However, it should be noted that the level of object ellipsis for morphological causatives is, by no means, negligible. Both the difficulty of generalising the causative marking as conveying causal meaning and the pervasive ellipsis in a naturalistic speech environment add more challenges to the learning condition. To understand morphological causatives, a child has to infer causality from the many variants of causative morphology situated in various contexts as well as the frequently incomplete argument structure. This might account for the slower development found in the present study. At the very least, it can be assured that the joint cues of argument structure and explicit marking do not show clear facilitation for the acquisition of morphological causatives. This is consistent with the previous finding that 3-year-old Turkish-learning children were as good at deriving causal meaning from a combination of argument structure and morphological cues as from only argument structure cues or only morphological cues (Ger et al., 2022a).

It cannot be ruled out, though, that the syntactic knowledge of argument structure might still be extracted from elliptical verbal constructions, at a certain stage of the learning process, thus indirectly helping with the causative acquisition. However, this requires a presumptive extra layer of syntactic learning, whereas many other cues support a more direct inference of causal semantics. For instance, there exist other morphological cues, such as the accusative marker (Göksun et al., 2008; Ural et al., 2009) in Turkish, that may explicitly support verb learning. Although patients are often elided in lexical

https://doi.org/10.1017/S0305000925000200 Published online by Cambridge University Press

causative constructions, their appearances are usually accompanied by an accusative marker, which highlights the existence of the patient argument and, therefore, helps children understand the causal meaning of causatives. Further, as others have previously pointed out (Huang, 2012; Narasimhan et al., 2005), discourse and extra-linguistic cues might be relevant to the acquisition of lexical causative meaning as well. As introduced earlier, objects retrieved in discourse (e.g., in previous utterances) or detected in the physical environment can remedy the missing essential units in the current linguistic expression.

For discourse contextual cues, in particular, supportive evidence includes the emergence of semantic and structural regularities from simple linear word co-occurrences (Mintz et al., 2002; Perfors et al., 2006; You et al., 2021). Inference of causative regularities might not necessarily rely on the explicit presence of the patient, but rather raw contextual linguistic units (i.e., co-occurrences of surface forms of linguistic tokens) in discourse without presumptive syntactic knowledge such as part-of-speech generalisation and hierarchical relations. For instance, the repetition of lexemes across situations in the same discourse,[3] often referred to as variation sets (Küntay & Slobin, 1996; Onnis et al., 2008), are prevalent across typologically diverse languages (Moran et al., 2019). These repetitive patterns, with the semantics conveyed by the linguistic tokens thereof, can facilitate statistical inference that helps children discriminate causatives from other semantic categories. Fully-specified argument structure on the surface might not necessarily be a prerequisite of causative learning.

Importantly, the present study, based on naturalistic data, has the advantage of revealing the acquisition of morphological causatives at an earlier age. Also, naturalistic recordings, especially in home settings where parents are free to engage in a wide range of activities, cover a large variety of speech contexts. This helps to maximally include causative verbs, the use of which largely depends on the type of context (Aktan-Erciyes & Göksun, 2023). By contrast, experimental studies, while helping uncover different factors at work, typically reveal later ages of achievement for a structure because they pose extra cognitive load with their task demands and can only be performed with a limited range of contexts. Nonetheless, the lag of morphological causative development we found in our corpus study well resonates with findings from previous experimental studies. For example, Göksun et al. (2008) found in their experimental study on causal enactment with causative markers that, up to at least 5 years of age, children acquiring Turkish are still developing their causative morphology, thus being not fully capable of comprehending the morphological constructions. This has been recently corroborated by Ger et al. (2022b), where they found late abstraction (at around age 4;10) of causative morphology for Turkish children in an experimental study using pseudo-verbs, and by Aktan-Erciyes et al. (2024), who showed that 5-year-old monolingual Turkish children, but not 7- and 9-year-olds, performed better with lexical causative verbs than with morphological causative verbs in an experimental causative verb production task. Similar late development of causative morphology has been observed in other morphologically complex languages as well, such as Quechua (Courtney, 2002) and Inuktitut (Allen, 1998).

