Martin, Fabienne & Ilić, Ivona & Schoenmakers, Gert-Jan & Alexiadou, Artemis. 2025. What do children do with *do*? Superfluous *do* in child English. *Glossa: a journal of general linguistics* 10(1). pp. 1–58.
DOI: https://doi.org/10.16995/glossa.10651

Open Library of Humanities

What do children do with *do*? Superfluous *do* in child English

Fabienne Martin, Utrecht University, f.e.martin@uu.nl

Ivona Ilić, Humboldt-Universität zu Berlin, ivona.ilic.1@hu-berlin.de

Gert-Jan Schoenmakers, Utrecht University, g.t.schoenmakers@uu.nl

Artemis Alexiadou, ZAS Berlin & Humboldt-Universität zu Berlin, artemis@zas-berlin.de

In English, superfluous *do* is phonologically unstressed, non-emphatic, and occurs in positive assertions. We propose that superfluous *do* has two different flavours in child English. First, auxiliary *do* is used by children to express Tense in a one-to-one relation with form. Second, children use *do* as a light verb to overtly express the concept of action, which is 'compressed' in the verbal structure and left silent in standard adult English. We report on studies on English CHILDES corpora showing that the *Aktionsart* of the verb used after *do* plays a role in the production of superfluous *do* in child English: in positive assertions, *do* is used before an agentive predicate (as in *I do play* or *I do painting*) more often in child speech than in child-directed speech. In the child data, superfluous *do* also occasionally appears after an auxiliary, which is to be expected if it behaves as a light verb. The results of the corpus studies also show that stative predicates are particularly frequent after *do* in positive assertions in both child and adult English, which we explain by the fact that stative predicates can have both generic and episodic uses in the present tense, whereas in the same tense, eventive predicates are much more restricted in their interpretation.

1 Introduction: overuse of functional verbs as undercompression errors in child language

Early work in language acquisition has often reported that across languages, children occasionally use *more* morphological material than expected to convey a given meaning (rather than less, as the case in omission errors). Overregularisations of irregular forms (plurals such as *mices* or simple past forms such as *broughted* in English, see Cazden 1968; Slobin 1971; Kuczaj 1977; 1978; Arnon 2009; Hein et al. 2024) are famous examples, but many other patterns have been described in the early literature on language acquisition (Slobin 1973; Karmiloff-Smith 1981; Bowerman 1982; Clark 1985; Slobin 1985) and references therein among others). The idea that children tend to overtly express underlying pieces of meaning typically compressed in the standard adult language is argued for by Slobin (1973). For Slobin, the children's tendency to overmark meaning is a generalised 'operating principle' (see (1)) and the children's temporary avoidance of the zero-morpheme constitutes a universal tendency (see (2)):[1]

(1) OPERATING PRINCIPLE E: Underlying semantic relations should be marked overtly and clearly. (Slobin 1973: 202)

(2) UNIVERSAL E2: There is a preference not to mark a semantic category by ∅ ("zero morpheme"). If a category is sometimes marked by ∅ and sometimes by some overt phonological form, the latter will, at some stage, also replace the ∅. (*ibid.*)

For Slobin, this tendency to overmark pieces of meaning normally left unexpressed in adult language reflects children's tendency to increase one-to-one mappings between meanings and forms. MacWhinney (1985) refers to Slobin's operating principles as underpinning his generalisation in (3):

(3) Children will overgeneralise non-zero allomorphs even when zero allomorphs are of much greater applicability. (MacWhinney 1985: 324)

Guasti et al.'s (2023) 'undercompression hypothesis' conveys the same idea cast in Sauerland & Alexiadou's (2020) Meaning First Approach.

Undercompression errors occasionally surface as superfluous or redundant use of a light or functional verb that expresses a piece of meaning normally elided (or compressed) in the verbal structure. An example of this kind of non-target use of light/functional verbs is documented in Bezinska et al. (2008) and Martin et al. (2022) about *faire* 'make' in L1 child French. Under the

[1] An anonymous reviewer notes that Slobin's universal E2 seems to predict that children go through a stage where the use of ∅ is *always* replaced by an overt morphology. This is obviously plainly false for most non-target productions by children. Under our interpretation, Slobin aims to describe a tendency which does not have to be always actualized. This also seems to be Bowerman's (2018) interpretation; for her, Slobin's principle states that there is a *preference* not to have a hole in the marking system, which does not entail that this preference will always be instantiated at the relevant stage (p. 10).

pattern of interest, French children appear to use overt causative marking superfluously, using the functional causative verb *faire* on top of a lexical causative verb (e.g., *montrer* 'show'), to express the very same meaning as the simple lexical causative. For instance, learners of French produce sentences such as (4a/b) to express the same meaning as the sentence without *faire*, as if *faire* were what Kittilä (2009) and Tyler (2022) call a non-valency increasing causative.

(4) French

a. tu peux me **faire** montrer? (Vanessa, 3;09)
tu peux me montrer? (target)
you can me make show?
'Can you show [it] to me?' (MTLN corpus, Le Normand 1986)

b. tu vas **faire** planter un arbre d'accord? (Antoine, 4;02)
tu vas planter un arbre d'accord? (target)
you will make plant a tree ok
'You'll plant a tree, OK?' (Paris corpus, Morgenstern et al. 2009)

Non-target uses of 'double causatives' (i.e. overt causative marking applied to a lexical causative) have also been reported in the production of L1 acquirers of Turkish (Aksu-Koç & Slobin 1985: 848). Similarly, Yamakoshi et al. (2018) show that children learning L1 Japanese understand double causatives as having the same meaning as simple causatives.

Another example of children's superfluous or redundant use of a light or functional verb expressing a concept that is typically elided or compressed in the verbal structure concerns light motion verbs. In French for instance, children occasionally use the verb *aller* 'go' in the context of a path verb (e.g., *descendre, monter* 'go down/up') to express exactly the same meaning as the path verb itself. To our knowledge, this pattern has not been systematically investigated yet, but we give some examples in (5)–(6), found in French CHILDES corpora.[2]

(5) French

là vais **aller** descendre avec ça
là je vais descendre avec ça (target)
there will go go.down with this
'I'm going to go down with this.'
(Madeleine, Paris corpus, Morgenstern & Parisse 2007)

[2] Example (5) is grammatical in standard French, but has a different meaning than the one that Madeleine aims to express, namely: 'I'm going to go and go down with this.' Sentence (6) is also grammatical in standard French without the pronoun *y* to mean 'I want to go and go upstairs'. But in standard French, the proform *y* doubling the locative adverbial (*en haut* 'upstairs') should then appear at the left of *monter* (*Je veux aller y monter, en haut* 'I want to go and go there, upstairs'). Here, the child seems to reanalyse the proform *y* as a part of the verb *y-aller*, not as a pro-form doubling the locative adverbial, which contributes to explain why it appears to the left of *aller*.

(6) French

je	veux	y	**aller**	monter	en haut	
je	veux			monter		(target)
I	wanted	there	go	go.up	upstairs	

'I want to go uptairs.' (Oli, GNP corpus, Paradis et al. 2000)

Martin et al. (2023) note that in most examples of overmarking in child language, the superfluous form added by the child to the target form (the verbs *faire* 'make' or *aller* 'go' in previous examples) is a separative/monoexponential form, that is, a form spelling out a single meaning unit in a particular instance, cf. Igartua (2015). In addition, Hein et al. (2023) note that in most cases, the superfluous form is either a *functional* element, or a form that shares with functional elements the property of expressing a grammatically encoded concept, that is, a concept exponed by closed-class items (prepositions, determiners, auxiliaries, light verbs, etc). We call those notions that are grammatically encoded across languages "functional concepts" (Cinque 2013, see also Cinque & Rizzi 2009). As Strickland (2017) argues, functional concepts are often core concepts, i.e. concepts that are innate and embedded in automatic cognitive processes. The concepts of cause, agency, action, event, animacy, core geometrical and numerical properties or the mass/count distinction are examples of core concepts (Spelke & Kinzler 2007; Baillargeon & Carey 2012). In the general spirit of Cinque 2013 and Cinque & Rizzi 2009, we take it that such concepts correspond to features associated with functional projections in the syntactic representation.

This paper focuses on the superfluous use of *do* in positive (present tense) utterances in child English, which is described in more detail in the next section. We argue on the basis of corpus studies that superfluous *do* in child English is another example of a functional verb used by the child to express functional concepts that are typically left silent in the target language.

Several scholars have argued that there is a (typically) unpronounced verbal head in agentive verbs, the content of which is closely related to the meaning of the overt light verb *do*, namely 'agentive event' or 'action' (see e.g. Folli & Harley 2008). Our proposal is that children initially overuse *do* as an exponent of this null verbal component with agentive verbs, as a result of their preference for overt exponence. This is consistent with a more general and cross-linguistic tendency of children to overuse light verbs.

The second part of our proposal is related to a previous hypothesis by Roeper (1991) and Hollebrandse & Roeper (1996). It concerns not only agentive sentences, but also stative sentences such as *We do like icecream*. We argue that in such sentences, the auxiliary *do* is occasionally used by children to express Tense with a separative free morpheme, in a one-to-one relation with form.

Adult English users also produce *do* in positive sentences like *I do like ice-cream*, but they favor *do* as an alternative to zero to signal emphasis (Bolinger 1972) via a manner implicature (Sauerland & Alexiadou 2020; Klein 2022: 16). Children are known to compute pragmatic

implicatures slowly (Guasti et al. 2005). We therefore expect the use by children of the zero morpheme and overt *do* to be up to some point in free variation rather than being governed by implicature-generating contexts as in adult language.

The idea, thus, is that superfluous *do* in child English comes in two guises; see **Figure 1**. It is used by the child to overtly express either the concept of action (in the context of an agentive verb), or the concept of tense (in positive sentences, with any kind of verbs). In adult language, the concept of action remains typically unpronounced in the context of an agentive verb, while the concept of tense is overtly spelled-out to convey emphasis only (in positive sentences, with any kind of verb).

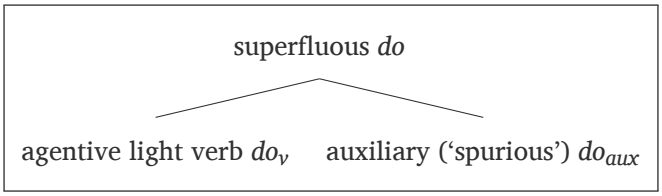


Figure 1: The two guises of superfluous *do* in child English.

The paper is structured as follows. Section 2 introduces emphatic and superfluous *do* in English. Section 3 presents the analysis we adopt for *do* in positive assertions in adult English. Section 4 spells out our hypotheses about superfluous *do* in child English. Section 5 presents the corpus studies that test these hypotheses. Section 6 reports similar uses of *do*-verbs to express the concept of action or agency in other languages, such as child Dutch and earlier varieties of English. Section 7 addresses the question of why stative predicates are particularly frequent after *do* in positive assertions in adult (and child) English.

2 Superfluous *do*

In standard adult English, auxiliary *do* must appear in several contexts when no other auxiliary is present. Negative sentences, questions with inversion, sentences with an emphatic meaning (phonological stress yielding verum focus), verb phrase ellipsis and VP topicalization are what Schütze (2004: 496) calls standard triggers for *do*-support; see (7).

(7) a. I don't know.
b. What do they say?
c. Well it's not nice when your favourite little boy tells you he doesn't like you. It [does]$_F$ make me very very sad. (child-directed speech, English CHILDES)
d. Ana likes Angelika, and Hamida does too.
e. We think we listen, but very rarely do we listen with real understanding, true empathy.

When none of the standard triggers for *do*-support is present, *do* does not need to appear, and in fact should not do so. However, it sometimes does; we are then dealing with what Schütze calls spurious *do*. Spurious *do* arises in the absence of the standard triggers of *do* listed in (7). It is phonologically unstressed and semantically non-emphatic. It appears in a positive assertion and is followed by a verb.

Spurious *do* is found in several varieties of English and in specific registers of standard English, where it is in free variation with zero. Spurious *do* is attested from the 16th to the 18th century (Lightfoot 1991; Warner 1993), as well as in many varieties of English (see Bohnacker 2013: 176 for an extensive list), such as Southwestern dialects of adult English from England (Klemola 1998) (see e.g. (8)).

(8) They [the sows] do come home at night and go in a shed see, when we do feed them. (Southwestern dialect, Klemola 1998)

Spurious *do* is also found in archaic legal English, as an alternative to the simple present or past tense (Quirk et al. 1985, who provide the example (9), cited by Schütze 2004):

(9) I, the undersigned, being of sound mind, do this day hereby bequeath…

Spurious *do* has also been claimed to exist in child (American and British) English. Like adults speaking dialectal or diachronic varieties just mentioned, learners of L1 English have been reported to produce *do* in contexts that do not contain standard triggers for *do*, i.e. unstressed, followed by a verb, in a positive context (Roeper 1991; Hollebrandse & Roeper 1996; Schütze 2004; 2013 and references therein):[3]

(10) You did make my bed a little fan. (Tim, 2;11, Roeper corpus, *apud* Schütze 2004)

For us, however, what Schütze (2004) calls 'spurious' *do*, which is the auxiliary *do*, is not the only guise of superfluous *do*. We propose that children also use the light verb *do* to express in a separative fashion the concept of action, which is 'compressed' in the verbal structure and left silent in adult modern standard English (as in many other languages). Differently from do_{aux}, do_{v} can appear in a sentence containing an auxiliary (since this second type of superfluous *do* is not itself an auxiliary), like in the following examples:[4]

[3] For child English data taken from CHILDES corpora, we give the name of the corpus after each utterance. The corpora investigated in our corpus studies are listed in **Table 1**.

[4] We keep the conventions of annotation used in the corpus data, and therefore include comments or corrections by the annotators (often in square brackets). Errors are either documented in the annotator's comments, or via the correct target form put in square brackets by the annotator. We put the relevant form of *do* followed by the verbal form in boldface font throughout this paper.

(11) a. he will **do** [*] **water** them. (Helen, 3;9, *Manchester*)

b. Context: Katla jumps up and down. Pleased, she says:
I did jump.
Adult: What?
Katla: I did jump.
Adult: You jumped, yeah.
Katla: Yeah, I DO did jump.[5] (Katla, 3;3.11, Bohnacker 2013: 188)

Corpus studies conducted on *do* in child English converge on the view that spurious *do* is clustered and very short-lived in child production. For instance, Bohnacker (2013) reports that 89% of the 129 occurrences of what she calls 'oversupplied' *do* she found in Katla's English were produced between age 3;1–3;3. During these two months, Katla uses *do* plus verb instead of finite verb in 25,2% of the cases (115/456). **Table 1** gives the time span during which spurious *do* is (mostly) attested in several studies (data are taken from Bohnacker 2013). As this Table shows, superfluous *do* seems to mostly be produced between 2 and 3 years old.

study	**child**	**time span for superfluous *do***
Zukowski (1996)	Ross	2;11–3;3
Bohnacker (2013)	Katla	3;0–3;6
Hollebrandse & Roeper (1996)	Tim	2;11–3;0
Knipschild (2007)	Joshua	2;4–3;1

Table 1: Time span reported for the production of superfluous *do* in longitudinal studies on children acquiring L1 English (data from Bohnacker 2013: 176).

This paper addresses the following questions. What do children do with *do*? What explains the *do* addition errors in early child English? Why do young children occasionally produce *do* in unstressed positive assertions? Does the *Aktionsart* of the verb used after *do* play a role in the production of superfluous *do* in child English? What is the distribution of verb classes after *do* in adult English? Does child speech reflect this distribution?

[5] Bohnacker (2013) analyses *jump* appearing after *do* in this utterance produced by Katla as a verb. An anonymous reviewer objects that it could also be a noun, in which case *do* is not an example of superfluous *do*. We observe that if Katla used *jump* as a noun here, she would probably have used a determiner at that age. Ute Bohnacker (p.c.) informed us that from a very early age, Katla showed productive use of both the definite and indefinite articles in English. There are thousands of articles before nouns in her data. She furthermore adds that many of the predicates after *do* produced by Katla unequivocally looked like verbs, and that she did not find any signs that pointed into the direction of them being nouns.

Our main hypothesis about superfluous *do* is that it has two different guises in child English. First, auxiliary *do* is used by children to express Tense with a separative free morpheme, i.e. in a one-to-one relation with form. This is related to the proposal of Hollebrandse & Roeper (1996); Roeper (1991) (see also Schütze 2013 on child spurious *do* as overtly realizing indicative Mood). Hollebrandse & Roeper (1996) argue that children prefer to express Tense with light verbs such as *do* instead of inflecting the thematic verb because this is easier — it involves fewer operations, such T to V lowering (Embick & Noyer 2001). Second, children use the light verb *do* to overtly express the concept of action, which is 'compressed' in the verbal structure and left silent in standard adult English.

This hypothesis makes two predictions, which we tested in corpus studies on English CHILDES. First, *do* should be used before an agentive predicate (or a predicate used agentively) more often in child English than in adult English, given that by assumption, children use *do* to express the concept of action, which is not expressed overtly in adult language. Analytic forms such as *do paint(ing)* rather than *paint(ing)* should be used more often by children than by adults. Second, since superfluous *do* in child English is not systematically auxiliary *do* but can also be used as a light verb, it should in principle also be found in child English after an auxiliary, in fact even *do* itself, as in Katla's last sentence in (11), repeated below for convenience.

(11) Yeah, I DO did jump. (Katla, 3;3, Bohnacker 2013: 188)

Before moving on to the corpus studies, in the next section we first present the analysis of *do* in positive assertions in adult and child English adopted in this study.

3 *do* in positive assertions

3.1 In adult English

In the literature, the standard analysis of *do* is that it is inserted in T(ense) to support affixes associated with this head in cases where T cannot combine with V (e.g. due to the presence of negation or due to T to C movement in matrix questions).

