



# Usage and acquisition patterns of the *how come*- and *why*-constructions by young English-speaking children\*

Jungsoo Kim

(Incheon National University)

**Kim, Jungsoo. 2024. Usage and acquisition patterns of the *how come*- and *why*-constructions by young English-speaking children. *Linguistic Research* 41(Special Edition): 109-155.** This paper explores young English-speaking children's usage and acquisition patterns of the *how come*- and *why*-constructions on the basis of authentic data from the CHILDES (Child Language Data Exchange System) subcorpora. In doing so, it particularly focuses on whether young children show the same or similar acquisition processes of the two constructions which have similar meanings and if not, in what aspects they display different behavior, and what implications they have in children's acquisition of *wh*-question constructions in general. A detailed analysis of the data shows that 1) children start to use and acquire the *how come*-construction later than the *why*-construction in general; 2) although they use and acquire the two constructions in a similar way to some extent, their usage and acquisition patterns differ rather significantly, regarding the dependent category/construction types, the illocutionary function distributions/preferences; 3) children's difficulty in employing SAI in the *why*-construction is demonstrated in a large number of underdeveloped/erroneous instances with a finite sentence dependent at early acquisition stages and frequent occurrences of the small clause dependent in spite of the lack of parents' input; 4) children use and acquire the *why*-construction earlier in the positive contexts than in the negative contexts; 5) parents' input does not play a sole determining role in accounting for children's usage and acquisition patterns of the two constructions. The findings of this paper thus make a significant contribution to the body of irregular *wh*-questions including the *how come*-construction and typical *wh*-questions such as the *why*-construction at the same time, inviting further subsequent empirical studies on the related constructions. (Incheon National University)

**Keywords** *how come*-construction, *why*-construction, acquisition, parents' input, CHILDES, dependent/construction types, illocutionary functions

---

\* This work was supported by Incheon National University (International Cooperative) Research Grant in 2023-0385.

## 1. Introduction

This paper discusses how young English-speaking children use and acquire the *how come*- and *why*-constructions, as exemplified below (Zwicky and Zwicky 1971; Stromswold 1990; Collins 1991; Culicover 1991; de Villers 1991; Rowland and Pine 2000; Berk 2003; Ochi 2004; Fitzpatrick 2005; Conroy 2006; Thornton 2008, 2016; Kim and Kim 2011; Duffield 2015; Kim and Kim 2017; Radford 2018):

- (1) a. How come John was hired?  
b. How come John left?
- (2) a. Why was John hired?  
b. Why did John leave?

The *how come*-construction is characterized as involving the expression *how come* at the beginning of the clause with a similar “causal” interpretation of *why* and lacking subject-auxiliary inversion (SAI) in matrix environments. Thus, the examples in (1) with *how come* can be paraphrased as in (2) with *why*, involving SAI. There are certain differences between these two constructions, including SAI; nonetheless, they can both be used to induce a causal interpretation.

In previous literature, young English-speaking children’s acquisition of the *why*-construction has been extensively discussed, in particular, in comparison with other *wh*-question constructions, based on corpus data (Stromswold 1990; de Villers 1991; Rowland and Pine 2000; Berk 2003; Thornton 2008, 2016). However, their acquisition of the *how come*-construction has not been examined thus far. In this regard, this paper attempts to show whether young children display the same acquisition processes of the two constructions which have similar meanings and if not, in what respects they exhibit different behavior, and to discuss implications of their acquisition of *wh*-question constructions in general.

The rest of this paper is structured as follows. Section 2 briefly reviews previous literature on children’s acquisition of the *why*-construction. Section 3 then discusses the widely-known basic grammatical properties of the *how come*- and *why*-constructions noted in previous literature. Section 4 provides a description of the data and the methodology used for analysis. Section 5 shows corpus findings observed from CHILDES (Child Language Data Exchange System) data (MacWhinney

and Snow 1985, 1990; MacWhinney 2000) and discusses how young children actually use the *how come*- and *why*-constructions in real life and what acquisition processes they undergo. Section 6 further discusses some major findings of this study in more detail with respect to how young children use and acquire the two constructions and what role their parents' input plays. The section then briefly summarizes the major contribution of this study and offers suggestions for subsequent research.

## 2. Previous approaches to children's acquisition of *why*-questions

Children's *why*-questions have long held the attention of many scholars in linguistics. The first studies date back to the early 20th century. For instance, Piaget (1929) and Isaacs (1930) argued that children begin to produce *why*-questions to gain more information in order to fill gaps in their knowledge and deal with anomalies, deviations, contrasts, or differences which stimulate a sense of unease or unsettledness (Bova and Arcidiacono 2013: 715).

In line with this, two major functions of children's *why*-questions have been found: 1) explanatory and 2) argumentative (Walton 2004; Bova 2011; Bova and Arcidiacono 2013). The first function is to ask for an explanation whereas the second one is to trigger the beginning of argumentative discussions with the other conversation participant(s). These two major functions of *why*-questions are presented in the dialogues below:

(3) FRA: Dad, why isn't it raining today?

DAD: Because today, the clouds are full of water. But they want to keep it just for themselves, a little longer!

(4) Situation: ALE touches and looks at the container with the drugs.

ALE: I'm going to take one of these. Yes!

MOM: You can't, Alessandro!

ALE: What?

MOM: You can't. [shakes her head]

ALE: Why not?

MOM: Because children have to take special drugs. They can't take drugs for adults. Otherwise, they will get sick.

Children ask *why*-questions with an explanatory function to find out the reasons that have caused an event already ascertained, as in (3). In this regard, there is no difference of opinion between children and the other conversation participants (parents, usually). On other other hand, as in (4), children also use *why*-questions with an argumentative function to seek to figure out the reasons on the basis of the other conversation participants' actions. Thus, there is a difference of opinion between children and the other conversation participants.

Previous literature focusing on the acquisition order of different types of *wh*-questions has shown that children start to produce and comprehend *why*-questions in later developmental stages than other *wh*-questions introduced by *who* and *what* (Ervin-Tripp 1970; Tyack and Ingram 1977; Rowland et al. 2003). Some previous studies have also shown that some English-speaking children fail to exhibit SAI in their *why*-questions even long after they have mastered it in other types of *wh*-questions (Klima and Bellugi 1966; Labov and Labov 1978; Stromswold 1990; de Villers 1991; Rowland and Pine 2000; Berk 2003; Rowland 2007; Thornton 2008).<sup>1</sup> As an illustration, consider the following *wh*-question sentences produced by one English-speaking child, AL, around the age of 3:

- (5) Data from Thornton (2008: 108, (1)):
- a. Why unicorns are pretend? (3;1)
  - b. Why you are going in that one? (3;2)
  - c. How did Tweetie get maked? (3;0)
  - d. When will we be big enough to climb up there? (3;1)
  - e. Who was it that you was talking to? (3;1)

Around this developmental stage, most of AL's *why*-questions lack SAI as in (5a) and (5b), as compared to questions with other *wh*-phrases as in (5c)-(5e).

Alongside the lack of SAI, it has been also shown that in children's *why*-questions various phrases can intervene between *why* and the subject, as exemplified by AL's utterances in (6) (Thornton 2008):

---

<sup>1</sup> More generally, previous literature has demonstrated an argument/adjunct asymmetry (Stromswold 1990; Sarma 1991; de Villers 1991). In other words, they have shown that when children successfully use SAI for *wh*-questions with argument *wh*-phrases (e.g., *who* and *what*), they fail to use it for those with adjunct *wh*-phrases (e.g., *why*, *how*, and *where*).

- (6) Data from Thornton (2008: 108, (2)):
- a. Why *SOME OF YOUR MAKE UP* I can't use [and some I can]? (5;2)
  - b. Why *this time* you're opening them like that? (4;2)
  - c. Why *when I was a baby* I loved Boomer's dog food? (3;6)
  - d. Why *always when he rings* I'm asleep? (5;3)

In (6a), a topicalized NP intervenes *why* and the subject. In (6b) and (6c), a temporal adjunct NP and a temporal adjunct clause introduced by *when* appear between *why* and the subject, respectively. In (6d), a frequency adjunct and a temporal adjunct clause both occur between the two constituents.

With respect to SAI, previous literature has identified some environments in which SAI takes place as opposed to matrix *why*-questions which frequently lack SAI (Conroy 2006; Crain et al. 2006; Conroy and Lidz 2007; Thornton 2008). For instance, it has been observed that when children's *why*-questions involve an embedded complement clause to induce a long-distance interpretation, SAI consistently occurs, unlike their simple matrix *why*-questions. Again, consider AL's *why*-question examples in (7):

- (7) Data from Thornton (2008: 135, (45)):
- a. Why do you think Santa's not coming this year? (3;10)
  - b. Why do you think that Boomer came in with us? (4;2)
  - c. Why do you think that Mommy would not wanna watch the show? (4;6)
  - d. Why do you not think there's going to be a Little Mermaid? (4;11)
  - e. Why do you think you're gonna have a bad afternoon? (5;0)
  - f. Why do you think that my electric car only goes backwards? (5;5)

In (7), each of the *why*-questions has to do with the embedded clause, giving rise to a long-distance interpretation, and SAI is found in the matrix clause.

In addition, Thornton (2008) noted that when children's matrix *why*-questions have a suggestion function, SAI is consistently employed, as in (8), as opposed to when they have a simple information-seeking question function. AL's *why*-questions with suggestion and rhetorical question functions are provided in (8):

- (8) a. Why don't you use this as a magic wand? (3;4) (suggestion)  
 b. Why would any witch not do spells? (3;11) (rhetorical question)

The *why*-question example in (8a) with a contracted negative dummy auxiliary verb *don't* and a second person pronoun subject is used to make a suggestion. The *why*-question example in (8b) is used not to ask for an answer but to emphasize the proposition that every witch would do spells. Examples like these are not true information-seeking questions and it has been found out that children systematically carry out SAI for these *why*-questions.

Conroy and Lidz (2007) also revealed on the basis of experimental data that some English-speaking children have problems with SAI for *why*-questions in production, but the same children maintain the adult-like ambiguity comprehension in *why*-questions with an embedded complement clause. For instance, observe the well-formed *why*-question examples with an embedded complement clause in (9) below:

- (9) a. Why did Joe think Monster ate his sandwich?  
 b. Why did you think Joe is wearing a sweater?

Some children in their production experiments could not successfully use the SAI strategy for examples as in (9); however, even those children had no problem in understanding such examples. Based on the experimental results, Conroy and Lidz (2007) argued that there is a production/comprehension asymmetry in young English-speaking children's *why*-questions in that they appear non-adult-like in production but they are fully adult-like in comprehension of *why*-questions with an embedded complement clause.

### **3. Grammatical properties of the *how come*-construction in comparison with the *why*-construction**

As noted in Section 1, the *how come*-construction is similar to the *why*-construction in that they both yield a causal interpretation, but they differ in SAI. Previous literature has shown several other intriguing grammatical properties of the two constructions. In this section, I discuss their grammatical properties that have been noted in literature, in particular, by making comparisons between them when possible and necessary.