We suggest that the synergy of cues and the development of different linguistic aspects is likely the key to verb learning. For instance, Göksun et al. (2008) found facilitation of

[3]For instance, with the verb "see," a child may hear "what did you *see*?," "did you *see* the dog?," "let's go and *see* the dog," and so on in a discourse.

https://doi.org/10.1017/S0305000925000200 Published online by Cambridge University Press

accusative marking for children learning Turkish in acting out causal actions, potentially supporting the overall causative acquisition. However, morphological markers can be acquired at different times due to the complexity of their syntactic or phonological forms (e.g., number of allomorphs) and the semantics being conveyed (e.g., number of meanings an affix bears) (Clark, 2017). In fact, Turkish-learning children are found to have not yet mastered the more complex nuances of case marking by the age of 6 years (Ketrez, 2005), making this cue less reliable. Even if case marking is fully acquired, cases alone might not suffice for causal interpretation. Murasugi et al. (2003), for example, looked into the Japanese morphological causative "-(s)ase" along with the case marking. While the accusative and dative markers are helpful for teasing apart the semantic roles, they are, in many cases, not sufficient for identifying the agent of the action conveyed by the causatively marked verb. Ultimately, the disambiguation relies on the semantics of all components in the context, such as the animacy of the arguments and the plausibility of the agent-patient scenario in real-world (e.g., a doll cannot act on itself to push a chair). In a similar vein, Booij (1996) attributed the late acquisition of contextual inflections, such as person and number agreement markers, to their strong correlation with semantics like agent and patient roles in context, as evidenced in experiments on children acquiring German inflections (Clahsen, 1986). In our present study, too, some children (Child 2, 5, 7 in Figure 5) only start to develop their causative marking when the lexical knowledge is established, as indicated by the almost stabilised development of all verbs. Also, many children (Child 1, 3, 4, 6, 7) keep the growth of morphological causatives at a moderate rate after the overall lexical developmental rate stabilises, before eventually reaching the adult-like level. It is likely that morphological development is heavily contingent on semantic grounds. Advanced semantic understanding of lexical words might have to precede the generalisation of morphological marking. Such lexical ground for grammatical development has been found in previous studies as well (Devescovi et al., 2005; Hoff et al., 2018; Marchman et al., 2004; Parra et al., 2011).

This link to semantics could, in turn, partly explain why morphological causative constructions in CDS render less object ellipsis. As Shibatani and Pardeshi (2002) suggested, the causal relation expressed by morphological constructions is less direct than lexical words, thus lacking semantic saliency and requiring structural amendment. This can be exemplified by the sentence *(bana) ayakkabımı düşürttü* (s/he made me drop my shoe) in example (4), where secondary causation is involved. Here, the causee "me" serves as the agent of the action of "drop," bridging two causal relations. The complete causal meaning in this sentence can be difficult to infer without specifying either the intermediary agent "me" or the patient "my shoe." Hence, both the intermediary agent and the patient could be less often elided to facilitate the inference and make the causal meaning more transparent. The scenarios with intermediary agents, in particular, could lead to greater difficulty in comprehending causal relations. However, due to the constraints of both the size of data (hence containing few occurrences of intermediary agents) and the accuracy of automatic taggers for complex sentences, our investigation was not extended to address the learning of secondary causation. Future studies can track the ellipsis of intermediary agents and that of patients to further examine the interplay between the multiple cues of semantic arguments.