If *do* may be associated with T under some conditions, it is conceivable that *do* may be viewed as being linked with T more generally and thus as an allomorph of Present T ∅ in the context of a verb. The idea that superfluous *do* may be an allomorph has been pursued by Schütze (2004; 2013). For Schütze, superfluous *do* is the marked allomorph of the Mood feature, whose unmarked allomorph is zero. The null allomorph occurs in nonemphatic, positive declaratives and imperatives. Therefore, syntax makes *do* available to express Mood, and Schütze (2004; 2013) points to superfluous *do* in different varieties of English as support for his account.

For Schütze (2013), what conditions this allomorphy is that *do* appears when T features (e.g., *-ed*) do not get affixed to a main verb, auxiliary or modal, and thus need an alternative host. This happens when some element blocks T lowering to V, that is, among others, verum

focus or negation. T lowering to V can be overridden, but is made (quasi-)obligatory because of an economy constraint repeated in (12), close in spirit to Siddiqi's (2006) *Minimise Exponence* principle; see also Sauerland & Alexiadou's (2020) compression principle.

(12) If two convergent derivations based on the same numeration differ in the number of overt morphemes needed to spell them out, the derivation using fewer overt morphemes blocks the one using more morphemes. (Schütze 2013)

On this view, superfluous *do* in child English is a symptom of the fact that the economy principle is not yet fully in place.

We adopt this analysis and view auxiliary *do* as an exponent of T in standard English. In other words, *do* and ∅ are in free variation as exponents of T (or free allomorphs of T). As in Schütze (2013), we predict the occurrence of 'superfluous' *do* in certain registers or varieties of English, in line with the observations reported in the previous section; see also Sauerland & Alexiadou (2020).

Assuming that *do* overtly realises Tense, we can explain the emphatic meaning of overt *do* in standard adult English: when Tense is overtly realised by a dedicated morpheme, verum focus may scope exclusively on this element. Since Tense is responsible for event quantification (Kratzer 1996), applying verum focus only on the element that existentially quantifies over the event yields the implicated emphasis on event actualisation. In contrast, with inflected thematic verbs, T is realised by overt or covert affixes. Null affixes cannot be focused, and if the affix is overt (as e.g. *-ed*), it is not possible to put an accent conveying verum focus on this affix while leaving the thematic verb unaccented (in order to avoid that the lexical information is also emphasised). So for instance, (13a) cannot convey emphasis on the actualisation of the working event; this utterance works at best as a *corrigens* sentence by which the speaker indicates that the addressee has not used the appropriate tense affix (Steube 2001). Furthermore, sentence (13b) cannot convey emphasis on event actualisation without also conveying emphasis on the lexical information conveyed by *work*. Sentence (13c) is the best choice for this, because *did* conveys only tense information, since dummy *do* is lexically empty.[6]

(13) a. She workED.
b. She WORKED.
c. She DID work.

Note that in this approach, there is a natural connection between emphatic *do* and other cases of overt realisation of an optional morpheme, the realisation of which typically triggers a manner implicature: the hearer understands that conditions are not normal and understands

[6] The explanation will have to be different for positive imperatives, where overt *do* is also optional but conveys politeness rather than emphasis (see Klein 2022, who also mentions Penhallurick 1985: 327 for experimental evidence in support of this claim). Klein (2022) tries to derive all extra pieces of meaning triggered by *do* in a variety of contexts via a core 'speaker-distancing' function of *do*.

that the more complex form is chosen to signal the presence of a more complex meaning. As Sauerland & Alexiadou point out, the same mechanism is observed with periphrastic causatives. For McCawley (1978), van Rooy (2004) or Benz (2006) for instance, lexical causatives and their *cause*-periphrastic counterparts can mean the same thing (even if the event chain is packaged differently by the two types of expressions, as Fodor 1970 made clear), and the directness constraint that is often associated with lexical causatives is derived as a manner implicature: the use of the marked (periphrastic) form triggers the inference that causation is non-prototypical (e.g., indirect), and consequently, the unmarked (lexical) form is associated with 'canonical' (e.g. direct) causation.

For Schütze (2013: 28), varieties with superfluous *do* (e.g., Early Modern English) do not include the economy principle (12): T lowering to V becomes optional in these varieties, and T does not need to combine with the main verb even when nothing blocks the combination. Failure to lower T leaves T features without a host, and *do* is chosen. Note, however, that it seems rather implausible that economy is not at play yet in earlier stages of the language. In Section 5, we refer to studies suggesting that superfluous *do* is not superfluous after all in Early Modern English and realises the concept of action.

But what about contemporary adult dialectal or register (e.g. legal) varieties? Again, it does not seem plausible that the economy principle is completely suspended in the grammar of all adult varieties that use superfluous *do*.

One way of explaining the existence of superfluous *do* in adult dialectal or register variants (which is compatible with Schütze's analysis) is to derive it from a difference in the ranking of communication principles across varieties (see Martin 2022 on expletive negation in French). Specifically, the idea is that standard varieties rank Gricean pragmatic principles such as *Avoid redundancy* and *Be Brief* higher than vernacular or colloquial varieties (see also Frei 1929). Because economy principles are ranked high in standard varieties, verbose or pleonastic forms such as *do* in the context of a verb must be avoided, since more economical ways of expressing the same meaning must be preferred. Vernacular or colloquial variants, on the other hand, are more lax with *Be Brief* and rank something like *Be clear* (i.e., favour morphological transparency) and *Be expressive* higher (or *Be extravagant*, to follow a suggestion from Haspelmath 2014). Dialectal and colloquial variants are often said to be "more expressive" (Frei 1929). Consequently, they are more tolerant of superfluous productions. It is also plausible that legal language (recall (9)) also ranks *Be clear* higher than in standard registers, since it is known that redundancy and clarity are a crucial tool in legal writing (Golden 2016).

3.2 In child English

We propose that superfluous *do* in child English is an example of the child language errors looked at by Slobin (1973) and Guasti et al. (2023) in which the child tends to prefer overt to covert forms to express an underlying piece of meaning (while adult language typically leaves this piece

of meaning unexpressed). We show in which respect superfluous *do* is a further example of such undercompression errors in the next subsections.

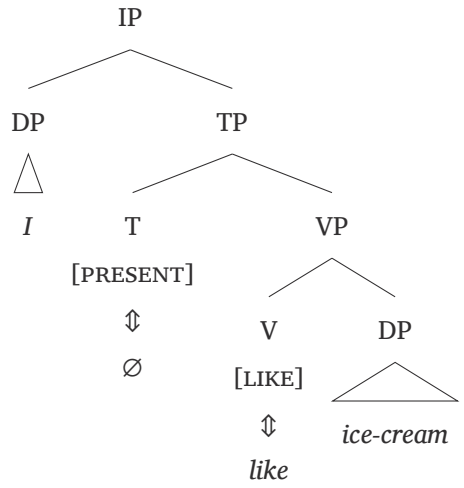
3.2.1 Superfluous *do* as the spell out of Tense

In a first case, the child uses *do* to realise Tense with a separative and free morpheme (cf. Hollebrandse & Roeper 1996; Roeper 1991, and Schütze 2013 for the related idea that in child English, superfluous *do* overtly realises indicative Mood). Auxiliary *do* systematically realises Tense when it embeds an unaccusative or a stative verb, as in (14) below (assuming that these examples are unstressed and noncontrastive, they are examples of superfluous *do* in early child English).[7]

(14) a. it does come out. (Becky, 2;6, *Manchester*)
b. The spider does like this car. (Fraser, 2;10, *MPI-Eva-Manchester*)

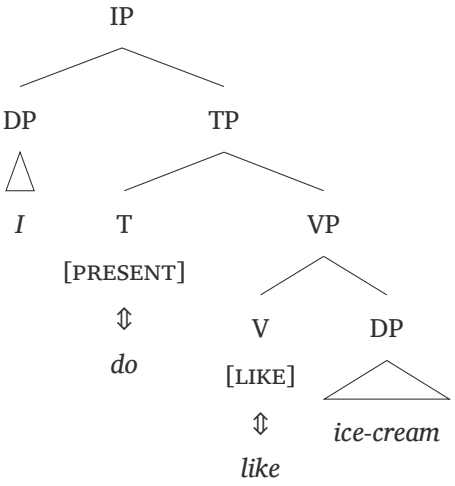
When children express Tense overtly without any additional emphasis, i.e. preferring the overt allormorph *do* to the zero allomorph affix to express 'just' PRESENT in the context of VP, they disregard economy or Sauerland & Alexiadou's (2020) principle of compression ('Elide whenever you can') (or the related 'Minimise Exponence' principle from Siddiqi 2006). We illustrate the difference between the realisation of the structure [I [PRESENT [LIKE [ICE-CREAM]]]] by adults and children in (15) and (16) (see Hein et al. 2023 for a similar treatment of undercompression errors across child language). Children occasionally overtly spell out PRESENT with *do* ignoring the fact that this choice triggers a manner implicature in the target language.

(15) *I like ice-cream* in standard adult English



[7] As Schütze (2013: 15) also notes, it is desirable to study the phenomenon of superfluous *do* on corpora containing sufficient prosodic information to be able to identify when *do* is used emphatically by the child. Although some English CHILDES corpora provide audio or video files, we did not have the time nor the expertise to analyse the prosodic profile of each child utterance containing an occurrence of *do* followed by a verb. But in the corpus study reported in Sections 4.1.3–4.1.4, we took advantage of the context to identify emphatic uses of *do* in child production.

(16) *I do*$_{superf}$ *like ice-cream* in child English



This occasional preference of children for a dedicated morpheme to express present Tense in a one-to-one fashion, that is, *not* to convey a manner implicature as in the target language, may be related to the difficulty of young children in deriving implicated meaning. This often involves some reasoning about why speakers use a particular variant rather than another (e.g. less verbose) variant, something very young children have been argued to struggle with, at least in some contexts (Guasti et al. 2005). Under Martin's (2022) hypothesis that varieties are characterised by different ranking of pragmatic principles, this type of non-target production can also be linked to a difficulty in ranking pragmatic principles as in the target language.

Manner implicatures in development are notoriously understudied compared to the acquisition of scalar implicatures (see discussion in Wilson 2017 and Wilson & Katsos 2020), and existing studies are mostly conducted with older children (5 to 6 in Wilson 2017, and 6 to 12 in Antoniou & Katsos 2017). The corpus study reported below and previous studies on superfluous *do* in child English can therefore be seen as informing the acquisition of manner implicatures by very young children.

3.2.2 Superfluous *do* as the spell-out of the concept of action

As **Figure 1** in the introduction illustrates, we propose that superfluous *do* in child English comes in two guises. While with non-agentive verbs, superfluous *do* can only be auxiliary *do*, in the context of agentive verbs, it can either be auxiliary *do* or the light verb *do*. That is, in the context of agentive verbs, the child uses superfluous *do* either to overtly realise Tense (as with non-agentive verbs), or, alternatively, to express the core concept of action 'incorporated' in agentive verbs, via a reanalysis of auxiliary *do* as the light verb *do*. Examples (17) are potential examples of the latter case.

(17) a. MOTHER: what're you doing ?
CHILD: I [/] I **do eat** it . (Helen, 3;6, *MPI-EVA-Manchester*
b. MOTHER: what do you do with a recorder ?
CHILD: I **do play** this one . (Becky, 2;10, *Manchester*)

In many different frameworks, statements such as *I eat* or *I play* are assumed to underlyingly involve the concept of action, which remains unexpressed in English as in many other languages. For instance, in his discussion of active vs. stative statements, Dowty (1979) builds the notion of agency in the meaning of an (unexpressed) predicate DO (see also Ross 1972). Similarly, in Harley (1995) and Folli & Harley (2008), the v_{do} head introducing the external argument in its specifier (roughly corresponding to Voice in Alexiadou et al. 2015, as Folli & Harley did not distinguish between *v* and Voice at the time) is itself also often not realised across languages.[8]

Our proposal is that the concept of action (or agentive event) encoded by these functional elements is occasionally overtly expressed by children under the guise of superfluous *do*. The notions of action and agent we assume here do not presuppose animacy. An action is an event which involves an entity occupying the thematic role of Agent as one of its participant. Such an entity is often an animate, but needs not be. For instance, the subjects of unergative verbs are agents, and some unergative verbs (like sound emission verbs) accept inanimate subjects (Folli & Harley 2008). In order to be an agent, an entity must minimally 'do' something (i.e. be an effector), which is sometimes translated in force-theoretic frameworks as the generation of a force vector (Gardenfors 2014; Copley & Harley 2015).

More concretely, we assume that activity verbs (unergatives and non-core transitives) start out with a structure that involves the light verb *do* and a root (Harley 1995; Folli & Harley 2005), the latter incorporating in the verbal head. The verbal head v_{do}, which is spelled-out by the light verb *do*,[9] associates the role of Agent to the external argument (which can be animate or

[8] Some languages have overt agent-marking morphologies, but as observed e.g. in Martin et al. (2023), these morphologies typically either express a weaker level of agency than usual (see e.g. the out-of-control or limited control morphologies across languages) or by contrast a stronger level of agency than expected, as for instance the specific verbal prefix in Ahcenese described in Legate (2014) or perfective markers in Tibetan dialects like Lhasa or Newari, which indicate that the event is under the *conscious* and *intentional* control of the subject (Delancey 1985: 52). Another relevant case is Dutch 'dummy' *doen*. Although the redundant use of *doen* is rather stigmatised (Schoenmakers 2023) and perceived as typical of child language, it occurs in adult L1 and then expresses either habituality, or what Sert et al. (2023) call 'intentional aspect'. Joo et al. (2025) offer experimental evidence for the idea that the 'normal' level of agency as encoded by agent Voice in English is achieved as soon as the agent is in control of their action, no matter whether they act consciously or intentionally. That overt agent-marking morphologies often express a higher or lower level of agency than expected is unsurprising under the view that languages rather mark less frequent meanings than more frequent ones (Haspelmath et al. 2014).

[9] An anonymous reviewer objects that if the verb *do* is the spell-out of v_{do}, we seem to predict that any verb formed with the light verb *do* should be obligatorily transitive, like light verb *do* is, contrary to the facts. We do not believe, however, that the transitivity of the light verb *do* is inherited by the verb whose structure it contributes to form.

inanimate).[10] Adopting Folli & Harley's (2005) analysis for the presentation, adults leave the agentive verbal head v_{do} unpronounced in the context of a root:

(18) Rules for Vocabulary Insertion in adult English

a. $\varnothing \Leftrightarrow [v_{do}]/\_\_[V]$

b. *do* $\Leftrightarrow [v_{do}]$

c. *play* $\Leftrightarrow$ [PLAY] (Harley 1995)

By contrast, children occasionally "disincorporate" the root from the verbal structure: they occasionally express the concept of action overtly by using the exponent 'do' in (18b) instead of the target null exponent in (18a), and the remaining conceptual structure is exponed by the predicate; compare also (20) to (21).

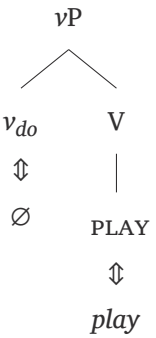
Adopting an idea of Hein et al. (2023; 2024), we propose that the redundant addition of *do* in the context of an agentive verb is the result of the child's occasional flouting of specificity when secondary features are involved:

(19) *Secondary feature negligence*
Children's representations of secondary features are unstable. Children therefore occasionally fail to consider them during Vocabulary Insertion. (Hein et al. 2024)

That is, the child fails to see that *do* should not be chosen as an overt spell-out v_{do} in the context of a verb, and occasionally prefers it to zero in this context (thereby also satisfying a preference for a more transparent mapping between semantics and morphology).

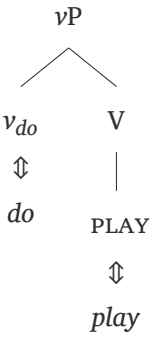
In (21), we extend the analysis of superfluous *faire* in child French by Hein et al. (2023) to superfluous *do* in child English.

(20) *play* in standard adult English



[10] While Folli & Harley (2005) still assume that v_{do} requires an animate agent, Folli & Harley (2008) drop this requirement and assume that agents (which they define through the notion of teleological capability) can also be inanimate.

(21) do_{superf} *play* in child English.



This proposal makes two predictions, supported by the corpus study presented in Section 4.1. First, we expect to find superfluous *do* with agentive verbs or verbs used agentively more often in child language than in adult language, because we assume that children sometimes use *do* to express the concept of action left silent in the target language. Second, superfluous *do* should sometimes appear after an auxiliary, and this only in the context of an agentive verb or a verb used agentively, since we assume that when *do* is used to express the concept of action, it is the light verb *do*.

3.3 The overuse of light verb constructions with *do*

While superfluous *do* is a good example of the overuse of *do* in child English, other patterns instantiating the same tendency do exist. One such pattern are light verb constructions built with *do*, as in the examples (22).

(22) a. **do drawing** . (Carl, 2;0, *Manchester*)
b. **do crayoning** , Mummy . (Becky, 2;1, *Manchester*)
c. I like to **do counting** [?] . (Eleanor, 2;2, *MPI-EVA-Manchester*)

Clark (1985) and Kambanaros & Grohmann (2015) argue that across languages, children with specific language impairment but also children with typical language development tend to overuse light verb constructions. Clark (1985: 737) reports that young learners of French, when describing activities, rely heavily on *faire* 'make' embedding a noun rather than using the more economical lexical form, e.g. say *faire les jardins* 'make the gardens' instead of *jardiner* 'garden', or even *faire la marche* 'make the walk' instead of *marcher* 'walk'. In a production task with children acquiring L1 Greek, Kambanaros & Grohmann (2015) similarly observe that children who are asked to name an action in a single verb tend to use a variety of multiword responses; e.g., instead of using the target verb *chtízo* 'build', children use the Greek counterparts of 'making bricks' or 'doing bricks' (e.g. producing *kano tuvla* 'I make bricks'). Alexiadou & Rizou (2023) noted that speakers of Heritage Greek in the US (as well as monolingual speakers of Greek in informal settings) also use periphrastic constructions instead of lexical verbs.