First, one salient grammatical property of the *how come*-construction involves the

invariable use of *come* in the construction, as shown in (10) (Kim and Kim 2011):<sup>2</sup>

- (10) a. How {come/\*came} John was hired?  
 b. How {come/\*comes} everyone hates John?

As in (10), only the bare form *come* can be used in the *how come* construction, but not the tense ones like *came* and *comes*.

Note also that no material can intervene between *how* and *come* in the *how come*-construction (Collins 1991; Kim and Kim 2011). Consider the following examples:

- (11) a. Why else would they reject you?  
 b. \*How else come they would reject you?  
 (12) a. Why the hell did he resign?  
 b. \*How the hell come he resigned?

As in (11a) and (12a), typical *why*-construction examples can occur with an adverbial expression like *else* and an emotive expression such as *the hell*, *the heck*, and *the devil*; on the other hand, the *how come*-construction does not allow for any intervening material between *how* and *come*, providing additional support for the idea that *how come* has an invariable, fixed form on its own.

One more piece of supporting evidence for the invariable form of *how come* comes from the impossibility to replace *how* with another *wh*-expression, as illustrated below

---

2 Kim and Kim (2017: 129-131) show that two variants of *how come* (i.e., *how came* and *how comes*) were used in the past but almost disappeared in Present-Day English (PDE), providing corpus examples from COHA (Corpus of Historical American English) like the following:

- (i) a. But how came you so wet? I've been in the river, answered John, ... (COHA 1864 FIC)  
 b. How came you here? (COHA 1870 FIC)  
 c. How comes Ned here? (COHA 1887 FIC)

Note, however, that in each of these examples *how came* or *how comes* occurs with a non-finite clause missing a finite verb. Notice also that it is not so clear whether each of these examples has a causal interpretation like *how come* and *why* without the surrounding context. To be more specific, all these examples can be understood as involving inversion of the main verb *came* or *comes* as a deictic motion verb and the manner meaning of *how*. Therefore, I agree with previous literature in that *how come* is a rather fixed form in PDE, but I would like to point out that Kim and Kim's (2017) claim needs further justification and investigation.

(Radford 2018: 221):

- (13) How/\*However/\*What/\*Who/\*When/\*Where come John resigned?

The *wh*-expression *how* in the *how come*-construction cannot be replaced by its *-ever* counterpart, *however*, and any other *wh*-expressions, showing a peculiarity of the construction.

However, Radford (2018: 218-219, 234-235) provides examples collected from some internet sources where *how come* does not form a single, invariable constituent on its own. Observe his data given below:

- (14) a. How came you never watched Sailor Moon? (turbomun.tumblr.com)  
 b. How else come they're in diamond? (reddit.com)  
 c. How the heck come I don't get my fair share?  
 (readersupportednews.com)  
 d. Why come you didn't call me last night? (urbandictionary.com)

These examples indicate that in certain varieties of English the expression may not be a fixed expression.

The next point to make about the *how come*-construction concerns its syntactic similarities to the *why*-construction. For instance, although they exhibit different behavior in terms of SAI when they take a finite sentence dependent, they behave the same in terms of the possibility to occur with a small clause (SC) dependent (Kim and Kim 2011; Kim and Kim 2017: 123).

- (15) a. Now, tell me the truth, {how come/why} you singing songs for me?  
 (COHA 2000 FIC)  
 b. {How come/Why} your head so hard? (COHA 1968 FIC)  
 c. {How come/Why} you in my daughters' chamber? (COHA 1934 FIC)

These examples indicate that both the *how come*-construction and the *why*-construction can take a non-finite small clause dependent, in addition to a finite sentence.

The two constructions also show the same behavior in some ellipsis environments



(Zwicky and Zwicky 1971: 183; Yoshida et al. 2015; Kim and Kim 2017: 123-124).<sup>3</sup>

Consider the following examples:

- (16) a. A: John left the party early. B: Why/How come? (Sluicing)  
 b. A: John was eating natto. B: Why/How come natto? (Stripping)

As in (16a), both *how come* and *why* can stand on their own but can be interpreted as a clause so long as relevant context is provided. In a similar manner, as in (16b), they can also occur with a fragment, inducing a clausal interpretation. In other words, the two constructions can be licensed in the sluicing and stripping contexts.

Moreover, both constructions can appear in embedded environments without SAI; however, they cannot appear in embedded environments as *to*-infinitival questions, as opposed to other *wh*-question constructions (Duffield 2015; Kim and Kim 2017):

- (17) a. I wonder {how come/why} she didn't come to the party last night.  
 b. I wonder {how come/why} John is not in school today.  
 (18) a. \*I don't know {how come/why} to leave.  
 b. We haven't decided {where to go/what to eat/who to vote for}.

These examples suggest that the *how come*-construction and the *why*-construction show the same behavior in embedded environments with respect to the absence of SAI and the impossibility to take a *to*-infinitival VP as their dependent.

Thus far, we have seen syntactic properties common in both the *how come*-construction and the *why*-construction. However, the two display different syntactic properties in a few respects as well. For instance, the *why*-construction is possible under the sluicing and stripping contexts along with the negative marker *not* while the *how come*-construction is not (Collins 1991: 34-35; Culicover 1991: 56; Merchant 2006: 22; Duffield 2015: 63; Yoshida et al. 2015: 362).<sup>4</sup>

<sup>3</sup> See Zwicky and Zwicky (1971), Fitzpatrick (2005), and Tsai (2008) for discussion on the different semantic interpretations for *why*- and *how*-questions and the *how come*-construction.

<sup>4</sup> The *why*-construction in the negative sluicing context, i.e., *why not?*, has two different functions, as shown in (i) below (Hofmann 2018; Stockwell 2022):

- (i) a. A: Jasmine didn't sleep. B: (Really?) Why not?  
 b. Let's dance. (Sure,) Why not?

- (19) a. A: John is not leaving. B: {Why/\*How come} not? (Negative sluicing)  
 b. A: Even John is respected. B: Then, {why/\*how come} not Mary?  
 (Negative stripping)

As shown by the contrast in the examples in (19), the *why*-construction is licensed in the negative sluicing and negative stripping contexts with the negation marker *not* in addition to their positive counterparts, but such negative ellipsis contexts are not possible with the *how come*-construction.

The two constructions under discussion also differ in the possibility to allow aggressively non-D-linked *wh*-expressions (Pesetsky 1987; Culicover 1991: 54).

- (20) a. Why the hell did John leave the party early?  
 b. \*How come the hell John left the party early?  
 (21) a. Why in the world does everyone hate John?  
 b. \*How come in the world everyone hates John?

As illustrated here, expressions like *the hell/heck/fuck* and *in the world* can occur with *why*, licensing the so-called aggressively non-D-linked construction; on the other hand, *how come* does not license the construction.

Another syntactic difference between the *why*-construction and the *how come*-construction concerns the availability of multiple *wh*-questions, as demonstrated below (Collins 1991: 37-38; Ochi 2004: 30; Conroy 2006: 5-6):

- (22) a. Why did John eat what?  
 b. \*How come John ate what?

As in (22a), the *why*-construction can introduce multiple *wh*-questions with other *wh*-expressions; however, as in (22b), this is not possible with the *how come*-construction.

---

In (ia), *why not* has an anaphoric function with a negative antecedent and it is paraphrased as *Why didn't Jasmine sleep?*; on the other hand, in (ib), the same expression has a rhetorical function and it is paraphrased as *There is no reason we shouldn't dance*. The former requires a negative linguistic antecedent whereas the latter can be licensed with a positive linguistic antecedent and even with a non-linguistic situational antecedent.

In addition to these syntactic differences, the two constructions show some different semantic/pragmatic properties as well, although both can trigger a causal interpretation. In particular, such semantic/pragmatic differences mostly arise due to their differences in factivity (Fitzpatrick 2005; Conroy 2006). That is, the *how come*-construction induces the presupposition of the truth of its dependent, while the *why*-construction does not. This contrast can be clearly seen in the following examples involving negative biased questions:

- (23) a. Why would John leave the party early?  
 b. \*How come John would leave the party early?

As pointed out by Collins (1991: 42-43) and Fitzpatrick (2005: 140), while examples as in (23a) introduced by *why* can yield a rhetorical question reading, examples as in (23b) introduced by *how come* cannot. For instance, the example in (23a) can be part of an answer to a question *Did John leave the party early?*, triggering a negatively-biased reading (i.e., John wouldn't leave the party early). This is possible because the *why*-construction does not presuppose the truth of its dependent. However, the example in (23b) cannot be used in the same context due to the conflict between the presupposition of the truth of the dependent of *how come* "John would leave the party early" and the negative bias reading of the rhetorical question "John wouldn't leave the party early".

This difference in factivity between the *why*-construction and the *how come*-construction can also account for the contrast observed in the examples below:

- (24) a. {Why/\*How come} go to the library?  
 b. {Why/\*How come} write a letter, when you could just call?  
 (25) {Why/\*How come} not do your homework now?  
 (26) a. Why don't we go out tonight?  
 b. \*How come we don't go out tonight?

Although *why* can occur with a tenseless base form VP, *how come* cannot (Collins 1991: 34-35; Conroy 2006: 5). Note here that the tenseless base form VP does not have a truth value and the *why*-construction examples in (24) and (25) evoke a non-factive reading. To be more specific, the examples in (24) with *why* convey a

negative suggestion, asking the addressee not to do the action denoted by the tenseless base form VP, meaning that there is no reason to go to the library and there is no reason to write a letter, respectively. The example in (25) with *why* gives rise to a non-factive, suggestive reading via a rhetorical question, meaning that the addressee should do their homework now (Huddleston and Pullum 2002: 834-835). Due to this characteristic of the tenseless base form VP in these environments and its clash with the factivity requirement of the *how come*-construction for its dependent, the ungrammaticality of examples as in (24) and (25) with *how come* is explained. In a similar manner, as in (26), whereas *why* can introduce a suggestion reading with a finite negative *do* verb, *how come* cannot (Conroy 2006: 5; Kim and Kim 2011: 7). The suggestion reading in (26a), of course, is a non-factive one. Since the *why*-construction can take a non-factive-denoting dependent whereas the *how come*-construction cannot, the contrast in grammaticality between (26a) and (26b) also naturally follows.

Factivity can explain the difference between the *why*-construction and the *how come*-construction with regard to the possibility to license NPIs as well (Fitzpatrick 2005: 140-141).

- (27) a. Why did John say anything?  
 b. \*How come John ever said anything?

As in (27a), the *why*-construction can license an NPI expression like *anything*; in contrast, as in (27b), the *how come*-construction cannot. Notice at this juncture that this kind of NPI constraint is seen in the dependent clause of factive predicates.

- (28) \*John didn't find out that John said anything.

The same behavior of the *how come*-construction and factive predicates in terms of the inability to occur with an NPI expression further confirms the factivity requirement of the *how come*-construction unlike the *why*-construction.