To summarise, we have found no hindrance of pervasive object ellipsis for Turkish children to acquire causatives. Moreover, the lower level of ellipsis for morphological causatives in CDS shows no facilitation in children's development of causative morphology. Together, these findings reveal that argument structure is an unreliable cue in naturalistic input due to a large ellipsis in Turkish and might be a less crucial facilitator for

https://doi.org/10.1017/S0305000925000200 Published online by Cambridge University Press

the development of causative constructions than previously claimed. We suggest that many other cues, often more explicit in their forms, could play equally important roles in the acquisition. Our general conclusion is that children learning Turkish, by the age of 3 years, acquire causative constructions despite the challenge posed by pervasive ellipsis. Future studies might benefit from paying more attention to discourse, morphological, and extra-linguistic cues that might better facilitate causative acquisition.

Abbreviations

1	first person
2	second person
3	third person
ABL	ablative
ACC	accusative
CAUS	causative
DAT	dative
FUT	future
GEN	genitive
N	neuter
NOM	nominative
POSS	possessive
PRS	present
PST	past
Q	question particle
SG	singular

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/S0305000925000200.

Data availability statement. Codes and anonymised data can be accessed at https://osf.io/u5f4m/.

Funding statement. The results of this research are part of the project "The role of causality in early verb learning: language-specific factors vs. universal strategies" (Swiss NSF Grant agreement No. 100015_169712; Stoll, Daum), the project "Acquisition processes in maximally diverse languages: Min(ding) the ambient languages (ACQDIV)" that has received funding from the European Research Council (ERC) under the European Union's Seventh Framework Programme FP7-2007-2013 (Grant agreement No. 615988; Stoll), and the NCCR Evolving Language, Swiss National Science Foundation Agreement No. 51NF40_180888.

Competing interests. The authors declare none.

References

Ackerman, F., & **Malouf, R.** (2013). Morphological organization: The low conditional entropy conjecture. *Language*, **89**(3), 429–464.

Aksu-Koç, A. A., & **Ketrez, F. N.** (2003). Early verbal morphology in Turkish: Emergence of inflections. In D. Bittner (Ed.), *Development of verb inflection in first language acquisition: A cross-linguistic perspective* (p. 27, Vol. **21**). Walter de Gruyter.

Aksu-Koç, A. A., & **Slobin, D. I.** (1985). The acquisition of Turkish. In D. I. I. Slobin (Ed.), *The crosslinguistic study of language acquisition* (pp. 839–878, Vol. **1**). Erlbaum.

Aktan-Erciyes, A., **Ger, E.**, & **Göksun, T.** (2024). Influences of early and intense l2 exposure on l1 causal verb production: Comparison of 5-, 7-, and 9-year-old bilingual and monolingual children. *First Language*, **44**(5), 474–492.

https://doi.org/10.1017/S0305000925000200 Published online by Cambridge University Press

Aktan-Erciyes, A., & **Göksun, T.** (2023). Early parental causal language input predicts children's later causal verb understanding. *Journal of Child Language*, **50**(1), 177–197.

Allen, S. E. M. (1998). Categories within the verb category: Learning the causative in inuktitut. *Linguistics*, **36**, 633–677.

Allen, S. E. M. (2000). A discourse-pragmatic explanation for argument representation in child inuktitut. *Linguistics*, **38**, 483–521.

Altan, A., **Kaya, U.**, & **Hohenberger, A.** (2016). Sensitivity of Turkish infants to vowel harmony in stem-suffix sequences: Preference shift from familiarity to novelty. In J. Scott, & D. Waughtal (Eds.), *BUCLD 40 online proceedings supplements*. Boston University Conference on Language Development (BUCLD).

Archer, E., **Park, I. M.**, & **Pillow, J. W.** (2014). Bayesian entropy estimation for countable discrete distributions. *The Journal of Machine Learning Research*, **15**(1), 2833–2868.

Baillargeon, R., **Kotovsky, L.**, & **Needham, A.** (1995). The acquisition of physical knowledge in infancy. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 79–116). Clarendon Press/Oxford University Press.

Bartsch, K., & **Wellman, H. M.** (1995). *Children talk about the mind*. Oxford University Press.