Wittenberg (2018) suggests that children might use light verb constructions as a productive device to make a verbal structure out of a noun, as observed by Wiese (2009) and Kotsinas (1988) in the production of multilingual speakers. Barner (2001), in contrast, claims on the basis of a study on child English that true light verbs appear late, suggesting that children prefer to use the lexical verb rather than the light verb construction (e.g., *kiss* rather than *give a kiss*). However, Barner also notes that among light verbs, *do* is the most frequent in early and late child English. We hypothesise that *do* presents a different profile than other verbs appearing in light verb constructions such as *give*, because *do* is by hypothesis simpler to acquire as it is a separative morpheme (in one-to-one relation between form and meaning unit): while *do* expresses a single concept (the action concept), this is not the case of *give*, which compresses more information in a single morpheme.

More specifically, we hypothesise that young children overuse the *do* + ing pattern as the result of a more general tendency to undercompress (Slobin 1973; Guasti et al. 2023). Again, children maximise morphological transparency: in examples such as (22), the abstract DO action concept and the more specific activity concept (e.g., the DRAW concept) are exponed in a one-to-one fashion by the light verb *do* and the root within the *-ing* form, rather than being compressed within a single morpheme (e.g., *draw*).

One question is whether in these child utterances, *do* is necessarily the light verb *do* or can also alternatively be the auxiliary *do*. This is rather implausible when the child uses the *do* + ing construction after an auxiliary like *do* or *can*, as in the examples (23).

(23) a. I **did do skipping** . (Eleanor, 2;4, *MPI-EVA-Manchester*)
b. **can** I **do cutting** ? (Anne, 2;7, *Manchester*)
c. I don't **wanna do drawing** , Mum . (Becky, 2;7, *Manchester*)

But examples (22a/b) are different in two respects. First, they do not contain an auxiliary, and second, they are not grammatical in adult English, while examples as in (23) are sometimes produced by adults (although as shown in our third corpus study reported in Section 4.2, adults very clearly prefer to introduce a determiner before the *-ing* form, a clear sign that it is used as a nominal gerund). One possibility is that examples like (22a/b) reflect the choice of *do* as a 'wrong' auxiliary (instead of *be* in (22a/b)). If it were the case, however, we should observe the same pattern in the context of stative verbs, which is disconfirmed in our data (as will be shown in Section 4.2, we hardly found examples of *do* followed by the gerund of a stative verb).

Another possibility is that in examples (22a/b), the child adds the light verb *do* to what is known as the 'primitive progressive' form. We know that L1 English learners first use the 'primitive progressive', i.e. the progressive form without auxiliary (Brown 1974: 318). In this first stage, they use the primitive progressive mainly with action verbs, but also with stative verbs (Brown 1974, *ibid.*).

In this perspective, these examples reflect children's tendency to 'decompress' the progressivised agentive verb into several parts: (i) the light verb *do* spells out the general concept of action, as we hypothesised to also be the case when *do* is combined with a *bare* agentive verb (see Section 3.2.2), (ii) the root expresses the more specific activity concept and (iii) the progressive carries some aspectual features.

At a later stage, an auxiliary can be added to this 'decompressed' form of the progressivized verbal form, as in (24):

(24) I'**m do** [*] **painting** . (Eleanor, 2;0, *MPI-EVA-Manchester*)

The hypothesis that young children overuse the *do* + ing light verb construction is supported by the corpus study reported in Section 4.2.

4 Corpus studies

4.1 Do+verb

4.1.1 First study: Methodology

In the CHILDES database (MacWhinney 2000), we extracted all utterances surfacing with *do* or *does* from 50 corpora of typically developing children acquiring North American and British English. We used the CLAN software to extract the data. For British English, we collected data from 210 children (age range: 1;0–7;0), and for North American English, we collected data from 1112 children (age range: 0;6–8;0). The corpora are listed in **Table 2**.[11]

We first extracted all utterances containing the verbs *do* or *does* (n = 46,231). From this set, we extracted those in which *do/does* embeds a verb (n = 710).[12] We included utterances with a null subject, which are common in early child English. In order to exclude potential adult-like emphatic *do*, we focused on the time span up to 3;6. We also discarded repetitions as well as questions or imperatives, arriving at a total of 438 occurrences.[13] We report the developmental trajectory of *do/does* + verb in positive assertions in Appendix 1.

[11] The data set is available at doi.org/10.5281/zenodo.15653004.

Stromswold (1990) already carried out computer searches on around 200,000 lines in the CHILDES data of 14 American English children (age 1;2–7;10). She found around 55.000 auxiliaries, including instances of *do*. She found very few instances of superfluous *do*, which she analyses as speech errors rather than some (weak) evidence of superfluous *do* in the child grammar. As Bohnacker (2013) observes, however, Stromswold mainly searched for doubly inflected VPs (*she did ate*). Therefore, she may have missed many potential cases of non-target unstressed instances *do* + infinitive in positive assertions.

[12] We discarded occurrences of *do* in the past to keep the set of data manageable, and leave an analysis of *did* followed by a verbal form for future research.

[13] Out of 50 corpora that served as the initial sample for data extraction, co-occurrences of *do* and a verb is attested in the following 30 corpora and is produced by around 50 children: Bloom, Bohannon, Braunwald, Brown, Clark, Demetras, Fletcher, Forrester, Garvey, Gelman, Gleason, Haggerty, Higginson, Kuczaj, Lara, MacWhinney, Manchester, MPI-EVA-Manchester, Nelson, New England, Post, Sachs, Snow, Suppes, Thomas, Tommerdahl, Valian, Van Houten, Weist, and Wells.

North American English corpora	British English corpora
Bates (Bates et al. 1991); Bernstein (Bernstein 1984); Bliss (Bliss 1988); Bloom (Bloom 1970); Bohannon (Bohannon III & Marquis 1977); Braunwald (Braunwald 1971); Brown (Brown 1974); Clark (Clark 1978); Demetras (Demetras 1986); Evans (MacWhinney 2000); Feldman (Feldman 1998); Garvey (Garvey 1979); Gathercole (Gathercole 1986); Gelman (Gelman et al. 1998; 2004); Gleason (Gleason 1980); Haggerty (Haggerty 1929); Hall (Hall & Nottenburg 1981); Higginson (Higginson 1985); HSLLD (Dickinson & Tabors 2001); Kuczaj (Kuczaj 1977); MacWhinney (MacWhinney 2000); McCune (McCune 1995); McMillan (MacWhinney 2000); Morisset (Morisset et al. 1995); Nelson (Nelson 1989); New England (Ninio et al. 1994); Newman Ratner (Newman et al. 2016); Post (Post 1992); Rollins (Rollins 2003); Sachs (Sachs 1983); Sawyer (Sawyer 1997); Snow (MacWhinney & Snow 1985); Sprott (Sprott 1992); Suppes (Suppes 1974); Tardif (MacWhinney 2000); Valian (Valian 1991); Van Houten (Van Houten 1986); Van Kleeck (MacWhinney 2000); Warren (Warren-Leubecker 1982); Weist (Weist et al. 2009)	Fletcher (Fletcher & Garman 1988); Forrester (Forrester 2002); Howe (Howe 1981); Kelly Quigley (Kelly et al. 2020); Lara (Rowland 2007); Manchester (Theakston et al. 2001); MPI-EVA-Manchester (Lieven et al. 2009); Thomas (Lieven et al. 2009); Tommerdahl (Tommerdahl & Kilpatrick 2014); Wells (Wells 1981)

Table 2: Summary of all corpora searched.

In the next step, we manually classified the verbal class of the predicate following *do/does* along the classification given in **Table 3**. Non-core transitive verbs are manner verbs without causative semantics (Levin 1999). Core transitive verbs are result verbs with causative semantics (Rappaport Hovav & Levin 1998 among many others). We give examples of verbs of child occurrences of *do* + verb for each aspectual type in (25).

(25) a. *stative*
because dolly **do** [*] **like** them . (Dominic, 2;6, *Manchester*)

b. *non-core transitive*
Baaee [= Mommy] **does do** poopee [: pee pee] . (Nathaniel, 2;5, *Snow*)

c. *core transitive*
I **do open** egg box . (Dominic, 2;6, *Manchester*)

d. *achievement*
Emmy **do find** the &bl blan(ket) . (Emily, 1;10, *Nelson*)

e. *anticausative*
that **does** [*] **open** . (Eleanor, 3;0, *MPI-Eva-Manchester*)

f. *unergative*
car **do** [*] **go** . (Dominic, 2;5, *Manchester*)

Verbal class	Example
non-core transitives (NCTs)	*drink*
core transitive (causatives)	*break*
transitive achievements	*lose*
anticausatives/unaccusatives	*break/fall*
unergatives	*walk*
statives	*know*

Table 3: Typology of aspectual verb types adopted for the annotation of the verb following *do/does* in the child and child-directed speech data.

We also annotated the thematic roles of the subjects (as well as their person features for the follow up study on the person/aspect split reported in Section 6.2) in the child data. We annotated the external argument of non-core transitive verbs, unergatives and causatives with an agentive subject as Agents. We also annotated the agents' animacy. Most agents were animate, but we did find inanimate agents in our sample, see (26)–(27) for some examples in child vs. child-directed speech respectively.[14] **Table 4** presents the distribution of inanimate and animate agents for each relevant class, but we did not include this information in our statistical analyses, because we have no specific hypotheses about animacy.

(26) Non-animate agents in child speech

a. **does bounce** . (Eleanor, 2;0, *MPI-Eva,Manchester*)

b. dolly **does got** her dress of . (Debbie, 2;3, *Wells*)

c. MOTHER: that's (.) the Furby isn't it ?
CHILD: yeah .
MOTHER: you put them together (.) or you put something metal by it (.) it makes a noise .
CHILD: oh it **does sing** . (Fraser, 3;0, *MPI-EVA-Manchester*)

(27) Non-animate agents in child-directed speech

a. with that pen? let me see how you do it. okay. does it oh it **does write** see?

b. It does fly. what what is it? **does fly**. it's it's an aeroplane isn't it?

c. the Spar shop there which is very near AuntieJack's house yeah? that **does sell** gas canisters because yeah?

[14] Following many authors, we therefore do not assume that animacy or volitionality is a prerequisite for agency as encoded in the grammar (on the idea that agents in language are not necessarily animates, see e.g. Cruse 1973; Delancey 1985; van Valin & Wilkins 1996; Alexiadou & Schäfer 2006; Koontz-Garboden 2009; Folli & Harley 2008 and Fauconnier 2012 among many others; and see Dowty 1979 and Demirdache 1997 among others on the view that agency in language does not presuppose volitionality). See Joo et al. (2025) for a formal definition of the role of agent, which encompasses both animate and inanimate agents.

We group under the broader label of Non-Agent subjects that are associated with the role Theme (such as the subject of unaccusative verbs or achievement verbs, see Piñón 1997 among others), the role Experiencer of subject-experiencer psych-verbs as well as the non-agentive (causer) subject of causative verbs.

One of our basic assumptions is that the distribution of verb classes in a given construction – here *do* followed by a bare, aspectless verb – in (child-directed) adult production can serve as a baseline for evaluating child production. Therefore, we applied the same methodology to annotate child-directed adult speech in the CHILDES database. We extracted all instances of present tense *do/does* followed by a verb using a CQL search in Sketch Engine on the enTenTen21 corpus (see Kilgarriff et al. 2014). The CQL search used to extract the data is repeated in (28) (the search was restricted to adult participants).[15]

(28) `[word = "does|do"][tag="VV.*"]`

This search returned 3915 tokens, from which we manually extracted all instances of positive assertions (N = 1886), excluding negative sentences, questions or *do*-imperatives, as well as 339 occurrences that did not instantiate the *do* + verb pattern or were not parsable. The final adult dataset contained 1547 items. We provide examples of *do* + verb utterances in child-directed speech for each aspectual class in (29).

(29) a. *stative*
I **do need** my coat.
b. *non-core transitive*
I drive that one. You **do drive** that.
c. *core transitive*
Yeah on the changing table I **do put** a diaper over her face and peek a boo with her.
d. *achievement*
It **does begin** with letter *t*.

[15] Recall that we exclusively looked at occurrences of *do/does* in child data, leaving the analysis of *did* followed by a verbal form for further research.

Furthermore, the adult data were not restricted to files in which the children were 3;6 or younger. An anonymous reviewer suggests that we should have done so, given the hypothesis that children by default are expected to match what they hear. Utterances they hear at later ages might not be relevant to their expected productions at earlier ages, as it is plausible that adults adapt their speech to children as children's language develops.

While we understand the reviewer's reasoning, we have not filtered the adult data according to the age of the child the adult is addressing for several reasons. The first is practical: identifying the age of the child addressed by the adult would have been extremely time-consuming. The second is that in a significant number of cases, an adult addresses several children. Thirdly, we believe that in natural contexts, children hear adult speech directed at children other than themselves, and that this also influences their production.

e. *anticausative*
Oh I thought it **does turn**. there look.

f. *unergative*
Oh no they **do fly**? They **do fly**. No they don't fly I don't think.

We annotated the child-directed speech occurrences along the same criteria as child data: (i) verb following *do/does*, (ii) class instantiated by this verb and (iii) thematic role (and person features) of the subject.[16]

We used the software JASP (version 0.19.3, JASP Team 2025) to run χ^2 tests on the data. The results inform us about potential associations between two categorical variables; the adjusted standardised residuals $\tilde{r}$ can then inform us about the relative contribution of each cell in the contingency tables and are therefore indicative of the source of a potential significant outcome. For this, we can make use of the standard critical values, since the adjusted Pearson's residuals are normally distributed. We applied a Bonferroni-correction to account for multiple comparisons.

4.1.2 First study: Results

In this first study, we compared the distribution of each verb class in the CHILDES database child data from children up to 3;6 years old and in child-directed speech, see **Table 4**.[17] We did not include animacy as a factor in our analyses. Compared to child-directed speech, the child data are characterized by a higher use of agentives (36.5% vs. 20.7%), in particular of unergatives (14.2% vs. 4.8%), and by a lower use of statives after *do* (53.4% vs. 67.3% with following statives). Apart from a slightly higher use of non-core transitives after *do* in child data (13.0% vs. 8.5%), other verbal classes have roughly the same distribution.

The results of a χ^2-test confirm the significant association between population and the agentivity of verbs ($\chi^2(1) = 46.30$, $p < .001$) as well as between population and the more fine-grained distribution of verb types ($\chi^2(6) = 67.60$, $p < .001$). Post-hoc hypothesis testing indicates that the significant association between population and agentivity is due to children

[16] We annotate as causer subjects DPs that refer to an eventuality (see e.g. Alexiadou et al. 2015) or a state-of-affair, as in the example (30) below, where *that* refers back to the eventuality of wearing a hat.

(30) he wants to wear the hat to work? yeah . yeah that's right that **does help** him work.

There are only few cases of causer subjects in the child and adult data. Most of them are found with object experiencer verbs or verbs of interaction like *help* (26 out of 38 cases in the adult data, 3 out of 5 in the child data; both classes of verbs are much less present in the child than adult data, which likely contributes to explain why causers are more frequent in the adult dataset). These verbs are known to welcome eventuality-denoting DPs in subject position (see Bouchard 1995; Engelberg 2005; Bott & Solstad 2014).

[17] Note that among transitive verbs, transitive achievements can only be non-agentive because since they denote instantaneous eventualities, these eventualities never amount to actions (given that any action takes *some* time, if only a short time). This also explains why achievement verbs are not compatible with agent-oriented adverbials such as *carefully* (Piñón 1997: Section 2.4).

Verbal class	***N* (child)**	**% (child)**	***N* (adult)**	**% (adult)**
non-core transitives	57	**13.0**	132	**8.5**
animate subject	*52*	*11.9*	*127*	*8.2*
inanimate subject	*5*	*1.1*	*5*	*0.3*
core transitive (causatives)	41	**9.4**	114	**7.4**
animate subject	*36*	*8.2*	*104*	*6.7*
inanimate subject	*5*	*1.1*	*10*	*0.6*
unergatives	62	**14.2**	75	**4.8**
animate subject	*37*	*8.4*	*52*	*3.4*
inanimate subject	*25*	*5.7*	*23*	*1.5*
Total agentives	160	**36.5**	321	**20.7**
core transitive (causatives)	5	1.1	49	3.2
transitive achievements	8	1.8	35	2.3
anticausatives/unaccusatives	31	7.1	101	6.5
statives	234	**53.4**	1041	**67.3**
Total non-agentives	278	**63.5**	1226	**79.3**
Grand total	438		1547	

Table 4: Distribution of verb types after *do/does* in child vs. child-directed adult data in CHILDES (child data for children up to 3;6 years old).

producing agentive verbs after *do* more often than adults, while they produce fewer non-agentive verbs after *do* often than adults ($\tilde{r}$ = ±6.80, p < .001). Further, post-hoc tests in which the child and adult data are compared per verb class reveal significant differences in the non-core transitives ($\tilde{r}$ = 2.82, p = .005), unergatives ($\tilde{r}$ = 6.78, p < .001), and statives ($\tilde{r}$ = -5.35, p < .001). The difference between child and adult frequencies was not significant in agentive causatives ($\tilde{r}$ = 1.37, p = .171), non-agentive causatives ($\tilde{r}$ = -2.30, p = .021),[18] transitive achievements ($\tilde{r}$ = –0.55, p = .582), or anticausatives/unaccusatives ($\tilde{r}$ = 0.41, p = .682). Our first prediction is thus partially confirmed: agentive verbs, or verbs used agentively, are more frequent after *do* in the child data than in the adult data, but this difference was only significant for non-core transitives and unergatives (i.e. not for causatives). Furthermore, adults produce (non-agentive) statives after *do* significantly more often than children.

[18] Note that the Bonferroni-corrected alpha level is adjusted to .007, after dividing it by the number of tests (seven). This result is therefore not significant, even though the *p*-value is lower than the usual value of .05.