Aside from the factivity-related differences between the *why*-construction and the *how come*-construction, the two can give rise to different interpretations with stative predicates. To illustrate, consider the examples in (29) (Bromberger 1992; Tsai 2008):

- (29) a. Why is the sky blue?  
 b. How come the sky is blue?

The *why*-construction example in (29a) can be most naturally uttered when the speaker is not aware of any scientific explanation and she simply presupposes that the sky is blue, so she wants to know the reason why the sky is blue. On the other hand, the *how come*-construction example in (29b) can be most naturally uttered when it was cloudy just this morning and the speaker presupposes that the sky is blue and something caused the sky to become blue (i.e., the sky should not be blue), so she wants to know what caused the sky to become blue. A causal answer “Because the clouds just all blew away” is appropriate only for (29b), but not for (29a). In other words, with a stative predicate, we have a counter-expectation for the *how come*-construction, but not for the *why*-construction.

One more semantic difference between the two constructions has to do with the availability of long distance question interpretations. That is, in contrast to the *why*-construction, the *how come*-construction does not induce a long-distance question interpretation (Zwicky and Zwicky 1971: 178-179; Collins 1991: 33-34; Culicover 1999; Ochi 2004: 30; Duffield 2015: 62).

- (30) a. Why did John think Mary ate his pizza?  
 b. How come John thought Mary ate his pizza?

The example in (30a) is ambiguous in that *why* can be associated either with the matrix clause denoting the “thinking” event or with the embedded clause expressing the “eating” event. However, the example in (30b) is not ambiguous because *how come* is only associated with the matrix clause denoting the “thinking” event. In other words, the *why*-construction permits a long-distance question whereas the *how come*-construction does not.

Thus far, I have reviewed grammatical properties of the *how come*-construction noted in previous literature and we have seen that it behaves the same as the *why*-construction in some respects but it shows different behavior from the *why*-construction in several other respects. In what follows, I provide an explanation of how I collected children’s and their parents’ *how come*- and *why*-construction data and coded them for analysis.

#### 4. Data and methodology

The corpus this study made use of is composed of longitudinal data from the CHILDES database in the form of transcripts of utterances of recorded interactions between American English-speaking children and their main caregiver(s) (mother or/and father) which satisfy the following criteria (MacWhinney and Snow 1985, 1990; MacWhinney 2000):

- (31) a. A corpus must at least include data produced by the child from the age of 3 to nearly 5 years, which marks the point where most researchers argue that children begin to display linguistic behavior similar to that of adults.
- b. A corpus must include at least 15 tokens of *how come* produced by the child.
- c. A corpus must include utterances by main caregiver(s) of the child (i.e., their parent(s)). This is to find out parents' role in children's acquisition of the *how come*- and *why*-constructions and also to avoid children's simple repetitions of their parents' utterances.

Data from six children satisfied the criteria above: Abe (Kuczaj 1977), Adam (Brown 1973), Laura (Braunwald 1971), Nathaniel (MacWhinney and Snow 1990), Roman (Weist and Zevenbergen 2008), and Ross (MacWhinney 2000). All these six children had language acquisition within the normal range and none of them were identified with language-related problems. Further detailed information about these children can be found on the CHILDES database. A number of other children in ENG-NA (North America) on the CHILDES database were considered but excluded simply because they did not satisfy the criterion in (31b). Only very few instances of the *how come*-construction were found in the production data by each child (fewer than five) in most subcorpora and no instance of the construction was found in them. This suggests that in general children begin to use and acquire the *how come*-construction at later stages than typical *wh*-question constructions including those introduced by *why*.

Even though all the subcorpora used for the present study satisfy all the criteria in (31), they are different in certain respects. The following table shows each child's age range, the number of utterances, and the MLU by each child and their parents:<sup>5</sup>

Table 1. Basic information of the subcorpora from the CHILDES database used for the current study

Child's name	Age range (years;months)	Child's utterances	MLU for child's data	Parents' utterances	MLU for parents' data
Abe	2;4-4;1	31,485	5.61	25,145	5.40
Adam	2;3-4;10	44,484	3.90	19,945	5.19
Laura	1;2-7;0	19,420	2.80	23,852	4.45
Nathaniel	2;5-3;9	12,865	3.19	19,354	5.49
Roman	2;2-4;7	10,867	4.57	6,041	6.24
Ross	1;4-8;0	34,726	5.17	61,195	5.54
Total		153,787	4.38	155,532	5.32

As demonstrated in Table 1, the corpora vary with respect to the children's age ranges, data sizes, and MLUs (mean length of utterance) to some extent. In spite of such differences, it is worth using all the data here to get a broad picture of the linguistic constructions under question.

Within each of the subcorpora used for the current study, first I searched all the tokens of *how come* and *why* produced by each target child and their parents, using the COMBO program in the CLAN software for the CHILDES (MacWhinney 2000). These simple search options provided me with the token numbers given in Table 2:<sup>6</sup>

Table 2. Token numbers of *how come* and *why* produced by the six children and their parents in the CHILDES subcorpora

Child's name	<i>How come</i>	<i>Why</i>
Abe	204 (276)	611 (582)
Adam	21 (0)	916 (528)
Laura	21 (66)	109 (243)
Nathaniel	50 (8)	604 (161)
Roman	15 (10)	67 (182)
Ross	18 (153)	818 (1,172)
Total	329 (513)	3,125 (2,868)

5 Detailed information of the subcorpora from the CHILDES used for the present study is provided in the Appendices.

6 Throughout the paper, the numbers within parentheses in tables correspond to the frequencies of examples produced by the children's parents.

Then, I filtered out examples like the following:

- (32) a. MOT: How come?  
 CHI: **How come.**  
 MOT: What's the matter with your nose? (Laura, 01;10.19)
- b. MOT: Oh, you've got the right number of fingers, Nathaniel.  
 CHI: How come?  
 MOT: Lucky you.  
 CHI: **How come he xxx.**  
 MOT: Oh, he likes sand. (Nathaniel, 03;09.04)
- (33) a. MOT: You were up late. Not early. Late.  
 CHI: Why late? But why <Doodledoo was up> [/] Doodledoo was up early?  
 MOT: **Why what?** Who Cockledoodledoo?  
 CHI: Yeah.  
 MOT: (Be)cause he's a rooster and roosters all get up early. (Nathaniel, 03;04.18)
- b. FAT: <Marky was> real scared of Dracula last night, wasn't he?  
 CHI: **So why didn't he xxx.** (Ross, 06;04.22)
- c. CHI: It gives me energy to yell. Then I'll start yelling. Mommy, <**why do**> [/] how do bones grow?  
 MOT: The cells keep expanding. And then your bones get longer. And milk helps the process.

As shown here, all instances that involve a simple repetition of the child's own previous utterance and another speech participant's as in (32a) and (33a), those with parts marked as unclear (e.g., marked with a question mark or xxx in accordance with CHAT convention) as in (32b) and (33b), and those that are switched by another construction in the middle of an utterance as in (33c) were excluded from the analysis. This filter-out process left 315 instances of the *how come*-construction and 2,994 instances of the *why*-construction from the six children's data; it also left 508 instances of the *how come*-construction and 2,774 instances of the *why*-construction from their parents' data. In what follows, I provide a quantitative and qualitative analysis of the children's acquisition and usage patterns of the two constructions in comparison



with their parents' input about them based on the data from the CHILDES subcorpora.

## 5. Corpus findings

In this section, I first show general distributions of the identified *how come*- and *why*-construction examples produced by six English-speaking children and their parents based on dependent categories/construction types and their functions, irrespective of the children's age. I then demonstrate the corpus-based observations about their development patterns for each construction type in detail with respect to its subtypes and their functions.

### 5.1 General distributions of the *how come*- and *why*-constructions

First, I classified the identified *how come*- and *why*-construction examples on the basis of their dependent categories/construction types. Table 3 below shows the overall frequencies of the attested examples and some representative examples are presented in (34) and (35):

Table 3. Distributional patterns of *how come*- and *why*-construction examples produced by the six children and their parents in the CHILDES subcorpora based on dependent categories/construction types

Dependent category/ Construction type	<i>How come</i>	<i>Why</i>
Finite S	157 (172)	1,440 (1,885)
SC	1 (0)	184 (0)
Sluicing	150 (333)	1,097 (634)
Negative sluicing	0 (0)	175 (235)
Stripping	7 (3)	79 (12)
Negative stripping	0 (0)	19 (8)
Total	315 (508)	2,994 (2,774)

(34) a. CHI: What's in here? **How come it has a big hole in the middle?**

MOT: for the sound. (Adam, 05;02.12) (Finite S)

b. CHI: **How come there too many belts?**

FAT: How come there are too many belts?

CHI: Because you know the big one just can't get in there and the little one + ... (Ross, 03;03.15) (SC)

c. MOT: Hey, Abe, why don't you help me pick up your little cards?

CHI: **How come?**

MOT: Because it's so messy in here with all of those cards lying on the floor. (Abe, 02;11.25) (Sluicing)

d. CHI: Yeah, I was wrong. It sure take a long time xxx.

FAT: Hm:?

CHI: It sure take a long time.

FAT: Oh, good deal.

CHI: **How come good deal?**

FAT: Because if you take a long time, you're going to do a good job, right? (Abe, 03;02.07) (Stripping)

(35) a. CHI: **Why do they call him a tough cookie?**

FAT: Because he's a strong man. (Ross, 04;09.02) (Finite S)

b. CHI: What are you doing?

MOT: Sitting.

CHI: **Why you sitting?**

MOT: Waiting for you to finish your water. (Adam, 03;00.11) (SC)

c. MOT: Tell the bus to come back.

CHI: **Why?**

MOT: Oh, (be)cause I like your company in here. (Laura, 03;10.00) (Sluicing)

d. CHI: Does Sherry have a brother?

MOT: Oh no.

CHI: **Why not?** (Adam, 03;05.01) (Negative sluicing)

e. MOT: They're all too tight (be)cause he's got so much stuff on under there.

CHI: **Why so much stuff?**

MOT: I guess he thought it was pretty cold and he wanted to dress very warmly. (Nathaniel, 03;04.10) (Stripping)

f. CHI: Why does that boy have a motorcycle?

FAT: I don't know.

CHI: **Why not a car?**

FAT: Because his daddy doesn't have enough money to buy him anything else. (Ross, 03;08.18) (Negative stripping)

A quick glance at the results here reveals that with respect to dependent categories and construction types, the six children's uses of the *how come*-construction are both similar to and different from their uses of the *why*-construction and also that their uses of the two constructions are similar to and different from their parents' uses of the constructions. First, in the children's uses of the two constructions, the most frequent dependent category/construction type is finite sentence as in (34a) (157 instances) and (35a) (1,440 instances), followed by sluicing, where the *wh*-expressions take no dependent at all, as in (34c) (150 instances) and (35c) (1,097 instances). In the children's data, both constructions are rather infrequent in the stripping context as in (34d) (7 instances) and (35e) (79 instances) as well. As noted in (19), only the *why*-construction is possible in the negative sluicing and negative stripping contexts, but not the *how come*-construction, and this distinction is confirmed in the children's data here (i.e., no instance of the *how come*-construction in the negative sluicing and negative stripping contexts vs. 175 and 19 instances of the *why*-construction in the two negative ellipsis contexts, respectively). Another striking difference regarding the children's uses of the *how come*- and *why*-constructions is related to the SC dependent is that the SC dependent is quite frequent only with the *why*-construction as in (35b) (184 instances), but not with the *how come*-construction as in (34b) (1 instance).