Bates, E., & **Goodman, J. C.** (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. *Language and Cognitive Processes*, **12**(5–6), 507–584. https://doi.org/10.1080/016909697386628

Bates, E., & **MacWhinney, B.** (1987). Competition, variation and language learning. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 157–193). Springer.

Bates, E., **MacWhinney, B.**, & **others**. (1989). Functionalism and the competition model. *The Crosslinguistic Study of Sentence Processing*, **3**, 73–112.

Bonawitz, E. B., **Ferranti, D.**, **Saxe, R.**, **Gopnik, A.**, **Meltzoff, A. N.**, **Woodward, J.**, & **Schulz, L. E.** (2010). Just do it? Investigating the gap between prediction and action in toddlers' causal inferences. *Cognition*, **115**(1), 104–117.

Booij, G. (1996). Inherent versus contextual inflection and the split morphology hypothesis. In G. Booij, J. van Marle (Eds.), *Yearbook of morphology 1995* (pp. 1–16). Springer.

Callaghan, T. C., **Moll, H.**, **Rakoczy, H.**, **Warneken, F.**, **Liszkowski, U.**, **Behne, T.**, & **Tomasello, M.** (2011). *Early social cognition in three cultural contexts*. Wiley-Blackwell.

Chao, A., & **Shen, T.-J.** (2003). Nonparametric estimation of Shannon's index of diversity when there are unseen species in sample. *Environmental and Ecological Statistics*, **10**(4), 429–443.

Clahsen, H. (1986). Verb inflections in German child language: Acquisition of agreement markings and the functions they encode. *Linguistics*, **24**(281), 79–121.

Clancy, P. M., **Du Bois, J.**, **Kumpf, L.**, & **Ashby, W.** (2003). Argument structure in Korean. In J. W. Du Bois, L. E. Kumpf, & W. J. Ashby (Eds.), *Preferred argument structure: Grammar as architecture for function* (pp. 81–108). John Benjamins Publishing.

Clark, E. V. (2017). Morphology in language acquisition. In *The handbook of morphology* (pp. 374–389). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781405166348.ch19

Comrie, B., & **Polinsky, M.** (1993). *Causatives and transitivity* (Vol. **23**). John Benjamins Publishing.

Corrigan, R., & **Denton, P.** (1996). Causal understanding as a developmental primitive. *Developmental Review*, **16**(2), 162–202.

Courtney, E. H. (2002). Child acquisition of quechua causatives and change-of-state verbs. *First Language*, **22**(1), 29–71.

Devescovi, A., **Caselli, M. C.**, **Marchione, D.**, **Pasqualetti, P.**, **Reilly, J.**, & **Bates, E.** (2005). A crosslinguistic study of the relationship between grammar and lexical development. *Journal of Child Language*, **32**(04), 759–786.

Dittmar, M., **Abbot-Smith, K.**, **Lieven, E.**, & **Tomasello, M.** (2008). German children's comprehension of word order and case marking in causative sentences. *Child Development*, **79**, 1152–1167.

Dixon, R. M. W. (2000). A typology of causatives: Form, syntax and meaning. In *Changing valency: Case studies in transitivity* (pp. 30–83). Cambridge: Cambridge University Press.

Dunn, J. (2010). Understanding feelings: The early stages. In J. S. Bruner & H. Haste (Eds.), *Making sense (Routledge Revivals): The child's construction of the world* (p. 17). Routledge.

Eryiğit, G. (2014). ITU Turkish NLP web service. In S. Wintner, M. Tadić, & B. Babych (Eds.), *Proceedings of the demonstrations at the 14th conference of the European Chapter of the Association for Computational Linguistics (EACL)*. Association for Computational Linguistics. https://doi.org/10.3115/v1/E14-2001

https://doi.org/10.1017/S0305000925000200 Published online by Cambridge University Press

Fisher, C. (2002). Structural limits on verb mapping: The role of abstract structure in 2.5-year-olds' interpretations of novel verbs. *Developmental Science*, **5**(1), 55–64.