While these results are suggestive, we are aware of the limitations of the methodology used. First, the data sample is limited; it would be desirable to examine more adult vs. child data (for instance, the survey could be extended to include *did* followed by a verbal form). Second, while we have limited the child sample to the time span known to be the window for superfluous *do* in order to avoid target-like emphatic *do* as much as possible, we did not distinguish between superfluous *do* and emphatic *do* in the child data (cf. also footnote 7). This issue is addressed in the second study reported in the following sections. Third, the annotators of the data were not blind to the hypothesis. Nevertheless, given the significant difference in the distribution of agentive (non-core transitive, unergative) and non-agentive (stative) verbs across child and adult data, we believe that *do* is plausibly used by young learners of L1 English to express the concept of action, as has been argued for in other Germanic child languages, such as child Dutch (see Section 5).

Additionally, we also tested our second prediction, namely that if *do* is reanalysed by the child as the light verb *do*, it should in principle occasionally occur after an auxiliary. Furthermore, *do* after an auxiliary should always embed an agentive verb (or a verb used agentively), which is the context in which *do* is assumed to be used by the child to overtly express the concept of action. This prediction was borne out: we found six examples of *do* embedding a verb in the context of an auxiliary (see (31)–(32)), and two after modal verbs like *wanna* or *hafta* in (32),[19] and the embedded verb was always agentive (see also example (11) from Bohnacker 2013 in Section 2).[20]

(31) a. I **can do mend** him. (Becky, 2;6, *Manchester*)
b. he **will do** [*] **water** them. (Helen, 3;9, *Manchester*)
c. it **can't do go** this way. (Carl, 2;2, *Manchester*)
d. **could do jump** over the footstool. (Abe, 3;3, *Kuczaj*)
e. it **can do walk**. (N.N, 4;8, *Gelman*)
f. I'**ll do [?] play**. (N.N, 5;10, *HSLLD*)

(32) a. you **hafta do jump** over then run around like that Mommy.
(Brooke, 3;5, *Van Houten*)
b. I **wanna do play** it. (Lara, 2;2, *Lara*)

[19] While, as noted by an anonymous reviewer, these modal verbs do not undergo subject-auxiliary inversion and trigger *do*-support (*Do you hafta jump?*) differently from auxiliaries such as *be* or *have*, their presence in these examples nevertheless rules out the possibility of *do* being an auxiliary itself and these modals seem to be fully developed auxiliairies.

[20] An anonymous reviewer suggests that (31d/e) could also alternatively involve a determinerless nominal. To settle the question, one would need to investigate whether the children producing these utterances still underuse determiners at the age of 3;3 or 4;8 in other contexts (as an indication, a recent study by Alhama et al. 2023 on the acquisition of determiners in English-speaking children show that they display productivity in their use of determiners from 30 months of age).

We also found 13 examples of the same type in child-directed speech, all of which except one (see (33b)) with an agentive infinitive (see (33) and (34) for examples). Although these examples are not many, their existence suggests that superfluous *do* may not be completely absent in child-directed speech.[21]

(33) a. okay we'**ll do check out** in that right? beep. put this in the bag. (Mother, *Feldman*)
b. right. hmm he **can do find** the letter. what else can you do? (Mother, *Belfast*)
c. no no get baby bear down it's come in here and we'**ll do get** your open your shape bag. (Mother, *Alex*)
d. I **need to do cook** some rice that you've helped get ready. (Mother; *Thomas*)
e. if I grown up like lossie I could do something like I really **could do go** on the bus. you really **could do go** on the bus right? (Father, *MacWhinney*)
f. want to see if you **can do guess** what the dessert is? (Mother, *Gleason*)
g. taste like ice cream . I gotta iron my clothes. sure do. sure **do gotta iron** my clothes. (Mother, *Brent*)
h. so you can either **do love** a bear or bear. bear. (Mother, *Providence*)
i. you know like we **could do put** the animals here. (ROZ, *Feldman*)
j. I now because you're lying down I **can do put** them on like this can't I? (Mother, *Lara*)
k. that was because if we're in here we **can do wash up** with my xxx hanging across the front door. (Mother, *Thomas*)

(34) did you guys want to um come here for lunch then we can decide if we **wanna do go** anywhere or anything. (Mother, *Providence*)

An anonymous reviewer notes that as documented by Quirk et al. (1985), British English allows a main verb *do* (that can, by definition, be preceded by an auxiliary) in elliptical contexts, including affirmative declaratives like in (35):

(35) Q: Will you go to the store?
A: I might/should/could do.

We are unsure that there is a direct connection between the latter use and the construals found in (33)–(34), since only three of the examples in (33)–(34) are produced by British English speakers (namely those in the Lara and Thomas corpora).

[21] According to the same anonymous reviewer, (33a) could perhaps contain a nominal (as in *The hotel requires check out by 11am*). Such an analysis is impossible in the other adult examples.

4.1.3 First study: Interim discussion

A possible alternative explanation for the observed overuse of *do* in agentive contexts is that it reflects a higher overall use of agentive predicates in early child English than in adult English. Perhaps young children tend to talk more about actions in general, and the overrepresentation of agentive predicates in the context of superfluous *do* simply reflects this tendency.

To assess whether children produce more *do* with agentive verbs simply because they produce more agentive verbs than adults, independently of *do*, we extracted the 100 most frequent verbal lemmas in child-directed speech on the one hand and child speech from children up to the age of 3 on the other in all English L1 CHILDES corpora accessible via an SQL query on Sketch Engine, see (36) (the data were restricted to adult participants vs. children aged 0 to 3). We extracted base verbs, thus excluding *be* and *have* used as auxiliaries but including them as main verbs.[22]

(36) `[tag = "VV|VB|VH"]`

These sets of verbs cover most of the concordance lines in both corpora (92% in the child corpora and 89% in the adult corpora). This means that around 10% of concordance lines contain a verb other than these 100 most frequent verbs. We then annotated these 100 most frequent verbal lemma according to verbal classes. We approximated the distribution of verb classes in child-directed speech vs. child speech via the frequency of each verb as given in Sketch Engine, see **Figures 1** and **2**. We provide the lists of these most frequent verbs together with their verb class ranked by frequency for child and child-directed speech in Appendix 2 (and *do* is a the top of these two lists).[23]

Figure 2 shows that the distribution of verb classes is roughly the same for adult and child speech. The main differences concern the frequency of non-core transitive verbs and statives: while in child production, stative verbs are more frequent than non-core transitives, the opposite is true for child-directed speech.

The general conclusion that can be drawn from this comparison is that young children do not tend to use more agentive verbs than adults in general. In fact, stative verbs are more frequent across child speech in these corpora. We therefore conclude that the higher proportion of agentive

[22] In the SQL language used in Sketch Engine, the VV tag includes the base form of lexical verbs, and excludes auxiliaries such as *be, have, do* and modal verbs. The VB tag is used for the verb *be* when used as a main verb (i.e. not as an auxiliary). The VH does the same for the verb *have*.

[23] A limitation of this procedure is that the verb is classified independently of any context. For example, verbs we classify as causatives alternate between intransitive (non-agentive) and transitive (potentially agentive) uses. However, since the annotation is the same for the adult and child data, this should not reduce the validity of the comparison much. Another case concerns verbs we classify as unergatives, which also have (non-core) transitive uses (*play*). But this is even less problematic for us, since the subject occupies the role of agent across all uses of these verbs.

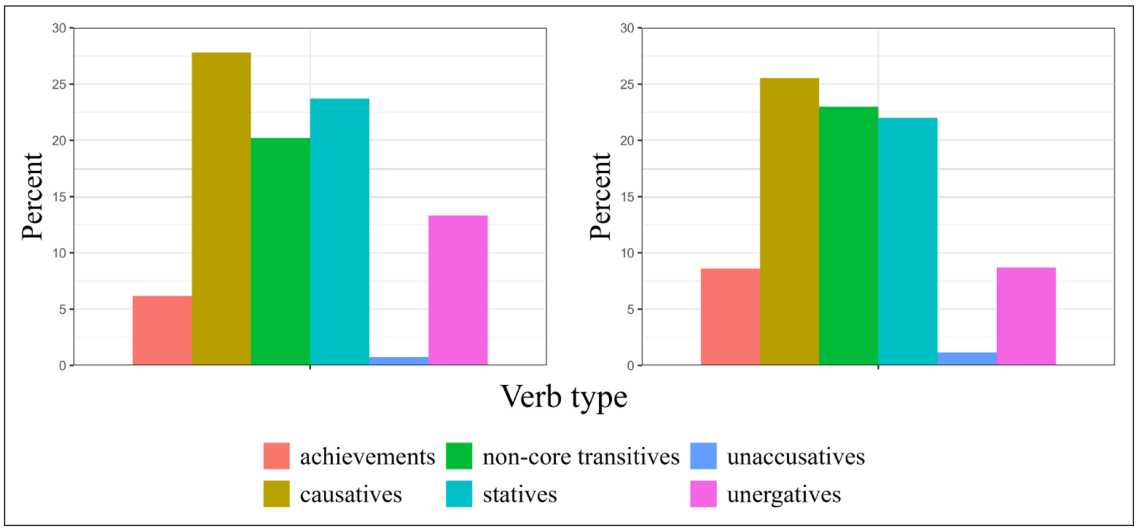


Figure 2: Distribution of verb types across the 100 most frequent verbs in the speech of children up to 3 years old (left panel) and child-directed speech (right panel) in CHILDES corpora accessible via Sketch Engine.

verbs after *do* in positive assertions in child speech does not reflect a higher use of agentive verbs by very young children independent of the particular *do* + verb construction.

4.1.4 Second study: Methodology

The first study informs us about the distribution of verb types after *do* in positive utterances in child vs. child-directed speech. A clear limitation of this study is that the child dataset includes occurrences of *do* + verb that represent a target-like use of emphatic *do*, or for which an analysis in terms of undercompression errors by which the child aims to expone Tense or the concept of action in a one-to-one fashion with form is not very plausible.

For instance, in some cases, these non-target-like child productions rather look like echoes of *do*-support built with the same verb in a preceding question or assertion of the caregiver, as may be the case in (37).

(37) MOTHER: does it hurt?
CHILD: 0it [*] **does hurt** . (Laura, 1;7, *Braunwald*)

In order to assess whether agentives after *do* are still more frequent in the child data than in the adult data once these potentially primed hits are put aside, we examined the context of all occurrences of *do/does* + verb in child data and divided them into four categories, according to whether the preceding utterance contained (i) the auxiliary *do* with the verb used by the child (which we will call the target verb), (ii) the auxiliary and/or light verb *do* but not the target verb, (iii) the target verb but not *do* and (iv) neither *do* nor the target verb. The

Lexical material in preceding utterances	*do/does/did* + target verb	*do/does/did* only	target verb only	neither *do/does/did* nor target verb
	105	72	78	183
value of *do/does* in child utterance				
emphatic	70 (66.7%)	6 (8.3%)	34 (43.6%)	–
repetition	6	0	–	–
syntactic context of preceding utterance				
questions	57	42	10	–
negation	42	25	9	–
elliptic *do*	6	5	–	–
verb type after *do/does* in child utterance				
agentive verbs	22 (**21.0%**)	37 (**51.4%**)	35 (**44.9%**)	–
non-core transitives	7	13	12	–
core transitives	7	13	11	–
unergatives	8	11	16	–
non-agentive verbs	83 (**79.0%**)	35 (**48.6%**)	43 (**55.1%**)	–
core transitives	0	1	0	–
transitive achievements	0	0	2	–
anticausatives/unaccusatives	12	3	4	–
statives	71	32	37	–

Table 5: Presence of auxiliary *do* and/or target verb in sentences preceding the *do/does* + verb occurrence in child data in CHILDES (for children up to 3;6 years old).

distribution of hits across the first three cases is shown in **Table 5**.[24] 183 instances belonged to the fourth category.

The same table also shows the syntactic context of *do* in the preceding utterance: *do*-support in questions, in negative sentences, or elliptic *do*. It also indicates whether the preceding context alone, without prosodic information, allows one to conclude with certainty that the child's production of *do* + verb is an instance of emphatic *do* (see below for the decision criteria and

[24] We thank an anonymous reviewer for suggesting us to investigate the preceding context of the child utterance.

We considered both present and past instances of auxiliary *do* in the preceding utterance, as the use of auxiliary *did* followed by the target verb can possibly prime the use of auxiliary *do* plus the same verb by the child.

some examples), or when *do* + verb in the child occurrence is a mere repetition of the preceding utterance. Moreover, the same table shows the verb class of the target verb used by the child after *do* across the first three cases (i)–(iii). In what follows, we look at these first three cases in more detail before presenting the results of a replication of the first study in which the potentially primed hits have been discarded.

24.0% of the child data set (105 occurrences out of 438) is preceded by an utterance containing both *do/does* and the target verb, see for instance (38):

(38) Auxiliary *do* and the target verb occurring in preceding utterance

a. MOTHER: do you like strawberrys ? CHILD: yes. MOTHER: do you?
CHILD: I **do like** this one . (Gail, 2;3, *Manchester*)

b. CAREGIVER: does it have a hole in it ?
CHILD: no it **does have** a hole in it . (Peter, 2;6, *Blom*)

c. MOTHER: did you have dinner ?
CHILD: I **do have** dinner. (Dominic, 2;7, *Manchester*)

d. CAREGIVER: but he has stripes .
CHILD: yeah he does . he **does have** stripes . (Nan, 2;10, *Garvey*)

In most of these cases (88 out of 105 hits, 83.1%), *do* in the preceding utterance is an occurrence of *do*-support in questions and/or negative sentences.[25] In other cases, elliptic *do* is used just before the child's *do/does* + verb production (cf. (38d)), or they are simply repetitions of *do* + verb in a preceding positive assertion. Examples such as (38b-c) look like non-target like occurrences of *do*-support: in (38b), the child intends to produce a negative sentence and instead produces a positive clause preceded by *no*, but uses *do*-support; in (38c), the child intends to express a past meaning and produces a present instead, but retains the occurrence of *do* bearing past tense features in the previous question, perhaps to convey emphasis.[26]

It is in this first subset of data that we find the most clear instances of emphatic *do* (70 out of 105 hits, 66.7%), see for instance examples (38a/c-d) and (40). We identified *do* followed by a verb as emphatic *do* when the sentence had a confirmatory value (more specifically, the English word *indeed* can be felicitously inserted into the sentence without changing its meaning, see Klein 2022: 10), as in (40a), or if it is in contrast with the previous assertion, as in as in (40b).

---

[25] An additional 11 preceding utterances are negative questions (*don't you like it?*).

[26] There are a few other occurrences of this type of present use of *do* to express a past meaning, see e.g. the following examples:

(39) a. What did you do in the kitchen? I **do wash** hands. (Adam, 2;8, *Brown*)
b. but did you like the Halloween party ? CHI: yeah . I **do** [?] **like** that . (Helen, 3;6, *Brown*)

(40) a. CHILD: I want a lolly . MOTHER: do you ?
I **do want** a lolly . (Lara, 3;1, *Lara*)
b. MOTHER: you don't like me ? CHILD: no . MOTHER: oh no .
I **do like** you now . (Becky, 2;10, *Manchester*)

In most instances of this set (71 out of 105 hits, 67.6%), the target verb is a stative verb, as in examples (40). Furthermore, in this set, most of the sentences with *do* followed by a stative verb are target-like instances of emphatic *do* (52 out of 71 hits, 73.2%). As an anonymous reviewer suggests, this strong correlation between emphatic *do* and statives may reflect a desire to emphasise mental states in response to a question in which the child contrasts or confirms their inner state with alternatives in reaction to a previous question or assertion. In support of this idea, we observe that out of the 57 preceding utterances with *do* plus the target verb in question format, 42 contain a stative verb (73.7%, see again (40) for examples). In contrast, only ten contain an agentive verb (17.5%). Six of these ten occurrences are clear cases of emphatic *do* e.g. (41a), while others such as (41b) are most likely merely primed by the use of *do* + verb in the question.

(41) a. MOTHER: we draw on paper , don't we ? MOTHER: -eh ? CHILD: yeah . MOTHER: yes . thank you .
CHILD: we do . We **do draw** on paper . (Liz, 2;8, *Manchester*)
b. FATHER: what does she do ?
she **does** just **do** Uncle Nwww . (Eleanor, 2;7, *MPI-Eva-Manchester*)

All instances of this first subset of data are discarded for Study 2, as these *do* + verb child productions may have been primed by the morphological material in the preceding utterance, are non-target attempts to use *do*-support (cf. (38b)), or are target-like occurrences of emphatic *do*.

In the second subset of data, the preceding utterance contains either the auxiliary and/or the light verb *do/does* but not the target verb ($n = 72$). In most cases ($n = 59$, 81.9%), the main verb of the preceding utterance is neither a hyponym nor a hyperonym of the target verb, as in (42).

(42) a. MOTHER: do you know what that one is ? CHILD: &-um pineapple . CAREGIVER: yeah .
CHILD: I **do eat** pineapple. (Dominic, 2;8, *Manchester*).
b. MOTHER: and did you look gorgeous? CHILD: yes.
CHILD: and Grandma <**does got**> [*] flower . (Fraser, 2;3, *MPI-Eva-Manchester*)
c. MOTHER: CAREGIVER: where did Michael go before?
Michael do dance. (Sarah, 2;6, *Brown*)

In others (n = 13, 18.1%), the main verb of the preceding utterance is the light verb *do*, and the child seems to use the verb following *do* to further specify the action already described by *do*:

(43) *do/does* occurring in preceding utterance (not the target verb)

a. CAREGIVER: what does it do ?
CHILD: a duh [: the] **does brush** some wecord [: record] wif when you put it . (Shem, 2;8, *Clark*)

b. MOTHER: what do you do with a recorder ?
CHILD: I **do play** this one . (Becky, 2;10, *Manchester*)

c. MOTHER: what do you do with that ?
CHILD: you **do** [*] **plug** it in this hole here . (Gina, 3;0, *MPI-EVA-Manchester*)

d. CAREGIVER: what d(o) you do with it ?
What I do ride. (Peter, 2;3, *Bloom*)

For us, these latter cases are good examples of undercompression errors, by which the child decomposes the verb into an agentive verbal head v_{do} and an action description.