As for the similarities between the children's uses of the *how come*-construction and their parents', the two most frequent dependent categories/construction types are finite S as in (34a) (157 instances in the children's data; 172 instances in the parents' data) and sluicing as in (34c) (150 instances in the children's data; 333 instances in the parents' data) and the construction is rarely found with an SC dependent as in (34b) (1 instance and no instance in the children's and parents' data) and in the stripping context as in (34d) (7 instances and 3 instances in the children's and parents' data). Nonetheless, the most frequent dependent category/construction type is finite S in the children's *how come*-construction data while it is sluicing in their parents' data. Statistical analysis using Fisher's Exact test shows that their distributions in these two contexts are significantly different ( $p < 0.0001$ ).

Concerning the similarities between the children's uses of the *why*-construction

and their parents', the two most frequent dependent categories/construction types are finite S as in (35a) (1,440 instances in the children's data; 1,885 instances in the parents' data) and sluicing as in (35c) (1,097 instances in the children's data; 634 instances in the parents' data). Negative sluicing as in (35d) is less frequently found in their data (175 instances in the children's data; 235 instances in the parents' data). Meanwhile, in their data, the *why*-construction is rather infrequent in the stripping (79 instances in the children's data; 12 instances in the parents' data) and negative stripping contexts (19 instances in the children's data; 8 instances in the parents' data) as in (35e) and (35f). One prominent difference is that the *why*-construction with an SC dependent as in (35b) is found only in the children's data, but not in their parents' data (184 instances vs. no instance).

Some notes about the identified *how come*- and *why*-construction data should be made at this point before moving on to further discussion about them. For instance, the examples taking a finite S dependent can be further classified, as shown in the following:

- (36) a. MOT: It's gone.  
 CHI: **How come it's gone?** How come?  
 MOT: Well, we gave it to Mwww [% friend]. (Laura, 02;07.10)
- b. CHI: I got my leg up. **That's how come you need to be out of the way.**  
 FAT: Why do I need to be out of the way? (Abe, 03;03.28)
- c. URS: Remember to put de [: the] cover on.  
 CHI: **How come is dat [: that]?**  
 URS: Because it get dry. (Adam, 04;10.02)
- (37) a. CHI: **Why did the chicken cross the road?**  
 MAR: Because the light turned green. (Ross, 05;10.01)
- b. CHI: **Do you know why I'm mad?**  
 FAT: Why?  
 CHI: Because my back hurts. (Ross, 03;06.16)
- c. URS: David made it.  
 CHI: **Why David made it for me?**  
 MOT: He didn't make it for you. (Adam, 02;11.28)
- d. MOT: He doesn't have a ball.

CHI: **Why not he have ball?**

MOT: I think it came off. (Adam, 02;11.28)

e. MOT: It went down the wrong hole did it?

CHI: **Why went down the wrong hole?**

MOT: (Be)cause I think that you were trying to breathe while you were trying to drink. (Nathaniel, 03;00.21)

As discussed at the beginning of this paper, the *how come*- and *why*-constructions differ in terms of SAI in matrix environments, and thus the examples in (36a) and (37a) are judged as well-formed. This distinction between the two constructions is no longer available in embedded environments; hence, the grammaticality of examples as in (36b) and (37b). However, in the matrix environment, the child's utterance of the *how come*-construction example in (36c) involves SAI while the *why*-construction example produced by the child in (37c) does not. In (37d), the intended use of the *why*-construction example uttered by the child is to mean *Why doesn't he have a ball?*, but he did not use the finite auxiliary verb. In (37e), the child's utterance of the *why*-construction example is missing the subject (i.e., *it*) without SAI. Examples like (36c), (37c), (37d), and (37e) are then taken to be unacceptable in standard English. Nevertheless, they are found in the identified examples from the CHILDES subcorpora and used for analysis. In particular, the frequencies of these underdeveloped or erroneous *how come*- and *why*-construction examples are as follows: 4 instances of the *how come*-construction with an inverted finite S dependent as in (36a) produced by the children; 644 instances of the *why*-construction with an incorrect use of a finite S dependent as in (37c)-(37e) uttered by the children vs. 12 instances of the *why*-construction with an incorrect use of a finite S dependent as in (37c)-(37e) uttered by their parents. This shows that overall these underdeveloped or erroneous examples are much more prevalent for the *why*-construction than for the *how come*-construction and most of them are from the children's data rather than their parents'. Notably, out of 1,440 *why*-construction examples with a finite S dependent produced by the children, 644 contain such errors (44.7%); in contrast, only 12 out of 1,885 *why*-construction examples with a finite S dependent produced by their parents involve such erroneous uses (0.6%). The distribution patterns of the *why*-construction examples with a finite S dependent by the children and their parents based on their correct and erroneous uses are statistically different ( $p < 0.0001$ ) and this clearly

indicates that young children undergo serious difficulty in using the *why*-construction with a finite S dependent in the right manner.

I then examined the attested corpus examples based on illocutionary functions. The two constructions can be used as pure inquiries; however, as noted earlier in (24)-(26), the *why*-construction can be used to make suggestions as well, but not the *how-come* construction, with a VP[base] dependent in the positive and negative stripping contexts and also with a finite S dependent containing a finite negative *do* verb. As expected, no *how come*-construction example from the CHILDES subcorpora has a suggestion function; on the other hand, certain numbers of identified *why*-construction examples involve a suggestion function. First, consider the following examples:

(38) a. CHI: Cheese.

FAT: You have a beautiful piece of cheese right on your table.

CHI: **Well, why don't I have some more toast please?**

FAT: You want this one?

CHI: Yes, please. (Laura, 02;07.08)

b. MOT: **Why don't you put that on the table?**

CHI: Okay. (Adam, 02;07.14)

As shown in these examples, the *why*-construction can be used with a suggestion function when it contains a first- or second-person pronoun and a finite negative *do* verb together.<sup>7</sup> The suggestion function is more frequently found in the parents'

<sup>7</sup> This does not mean that every instance of the *why*-construction with a first- or second-person pronoun and a finite negative *do* verb together is always used to make a suggestion.

(i) a. CHI: Where is Pest?

FAT: Pest is on the other side of the river.

CHI: **Well, why don't we live in Pest?**

FAT: Because we live in Buda.

CHI: I like Buda best. (Ross, 03;06.25)

b. CHI: I doesn't know how to put this together.

MOT: **Why don't you know how?**

CHI: I don't know. She is making a dessert, Mommy. (Adam, 03;08.01)

Examples of the *why*-construction like these are used with a pure inquiry function, although the subject is a first- or second-person pronoun and the negative *do* verb is involved. When these examples are used to make pure questions, the main verb typically describes a state rather than an action.

data than the children's. Regarding the correct *why*-construction with a finite S dependent in the matrix environment, among 637 instances produced by the children 66 have a suggestion function (10.4%) while among 1,636 instances produced by their parents 493 involve a suggestion function (30.1%). Their distributional difference here is indeed statistically significant ( $p < 0.0001$ ) and this indicates that although the pure inquiry function is more salient than the suggestion function for the *why*-construction with a finite S dependent in the matrix environment both in the children's and their parents' data, the suggestion function is significantly more prominent in the parents' data than in the children's data.

As for *why*-construction examples with a VP[base] dependent in the positive and negative tripping contexts, pure inquiry and suggestion functions are found as demonstrated in (39):

(39) a. MOT: Will you get off the cake Squeaky and sit at the table if you want a piece of cake.

CHI: **Why sit at the table havin(g) piece of cake?**

MOT: (Be)cause that's the proper way to eat cake. (Nathaniel, 03;04.09)  
(Positive *why*-stripping with a VP[base] dependent and a pure inquiry function)

b. MOT: We only need the lights on at night.

CHI: **Why not keep a light on?**

MOT: Because we have the sun. (Adam, 02;10.30) (Negative *why*-stripping with a VP[base] dependent and a pure inquiry function)

c. CHI: It's so hard. I can't.

URS: No?

CHI: No.

MOT: Here's a little hole. **Why not open it from there?** (Adam, 03;04.18) (Negative *why*-stripping with a VP[base] dependent and a suggestion function)

In the parents' positive *why*-stripping construction examples, none takes a VP[base] dependent and all the children's positive *why*-stripping construction examples have a pure inquiry function (29 instances). In their negative *why*-stripping construction

examples, 10 take a VP[base] dependent in the children's data and 6 contain a VP[base] dependent in the parents' data. In the children's negative *why*-stripping construction examples with a VP[base] dependent, 6 involve a pure inquiry function while 4 have a suggestion function. In the parents' negative *why*-stripping construction examples with a VP[base] dependent, 3 have pure inquiry function and the remaining 3 involve a suggestion function. These results show that the positive *why*-stripping construction with a VP[base] dependent is found in the children's data, but not in the parents' data, and the pure inquiry function is dominant rather than the suggestion function. They also demonstrate that the negative *why*-stripping construction with a VP[base] dependent is found in both the children's data and their parents' and that their illocutionary functional distributions are similar.

## 5.2 Overall acquisition patterns of the *how come*- and *why*-constructions

Let us now consider the overall frequencies of the attested *how come*- and *why*-construction examples produced by the six children and their parents based on the children's age periods in Table 4 below:

Table 4. Distributions of *how come*- and *why*-construction examples produced by the six children and their parents in the CHILDES subcorpora based on children's age periods

Age period	<i>How come</i>	<i>Why</i>	Age period	<i>How come</i>	<i>Why</i>
1;00-1;05	0 (2)	3 (12)	4;06-4;11	32 (42)	169 (184)
1;06-1;11	0 (7)	7 (42)	5;00-5;05	3 (8)	59 (80)
2;00-2;05	2 (26)	33 (100)	5;06-5;11	0 (11)	50 (106)
2;06-2;11	29 (81)	587 (528)	6;00-6;05	0 (30)	33 (115)
3;00-3;05	61 (114)	1,273 (614)	6;06-6;11	6 (16)	34 (118)
3;06-3;11	177 (102)	489 (461)	7;00-7;05	1 (9)	10 (61)
4;00-4;05	4 (60)	237 (331)	7;06-7;11	0 (0)	10 (22)
			Total	315 (508)	2,994 (2,774)

As can be seen here, the *why*-construction is not only more frequently used than the *how come*-construction by the six children, but the former is also found earlier than the latter. In each age period, the frequency of the children's *why*-construction examples is higher than that of their *how come*-construction examples. In addition, all the six children produced the first occurrence of the *why*-construction earlier than that of the *how come*-construction. The time span between the first occurrences of



the two constructions varies among the six children, from about 5 months to 2 years and 1 month. The children's data here then show that in general young English-speaking children start to use and acquire the *how come*-construction later than the *why*-construction and they use the *why*-construction more frequently than the *how come*-construction.