Fisher, C., **Gertner, Y.**, **Scott, R. M.**, & **Yuan, S.** (2010). Syntactic bootstrapping. *Wiley Interdisciplinary Reviews: Cognitive Science*, **1**(2), 143–149.

Ger, E., **Küntay, A. C.**, **Göksun, T.**, **Stoll, S.**, & **Daum, M. M.** (2022a). Do typological differences in the expression of causality influence preschool children's causal event construal? *Language and Cognition*, **14**(2), 161–184.

Ger, E., **Stuber, L.**, **Küntay, A. C.**, **Göksun, T.**, **Stoll, S.**, & **Daum, M. M.** (2021). Influence of causal language on causal understanding: A comparison between swiss German and Turkish. *Journal of Experimental Child Psychology*, **210**, 105182.

Ger, E., **You, G.**, **Küntay, A. C.**, **Göksun, T.**, **Stoll, S.**, & **Daum, M. M.** (2022b). Gradual route to productivity: Evidence from Turkish morphological causatives. *Cognitive Science*, **46**(12), e13210.

Gleitman, L. R. (1990). The structural sources of verb meanings. *Language Acquisition*, **1**, 3–55.

Göksun, T., **Küntay, A. C.**, & **Naigles, L. R.** (2008). Turkish children use morphosyntactic bootstrapping in interpreting verb meaning. *Journal of Child Language*, **35**(02), 291–323.

Haspelmath, M. (2001). The European linguistic area: Standard average European. In M. Haspelmath, E. König, W. Oesterreicher, & W. Raible (Eds.), *Language typology and language universals: An international handbook* (pp. 1492–1510). Mouton de Gruyter.

Hoff, E., **Quinn, J. M.**, & **Giguere, D.** (2018). What explains the correlation between growth in vocabulary and grammar? New evidence from latent change score analyses of simultaneous bilingual development. *Developmental Science*, **21**(2), e12536.

Hohenberger, A., **Kaya, U.**, & **Altan, A.** (2017). Discrimination of vowel-harmonic vs vowel-disharmonic words by monolingual Turkish infants in the first year of life. In *Proceedings of the 41st Annual Boston University conference on language development* (pp. 309–322).

Huang, C.-c. (2012). Referential choice and informativeness in mother–child conversation: A focus on the mother. *Language Sciences*, **34**(3), 319–338.

Kempe, V., & **MacWhinney, B.** (1998). The acquisition of case marking by adult learners of Russian and German. *Studies in Second Language Acquisition*, **20**(4), 543–587.

Ketrez, F. N. (1999). *Early verbs and the acquisition of Turkish argument structure* (Master's thesis). Boğaziçi University, Istanbul.

Ketrez, F. N. (2005). *Children's* scope of indefinite objects*. University of Southern California.

Küntay, A. C., **Koçbaş, D.**, & **Sabri Taşçı, S.** (Unpublished). Koç university longitudinal language development database on language acquisition of 8 children from 8 to 36 months of age.

Küntay, A. C., & **Slobin, D. I.** (1996). Listening to a Turkish mother: Some puzzles for acquisition. In D. I. Slobin, J. Guo, & A. Kyratzis (Eds.), *Social interaction, social context, and language: Essays in honor of Susan Ervin-Tripp* (pp. 265–286). Erlbaum.

Legare, C. H., **Sobel, D. M.**, & **Callanan, M.** (2017). Causal learning is collaborative: Examining explanation and exploration in social contexts. *Psychonomic Bulletin & Review*, **24**(5), 1548–1554.

Leslie, A. M., & **Keeble, S.** (1987). Do six-month-old infants perceive causality? *Cognition*, **25**(3), 265–288.