In this second dataset, there are few clear cases of emphatic *do* (6 out of 72 occurrences, 8.3%) and target verbs are rather equally divided between agentive and non-agentive verbs. We kept all instances of this set except for those containing emphatic *do*.

In the third and last subset of data, the preceding utterance contains the target verb, but not in the context of the auxiliary *do* (n = 78), see (44) for some examples:

(44) the target verb occurring in preceding utterance (not *do/does*)

a. CHILD: -um I want draw orange .
CHILD: Mummy **do draw** orange . (Carl, 2;2, *Manchester*)

b. MOTHER: what's too big ? CHILD: crash . MOTHER: the what ?
CHILD: I **do crash** . (Warren, 1;11, *Manchester*)

c. MOTHER: where's Pooh bear gone ?
CHILD: there . Pooh bear a@p work . Pooh bear **do work** . (Fraser, 2;0, *MPI-EVA-Manchester*)

d. CHILD: it's mine . I made it .
CHILD: I **do made** [*] the shopping . (Becky, 2;9, *Manchester*)

Unsurprisingly, this set contains fewer questions or negative sentences in the preceding utterance (ten and nine, or 12.8% and 11.5%, respectively, out of 78 hits). Target verbs are again almost equally divided between agentive and non-agentive predicates.

Since the child uses the same verb as in the preceding occurrences, the number of cases of clear emphatic *do* is high: 34 of these occurrences (43.6%) are clear instances of emphatic *do*, as in (44d) or the examples in (45) below.

(45) a. CAREGIVER: I thought she wore a diamond necklace .
CHILD: yeah ! Mommy **does wear** a diamond necklace .
(Helen, 3;6, *MPI-EVA-Manchester*)

b. CHILD: he was sparklin(g) . CAREGIVER: oh wow !
CHILD: yeah, he **does spar(k)le** . (Helen, 3;6, *MPI-EVA-Manchester*)

For Study 2, we discarded these clear cases of emphatic *do* and retained the other instances.

When the preceding utterance contains neither *do* nor the target verb, it is usually very difficult to determine whether the child used *do/does* is used emphatically simply based on the context, see e.g. (46).

(46) a. MOT: know who's dress this was ?
CHI: -um . I xxx dress I **do grow** . (Sarah, 2;11, *Brown*)

b. MOT: not seen this one in here before .
because I **do eat** it . (Dominic, 2;7, *Manchester*).

For Study 2, we thus included all of these remaining cases ($n = 183$, for a total of 293 cases for this second study). The statistical analysis was the same as in Study 1, but this time around the occurrences of *do* + verb that were either possibly primed by the use of *do* with the same verb in the preceding utterance or target-like instances of emphatic *do* were excluded. Recall that the *p*-value for statistical significance in the post-hoc tests with the verb class categorization has been Bonferroni-corrected to .007, so as to account for multiple comparisons. The adult data were identical to those in Study 1.

4.1.5 Second study: Results

As can be seen in **Table 6** (from which the information about animacy has been removed), once emphatic and primed uses of *do* are excluded from the child dataset, the proportion of agentive verbs (or verbs used agentively) in the set of *do* + verb instances is much higher in the child data than in the adult data (45.4% vs. 20.7%). Compared to Study 1, statives are much less present in the child data, because many are used after target-like emphatic uses of *do*, and the proportion of all agentive subtypes is consequently higher in the child data than in Study 1. We take this to confirm that very young children occasionally use *do* as an exponent of the action concept, which is 'compressed' in the verb used after *do*.

The results of a χ^2-test again indicate that the association between population on agentivity is significant ($\chi^2(1) = 80.49, p < .001$), with children producing more agentive verbs and fewer non-agentive verbs after *do* than adults ($\tilde{r} = \pm 8.97, p < .001$), as well as between population and the more fine-grained distribution of verb types ($\chi^2(6) = 97.62, p < .001$). Post-hoc tests again show significant differences in the non-core transitives ($\tilde{r} = 4.32, p < .001$), unergatives ($\tilde{r} = 7.25, p < .001$), and statives ($\tilde{r} = -6.83, p < .001$), but also in the agentive causatives

Verbal class	*N* (child)	% child)	*N* (adult)	% (adult)
non-core transitives	49	**16.7**	132	**8.5**
core transitive (causatives)	36	**12.3**	114	**7.4**
unergatives	48	**16.4**	75	**4.8**
Total agentives	133	**45.4**	321	**20.7**
core transitive (causatives)	1	0.3	49	3.2
transitive achievements	7	2.4	35	2.3
anticausatives/unaccusatives	16	5.5	101	6.5
statives	136	**46.4**	1041	**67.3**
Total non-agentives	160	**54.6**	1226	**79.3**
Grand total	293		1547	

Table 6: Distribution of verb types after *do/does* in child data (up to 3;6 year old and excluding contextually primed and emphatic use of *do*) vs. child-directed adult data in CHILDES.

($\tilde{r}$ = 2.82, p = .005) and non-agentive causatives ($\tilde{r}$ = -2.73, p = .006). The difference between child and adult frequencies was not significant in transitive achievements ($\tilde{r}$ = 0.13, p = .897) or anticausatives/unaccusatives ($\tilde{r}$ = –0.69, p = .490). Thus, the results of our second study fully corroborate the findings of Study 1 and add that causatives are used by children more often after *do* than by adults if they are agentive and less often than by adults when they are non-agentive. These findings are in line with our hypotheses.

4.2 Do+gerund: Methodology and results

We now turn to our second hypothesis according to which children acquiring English use *do* followed by a gerund more often than adults. To test this, we extracted all utterances with *do/does* followed by a gerund form in the same North American and British child corpora as before (see **Table 2**). We annotated the data from children up to the age of five (max. 59 months) according to the same criteria as in the previous corpus study: age of the child, verb used after *do* (in a gerund form), verb type and thematic role of the subject. We also indicated when intervening material between *do* and the gerund is present (mostly a determiner, in rare cases an adjective, a noun or an adverb). We thus collected 349 relevant utterances from children; examples are given in (47) for several verb types.

(47) a. *non-core transitive*
I like to **do counting** [?] . (Eleanor, 2;2, *MPI-Eva-Manchester*).
b. *causative*
I can't **do building** a bridge. (Thomas, 3;1, *Thomas*)

c. *unergative*
do running fast. (Eve, 1;9, *Brown*)
d. *unaccusative*
that doesn't **do falling**. (Lara, 3;2, *Lara*)

To assess the relative frequency of the *do* + gerund construction in child speech, we then computed the ratio of these forms in the set of all occurrences of *do* followed by a verbal form (a gerund, an infinitive or in rare cases an inflected verbal form). We collected and annotated data from child-directed speech following the same procedure. We manually extracted all sentences that instantiated a light verb construction headed by *do/does* and followed by a gerund, with or without a determiner intervening between the two; the SQL query was the same as for the previous corpus study for tokens without an intervening determiner. For the cases where the gerund is used with a determiner, we used the query in (48) (again limiting the results to adult participants). This procedure yielded a total of 166 constructions.

(48) `[word = "does|do"][tag="DT.*"][tag="VV.*" & word = ".*ing" ]`

Examples of child-directed speech occurrences of *do* + gerund in CHILDES corpora are given in (49).

(49) a. *non-core transitive*
you have to keep letting me to **do** some **studying** okay?
b. *unergative*
Let's take a listen to you. **Do** some **couching**.
c. *stative*
yes you just like your mum to be daft. You **do seeing** me daft so.
d. *unaccusative*
you think it's going to **do** some **raining**. yes.

The distribution of the data is presented in **Table 7**. This table shows that the proportion of gerunds following *do/does* in the set of all occurrences of *do/does* followed by a verbal form – gerunds or infinitives (or, in a few cases, inflected forms) – is much higher in the child data than in the adult data (44.3% vs. 9.7%). These proportions are significantly different from each other ($z = 19.87, p < .001$).

Overall, light verb constructions headed by *do* occur overwhelmingly with agentive verbal forms, not only in the adult data, but also in child language production. This supports the idea that children overuse the *do* + gerund pattern with an agentive verb. For us, this is because the child aims to express the concept of action compressed in the verbal form in a one-to-one fashion.

construction	***N* (child)**	**% total (child)**	***N* (adult)**	**% total (adult)**
do + gerund	**349**	**44.3**	**166**	**9.7**
with determiner	142	18.0	137	8.0
countable	*67*		*41*	
non-countable	*75*		*96*	
without determiner	207	26.3	29	1.7
verb class of gerund	***N* (child)**	**% gerunds (child)**	***N* (adult)**	**% gerunds (adult)**
non-core transitive	268	76.8	91	54.8
core transitives (causatives)	20	5.7	32	19.3
unergatives	58	16.6	40	24.1
transitive achievements	1	0.3	1	0.6
anticausatives/unaccusatives	2	0.6	0	0
statives	0	0	2	1.2
do + base verb	**438**	55.7	**1547**	90.3
Grand total	**787**	100	**1713**	100

Table 7: *do/does* followed by an infinitive/gerund in positive assertions in child (up to 59 months) vs. child-directed adult data in English CHILDES corpora.

Furthermore, while 82.5% of *do* + gerund constructions occur with a determiner (usually those heading non-countable predicates, such as *a/the*) in child-directed speech, only 40.7% do so in child speech.

Finally, the fact that the vast majority of predicates used in the *do* light verb construction in child English are agentive verbs suggests that children master the distinction between agentive verbs and non-agentive ones (such as unaccusatives or statives) at an early age. This is a further confirmation of previous studies showing that children master the distinction between eventive and stative predicates (van Hout 2016 among others).

4.3 Overall discussion

Our corpus studies have yielded two new results. The results of the first two corpus studies are consistent with our hypothesis that young children learning English tend to produce *do* before an agentive verb more often than adults do. Our second corpus study shows that young L1 learners of English also tend to use light verb constructions headed by *do* more often than adults do. We take these data to support our hypothesis that the overuse of *do* in early child English

reflects children's tendency to occasionally prefer a non-targetlike overt exponent to explicitly express some underlying semantic structure (Slobin 1973; Guasti et al. 2023). Superfluous *do* is used by the child to express in a one-to-one fashion either the concept of Tense (Hollebrandse & Roeper 1996; Roeper 1991), or the core action concept, which usually remains silent in adult English.

5 Superfluous agentive *do* elsewhere

For other West Germanic languages, it has been noticed that young children go through a phase in which they overuse functional or light verbs (see Blom et al. 2013). An example is the case of superfluous *doen* 'do' or *gaan* 'go' in child Dutch (Hollebrandse & Roeper 1996; Blom 2003; Hollebrandse et al. 2013; Jordens 2013). These verbs are produced by children in contexts where they do not normally occur. In Dutch, *doen*-periphrasis as in (50b) does not belong to prescriptive Dutch, but occurs frequently in non-standard speech and is geographically widespread (Sert et al. 2023 and citations therein).

(50) Dutch

a. doetie viesmake. (Andrea, 2;1)
does.he dirty-make.INF
'He's making (it) dirty.' (Jordens 2013: 352)

b. doe ook praten. (Sarah, 2;5)
do also talk
'I'm also talking.' (van Kampen 1997)

For Blom (2003), superfluous *doen* in child Dutch is a strategy used by children to avoid the problem raised by the marking of Tense by verbal inflection. But another account of superfluous *doen* has been proposed for child Dutch (Jordens 2013), in line with what we propose for child English *do*. For Jordens, in a first stage that he calls the 'lexical stage' (spanning from 1 to 2;2 years old), superfluous *doen* is not a dummy auxiliary, but serves as the lexical head of the clause, and only appears in the context of agentive verbs (at this stage, he found no examples in his corpus with a stative verb or a non-agentive change-of-state verb). For him, *doen* is used by the child to express the default function of (agent) 'control'. In adult Dutch too, superfluous *do* has been claimed to put more emphasis on the action in one of its uses, the other being the expression of habituality (see discussion in Sert et al. 2023).[27]

[27] The case of superfluous *doen* in Dutch remains different from the one of superfluous *do* in English: while Dutch superfluous *doen* mainly appears with agentives, English superfluous *do* also appears in the context of non-agentives (as an exponent of Tense). Furthermore, superfluous *doen* is known to appear in child-directed speech and adult varieties, while this is not reported for English.

In the diachrony of English it has been noted that superfluous *do* in Middle and Early Modern English mostly appears with agentive transitive verbs (see Kroch et al. 1982; Denison 1985; Bohnacker 2013: 202). Ecay (2015) describes the grammaticalisation path of *do* as follows: in Middle English, '*do*' is a causative predicate, which is merged in V and takes a clausal object of some (small) size. By contrast, in Early Modern English, it is reanalyzed as the head of a projection relatively low in the functional hierarchy, which is associated with the layer introducing agents. According to Ecay, the most robust generalisation is that *do*-support in positive assertions is generated by a grammar which uses this form to mark the presence of an agentive external argument. In some sense, children producing superfluous *do* in the context of an agentive predicate come back to the starting point of a *do*-cycle in the history of English.

6 The prevalence of statives after *do* in positive assertions

6.1 Statives after *do*: child vs. adult English

Our first corpus study on the use of *do* in positive assertions shows that stative predicates are very common in the context of emphatic *do/does* in the adult data. Statives appear in two thirds of the cases in this context.[28]

This high prevalence of statives after emphatic *do/does* (and more generally with simple present tense forms in adult speech) can been seen as the result of the restriction of eventive predicates in simple present sentences to generic or habitual statements (Comrie 1976). That is, emphatic *do/does* and simple present tense in general are very often found with lexically stative verbs, because statives are more flexible in their interpretations in sentences with simple present tense, as they can have both episodic and generic uses. By contrast, with eventives, simple present sentences are typically generic or habitual statements, as shown in (51).

(51) a. Esther does play tennis. (generic/habitual only)
b. Esther plays tennis. (generic/habitual only)

Turning to child data, stative predicates are also very frequent in children production (53.4% for children up to 3;6 years of age in our first corpus study). This reflects the frequency of statives in the input: the prevalence of statives with (present tense) superfluous *do/does* observed in our first corpus study can be seen as a confirmation of the Aspect Hypothesis (Andersen & Shirai 1994; Shirai & Andersen 1995). This hypothesis predicts that the present tense is more likely to be used with verbs that are often used in the present tense in the adult input, such as statives.

[28] These results are consistent with previous research on the distribution of the simple present tense morpheme *-s* in child-directed adult English. Wang & Shirai (2015) conducted a corpus study on the ratio of stative verbs used with third person *-s* to the total verbs combining with *-s* for three caregivers and the average percentage of stative verbs was above 50% for each caregiver (up to 58% for Naomi's mother, see Wang & Shirai 2015: 11). So the prevalence of statives holds not only with *do/does*, but also more generally with simple present forms.

In our child data, however, we find occurrences of *do/does* followed by an eventive predicate that amount to episodic statements, see e.g. (52). Here they use the present tense in a non-adult like way. This may be related to the difficulty raised by generic and habitual reference for children (Brown 1974; Shirai 1991), which in turn may be due to the fact that generic and habitual statements compress a more complex conceptual structure than episodic sentences. It may also be related to the tendency of early L1 English learners to use the predicate bare, without aspect marking (see also the so-called *Root infinitive stage* documented in Hoekstra & Hyams 1998 and Rizzi 1993).

(52) a. Mommy, I **do wash** my hand. (Laura, 2;1, *Brauwald*)
b. MOTHER: what're you doing ?
CHILD: I **do** [*] **crash** Edward . (Dominic, 2;5, *Manchester*)

The prevalence of statives in the context of *do* in child English seems to extend to *do*-support in questions. We examined all occurrences of *do* in the North American and British English CHILDES child corpora (n = 4459) and annotated them for verb types. As can be seen in **Figure 3**, children very often ask questions with a stative verb, as in the example (53).[29]

(53) d(o) yo(u) want the bottle? (Marjorie, 2;3, *Bliss*)

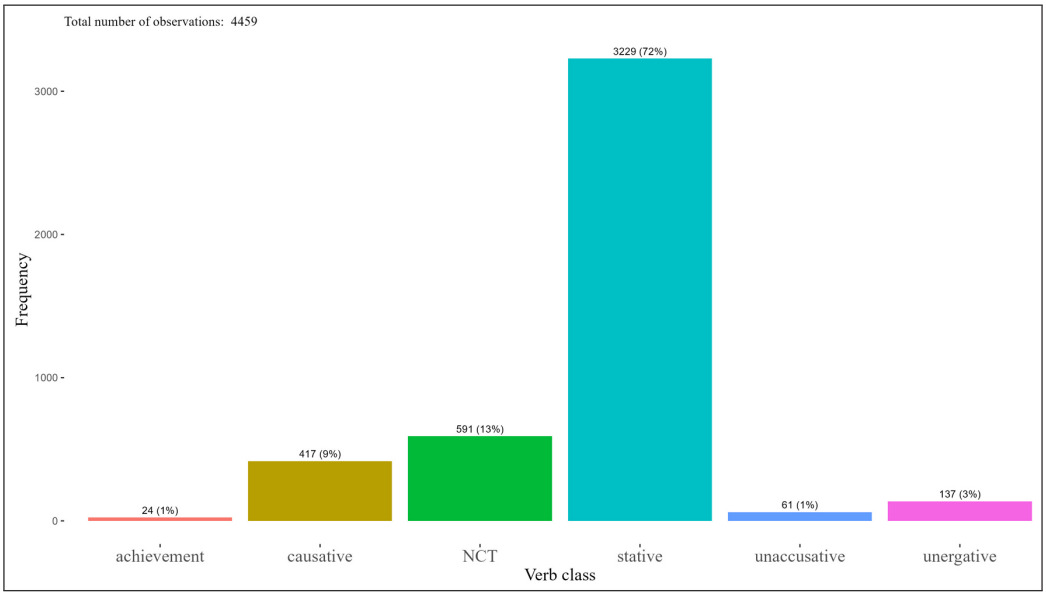


Figure 3: Distribution of verb types after *do* in questions in child data (CHILDES corpora).