Observe then that overall the frequency distributions of the two constructions produced by the six children based on their age periods display similar behavior to those produced by their parents. Generally, when the frequency of the *how come*- and *why*-construction examples uttered by the children is relatively high, the frequency of their parents' counterparts is also high (2;06-3;11 for the *how come*-construction and 2;06-4;11 for the *why*-construction, in particular). This implies that parents' input influences young children's uses and acquisition of the *how come*- and *why*-constructions to some extent.

Note also that the frequency of the children's *how come*-construction examples is lower than that of their parents' in each age period, except one age period of 3;06-3;11; similarly, in general, the frequency of the children's *why*-construction examples is lower than that of their parents', with the exception of 2;06-3;11. This suggests that young children acquire the two constructions in a similar way, with their parents' input playing a certain role.

### **5.3 Acquisition patterns of the *how come*-construction by dependent category/construction types**

I then checked the uses and acquisition patterns of the *how come*-construction by dependent category/construction types. The distributions of the identified *how come*-construction examples depending on the children's age periods and the dependent category/construction types are given in the following table:

Table 5. Distributions of *how come*-construction examples produced by the six children and their parents in the CHILDES subcorpora based on children's age periods and dependent category/construction types

Age period	Finite S	SC	Sluicing	Stripping	Total
1;00-1;05	0 (2)	0 (0)	0 (0)	0 (0)	0 (2)
1;06-1;11	0 (5)	0 (0)	0 (2)	0 (0)	0 (7)
2;00-2;05	0 (11)	0 (0)	1 (15)	1 (0)	2 (26)
2;06-2;11	5 (19)	0 (0)	24 (62)	0 (0)	29 (81)
3;00-3;05	5 (27)	1 (0)	53 (85)	2 (2)	61 (114)
3;06-3;11	110 (35)	0 (0)	63 (66)	4 (1)	177 (102)
4;00-4;05	4 (15)	0 (0)	0 (45)	0 (0)	4 (60)
4;06-4;11	26 (16)	0 (0)	6 (26)	0 (0)	32 (42)
5;00-5;05	1 (2)	0 (0)	2 (6)	0 (0)	3 (8)
5;06-5;11	0 (8)	0 (0)	0 (3)	0 (0)	0 (11)
6;00-6;05	0 (18)	0 (0)	0 (12)	0 (0)	0 (30)
6;06-6;11	5 (8)	0 (0)	1 (8)	0 (0)	6 (16)
7;00-7;05	1 (6)	0 (0)	0 (3)	0 (0)	1 (9)
Total	157 (172)	1 (0)	150 (333)	7 (3)	315 (508)

As noted in Table 3 above and as shown here again, overall, in the children's uses of the *how come*-construction, the two most frequent dependent category/construction types are finite S and sluicing, with similar frequencies. The results in Table 5 show that in each age period up to the age of 3 years and 5 months the frequency of the *how come*-construction is higher in the sluicing context than with a finite S dependent; on the other hand, the reverse tendency is the case since then. One may then assume that children acquire the *how come*-construction earlier in the sluicing context than with a finite S dependent. A closer look at the data indeed reveals that it is generally the case. For four of the six children, the *how come*-construction is found earlier in the sluicing context than with a finite S dependent. For one of them, the opposite pattern is found and for the remaining child, the two dependent category/construction types are found in the same year;month.day dataset. Their first utterances of the *how come*-construction in these two contexts are given below:

(40) a. CHI: Some pomegranates are spicy.

MOT: Gosh, this is really juicy.

CHI: **How come?**

MOT: Because it must be just right. (Abe, 02;11.18)

- b. URS: She goes to sleep all day while we're not there.  
 CHI: All day. **How come?** She's not (sup)posed to go to sleep all day.  
 MOT: She doesn't want to be tired when they do get home. (Adam, 04;09.02)
- c. ADU: No, I'm not making my Mummy's name.  
 CHI: **How come?**  
 ADU: (Be)cause I don't want to. (Laura, 02;05.01)
- d. MOT: Well, how come you can't get any puzzles out?  
 CHI: Well, **how come?**  
 MOT: How come?  
 CHI: I will get other all so much other toys. (Nathaniel, 03;04.10)
- e. CAR: My doggy just died. He makes me sad.  
 CHI: **Daddy, how come?** How come he died? (Roman, 04;06.11)
- f. CHI: You don't want to eat dinner? Hmhm. **How come?** Because.  
 Mhm. Okay, we'll turn this off. (Ross, 03;03.27)
- (41) a. GRA: Why don't you use another body? Then you can put shoes on.  
 CHI: **How come Grandma you call my daddy Sonny?**  
 GRA: Oh, that's his name to me. (Abe, 02;11.30)
- b. MOT: How about this?  
 CHI: Ahhah, so you brought dis [: this] one back.  
 URS: That's a new one.  
 CHI: <How do you get> [/]. **How come you got it?** Can't open it.  
 URS: I'll start it for you. (Adam, 04;04.01)
- c. MOT: Why do I have stockings on? (Be)cause it's cold today.  
 CHI: And Swwww [% Mother] **how come you have Slacks too?**  
 MOT: How come I have Slacks over the stocking?  
 CHI: Yeah.  
 MOT: (Be)cause I was cold. (Laura, 02;06.24)
- d. MOT: That's a strange thing. Stuck on a road.  
 CHI: **How come dat's [: that's] on a road?** (Nathaniel, 03;09.02)
- e. FAT: Hm:?  
 CHI: Um, **how come he can be good if he had black eyes? How come he had black? How come he smelled blood from Dori?** Um <and> [/] and then um his eyes turned that color?

FAT: Does that mean he's mean? (Roman, 04;06.11)

- f. CHI: **How come people don't go to preschool when cartoons come on [= on Saturday]?** (Ross, 03;09.01)

The time difference between the first uses of the *how come*-construction in the sluicing context and the one with a finite S dependent varies depending on the child but it is shorter than 5 months. Therefore, the results in Table 5 and the examples in (40)-(41) indicate that in general young English-speaking children use and acquire the *how come*-construction a little earlier in the sluicing context than with a finite S dependent.

Notice also that in the children's data, the *how come*-construction is scarce with an SC dependent and in the stripping context. In their data, only one instance of the *how come*-construction is found with an SC dependent as already given in (34b) and seven instances of the construction are observed in the stripping context including the one in (34d). Although *how come*-construction examples with a finite S dependent and in the sluicing context are found for all the six children, those with an SC dependent and in the stripping context are found only for some of them (i.e., one child for the *how come*-construction with an SC dependent and three children for the one in the stripping context). This along with the results in Table 5 then indicates that young English-speaking children use and acquire the *how come*-construction later with an SC dependent and in the stripping context than with a finite S dependent and in the sluicing context.

Next, in the parents' data the frequency of the *how come*-construction is higher in the sluicing context than with a finite S dependent up to the children's age of 5;05 and since then the opposite is true. This seems to imply that parents tend to use the *how come*-construction more preferably in the sluicing context than with a finite S dependent when their children are at the early acquisition stages and as the children grow older, they tend to use the construction more favorably with a finite S dependent than in the sluicing context.

Furthermore, in general, parents' input seems to play a certain role in children's uses and acquisition patterns of the *how come*-construction. Overall, the frequency tendency of the children's *how come*-construction examples correlates with that of their parents' depending on the dependent category/construction types. For instance, the frequencies of the children's *how come*-construction examples with an SC dependent and in the stripping context are very low just as those of their parents'.

When the frequency of the children's *how come*-construction examples with a finite S dependent and in the sluicing context is high, that of their parents' is also high; on the other hand, when the frequency of their *how come*-construction in the two environments is low, that of their parents' is low as well. Nonetheless, there are times when the frequency of the children's *how come*-construction examples does not correlate with that of their parents' when they take a finite S dependent and they occur in the sluicing context. In particular, in each six-month age period between 3;06 and 4;11, the frequency of the children's *how come*-construction examples is higher with a finite S dependent than in the sluicing context, but the reverse pattern is found in their parents' data. These together show that although parents' input generally plays an important role in children's uses and acquisition patterns of the *how come*-construction, it does not play an absolute role.

#### 5.4 Acquisition patterns of the *why*-construction by dependent category/construction types

I also looked into the uses and acquisition patterns of the *why*-construction by dependent category/construction types. The distributions of the observed *why*-construction examples based on the children's age periods and the dependent category/construction types are presented in Table 6 below:

Table 6. Distributions of *why*-construction examples produced by the six children and their parents in the CHILDES subcorpora based on children's age periods and dependent category/construction types

Age period	Finite S	SC	Sluicing	Neg sluicing	Stripping	Neg stripping	Total
1;00-1;05	1 (12)	0 (0)	2 (0)	0 (0)	0 (0)	0 (0)	3 (12)
1;06-1;11	4 (31)	0 (0)	2 (5)	0 (5)	1 (1)	0 (0)	7 (42)
2;00-2;05	4 (84)	0 (0)	16 (5)	9 (11)	3 (0)	1 (0)	33 (100)
2;06-2;11	221 (342)	69 (0)	190 (120)	82 (64)	19 (2)	6 (0)	587 (528)
3;00-3;05	502 (415)	73 (0)	585 (144)	65 (50)	45 (4)	3 (1)	1,273 (614)
3;06-3;11	305 (320)	28 (0)	142 (103)	6 (34)	5 (2)	3 (2)	489 (461)
4;00-4;05	162 (232)	8 (0)	52 (81)	8 (18)	5 (0)	2 (0)	237 (331)
4;06-4;11	118 (114)	8 (0)	45 (55)	1 (13)	1 (0)	0 (2)	169 (184)
5;00-5;05	41 (59)	4 (0)	16 (16)	0 (5)	0 (0)	0 (0)	59 (80)
5;06-5;11	36 (69)	2 (0)	14 (28)	0 (7)	0 (1)	0 (1)	50 (106)

6;00-6;05	20 (75)	0 (0)	12 (29)	0 (11)	0 (0)	1 (0)	33 (116)
6;06-6;11	19 (84)	0 (0)	12 (29)	2 (4)	0 (1)	1 (0)	34 (118)
7;00-7;05	4 (34)	0 (0)	3 (15)	1 (9)	0 (1)	2 (2)	10 (61)
7;06-7;11	3 (14)	0 (0)	6 (4)	1 (4)	0 (0)	0 (0)	10 (22)
Total	1,440 (1,885)	184 (0)	1,097 (634)	175 (235)	79 (12)	19 (8)	2,994 (2,774)

As discussed in Table 3 above and as can be seen here once again, in general, in the children's uses of the *why*-construction, the two most frequent dependent category/construction types are finite S and sluicing. The results show that up to the age of 2;05, the frequency of the *why*-construction is higher in the sluicing context than with a finite S dependent; however, since then, the reverse pattern is true. One may naturally assume that children acquire the *why*-construction earlier in the sluicing context than with a finite S dependent. However, as noted earlier in (37), the children's data contain numerous underdeveloped or erroneous *why*-construction examples with a finite S dependent. Thus, I checked the correct and incorrect uses of the children's *why*-construction examples depending on the children's age periods, whose distributions are given in the following table:

Table 7. Distributions of *why*-construction examples with a finite S dependent produced by the six children and their parents in the CHILDES subcorpora based on children's age periods and their correct and incorrect uses

Age period	Correct	Incorrect	Total
1;00-1;05	0 (12)	1 (0)	1 (12)
1;06-1;11	3 (31)	1 (0)	4 (31)
2;00-2;05	4 (84)	0 (0)	4 (84)
2;06-2;11	36 (341)	185 (1)	221 (342)
3;00-3;05	187 (408)	315 (1)	502 (415)
3;06-3;11	196 (319)	109 (1)	305 (320)
4;00-4;05	139 (231)	23 (1)	162 (232)
4;06-4;11	111 (114)	7 (0)	118 (114)
5;00-5;05	40 (59)	1 (0)	41 (59)
5;06-5;11	34 (68)	2 (1)	36 (69)
6;00-6;05	20 (75)	0 (0)	20 (75)
6;06-6;11	19 (83)	0 (1)	19 (84)
7;00-7;05	4 (34)	0 (0)	4 (34)
7;06-7;11	3 (14)	0 (0)	3 (14)
Total	796 (1,873)	644 (12)	1,440 (1885)

As demonstrated here, if we disregard the very few instances up to the age of 2;05, the children's correct uses of the *why*-construction with a finite S dependent begin to outnumber their incorrect ones from the age of 3;06 and since the age of 4;06 the correct uses account for more than 90%. The results in Tables 6 and 7 together then imply that children acquire the *why*-construction earlier in the sluicing context than with a finite S dependent. The children's first occurrences of the *why*-construction in these two environments further confirm that they generally acquire the *why*-construction earlier in the sluicing context than with a finite S dependent. The following are the first utterances of the *why*-construction in the two environments produced by the children:<sup>8</sup>

- (42) a. CHI: Jean and Jean and Jean and Sherry's daddy?  
 FAT: Uhhuh.  
 CHI: **Why?**  
 FAT: Probably because the hall is dirty. (Abe, 02;07.00)
- b. MOT: Sit here on the side if you wanna.  
 CHI: Side wanna. **Why?**  
 MOT: No, down there. (Adam, 02;04.30)
- c. MOT: Hi.  
 CHI: **Why?**  
 MOT: Hi, Laura.  
 CHI: Hi. (Laura, 01;05.13)
- d. MOT: Nathaniel sings along in this book.  
 CHI: **Why?**  
 FAT: Okay. Get up here and sing along. (Nathaniel, 02;06.16)
- e. CAT: That doesn't belong there.  
 CHI: **Why?**  
 CAT: Why, because that's a refrigerator. (Roman, 02;02.20)
- f. FAT: No walk now.  
 CHI: **Why?**  
 FAT: We are going to do this first. (Ross, 02;04.09)
- (43) a. CHI: I wanna watch TV.

---

<sup>8</sup> The examples in (43) are the children's first occurrences of the *why*-construction with a finite S dependent in the matrix environment involving SAI.

FAT: There's nothing on TV.

CHI: **Why don't we play a game, okay?**

FAT: Let's play hide and seek. (Abe, 02;06.06)

b. CHI: What kind of box is this?

MOT: Alphabet basket.

CHI: Alphabet box? **Why d(o) you carry it by de [: the] handle?** (Adam, 03;03.18)

c. CHI: Shout. Shout [/] shout. **Why are you shouting?** (Laura, 01;11.13)

d. CHI: Then something went bump. How that bump made us jump. And look, eh, eh, eh, it was the cat in the hat. **Why would you sit there like that?** (Nathaniel, 03;04.08)

e. JEN: Ready?

CHI: **Why don't you show me it?**

JEN: What's those?

CHI: I don't know. (Roman, 02;05.21)

f. FAT: Are the kids running?

CHI: Yeah, the kids run. **Why are we going home?**

FAT: Because Marky hasta go to sleep. (Ross, 02;07.18)

As illustrated here, except for one child, the *why*-construction is found earlier in the sluicing context than with a finite S dependent, in particular, in the matrix environment with SAI, suggesting the earlier acquisition and correct uses of the *why*-construction in the sluicing context than with a finite S dependent. Note at this juncture that these two syntactic environments are where the children receive the most input by their parents regarding the *why*-construction. Interestingly, the parents' input is much higher for the finite S dependent (1,885 instances) than for the sluicing context (634 instances). This then suggests that young children have difficulty mastering the *why*-construction with a finite S dependent, which requires SAI in the matrix environment, further indicating that although parents' input plays some role in children's acquisition of the *why*-construction, it is not a sole determining factor.<sup>9</sup>

---

<sup>9</sup> A reviewer points out that if young children acquire SAI in the *why*-construction at later developmental stages, it is predicted that their correct uses with SAI should have high frequencies in the later acquisition periods; however, the frequencies of their correct uses with SAI indeed decrease in Table 7. One may find this puzzling, but it is due to the organization of the data used for the present study.



As noted above, the *why*-construction examples with an inverted finite S dependent can have either a pure inquiry function or a suggestion function. Then, now consider the distribution patterns of the *why*-construction examples with an inverted finite S dependent based on the children's age periods and illocutionary functions in Table 8 below:

Table 8. Distributions of *why*-construction examples with an inverted finite S dependent produced by the six children and their parents in the CHILDES subcorpora based on children's age periods and illocutionary functions

Age period	Pure inquiry	Suggestion	Total
1;00-1;05	0 (5)	0 (6)	0 (11)
1;06-1;11	3 (20)	0 (7)	3 (27)
2;00-2;05	3 (39)	1 (40)	4 (79)
2;06-2;11	21 (181)	4 (121)	25 (302)
3;00-3;05	143 (245)	6 (115)	149 (360)
3;06-3;11	165 (198)	7 (85)	172 (283)
4;00-4;05	77 (159)	15 (42)	92 (201)
4;06-4;11	77 (63)	16 (35)	93 (98)
5;00-5;05	29 (42)	5 (11)	34 (53)
5;06-5;11	28 (54)	4 (7)	32 (61)
6;00-6;05	13 (51)	3 (8)	16 (59)
6;06-6;11	8 (55)	3 (9)	11 (64)
7;00-7;05	2 (23)	1 (4)	3 (27)
7;06-7;11	2 (8)	1 (3)	3 (11)
Total	571 (1,143)	66 (493)	637 (1,636)

In the children's data, up to the age of 3;11, in general, the *why*-construction with an inverted finite S dependent is predominantly used with an inquiry function but since the age of 4, the frequency difference between those with an inquiry function and the ones with a suggestion function is much smaller. Statistical analysis with Fisher's Exact test shows that their distributions based on illocutionary functions indeed are significantly different ( $p < 0.0001$ ), indicating that although the inquiry function is always preferred over the suggestion function in the children's uses of the *why*-construction with an inverted finite S dependent, it is less stronger since

---

The numbers of children's utterances between the age periods of 2;00 and 4;11 are significantly higher than those after the age of 5;00, as can be seen in Appendix A. That is why I do not attempt to compare between simple raw frequencies for the grammatical properties discussed in this paper; instead, I discuss distribution patterns for them based on Fisher's Exact test results.

the age of 4. Interestingly, the reverse pattern is found in their parents' data. That is, the illocutionary functional difference of their parents' *why*-construction examples is smaller until the children's age of 3;11 than since their age of 4 and their difference is also statistically attested ( $p < 0.0001$ ). The results here, therefore, mean that children in general use and acquire the *why*-construction with an inverted finite S dependent earlier with an inquiry function than with a suggestion function and in doing so, parents' input does not play a crucial role in determining children's uses and acquisition of the *why*-construction.

Next, other relatively frequent dependent category/construction types found in the *why*-construction produced by the children are small clause (SC) and negative sluicing. The children's relatively frequent uses of the *why*-construction with an SC dependent are surprising, considering no input by their parents. Again, this means that there is something else in accounting for children's uses and acquisition of the *why*-construction aside from parents' input. Then, observe the children's first uses of the *why*-construction with an SC dependent in the following:<sup>10</sup>

- (44) a. FAT: I sure am.  
       CHI: **Why you doing work at home?**  
       FAT: Why am I? (Abe, 02;07.00)
- b. MOT: I guess I'm not looking in the right place.  
       CHI: **Why not you looking right place?** (Adam, 02;01.30)
- c. CHI: **Why he sweepin(g)?**  
       MOT: He's not sweeping. He's riding the broomstick. (Nathaniel, 03;02.27)
- d. JEN: He's in the hole huh ?  
       CHI: Yeah.  
       JEN: Yeah.  
       CHI: **Hey, why him getting black?**  
       JEN: (Be)cause you know what this is? (Roman, 02;06.11)
- e. CHI: **Why you putting on your bed clothes on?** (Ross, 03;01.05)

The children's age at which they begin to use the *why*-construction with an SC dependent

---

<sup>10</sup> In Laura's dataset, no *why*-construction example with an SC dependent is found.

varies. It is somewhere between 2;01.30 and 3;02.27 and it is not found after the age period of 5;06-5;11. This suggests that they generally use the *why*-construction with an SC dependent, because they do not master the uses of *be* by this age. In other words, since they do not fully acquire how to use *be* properly, they seem to use the *why*-construction with an SC dependent instead of a finite S dependent.

As for the negative *why*-sluicing construction, overall it is much less frequently used than its positive counterpart by the children (1,097 instances vs. 175 instances), suggesting that young children acquire the negative *why*-sluicing construction later than the positive one. A couple of first occurrences of the negative *why*-sluicing construction produced by the children are given in (45):

(45) a. MOT: Don't take those out. Leave them in there.

CHI: **Why not?**

MOT: Why not? Because they're Daddy's. (Adam, 02;03.04)

b. JEN: Roman, don't ever put bees in your mouth, okay?

CHI: **Why not?** (Roman, 02;10.15)

The data indeed reveal that for five out of the six children the negative *why*-sluicing construction is found later than its positive counterpart, confirming the later uses and acquisition of the former than the latter.

Another difference between the negative and positive *why*-sluicing constructions produced by the children concerns their frequency relative to their parents' input. As discussed above, in general, the frequency of the positive *why*-sluicing construction examples is higher for the children than for their parents; on the other hand, the opposite is the case for the negative *why*-sluicing construction. Nonetheless, the two types of sluicing constructions show similar behavior in that the children's frequency is similar to or higher than their parents' in 2;00-3;05.