Lidz, J., **Waxman, S.**, & **Freedman, J.** (2003). What infants know about syntax but couldn't have learned: Experimental evidence for syntactic structure at 18 months. *Cognition*, **89**(3), 295–303.

Marchman, V. A., **Martínez-Sussmann, C.**, & **Dale, P. S.** (2004). The language-specific nature of grammatical development: Evidence from bilingual language learners. *Developmental Science*, **7**(2), 212–224.

Mazara, J., & **Stoll, S.** (2019). Development of verb morphology: From item-specificity to proficient use. In A. K. Goel, C. M. Seifert, & C. Freksa (Eds.), *Proceedings of the 41st annual conference of the cognitive science society* (pp. 2325–2331). Montreal, QC: Cognitive Science Society.

Messenger, K., **Yuan, S.**, & **Fisher, C.** (2015). Learning verb syntax via listening: New evidence from 22-month-olds. *Language Learning and Development*, **11**(4), 356–368.

Mintz, T. H., **Newport, E. L.**, & **Bever, T. G.** (2002). The distributional structure of grammatical categories in speech to young children. *Cognitive Science*, **26**, 393–424.

Moran, S., **Lester, N. A.**, **Gordon, H.**, **Küntay, A. C.**, **Pfeiler, B.**, **Allen, S. E. M.**, & **Stoll, S.** (2019). Variation sets in maximally diverse languages. In M. M. Brown & B. Dailey (Eds.), *Proceedings of the 43rd Boston University conference on language development* (pp. 427–440). Cascadilla Press.

https://doi.org/10.1017/S0305000925000200 Published online by Cambridge University Press

Moran, S., **Schikowski, R.**, **Pajović, D.**, **Hysi, C.**, & **Stoll, S.** (2016). The ACQDIV Database: Min(d)ing the Ambient Language. In N. Calzolari, K. Choukri, T. Declerck, S. Goggi, M. Grobelnik, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 16)* (pp. 4423–4429). European Language Resources Association (ELRA).

Muggeo, V. M. (2020). *Selecting number of breakpoints in segmented regression: Implementation in the R package segmented.* Technical report.

Muggeo, V. M. (2008). Segmented: An R package to fit regression models with broken-line relationships. *R News*, **8**(1), 20–25.

Murasugi, K., **Hashimoto, T.**, & **Kato, S.** (2003). On the acquisition of causatives in Japanese. *Online Proceeding Supplement of BUCLD 28*.

Naigles, L. R. (1996). The use of multiple frames in verb learning via syntactic bootstrapping. *Cognition*, **58**(2), 221–251.

Naigles, L. R., **Gleitman, H.**, & **Gleitman, L.** (1993). Children acquire word meaning components from syntactic evidence. *Language and Cognition: A Developmental Perspective*, **5**, 104–140.

Nakipoğlu, M., **Uzundağ, B. A.**, & **Sarigül, Ö.** (2022). Young minds' quest for regularity: Evidence from the Turkish causative. *Journal of Child Language*, **49**(6), 1214–1241.

Narasimhan, B., **Budwig, N.**, & **Murty, L.** (2005). Argument realization in hindi caregiver-child discourse. *Journal of Pragmatics*, **37**, 461–495.

Nemenman, I., **Shafee, F.**, & **Bialek, W.** (2001). Entropy and inference, revisited. In T. Dietterich, S. Becker, & Z. Ghahramani (Eds.), *Advances in neural information processing systems* (Vol. **14**). MIT Press.

Onnis, L., **Waterfall, H. R.**, & **Edelman, S.** (2008). Learn locally, act globally: Learning language from variation set cues. *Cognition*, **109**(3), 423–430.