[29] With respect to age, dummy *do* in questions is most frequent between 45–50 months (perhaps because children ask more questions at that age); i.e., its peak does not overlap with the peak of *do* in positive assertions.

verb	frequency	% of concordance lines
want	19,057	22.57
think	13,935	16.50
know	9,423	11.16
remember	4,985	5.90
do	4,827	5.72
wanna	4,108	4.87
see	3,231	3.83
say	3,013	3.57
need	2,353	2.79
go	1,838	2.18
mean	1,803	2.14
like	1,667	1.97

Table 8: Most frequent verbs used after *do/does* followed by a pronoun in child-directed speech in CHILDES corpora accessible via Sketch Engine.

We believe that the prevalence of statives with dummy *do* in questions again reflects a tendency in the input. In our first corpus study, when the child answered with a *do* + verb positive assertion to a *do*-question with the same verb, the verb is stative in 75.4% of cases (43 out of 57 cases). We also looked at the twelve most frequent verbs after *do/does* followed by a pronominal subject in child-directed speech in English CHILDES, and most of them are statives. **Table 8** lists these verbs in order of frequency and gives the total number of occurrences followed by the percentage of concordance lines of the corpus in which they occur. The CQL query we employed is given in (54) and returned 84,433 items (the query was restricted to restricted to participants aged 18 and over).

(54) `[word = "does|do"][lemma = "I|you|it|she|he|we|they"][tag="VV.*"]`

The high frequency of stative predicates after *do/does* + pronoun in questions in child-directed speech is, again, unsurprising, since eventive predicates are again more restricted in their interpretation in such environments, given that they typically only have generic interpretations in questions of this form, see e.g. (55).[30]

[30] While we did not systematically compare the distribution of agentive and non-agentive predicates in questions containing dummy *do* in child vs. child-directed speech, it is interesting to spell out the predictions of our analysis for superfluous *do* in child English. As an anonymous reviewer observes, one could argue that we expect dummy *do* in questions to be reanalyzed as 'agentive' *do* by children (i.e., what is dummy *do* in questions in adult English could in some cases be reanalyzed as agentive *do* by children). Let us assume that this is indeed the case. Do we then predict agentive verbs to be more frequent in *do*-questions? We do not believe so, since independently of that, dummy *do* is necessary to spell out Tense in Y/N questions containing a full lexical verb, be it agentive or not. We therefore do not predict agentives to be more frequent in *do*-questions even under the assumption that in such contexts, *do* was also endowed with agentivity-semantics by young talkers.

(55) Which writers do you read?
  a. i. Which writers do you habitually/generally read? (generic/habitual)
  b. ii. #Which writers are you reading right now? (episodic)

6.2 A person/aspect split

An intriguing tendency in the data from our first corpus study is that the proportion of non-agentive (mostly stative) predicates is higher in first-person singular assertions (145 out of 200, 72.5%) than in other persons in the child data, see the top panel in **Figure 4**. 141 of these 145 cases (97.2%) were stative verbs. Examples of first person singular assertions with *do* followed by a stative verb are given in (56).

(56) a. I **do want** this biscuit. (Lara, 2;11, *Lara*)
  b. I **do want** watermelon. I just want watermelon. (child-directed speech)

In contrast, the relative proportion of non-agentive predicates is much smaller for assertions with a third person pronoun (58.9% non-agentive), a full DP (54.7%) or a null subject (51.0%); see (57) for some illustrative examples in child English.

(57) a. they **do eat** water. (Becky, 2;8, *Manchester*)
  b. Allen **does play** basketball. (April, 2;10, *Higginson*)
  c. MOTHER: it'll go in the garage if you want . CHILD: yeah . MOTHER: just about . CHILD: **does go** in . (Aran, 2;2, *Manchester*).

At first sight, one might think that the association between *do* + statives and first person utterances is due to a high proportion of *do* + stative child assertions in response to a question. But this is not the case: in the set of 438 occurrences of *do/does* + verb produced by children up to the age of 3;6, 234 instances contain a stative verb, of which only a third (75, i.e. 32.1%) are answers to a question.

In our view, this association instead reflects the fact that very young children are more likely to report their own mental states rather than those of others. Indeed, most of the non-agentive verbs used after *do* are attitude verbs such as *want* and *like*, which express mental states (171 out of 279, 61.2%).

The distribution of subject types in *do/does* + verb positive assertions is very different in the adult data, see the bottom panel in **Figure 4** (based on child-directed speech data collected for the corpus study on *do* + verb). *Do* + verb configurations are preceded by first-person singular forms relatively less often than in child data; instead, second-person pronouns and third person singular forms are used more commonly. In child-directed speech, verbs that appear after *do/does* in positive assertions are mostly statives across all persons (1041 out of 1547 for all persons, 67.3%; see **Table 4**), with the exception of the third plural (50 out of 115, 43.5%) and DPs

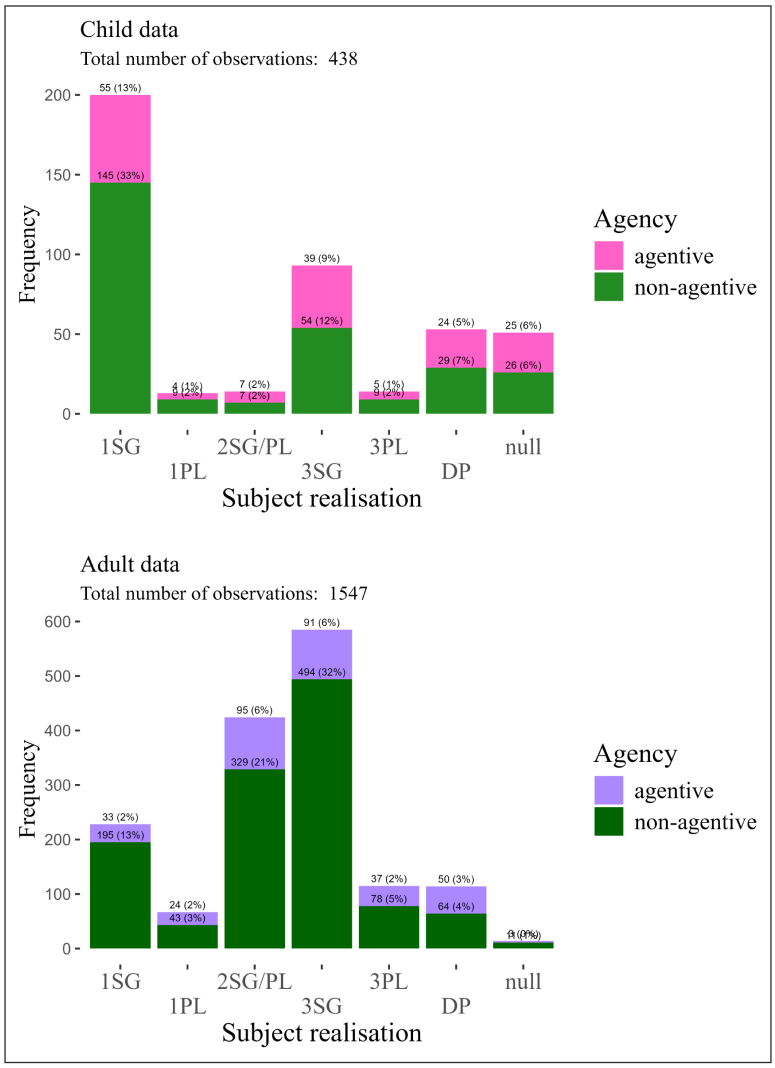


Figure 4: Distribution of subject types with *do* + verb in child data (up to 3;6 years old, top panel) and in child-directed speech (bottom panel) in CHILDES corpora. Note that the percentages reflect proportions of the total number of observations.

(47 out of 114, 41.2%).[31] In other words, there is no person/aspect split in child-directed speech, in contrast with early child speech: most of the *do/does* + verb positive assertions contain a stative verb, for reasons explained in the previous section.

[31] The proportions of the other categories were notably larger: first singular (171 out of 228, 58.2%), first plural (39 out of 67, 58.2%), second-person (299 out of 424, 70.5%), third singular (408 out of 585, 69.7%), and null (10 out of 14, 71.4%). Note that with a third plural subject, *do/does* + verb positive assertions are often generalisations about activities characteristic of some kind, as in (58):

(58) a. what do they eat ? yy what . worms . worms ? I bet they do eat worms .
b. I don't think cows do wag their tails darling . they swish them to get rid of the flys .

7 Conclusions

We gathered corpus data suggesting that superfluous *do* has a dual function in child English: it realises Tense with a dedicated morpheme (Hollebrandse & Roeper 1996) or overtly realises the concept of action that is left silent in adult language, by 'disincorporating' the verbal head *do* from the verbal structure. From this perspective, early child English is a kind of window into the internal structure of verbal predicates. An Internet English variety has been analysed in exactly the same way: Punske & Butler (2019) show that in the Doggo meme language, speakers make a heavy use of the *do* construction, producing structures such as *do me a beam up* where standard English expects *beam me up*.[32] Punske & Butler (2019) analyse this Doggo construction in exactly the same way as we do, via the non-incorporation of the root into *v* or Voice. They suggest that language games provide insights into the internal structure of language, as Slobin (1973) or Guasti et al. (2023) claim that child languages also do.

We have seen that in child English, agentive predicates increase the use of superfluous *do* in positive assertions: the proportion of agentive verbs after *do* is significantly higher in the child data than in the adult data. This supports our general hypothesis that with superfluous *do*, children sometimes tend to overtly express the core concept of action with a dedicated (agentive) morpheme (whereas in adult language, this concept is compressed in the verbal structure itself).

According to our proposal, children and adults share the same grammar when auxiliary *do* is used to optionally realise Tense (Sauerland & Alexiadou 2020). However, while this is associated with an additional manner implicature in the target language, children between the ages of 2 and 3 do not yet compute this implicature, but use *do* to realise Tense overtly via a dedicated morpheme (a kind of error due to their preference for one-to-one mapping between meanings and forms). But apart from the lack of manner implicature, children's use of auxiliary *do* is adult-like. They tend to use auxiliary *do* in positive assertions more often with stative predicates because emphatic *do* is very often produced with statives by adults, supporting the Aspect Hypothesis.

In addition to this first non-target but nonetheless grammatical use of *do*, children also *reanalyse* the auxiliary *do* as the light verb *do* in positive assertions. Differently from the previous use, this use of *do* does not seem to belong to the Standard English grammar, where *do* is not used as a light verb before another verb in positive assertions. Children use *do* in this way to overtly express the core concept of action that is already expressed by the embedded verbal structure, as also reported in child Dutch (Jordens 2013). An additional piece of evidence for the proposal that young children sometimes use *do* as a light verb in positive assertions is that they sometimes use it after an auxiliary, as in e.g. *I can do walk*. We found a few examples of this type in child-directed speech, reported in (33).

[32] We thank an anonymous reviewer for this reference.

The light verb *do* is also used by children to express the action concept overtly in another construction, namely the light verb construction *do* + ing. The fact that *do* + ing is used almost exclusively with agentive items in early child speech suggests that young children recognise that *do* spells out the concept of action, and also confirms that they distinguish agentive from non-agentive predicates at an early stage.

This non-target use of light verb *do* in child language brings a new source of evidence to the debate about verbal structures, and can be seen as supporting the idea that there are different flavors of *v* (Folli & Harley 2005).

How, then, do children move out of decompression? According to the analysis proposed in this paper, achieving target-like behavior does not mean stopping the production of *do* + verb(ing) patterns altogether—these *do* belong to some adult varieties and are therefore target in the right social context. Rather, it boils down to understanding that the use of these patterns often triggers manner implicatures via the *Be Brief* principle: the most verbose variant is often understood as signalling emphasis or intensification on a particular unit of meaning. In order to achieve target-like behavior, the child must therefore understand that *Be Brief* allows them to be verbose only when they have a good reason to be so. This may not be easy for a child.

Appendix 1

In line with previous reports of longitudinal studies documenting superfluous *do* in child English, *do/does* followed by a verb appears in our data most frequently between the age of 2;3 and 3;4. However, this trajectory should be looked at with caution, since the largest number of files in CHILDES comes from 2 to 3 years old children. Therefore, whatever analysis one does, the data will always turn out most frequent at this stage and then decrease, without necessarily reflecting a decreasing tendency of the structure observed.

To test whether the higher number of occurrences of *do* + verb in the 2 to 3 years old range is the consequence of the higher number of utterances in the CHILDES database for this group age or not, we compared the total number of *do* + verb utterances with the total number of verbal utterances in the database. An evaluation search was conducted in order to access the total number of verbs. Furthermore, the latter set of utterances was split into ten age groups with 6-months differences. Under group 1, we subsumed all the utterances up to the age of 2, and under group 10 all the utterances from the age of 6. The rest is distributed as indicated in **Table 3**.

	age	**all V**	**do/does + V**	**ratio**
Group 1		17380	5	0.00029
Group 2	2;0–2;5	145296	79	0.00054
Group 3	2;3–2;11	209912	234	0.00111
Group 4	3;0–3;5	180640	144	0.00080
Group 5	3;6–3;11	85303	40	0.00047
Group 6	4;0–4;5	89923	80	0.00089
Group 7	4;6–4;11	122520	68	0.00056
Group 8	5;0–5;5	36169	35	0.00097
Group 9	5;6–5;11	14421	13	0.00090
Group 10		27011	12	0.00044

Table 9: Ratio of do/does + V in the set of verbal utterances in English CHILDES child data.

As the data show, the proportion reaches the highest value in the age group 3 (2;5–2;11) and then decreases. This implies that the large number of occurrences of *do* + V in this age group is not merely due to the number of utterances from this age group being the largest overall.

Appendix 2

	CHILD		ADULT	
FREQ	VERB	CLASS	VERB	CLASS
1	do	NCT	do	NCT
2	go	unergative	see	achievement
3	want	stative	go	unergative
4	get	causative	put	causative
5	have	stative	get	causative
6	put	causative	be	stative
7	see	achievement	have	stative
8	play	unergative	let	causative
9	let	causative	look	NCT
10	make	causative	know	stative
11	know	stative	like	stative
12	be	stative	make	causative
13	look	NCT	take	causative
14	sit	stative	come	achievement
15	take	causative	tell	NCT
16	like	stative	eat	NCT
17	come	achievement	want	stative
18	eat	NCT	say	NCT
19	help	causative	play	unergative
20	find	achievement	think	stative
21	thank	NCT	sit	stative
22	read	NCT	give	causative
23	need	stative	remember	stative
24	open	causative	find	achievement
25	please	stative	wanna	stative
26	hold	causative	thank	NCT
27	show	causative	help	causative
28	watch	NCT	try	NCT
29	lie	stative	turn	causative
30	hide	causative	wait	unergative
31	cut	causative	read	NCT
32	give	causative	show	causative
33	think	stative	hold	causative

(Contd.)

	CHILD		ADULT	
FREQ	VERB	CLASS	VERB	CLASS
34	turn	causative	open	causative
35	work	unergative	watch	NCT
36	stay	stative	please	stative
37	draw	NCT	hear	stative
38	buy	causative	keep	causative
39	wanna	stative	stop	causative
40	leave	achievement	bring	causative
41	write	NCT	throw	causative
42	say	NCT	leave	achievement
43	stand	stative	use	NCT
44	move	causative	pick	causative
45	gotta	causative	push	NCT
46	wait	unergative	pull	NCT
47	sing	NCT	need	stative
48	catch	causative	sleep	unergative
49	fall	unaccusative	draw	NCT
50	pick	causative	talk	unergative
51	push	NCT	buy	causative
52	pour	NCT	move	causative
53	sleep	unergative	work	unergative
54	fit	stative	stay	stative
55	stop	causative	listen	unergative
56	tell	NCT	write	NCT
57	wash	NCT	fit	stative
58	moo	unergative	gotta	stative
59	fly	unergative	ask	NCT
60	shut	causative	sing	NCT
61	keep	causative	drink	NCT
62	try	NCT	fix	causative
63	break	causative	stand	stative
64	bring	causative	wash	NCT
65	hear	stative	fall	unaccusative
66	throw	causative	build	NCT
67	stick	causative	finish	achievement
68	mend	causative	catch	causative

(Contd.)

	CHILD		ADULT	
FREQ	VERB	CLASS	VERB	CLASS
69	reach	achievement	start	unaccusative
70	hit	NCT	close	causative
71	use	NCT	cut	causative
72	excuse	causative	touch	NCT
73	drive	NCT	call	NCT
74	build	NCT	wear	stative
75	pardon	causative	clean	causative
76	boo	unergative	break	causative
77	fix	causative	hurt	causative
78	wipe	NCT	excuse	causative
79	woo	unergative	feel	stative
80	lift	causative	mean	stative
81	got	causative	hang	causative
82	carry	causative	wipe	NCT
83	cook	NCT	stick	causative
84	pull	causative	walk	NCT
85	pretend	NCT	roll	causative
86	drink	NCT	mind	stative
87	ride	NCT	round	causative
88	roll	causative	drive	causative
89	knock	NCT	count	NCT
90	run	unergative	happen	achievement
91	walk	unergative	run	unergative
92	touch	NCT	pop	unaccusative
93	hurt	causative	pretend	NCT
94	rain	unaccusative	blow	unaccusative
95	ask	NCT	pour	NCT
96	jump	unergative	bite	NCT
97	mind	stative	ok	stative
98	hire	causative	boo	unergative
99	meow	unergative	love	stative
100	talk	unergative	ride	NCT

Table 10: 100 most frequent verbs in child vs. adult data (CHILDES corpora accessible via Sketch Engine, data from children aged 0–3).