One more note to make concerning the identified negative *why*-sluicing construction examples has to do with its rhetorical use. The negative *why*-sluicing construction can have a rhetorical function, not an inquiry function, in particular, when it has a positive antecedent. In the children's data, no negative *why*-sluicing construction example is found with a rhetorical function. Even in their parents' data, only 10 negative *why*-sluicing construction examples are found to have a rhetorical function, some of which are presented in (46):

- (46) a. FAT: I work in the other part.  
 CHI: Why do you work there?  
 FAT: **Why not?** (Ross, 03;08.05)
- b. FAT: We should put them up, right?  
 MAR: Well we hafta hurry.  
 FAT: **Why not, right?**  
 MAR: Yeah.  
 FAT: Right.  
 MAR: We should hurry. (Ross, 06;02.22)

All these 10 negative *why*-sluicing examples come from the parents of one child, Ross. Then, the results here suggest that young children predominantly use the negative *why*-sluicing construction with an inquiry function and they use and acquire the construction with a rhetorical function at later acquisition stages and that parents' scarce input seems to account for this usage and acquisition fact.

The two remaining infrequent dependent category/construction types for the *why*-construction produced by the six children are positive and negative stripping constructions. In the children's data, the *why*-stripping construction is more frequent in the positive context than in the negative one, although the construction is in general much less frequent than the *why*-construction with other dependent category/construction types. The results in Table 6 also seem to demonstrate that children use and acquire the positive *why*-stripping construction earlier than its negative counterpart and this is indeed the case as can be seen in the following examples, which are the first occurrences of the positive and negative *why*-stripping constructions produced by the children:

- (47) a. MOT: No, soy sauce is for rice.  
 CHI: Soy sauce?  
 MOT: Uh-huh.  
 CHI: **Why soy sauce for rice?**  
 MOT: Because of its unusual flavor. (Abe, 02;07.07)
- b. MOT: Try it now.  
 CHI: What dat [: that] come from. **Why dat [: that]?**  
 MOT: I just closed the door. (Adam, 02;06.03)

- c. FAT: No, don't spill it. Please don't spill it.  
 CHI: **Why hot?** (Laura, 01;11.24)
- d. MOT: No peanuts in it.  
 CHI: **Why no peanuts in it?**  
 MOT: **(Be)cause it's not peanut butter.** (Nathaniel, 03;00.01)
- e. FAT: Is that what froggies do?  
 CHI: Yes.  
 JEN: Huh.  
 CHI: **That why go like this, ribbit, ribbit, then I have a green mouth.**  
 JEN: Then you have a green mouth? (Roman, 02;10.01)
- f. CHI: **Why folk dancing there?**  
 FAT: Well, because we live here honey. (Ross, 03;02.21)
- (48) a. MOT: No milk, just a dish with blueberries.  
 CHI: **Why not a big dish? Why not a big dish, Mom?**  
 MOT: I don't know. Just because I always give you one of these tiny bowls. (Abe, 02;08.06)
- b. MOT: Oh, no, you don't turn anything.  
 CHI: Turn? Why not? **Why not turn?** (Adam, 02;04.15)
- c. MOT: Which glass are you going to drink the malt out of? That one?  
 (Be)cause it has more? The tall thin cup has more?  
 CHI: **Why not this?** (Laura, 04;03.11)
- d. CHI: **Why / why little? Why not big?**  
 REN: (Be)cause it's special. (Nathaniel, 03;00.21)
- e. FAT: Just this one, the big one.  
 CHI: **Why not the little one?**  
 FAT: It won't work. (Ross, 03;06.07)

The positive *why*-stripping construction is found earlier than its negative counterpart for five out of the six children with the exception of Adam.<sup>11</sup>

In addition, a wider variety of syntactic categories are found as the dependent in the positive *why*-stripping construction than in the negative one.

---

11 In one child's data, no negative *why*-stripping construction example is observed.

- (49) a. CHI: **Why broken?** Huh?  
 MOT: I guess it's because you weren't playing with it gently enough.  
 (AdjP) (Adam, 03;04.01)
- b. MOT: You were up late. Not early. Late.  
 CHI: **Why late?** But why Doodledoo was up / Doodledoo was up early?  
 MOT: Why what? (AdvP) (Nathaniel, 03;04.18)
- c. CHI: A marble bag for what?  
 URS: For marbles.  
 CHI: For marbles. **Why a marble bag?**  
 URS: It would be good to carry tiny cars. (NP) (Adam, 03;02.21)
- d. MOT: That's as far as it'll go honey.  
 CHI: Why?  
 MOT: (Be)cause there's something else there. Okay.  
 CHI: **Why (be)cause something else there?** (Sub-clause) (Nathaniel, 03;00.21)
- e. MOT: Will you get off the cake Squeaky and sit at the table if you want a piece of cake.  
 CHI: **Why sit at the table havin(g) piece of cake?**  
 MOT: (Be)cause that's the proper way to eat cake. (VP[base]) (Nathaniel, 03;04.09)
- f. CHI: **Why shaking her head // your head?**  
 URS: What are you going to do to that one? (VP[-ing]) (Adam, 03;01.09)
- (50) a. CHI: **Why / why little? Why not big?**  
 REN: (Be)cause it's special. (AdjP) (Nathaniel, 03;00.21)
- b. CHI: Why does that boy have a motorcycle?  
 FAT: I don't know.  
 CHI: **Why not a car?**  
 FAT: Because his daddy doesn't have enough money to buy him anything else. (NP) (Ross, 03;08.18)
- c. MOT: Try and not mix (th)em up.  
 CHI: **Why not mix (th)em up?**  
 MOT: Just put the right things in each container. (VP[base]) (Abe, 04;01.20)

d. CHI: **Why not doing this brush?**

MOT: Well I think he already used that brush. (VP[-*ing*]) (Nathaniel, 03;04.18)

As illustrated here, the positive *why*-stripping construction examples produced by the children are found with an AdjP, AdvP, NP, subordinate clause, VP[base], and VP[-*ing*] as their dependent; on the other hand, their negative *why*-stripping examples are observed only with a subset of those dependent categories. The results in Table 6 and the examples as in (47)-(50) imply that young children acquire the positive *why*-stripping construction earlier than its negative counterpart and they use the former in a more productive way than the latter.

Moreover, as discussed above in (39), the children's negative *why*-stripping construction examples with a VP[base] dependent have both pure inquiry and suggestion functions, unlike their positive *why*-stripping construction examples with a VP[base] dependent which are found only with a pure inquiry function. A closer look at their negative *why*-stripping construction examples with a VP[base] dependent shows that young children acquire its two functions in different acquisition stages, as shown in Table 9 below:<sup>12</sup>

Table 9. Distributions of negative *why*-stripping construction examples with a VP[base] dependent produced by the six children in the CHILDES subcorpora based on children's age periods and illocutionary functions

Age period	Pure inquiry	Suggestion	Total
2;00-2;05	1	0	1
2;06-2;11	3	1	4
4;00-4;05	1	0	1
6;00-6;05	1	0	1
6;06-6;11	0	1	1
7;00-7;05	0	2	2
Total	6	4	10

Five out of six negative *why*-stripping construction examples with a VP[base] dependent and a pure inquiry function uttered by the six children are found before the age of 4;05; in contrast, three out of four negative *why*-stripping construction examples with a VP[base] dependent and a suggestion function produced by the

<sup>12</sup> The rows where no instances are found in Table 9 are omitted to save space.

children are found after the age of 6;06. The results here then suggest that young children acquire the negative *why*-stripping construction later with a suggestion function than with a pure inquiry function.

## 6. General discussion and conclusion

In this section, I discuss some major findings of the present study regarding young children's usage and acquisition patterns of the *how come*- and *why*-constructions noted in the previous section. First, it has been shown in this study that in general children use and acquire the *how come*-construction later than the *why*-construction. Quirk et al. (1985: 839-841) note that English employs several irregular *wh*-question constructions including the *how come*-construction (e.g., *How come you're so late?*), the *why (not)* directive construction (e.g., *Why (not) listen to him?*), the verbless *why*-existential construction (e.g., *Why no classes today?*), the subjectless infinitival *wh*-question construction (e.g., *What to do next?*), the *what if*-construction (e.g., *What if it rains?*), and the *how about/what about*-constructions (e.g., *How about another kiss?/What about following us in your car?*). These irregular *wh*-question constructions have received little attention in children's acquisition. Nonetheless, Kim (2022) shows that children begin to use and acquire the *how about/what about*-constructions later stages than prototypical *wh*-question constructions. The finding of the current study seems to support the idea that in general children use and acquire the irregular *wh*-question constructions later than canonical *wh*-question constructions. The fact that the *why*-construction taking certain dependent categories such as finite inverted S and VP[base] is acquired later with a suggestion function than with a pure inquiry function appears to further support this idea.

Next, the present study has shown that young English-speaking children use and acquire the *how come*- and *why*-constructions in a similar way to some extent. The most prominent similarity is related to the two most frequent dependent category/construction types, namely, the finite S dependent and the sluicing context. Nevertheless, a closer examination has revealed that their usage and acquisition patterns of the two most frequent dependent/construction types are rather different. As for the *how come*-construction, the two most frequent dependent/construction types have similar frequencies and they do not involve erroneous uses. In contrast, superficially,



the *why*-construction is more frequently found with a finite S dependent than in the sluicing context; however, young children's *why*-construction examples with a finite S dependent contain quite a large number of erroneous uses. This implies that young children acquire SAI in the *why*-construction in later developmental stages, lending additional support to the claim made in the previous literature (Klima and Bellugi 1966; Labov and Labov 1978; Stromswold 1990; de Villers 1991; Rowland and Pine 2000; Berk 2003; Rowland 2007; Thornton 2008). The finding that children's frequent uses of the SC dependent in the *why*-construction, not in the *how come*-construction, further suggests their later acquisition of SAI in the *why*-construction.

Another notable finding of this study regarding young children's usage and acquisition patterns of the *how come*- and *why*-constructions concerns their occurrences in the negative contexts as compared to their positive counterparts. As for the *how come*-construction, no occurrence is found in the negative sluicing and negative stripping contexts, confirming the previous observation (Collins 1991; Culicover 1991; Merchant 2006; Duffield 2015; Yoshida et al. 2015). As for the *why*-construction, fewer occurrences are found in the negative sluicing and negative stripping contexts than in their positive counterparts. This points towards the idea of young children's production and comprehension difficulties with negation form and meaning. Since negative utterances not only are longer but they also involve more complex concepts or ideas than their positive counterparts, it takes more time to acquire the former than the latter, and our finding here is then in line with the observations of previous literature (e.g., Klima and Bellugi 1966; Cameron-Faulkner et al. 2007; Nordmeyer and Frank 2014).