Orlitsky, A., **Suresh, A. T.**, & **Wu, Y.** (2016). Optimal prediction of the number of unseen species. *Proceedings of the National Academy of Sciences*, **113**(47), 13283–13288. https://doi.org/10.1073/pnas.1607774113

Ozturk, O., & **Papafragou, A.** (2016). The acquisition of evidentiality and source monitoring. *Language Learning and Development*, **12**(2), 199–230. https://doi.org/10.1080/15475441.2015.1024834

Parra, M., **Hoff, E.**, & **Core, C.** (2011). Relations among language exposure, phonological memory, and language development in Spanish–English bilingually developing 2-year-olds. *Journal of Experimental Child Psychology*, **108**(1), 113–125.

Pearl, J. (2009). *Causality*. Cambridge University Press.

Perfors, A., **Regier, T.**, & **Tenenbaum, J. B.** (2006). Poverty of the stimulus? A rational approach. In *Proceedings of the 28th annual conference of the cognitive science society*. Cognitive Science Society.

Piaget, J. (1927). *La causalité physique chez l'enfant* (Vol. **4**). Librairie Félix Alcan.

Piaget, J. (1955). *The child's construction of reality*. Routledge. https://doi.org/10.4324/9781315009650

Prado Martín, F. M. del, **Kostić, A.**, & **Baayen, R. H.** (2004). Putting the bits together: An information theoretical perspective on morphological processing. *Cognition*, **94**(1), 1–18.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, **27**(3), 379–423.

Shibatani, M. (2002). *The grammar of causation and interpersonal manipulation* (Vol. **48**). John Benjamins Publishing.

Shibatani, M., & **Pardeshi, P.** (2002). The causative continuum. *Typological Studies in Language*, **48**, 85–126.

Sofu, H., & **Diser, S.** (2018). Acquisition of ditransitive verbs in Turkish. *Çukurova Üniversitesi Eğitim Fakültesi Dergisi*, **47**(2), 833–843.

Stoll, S., **Bickel, B.**, **Lieven, E.**, **Banjade, G.**, **Bhatta, T. N.**, **Gaenszle, M.**, **Paudyal, N. P.**, **Pettigrew, J.**, **Rai, I. P.**, **Rai, M.**, & **Rai, N. K.** (2012). Nouns and verbs in chintang: Children's usage and surrounding adult speech. *Journal of Child Language*, **39**, 284–321.

Şener, S., & **Takahashi, D.** (2010). Ellipsis of arguments in Japanese and Turkish. *Nanzan Linguistics*, **6**, 79–99.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Harvard University Press.

https://doi.org/10.1017/S0305000925000200 Published online by Cambridge University Press

Trabasso, T., & **Van Den Broek, P**. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language*, **24**(5), 612–630.

Ural, A. E., **Yuret, D.**, **Ketrez, F. N.**, **Koçbaş, D.**, & **Küntay, A. C.** (2009). Morphological cues vs. number of nominals in learning verb types in Turkish: The syntactic bootstrapping mechanism revisited. *Language and Cognitive Processes*, **24**, 1393–1405.

Yang, C. D. (2004). Universal grammar, statistics or both? *Trends in Cognitive Sciences*, **8**(10), 451–456.

You, G., **Bickel, B.**, **Daum, M. M.**, & **Stoll, S.** (2021). Child-directed speech is optimized for syntax-free semantic inference. *Scientific Reports*, **11**(1), 16527.

You, G., **Rüst, O.**, **Stoll, S.**, & **Baroni, M.** (2024). *Entropy correction in corpus studies* (In preparation).

Yuan, S., **Fisher, C.**, & **Snedeker, J.** (2012). Counting the nouns: Simple structural cues to verb meaning. *Child Development*, **83**(4), 1382–1399.

Cite this article: You, G., Ger, E., Daum, M.M., & Stoll, S. (2026). Children Learn Causatives Despite Pervasive Ellipsis: Evidence from Turkish. *Journal of Child Language* **53,** 950–971, https://doi.org/10.1017/S0305000925000200

https://doi.org/10.1017/S0305000925000200 Published online by Cambridge University Press