Funding information

This project was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (LeibnizDream, grant agreement No 856421).

Acknowledgements

We thank our anonymous reviewers for very valuable feedback and constructive criticisms on previous versions of this paper. We are also grateful to the audience of LAGB 2022 and Universität Leipzig for their comments, to Ute Bohnacker for discussion about the data she collected for her study on superfluous *do*, and to Isabel Oltra-Massuet and Víctor Acedo-Matellán for their editorial work throughout. We are also very grateful to Fabienne Salfner for her help with the data management.

Competing interests

The authors have no competing interests to declare.

References

Aksu-Koç, Ayhan & Slobin, Dan. 1985. The acquisition of Turkish. In Slobin, Dan (ed.), *The cross-linguistic studies of language acquisition, vol. 1: The data*, 839–876. Hillsdale: Lawrence Erlbaum Associates.

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

Alexiadou, Artemis & Rizou, Vasiliki. 2023. The use of periphrasis for the expression of aspect by Greek heritage speakers: A case study of register variation narrowing. *Register Studies* 5(1). 82–110. DOI: https://doi.org/10.1075/rs.20022.ale

Alexiadou, Artemis & Schäfer, Florian. 2006. Instrument subjects are agents or causers. In Baumer, Donald & Montero, David & Scanlon, Michael (eds.), *West Coast Conference on Formal Linguistics (WCCFL) 25*, 40–48. Somerville, MA: Cascadilla Proceedings Project.

Alhama, Raquel G. & Foushee, Ruthe & Byrne, Daniel & Ettinger, Allyson & Goldin-Meadow, Susan & Alishahi, Afra. 2023. Linguistic productivity: the case of determiners in English. In Park, Jong C. & Arase, Yuki & Hu, Baotian & Lu, Wei & Wijaya, Derry & Purwarianti, Ayu & Krisnadhi, Adila Alfa (eds.), *Proceedings of the 13th international joint conference on natural language processing and the 3rd conference of the Asia-Pacific chapter of the association for computational linguistics (volume 1: Long papers)*, 330–343. Nusa Dua, Bali: Association for Computational Linguistics. aclanthology.org/2023.ijcnlp-main.21/. DOI: https://doi.org/10.18653/v1/2023.ijcnlp-main.21

Andersen, Roger W. & Shirai, Yasuhiro. 1994. Discourse motivations for some cognitive acquisition principles. *Studies in Second Language Acquisition* 16(2). 133–156. DOI: https://doi.org/10.1017/S0272263100012845

Antoniou, Kyriakos & Katsos, Napoleon. 2017. The effect of childhood multilingualism and bilectalism on implicature understanding. *Applied Psycholinguistics* 38(4). 787–833. DOI: https://doi.org/10.1017/S014271641600045X

Arnon, Inbal. 2009. *Starting big – The role of multi-word phrases in language learning and use*. Palo Alto: Stanford dissertation.

Baillargeon, Renée & Carey, Susan. 2012. Core cognition and beyond: The acquisition of physical and numerical knowledge. In Pauen, Sabina (ed.), *Early childhood development and later outcome*, 33–65. Cambridge: Cambridge University Press.

Barner, David. 2001. Light verbs and the flexible use of words as noun and verb in early language learning. McGill University MSc Thesis, Montreal.

Bates, Elizabeth & Bretherton, Inge & Snyder, Lynn Sebestyen. 1991. *From first words to grammar: Individual differences and dissociable mechanisms*, vol. 20. Cambridge, MA: Cambridge University Press.

Benz, Anton. 2006. Partial blocking and associative learning. *Linguistics and Philosophy* 29(5). 587–615. DOI: https://doi.org/10.1007/s10988-006-9005-3

Bernstein, Nan. 1984. Patterns of vowel modification in motherese. *Journal of Child Language* 11(3). 557–578. DOI: https://doi.org/10.1017/S030500090000595X

Bezinska, Yanka & Chevrot, Jean-Pierre & Novakova, Iva. 2008. Le prédicat complexe *faire+Vinf* dans le langage enfantin entre 4 et 6 ans. In Durand, Jacques & Habert, Benoit & Laks, Bernard (eds.), *Congrès mondial de linguistique française*, 1707–1721. EDP Sciences. DOI: https://doi.org/10.1051/cmlf08224

Bliss, Lynn. 1988. The development of modals. *Journal of Applied Developmental Psychology* 9(3). 253–261. DOI: https://doi.org/10.1016/0193-3973(88)90028-7

Blom, Elma. 2003. *From root infinitive to finite sentence*. Utrecht: Universiteit Utrecht dissertation.

Blom, Elma & van de Craats, Ineke & Verhagen, Josje. 2013. *Dummy auxiliaries in first and second language acquisition*. Boston/Berlin: De Gruyter. DOI: https://doi.org/10.1515/9781614513476

Bloom, Lois. 1970. *Language development: Form and function in emerging grammars*. Cambridge, MA: MIT Press.

Bohannon III, John Neil & Marquis, Angela Lynn. 1977. Children's control of adult speech. *Child Development* 48(3). 1002–1008. DOI: https://doi.org/10.1111/j.1467-8624.1977.tb01259.x

Bohnacker, Ute. 2013. Reflections on dummy DO in child language and syntactic theory. In Blom, Elma & van de Craats, Ineke & Verhagen, Josje (eds.), *Dummy auxiliaries in first and second language acquisition*, vol. 49, 171. Berlin & New York: Walter de Gruyter. DOI: https://doi.org/10.1515/9781614513476.171

Bolinger, Dwight. 1972. *That's that*, vol. 155 (Janua Linguarum. Series Minor). Berlin & New York: De Gruyter Mouton.

Bott, Oliver & Solstad, Torgrim. 2014. From verbs to discourse: A novel account of implicit causality. In Hemforth, Barbara & Mertins, Barbara & Fabricius-Hansen, Cathrine (eds.), *Psycholinguistic approaches to meaning and understanding across languages*, 213–251. Dordrecht: Springer. DOI: https://doi.org/10.1007/978-3-319-05675-3_9

Bouchard, Denis. 1995. *The semantics of syntax*. Chicago: Chicago University Press.

Bowerman, Melissa. 1982. Starting to talk worse: Clues to language acquisition from children's late speech errors. In Strauss, Sidney (ed.), *U shaped behavioral growth*, 101–145. Academia Press. DOI: https://doi.org/10.1016/B978-0-12-673020-3.50012-4

Bowerman, Melissa. 2018. *Ten lectures on language, cognition, and language acquisition*. Leiden, The Netherlands: BRILL. DOI: https://doi.org/10.1163/9789004362826

Braunwald, Susan. 1971. Mother-child communication: the function of maternal-language input. *Word* 27(1–3). 28–50. DOI: https://doi.org/10.1080/00437956.1971.11435613

Brown, Roger. 1974. *A first language: The early stages*. Cambridge, MA: Harvard University Press. DOI: https://doi.org/10.4159/harvard.9780674732469

Cazden, C. 1968. The acquisition of noun and verb inflections. *Child Development* 39(2). 433–448. DOI: https://doi.org/10.1111/j.1467-8624.1968.tb04436.x

Cinque, Guglielmo. 2013. Cognition, universal grammar, and typological generalizations. *Lingua* 130. 50–65. DOI: https://doi.org/10.1016/j.lingua.2012.10.007

Cinque, Guglielmo & Rizzi, Luigi. 2009. The cartography of syntactic structures. In Heine, Bernad & Narrog, Heiko (eds.), *Oxford handbook of linguistic analysis*, 51–65. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780199544004.013.0003

Clark, Eve. 1978. Awareness of language: Some evidence from what children say and do. In Sinclair, Anne & Jarvella, Robert & Levelt, Willem (eds.), *The child's conception of language*, 17–43. Berlin: Springer. DOI: https://doi.org/10.1007/978-3-642-67155-5_2

Clark, Eve. 1985. The acquisition of Romance, with special reference to French. In Slobin, Dan (ed.), *The cross-linguistic studies of language acquisition. vol. 1: The data*, 687–782. Hillsdale, NJ: Lawrence Erlbaum Associates.

Comrie, Bernard. 1976. *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge University Press.

Copley, Bridget & Harley, Heidi. 2015. A force-theoretic framework for event structure. *Linguistics and Philosophy* 38(2). 103–158. DOI: https://doi.org/10.1007/s10988-015-9168-x

Cruse, D. 1973. Some thoughts on agentivity. *Journal of Linguistics* 47(1). DOI: https://doi.org/10.1017/S0022226700003509

Delancey, Scott. 1985. On active typology and the nature of agentivity. In Plank, Frans (ed.), *Relational typology*, 47–60. Berlin, New York & Amsterdam: Mouton. DOI: https://doi.org/10.1515/9783110848731.47

Demetras, Martha Jo-Ann. 1986. *Working parents' conversational responses to their two-year-old sons (linguistic input, language acquisition)*. Tucson, Arizona: The University of Arizona.

Demirdache, Hamida. 1997. 'Out of control' in Salish and event (de)composition. Manuscript, University of British Columbia.

Denison, David. 1985. The origins of periphrastic do: Ellegård and Visser reconsidered. In Eaton, Roger & Leek, Frederike & Koopman, Willem & Fischer, Olga (eds.), *4th International Conference*

on English Historical Linguistics. 45–60. Amsterdam & Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/cilt.41.07den

Dickinson, David & Tabors, Patton. 2001. *Beginning literacy with language: Young children learning at home and school*. Baltimore: Paul Brookes Publishing.

Dowty, David R. 1979. *Word meaning and Montague grammar*. Dordrecht: D. Reidel. DOI: https://doi.org/10.1007/978-94-009-9473-7

Ecay, Aaron. 2015. *A multi-step analysis of the evolution of English do-support*. Philadelphia: University of Pennsylvania dissertation.

Embick, David & Noyer, Rolf. 2001. Movement operations after syntax. *Linguistic Inquiry* 32(4). 555–595. DOI: https://doi.org/10.1162/002438901753373005

Engelberg, Stefan. 2005. Stativity, supervenience, and sentential subjects. In Maienborn, Claudia & Wöllstein, Angelika (eds.), *Event arguments: Foundations and applications*, 45–68. Tübingen: Niemeyer. DOI: https://doi.org/10.1515/9783110913798.45

Fauconnier, Stefanie. 2012. *Constructional effects of involuntary and inanimate agents: A cross-linguistic study*. Leuven: Katholieke Universiteit Leuven dissertation.

Feldman, Andrea. 1998. *Constructing grammar: Fillers, formulas, and function*. Boulder, CO: University of Colorado at Boulder.

Fletcher, Paul & Garman, Michael. 1988. Normal language development and language impairment: Syntax and beyond. *Clinical Linguistics and Phonetics* 2(2). 97–114. DOI: https://doi.org/10.3109/02699208808985246

Fodor, Jerry. 1970. Three reasons for not deriving *kill* from *cause to die*. *Linguistic Inquiry* 1(4). 429–438.

Folli, Raffaella & Harley, Heidi. 2005. Flavors of v. In *Aspectual inquiries*, vol. 62 (Studies in Natural Language and Linguistic Theory), 95–120. Dordrecht: Springer. DOI: https://doi.org/10.1007/1-4020-3033-9_5

Folli, Raffaella & Harley, Heidi. 2008. Teleology and animacy in external arguments. *Lingua* 118(2). 190–202. DOI: https://doi.org/10.1016/j.lingua.2007.02.004

Forrester, Michael. 2002. Appropriating cultural conceptions of childhood: Participation in conversation. *Childhood* 9(3). 255–278. DOI: https://doi.org/10.1177/0907568202009003043

Frei, Henri. 1929. *La grammaire des fautes*. Paris: P. Geuthner.

Gardenfors, Peter. 2014. *The geometry of meaning: Semantics based on conceptual spaces*. Cambridge, MA: MIT press. DOI: https://doi.org/10.7551/mitpress/9629.001.0001

Garvey, Catherine. 1979. An approach to the study of children's role play. *The Quarterly Newsletter of the Laboratory of Comparative Human Cognition* 1(4). 69–73.

Gathercole, Virginia. 1986. The acquisition of the present perfect: Explaining differences in the speech of Scottish and American children. *Journal of Child Language* 13(3). 537–560. DOI: https://doi.org/10.1017/S0305000900006875

Gelman, Susan & Coley, John & Rosengren, Karl & Hartman, Erin & Pappas, Athina & Keil, Frank. 1998. Beyond labeling: The role of maternal input in the acquisition of richly structured categories. *Monographs of the Society for Research in Child Development* 63(1). 1–148. DOI: https://doi.org/10.2307/1166211

Gelman, Susan A. & Taylor, Marianne G. & Nguyen, Simone P. 2004. Mother-child conversations about gender: Understanding the acquisition of essentialist beliefs. *Monographs of the Society for Research in Child Development* 69(1). 1–142. DOI: https://doi.org/10.1111/j.0037-976X.2004.00276.x

Gleason, Jean Berko. 1980. The acquisition of social speech routines and politeness formulas. In Giles, Robinson (ed.), *Language: Social psychological perspectives*, 21–27. Oxford: Pergamon. DOI: https://doi.org/10.1016/B978-0-08-024696-3.50009-0

Golden, John. 2016. Redundancy: When law repeats itself. *Texas Law Review* 94(4). 629–711.

Guasti, Maria Teresa & Alexiadou, Artemis & Sauerland, Uli. 2023. Undercompression errors as evidence for conceptual primitives. *Frontiers in Psychology* 14. DOI: https://doi.org/10.3389/fpsyg.2023.1104930

Guasti, Maria Teresa & Chierchia, Gennaro & Crain, Stephen & Foppolo, Francesca & Gualmini, Andrea & Meroni, Luisa. 2005. Why children and adults sometimes (but not always) compute implicatures. *Language and Cognitive Processes* 20(5). 667–696. DOI: https://doi.org/10.1080/01690960444000250

Haggerty, Laura CG. 1929. What a two-and-one-half-year-old child said in one day. *The Pedagogical Seminary and Journal of Genetic Psychology* 37(1). 75–101. DOI: https://doi.org/10.1080/08856559.1930.10532224

Hall, Nagy, William & Nottenburg. 1981. *Situational variation in the use of internal state words. Technical report no. 212.* Champaign, IL: University of Illinois.

Harley, Heidi. 1995. *Subjects, events, and licensing*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Haspelmath, Martin. 2014. Why is grammaticalization irreversible? *Linguistics* 37(6). 1043–1068. DOI: https://doi.org/10.1515/ling.37.6.1043

Haspelmath, Martin & Calude, Andreea & Spagnol, Michael & Narrog, Heiko & Bamyaci, Elif. 2014. Coding causal–noncausal verb alternations: A form–frequency correspondence explanation. *Journal of Linguistics* 50(3). 587–625. DOI: https://doi.org/10.1017/S0022226714000255

Hein, Johannes & Driemel, Imke & Martin, Fabienne & Nie, Yining & Alexiadou, Artemis. 2023. Errors of multiple exponence in child language. In Lu, Jiayi & Petersen, Erika & Zaitsu, Anissa & Harizanov, Boris (eds.), *Proceedings of the 40th West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Proceedings Project.

Hein, Johannes & Driemel, Imke & Martin, Fabienne & Nie, Yining & Alexiadou, Artemis. 2024. Errors of multiple exponence in child English: A study of past tense formation. *Morphology* 35. 143–183. DOI: https://doi.org/10.1007/s11525-024-09434-x

Higginson, Roy Patrick. 1985. *Fixing: Assimilation in language acquisition*. Washington: Washington State University dissertation.

Hoekstra, Teun & Hyams, Nina. 1998. Aspects of root infinitives. *Lingua* 106(1–4). 81–112. DOI: https://doi.org/10.1016/S0024-3841(98)00030-8

Hollebrandse, Bart & Roeper, Thomas. 1996. The concept of DO-insertion and the theory of INFL in acquisition. In Koster, Charlotte & Wijnen, Frank (eds.), *Proceedings of the Groningen Assembly on Language Acquisition*, 261–272.

Hollebrandse, Bart & van Koert, Margreet & van Hout, Angeliek. 2013. Semantic dummy verbs in child Dutch. In Blom, Elma & van de Craats, Ineke & Verhagen, Josje (eds.), *Dummy auxiliaries in first and second language acquisition*, 75–100. Boston/Berlin: De Gruyter. DOI: https://doi.org/10.1515/9781614513476.75

Howe, Christine. 1981. *Acquiring language in a conversational context*. New York: Academic Press.

Igartua, I. 2015. From cumulative to separative exponence in inflection: Reversing the morphological cycle. *Language* 91(3). 676–722. DOI: https://doi.org/10.1353/lan.2015.0032

JASP Team. 2025. JASP (Version 0.19.3) [Computer software]. https://jasp-stats.org/.

Joo, Sehrang & Yousif, Sami & Martin, Fabienne & Keil, Frank & Knobe, Joshua. 2025. No privileged link between intentionality and causation: Generalizable effects of agency in language. *Cognition*.

Jordens, Peter. 2013. Dummies and auxiliaries in the acquisition of L1 and L2 Dutch. In Blom, Elma & van de Craats, Ineke & Verhagen, Josje (eds.), *Dummy auxiliaries in first and second language acquisition*, 341–368. Boston/Berlin: De Gruyter. DOI: https://doi.org/10.1515/9781614513476.341

Kambanaros, Maria & Grohmann, Kleanthes K. 2015. More general all-purpose verbs in children with specific language impairment? Evidence from Greek for not fully lexical verbs in language development. *Applied Psycholinguistics* 36(5). 1029–1057. DOI: https://doi.org/10.1017/S0142716414000034

Karmiloff-Smith, Annette. 1981. *A functional approach to child language: A study of determiners and reference*. Cambridge, MA: Cambridge University Press.

Kelly, Linda & Quingley, Jean & Nixon, Elizabeth. 2020. Father-child shared book-reading in light- and heavy-text conditions. DOI: https://doi.org/doi:10.21415/0CHV-HZ36

Kilgarriff, Adam & Baisa, Vít & Bušta, Jan & Jakubíček, Miloš & Kovář, Vojtěch & Michelfeit, Jan & Rychlỳ, Pavel & Suchomel, Vít. 2014. The Sketch Engine: ten years on. *Lexicography* 1(1). 7–36. DOI: https://doi.org/10.1007/s40607-014-0009-9

Kittilä, Seppo. 2009. Causative morphemes as non-valency increasing devices. *Folia Linguistica* 43(1). 67–94. DOI: https://doi.org/10.1515/FLIN.2009.002

Klein, Michael. 2022. *What does* do *do? An investigation into the function of the English "dummy auxiliary"*. Fairfax, VA: George Mason University dissertation.

Klemola, Juhani. 1998. Semantics of DO in South-Western dialects of English. In Tieken-Boon van Ostade, Ingrid & van der Wal, Marijke & van Leuvensteijn, Arjan (eds.), *Do in english, dutch and german: History and present-day variation. munster*, 25–52. Stichting Neerlandistiek VU Amsterdam and Nodus Publikationen Münster.

Knipschild, Bettina. 2007. *Verb placement, DO-insertion and object shift: A case study of a bilingual German/English child*. Essex: University of Essex dissertation.

Koontz-Garboden, Andrew. 2009. Anticausativization. *Natural Language & Linguistic Theory* 27(1). 77–138. DOI: https://doi.org/10.1007/s11049-008-9058-9

Kotsinas, Ulla-Britt. 1988. Immigrant children's Swedish — a new variety? *Journal of Multilingual & Multicultural Development* 9(1–2). 129–140. DOI: https://doi.org/10.1080/01434632.1988.9994324

Kratzer, Angelika. 1996. Severing the external argument from its verb. In Rooryck, Johan (ed.), *Phrase structure and the lexicon*, 109–137. New York & Berlin: Springer. DOI: https://doi.org/10.1007/978-94-015-8617-7_5

Kroch, Anthony S & Myhill, John & Pintzuk, Susan. 1982. Understanding do. In Tuite, Kevin & Schneider, Robinson & Chametzky, Robert (eds.), *Papers from the Eighteenth Regional Meeting of Chicago Linguistics Society*. 282–294. Chicago: Chicago Linguistics Society.

Kuczaj, Stan A. 1977. The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior* 16(5). 589–600. DOI: https://doi.org/10.1016/S0022-5371(77)80021-2

Kuczaj, Stan A. 1978. Children's judgments of grammatical and ungrammatical irregular past-tense verbs. *Child Development* 49(2). 319–326. DOI: https://doi.org/10.1111/j.1467-8624.1978.tb02318.x

Le Normand, Marie-Thérèse. 1986. A developmental exploration of language used to accompany symbolic play in young, normal children (2–4 years old). *Child: Care, Health and Development* 12(2). 121–134. DOI: https://doi.org/10.1111/j.1365-2214.1986.tb00492.x

Legate, Julie Anne. 2014. *Voice and v: Lessons from Acehnese*. Cambridge, MA: MIT Press. DOI: https://doi.org/10.7551/mitpress/9780262028141.001.0001

Levin, Beth. 1999. Objecthood: An event structure perspective. In Billings, Sabrina & Boyle, John & Griffith, Aaron (eds.), *Proceedings of the 35th Meeting of the Chicago Linguistics Society*, vol. 1, 223–247. Chicago: Chicago Linguistic Society.

Lieven, Elena & Salomo, Dorothé & Tomasello, Michael. 2009. Two-year-old children's production of multiword utterances: A usage-based analysis. *Cognitive Linguistics* 20(3). 481–508. DOI: https://doi.org/10.1515/COGL.2009.022

Lightfoot, David. 1991. *How to set parameters: Arguments from language change*. Cambridge, MA: MIT Press.

MacWhinney, Brian. 1985. Hungarian language acquisition as an exemplification of a general model of grammatical development. In Slobin, Dan (ed.), *The cross-linguistic studies of language acquisition. vol. 2: Theoretical issues*. Hillsdale: Lawrence Erlbaum Associates.

MacWhinney, Brian. 2000. *The CHILDES project: Tools for analyzing talk*. Cambridge, MA: MIT Press.

MacWhinney, Brian & Snow, Catherine. 1985. The child language data exchange system. *Journal of Child Language* 12(2). 271–295. DOI: https://doi.org/10.1017/S0305000900006449

Martin, Fabienne. 2022. Jean-Pierre Koenig and Yanwei Jin's analysis of expletive negation: commentary. Talk presented at *Time in Translation workshop*, Universiteit Utrecht, September 23, 2022.

Martin, Fabienne & Nie, Yining & Alexiadou, Artemis & Guasti, Maria Teresa. 2022. Wearing Causation on Its Sleeve: Overt *cause* in Child French Causatives. In Gong, Ying & Kpogo, Felix (eds.), *BUCLD 46: Proceedings of the 46th Annual Boston University Conference on Language Development*, 497–510. Somerville, MA: Cascadilla Press.

Martin, Fabienne & Nie, Yining & Driemel, Imke & Hein, Johannes & Alexiadou, Artemis. 2023. Morphological transparency and redundancy in child language. Manuscript, Utrecht University, San Jose State University, York University, Humboldt-Universität zu Berlin and ZAS Berlin.

McCawley, James D. 1978. Conversational implicature and the lexicon. In Cole, Peter (ed.), *Pragmatics*, 245–259. Brill. DOI: https://doi.org/10.1163/9789004368873_009

McCune, Lorraine. 1995. A normative study of representational play in the transition to language. *Developmental Psychology* 31(2). 198–206. DOI: https://doi.org/10.1037//0012-1649.31.2.198

Morgenstern, A. & Parisse, C. 2007. Codage et interprétation du langage spontané d'enfants de 1 à 3 ans. *Corpus* 6. 55–78. DOI: https://doi.org/10.4000/corpus.922

Morgenstern, Aliyah & Benazzo, S. & Leroy, M. & Mathiot, E. & C., Parisse. & Sekali, M. 2009. *L'enfant dans la langue. De l'observation du naturaliste à l'analyse du linguiste*. Paris: Presses de la Sorbonne Nouvelle.

Morisset, Colleen & Barnard, Kathryn & Booth, Cathryn. 1995. Toddlers' language development: Sex differences within social risk. *Developmental Psychology* 31(5). 851–865. DOI: https://doi.org/10.1037//0012-1649.31.5.851

Nelson, Katherine. 1989. *Narratives from the crib*. Cambridge, MA: Harvard University Press.

Newman, Rochelle & Rowe, Meredith & Ratner, Nan Bernstein. 2016. Input and uptake at 7 months predicts toddler vocabulary: The role of child-directed speech and infant processing skills in language development. *Journal of Child Language* 43(5). 1158–1173. DOI: https://doi.org/10.1017/S0305000915000446

Ninio, Anat & Snow, Catherine & Pan, Barbara & Rollins, Pamela. 1994. Classifying communicative acts in children's interactions. *Journal of Communication Disorders* 27(2). 157–188. DOI: https://doi.org/10.1016/0021-9924(94)90039-6

Paradis, Johanne & Nicoladis, Elena & Genesee, Fred. 2000. Early emergence of structural constraints on code-mixing: Evidence from French–English bilingual children. *Bilingualism: Language and Cognition* 3(3). 245–261. DOI: https://doi.org/10.1017/S1366728900000365

Penhallurick, John. 1985. The semantics of auxiliary do. *Studies in Language* 9(3). 311–333. DOI: https://doi.org/10.1075/sl.9.3.02pen

Piñón, Christopher. 1997. Achievements in an event semantics. In Lawson, Aaron & Cho, Eun (eds.), *Semantics and linguistic theory (SALT) 7*, 273–296. Ithaca, New York: CLC Publications, Cornell University. DOI: https://doi.org/10.3765/salt.v7i0.2781

Post, Kathryn Nolan. 1992. *The language learning environment of laterborns in a rural Florida community*. Cambridge, MA: Harvard University dissertation.

Punske, Jeffrey & Butler, Elizabeth. 2019. Do me a syntax: Doggo memes, language games and the internal structure of English. *Ampersand* 6. Article 100052. DOI: https://doi.org/10.1016/j.amper.2019.100052

Quirk, Randolph & Greenbaum, Sidney & Leech, Geoffrey & Svertvik, Jan. 1985. *A comprehensive grammar of the English language*. London & New York: Longman.

Rappaport Hovav, Malka & Levin, Beth. 1998. Building verb meanings. In Butt, Miriam & Geuder, Wilhelm (eds.), *The projection of arguments: Lexical and compositional factors*, 97–134. Chicago: CSLI Publications.

Rizzi, Luigi. 1993. Some notes on linguistic theory and language development: The case of root infinitives. *Language Acquisition* 3(4). 371–393. DOI: https://doi.org/10.1207/s15327817la0304_2

Roeper, Thomas. 1991. How a marked parameter is chosen: Adverbs and *do*-insertion in the IP of child grammar. *University of Massachusetts Occasional Papers in Linguistics* 17(1). 9.

Rollins, Pamela Rosenthal. 2003. Caregivers' contingent comments to 9-month-old infants: Relationships with later language. *Applied Psycholinguistics* 24(2). 221–234. DOI: https://doi.org/10.1017/S0142716403000110

Ross, John Robert. 1972. Act. In Davidson, Donald & Harman, Gilbert (eds.), *Semantics of natural language*, vol. 40 (Synthese Library), 70–126. Dordrecht: Springer. DOI: https://doi.org/10.1007/978-94-010-2557-7_4

Rowland, Caroline. 2007. Explaining errors in children's questions. *Cognition* 104(1). 106–134. DOI: https://doi.org/10.1016/j.cognition.2006.05.011

Sachs, Jacqueline. 1983. Talking about the *there* and *then*: The emergence of displaced reference in parent-child discourse. *Children's Language* 4. 1–28.

Sauerland, Uli & Alexiadou, Artemis. 2020. Generative grammar: A meaning first approach. *Frontiers in Psychology* 11. 3104. DOI: https://doi.org/10.3389/fpsyg.2020.571295

Sawyer, Keith. 1997. *Pretend play as improvisation: Conversation in the preschool classroom*. New York: Psychology Press.

Schoenmakers, Gert-Jan. 2023. Linguistic judgments in 3D: The aesthetic quality, linguistic acceptability, and surface probability of stigmatized and non-stigmatized variation. *Linguistics* 61(3). 779–824. DOI: https://doi.org/10.1515/ling-2021-0179

Schütze, Carson T. 2004. Synchronic and diachronic microvariation in English *do*. *Lingua* 114(4). 495–516. DOI: https://doi.org/10.1016/S0024-3841(03)00070-6

Schütze, Carson T. 2013. Superfluous *do* and comparison of spell-outs. In Blom, Elma & van de Craats, Ineke & Verhagen, Josje (eds.), *Dummy auxiliaries in first and second language acquisition*, 13–38. Berlin: Mouton de Gruyter.

Sert, Cansel & Hubers, Ferdy & Redl, Theresa & Hoop, Helen de. 2023. On the acceptability of the not so dummy auxiliary 'do' in Dutch. In Leufkens, Sterre & Bril, Marco (eds.), *Linguistics in the Netherlands* 40, 210–229. DOI: https://doi.org/10.1075/avt.00088.ser

Shirai, Yasuhiro. 1991. *Primacy of aspect in language acquisition: Simplified input and prototype*. Los Angeles: University of California dissertation.

Shirai, Yasuhiro & Andersen, Roger W. 1995. The acquisition of tense-aspect morphology: A prototype account. *Language* 71(4). 743–762. DOI: https://doi.org/10.2307/415743

Siddiqi, Daniel. 2006. *Minimize exponence: Economy effects on a model of the morphosyntactic component of the grammar*. Tucson, AZ: The University of Arizona dissertation.

Slobin, Dan. 1971. On the learning of morphological rules: A reply to Palermo and Eberhart. In Slobin, Dan (ed.), *The ontogenesis of grammar: A theoretical symposium*, 215–223. New York: Academia Press.

Slobin, Dan. 1973. Cognitive prerequisites for the development of grammar. In Ferguson, Charles A. & Slobin, Dan (eds.), *Studies of child language development*, 175–208. New York: Holt, Rinehart, & Winston.

Slobin, Dan. 1985. Crosslinguistic evidence for the language-making capacity. In Slobin, Dan (ed.), *The cross-linguistic studies of language acquisition. vol. 2: Theoretical issues*, 406–605. Hillsdale: Lawrence Erlbaum Associates.

Spelke, Elizabeth & Kinzler, Katherine. 2007. Core knowledge. *Developmental Science* 10(1). 89–96. DOI: https://doi.org/10.1111/j.1467-7687.2007.00569.x

Sprott, Richard. 1992. Children's use of discourse markers in disputes: Form-function relations and discourse in child language. *Discourse Processes* 15(4). 423–440. DOI: https://doi.org/10.1080/01638539209544821

Steube, Anita. 2001. Correction by contrastive focus. *Theoretical Linguistics* 27(2–3). 215–250. DOI: https://doi.org/10.1515/thli.2001.27.2-3.215

Strickland, Brent. 2017. Language reflects "core" cognition: A new theory about the origin of cross-linguistic regularities. *Cognitive Science* 41(1). 70–101. DOI: https://doi.org/10.1111/cogs.12332

Stromswold, Karin. 1990. *Learnability and the acquisition of auxiliaries*. Cambridge: Massachusetts Institute of Technology dissertation.

Suppes, Patrick. 1974. The semantics of children's language. *American Psychologist* 29(2). 103–114. DOI: https://doi.org/10.1037/h0036026

Theakston, Anna & Lieven, Elena & Pine, Julian & Rowland, Caroline. 2001. The role of performance limitations in the acquisition of verb-argument structure: An alternative account. *Journal of Child Language* 28(1). 127–152. DOI: https://doi.org/10.1017/S0305000900004608

Tommerdahl, Jodi & Kilpatrick, Cynthia. 2014. The reliability of morphological analyses in language samples. *Journal of Language Testing* 31(1). 3–18. DOI: https://doi.org/10.1177/0265532213485570

Tyler, Matthew. 2022. Non-valency-increasing causatives in a Distributive morphology perspective. Talk presented at LAGB 2022, 15–18 September.

Valian, Virginia. 1991. Syntactic subjects in the early speech of American and Italian children. *Cognition* 40(1–2). 21–81. DOI: https://doi.org/10.1016/0010-0277(91)90046-7

van Hout, Angeliek. 2016. Lexical and grammatical aspect. In Lidz, Jeff & Pater, Joe & Snyder, William (eds.), *Handbook of developmental linguistics*, 587–610. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780199601264.013.25

Van Houten, Lori. 1986. Role of maternal input in the acquisition process: The communicative strategies of adolescent and older mothers with their language learning children. Paper presented at the 11th Annual Boston University Conference on Language Development.

van Kampen, Jacqueline. 1997. *First steps in wh-movement*. Utrecht: Universiteit Utrecht dissertation.

van Rooy, Robert. 2004. Signalling games select Horn strategies. *Linguistics and Philosophy* 27(4). 493–527. DOI: https://doi.org/10.1023/B:LING.0000024403.88733.3f

van Valin, Robert & Wilkins, David. 1996. The case for "effector": Case roles, agents, and agency revisited. In Shibatani, Masayoshi & Thompson, Sandra (eds.), *Grammatical constructions*, 289–322. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oso/9780198235392.003.0011

Wang, Jing & Shirai, Yasuhiro. 2015. The development of third person singular present form *-s*. In Ayoun, Dalila (ed.), *The acquisition of the present*, 1–20. Amsterdam & Philadelphia: Benjamins. DOI: https://doi.org/10.1075/z.196.01wan

Warner, Anthony R. 1993. *English auxiliairies. Structure and history*, vol. 66 (Cambridge studies in linguistics).

Warren-Leubecker, Amye. 1982. *Sex differences in speech to children*: Georgia Institute of Technology dissertation.

Weist, Richard M & Pawlak, Aleksandra & Hoffman, Karen. 2009. Finiteness systems and lexical aspect in child Polish and English. *Linguistics* 47(6). 1321–1350. DOI: https://doi.org/10.1515/LING.2009.046

Wells, Gordon. 1981. *Learning through interaction, volume 1: the study of language development*. Cambridge: Cambridge University Press.

Wiese, Heike. 2009. Grammatical innovation in multiethnic urban Europe: New linguistic practices among adolescents. *Lingua* 119(5). 782–806. DOI: https://doi.org/10.1016/j.lingua.2008.11.002

Wilson, Elspeth & Katsos, Napoleon. 2020. Acquiring implicatures. In Schneider, Klaus P. & Ifantidou, Elly (eds.), *Handbooks of pragmatics. Developmental and clinical pragmatics*, 118–148. Berlin & New York: Walter de Gruyter. DOI: https://doi.org/10.1515/9783110431056-005

Wilson, Elspeth Amabel. 2017. *Children's development of quantity, relevance and manner implicature understanding and the role of the speaker's epistemic state*: University of Cambridge dissertation.

Wittenberg, Eva. 2018. *He Gave My Nose a Kick* or *He Kicked My Nose*? Argument structure alternations and event construal. *Psychology of Learning and Motivation* 68. 337–360. DOI: https://doi.org/10.1016/bs.plm.2018.08.008

Yamakoshi, Kyoko & Miura, Kaori & Jorinbo, Hanako & Angata, Kayoko & Yamasaki, Kaori. 2018. An experimental study of children's comprehension of lexical and productive causatives in

Japanese. In Nishiyama, Kunio & Kishimoto, Hideki & Aldridge, Edith (eds.), *Topics in theoretical Asian linguistics: Studies in honor of John B. Whitman*, 229–250. Amsterdam & Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/la.250.12yam

Zukowski, Andrea. 1996. Auxiliary auxiliaries: Spare parts in early speech. Manuscript, Boston University.