The results in the present study overall indicate that parents' input plays some role in children's usage and acquisition patterns of the *how come*- and *why*-constructions, but obviously not an absolute role. For instance, distributions of the finite S dependent and the sluicing context in the *how come*-construction are significantly different in the children's and their parents' data. In a similar manner but in the opposite direction, distributions of the correct S dependent and the sluicing context in the *why*-construction are significantly different in the children's and their parents' data. In addition, functional preference differences of the correct *why*-construction with a finite S dependent in the matrix environment produced by the children and their parents are observed. Furthermore, despite no parents' input for the *why*-construction with an SC dependent, frequent uses of the construction

with an SC dependent are found in the children's data. All these together commonly show that parents' input is not a sole determining factor in accounting for children's uses and acquisition of the two constructions.<sup>13</sup>

As pointed out at the beginning of this paper, young children's acquisition of the *how come*-construction has not been investigated. In addition, although young children's acquisition of the *why*-construction has received much attention in the previous literature, none of it has discussed the topic as extensively as is done in the present study, referring to diverse dependent category/construction types, functional distributions/preferences, and the role of their parents' input. Needless to say, none of the previous literature has discussed the usage and acquisition patterns of the two constructions at the same time, although they both can be used to convey a causal meaning. Therefore, the current study contributes to the body of literature on both typical and irregular *wh*-question constructions, inviting further empirical studies on children's usage and acquisition patterns of other irregular *wh*-question constructions. Furthermore, as suggested by a reviewer, it would be interesting to explore how students learning English as a second language use and acquire the *how come*- and *why*-constructions, with a particular focus on how similar and different patterns they show in using the two constructions and whether or not they exhibit similar behavior to young English-speaking children, which I leave to future research.

## References

- Berk, Stephanie. 2003. Why "why" is different. In Barbara Beachley, Amanda Brown, and Frances Conlin (eds.), *BUCLD 27: Proceedings of the 27th Annual Boston University Conference on Language Development*, 127-137. Somerville, MA: Cascadilla Press.
- Bova, Antonio. 2011. Functions of "why" questions used by children in family conversations. *Procedia - Social and Behavioral Sciences* 30: 776-782.
- Bova, Antonio and Francesco Arcidiacono. 2013. Investigating children's *why*-questions: A study

---

13 As a reviewer points out, some more explanation is required as to why the parents' input does not play an absolute role in determining the children's usage and acquisition patterns of the *how come*- and *why*-constructions. Although I cannot do full justice to this here, I assume that some of the issue is related to the children's later acquisition of SAI (e.g., for the *why*-construction examples with an SC dependent) and the functional preference difference between the children and their parents in uses of the two constructions. Needless to say, a further detailed analysis is needed to discuss this issue and I leave it to future research.

- comparing argumentative and explanatory function. *Discourse Studies* 15(6): 713-734.
- Braunwald, Susan R. 1971. Mother-child communication: The function of maternal-language input. *Word* 27(1-3): 28-50.
- Bromberger, Sylvain. 1992. *On what we know we don't know: Explanation, theory, linguistics, and how questions shape them*. Chicago, IL: University of Chicago Press.
- Brown, Roger. 1973. *A first language: The early stages*. Cambridge, MA: Harvard University Press.
- Cameron-Faulkner, Thea, Elena Lieven, and Anna Theakston. 2007. What part of *no* do children not understand? A usage-based account of multiword negation. *Journal of Child Language* 34(2): 251-282.
- Collins, Chris. 1991. *Why and how come*. In Lisa L.S. Cheng and Hamidae Demirdache (eds.), *MIT working papers in linguistics*, volume 15, 31-45. Cambridge, MA: MITWPL.
- Conroy, Anastasia. 2006. The semantics of *how come*: A look at how factivity does it all. In Nina Kazanina, Utako Minai, Philip J. Monahan, and Heather L. Taylor (eds.), *University of Maryland working papers in linguistics*, volume 14, 1-24. College Park, MD: UMWPL.
- Conroy, Anastasia and Jeffrey Lidz. 2007. Production/comprehension asymmetry in children's *why* questions. In Alyona Belikova, Luisa Meroni, and Mari Umeda (eds.), *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)*, 73-83. Somerville, MA: Cascadilla Proceedings Project.
- Crain, Stephen, Takuya Goro, and Rosalind Thornton. 2006. Language acquisition is language change. *Journal of Psycholinguistic Research* 35(1): 31-49.
- Culicover, Peter. 1991. Polarity, inversion, and focus in English. In German Westphal, Yongkyoon No, Benjamin Ao, Mark Libucha, and Hee-Rahk Chae (eds.), *Proceedings of the 8th Eastern States Conference on Linguistics*, 46-68. Baltimore, MD: University of Maryland.
- Culicover, Peter. 1999. *Syntactic nuts: Hard cases, syntactic theory, and language acquisition*. New York, NY: Oxford University Press.
- Duffield, Nigel. 2015. Where not to put *why*, and *why not*. *The Journal of Konan University Faculty of Letters* 165: 57-68.
- Ervin-Tripp, Susan M. 1970. Discourse agreement: How children answer questions. In John R. Hayes (ed.), *Cognition and the development of language*, 79-107. New York, NY: Wiley and Son.
- Fitzpatrick, Justin. 2005. The *whys* and *how comes* of presupposition and NPI licensing in questions. In John Alderete, Chung-hye Han, and Alexei Kochetov (eds.), *Proceedings of the 24th West Coast Conference on Formal Linguistics*, 138-145. Somerville, MA: Cascadilla Proceedings Project.
- Hofmann, Lisa. 2018. '*Why not?*' - Polarity ellipsis in *why*-questions. Ms. University of California, Santa Cruz.
- Huddleston, Rodney and Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English*

- language*. Cambridge: Cambridge University Press.
- Isaacs, Nathan. 1930. Children's "why" questions. In Susan Isaacs (ed.), *Intellectual growth in young children*, 291-349. London: Routledge and Sons.
- Kim, Jong-Bok and Okgi Kim. 2011. English *how come* construction: A double life. *Studies in Generative Grammar* 21(4): 587-607.
- Kim, Jungsoo. 2022. Children's acquisition and uses of *what about* and *how about* constructions: A corpus-based perspective. *Linguistic Research* 39(Special Edition): 51-96.
- Kim, Okgi and Jong-Bok Kim. 2017. English *how come* constructions: A diachronic perspective. *Language and Information* 21(1): 121-134.
- Klima, Edward S. and Ursula Bellugi. 1966. Syntactic regularities in the speech of children. In John Lyons and Richard J. Wales (eds.), *Psycholinguistic papers*, 183-208. Edinburgh: University of Edinburgh Press.
- Kuczaj, Stan. 1977. The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior* 16(5): 589-600.
- Labov, William and Teresa Labov. 1978. Learning the syntax of questions. In Robin Campbell and Philip Smith (eds.), *Recent advances in the psychology of language*, volume 4b, 1-44. New York, NY: Plenum Press.
- MacWhinney, Brian. 2000. *The CHILDES project: Tools for analyzing talk*. Mahwah, NJ: Lawrence Erlbaum Associates.
- MacWhinney, Brian and Catherine Snow. 1985. The Child Language Data Exchange System. *Journal of Child Language* 12(2): 271-296.
- MacWhinney, Brian and Catherine Snow. 1990. The Child Language Data Exchange System: An update. *Journal of Child Language* 17(2): 457-472.
- Merchant, Jason. 2006. *Why no(t)?* *Style* 40: 20-23.
- Nordmeyer, Ann and Michael Frank. 2014. The role of context in young children's comprehension of negation. *Journal of Memory and Language* 77: 25-39.
- Ochi, Masao. 2004. *How come* and other adjunct *wh*-phrases: A cross-linguistic perspective. *Language and Linguistics* 5(1): 29-57.
- Pesetsky, David. 1987. *Wh*-in-situ: Movement and unselective binding. In Eric J. Reuland and Alice ter Meulen (eds.), *The representation of (in)definiteness*, 98-129. Cambridge, MA: The MIT Press.
- Piaget, Jean. 1929. *The child's conceptions of the world*. London: Routledge and Kegan Paul.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.
- Radford, Andrew. 2018. *Colloquial English: Structure and variation*. Cambridge: Cambridge University Press.
- Rowland, Caroline. 2007. Explaining errors in children's questions. *Cognition* 104(1): 106-134.
- Rowland, Caroline and Julian Pine. 2000. Subject-auxiliary inversion errors and *wh*-question acquisition: 'What do children know'? *Journal of Child Language* 27(1): 157-181.

- Rowland, Caroline, Julian Pine, Elena Lieven, and Ana Theakston. 2003. Determinants of acquisition order in *wh*-questions: Re-evaluating the role of caregiver speech. *Journal of Child Language* 30(3): 609-635.
- Sarma, Jaya. 1991. *The acquisition of wh-questions in English*. PhD Dissertation. University of Connecticut.
- Stockwell, Richard. 2022. Elliptical *why not*. Ms. University of Oxford.
- Stromswold, Karin J. 1990. *Learnability and the acquisition of auxiliaries*. PhD Dissertation. Massachusetts Institute of Technology.
- Thornton, Rosalind. 2008. Why continuity. *Natural Language and Linguistic Theory* 26(1): 107-146.
- Thornton, Rosalind. 2016. Acquisition of questions. In Jeffrey L. Lidz, William Snyder, and Joe Pater (eds.), *The Oxford handbook of developmental linguistics*, 1-38. Oxford: Oxford University Press.
- Tsai, Wei-Tien Dylan. 2008. Left periphery and *how-why* alternations. *Journal of East Asian Linguistics* 17(2): 83-115.
- Tyack, Drothy and David Ingram. 1977. Children's production and comprehension of questions. *Journal of Child Language* 4(2): 211-224.
- de Villers, Jill. 1991. Why questions? In Thomas Maxfield and Bernadette Plunkett (eds.), *UMOP special edition: Papers in the acquisition of wh*, 155-175. Amherst, MA: Graduate Linguistics Students Association (GLSA).
- Walton, Douglas. 2004. A new dialectical theory of explanation. *Philosophical Explorations* 7(1): 71-89.
- Weist, Richard M. and Andrea A. Zevenbergen. 2008. Autobiographical memory and past time reference. *Language Learning and Development* 4(4): 291-308.
- Yoshida, Masaya, Chizuru Nakao, and Iván Ortega-Santos. 2015. The syntax of *why*-stripping. *Natural Language and Linguistic Theory* 33(1): 323-370.
- Zwicky, Arnold and Ann Zwicky. 1971. *How come* and *what for*. In Dale Eliot, Michael Geis, Alexander Grosu, Barry Nobel, Ann Zwicky, and Arnold Zwicky (eds.), *Working papers in linguistics*, volume 8, 173-185. Columbus, OH: Ohio State University.

**Appendix A.**

Children's detailed information of the subcorpora from the CHILDES database used for the current study based on childrens' age periods, utterance numbers, and MLUs

Age period	Abe	Adam	Laura	Nathaniel	Roman	Ross	Total
1;00-1;05			782 (1.18)			779 (1.63)	1,561 (1.40)
1;06-1;11			3,985 (1.88)			177 (1.63)	4,162 (1.87)
2;00-2;05	951 (3.97)	5,762 (2.29)	4,489 (2.54)	2,175 (2.66)	1,205 (3.48)	542 (3.16)	15,124 (2.65)
2;06-2;11	9,101 (4.84)	10,128 (2.88)	1,580 (2.99)	5,199 (3.17)	2,590 (3.56)	5,602 (3.77)	34,200 (3.65)
3;00-3;05	8,113 (5.72)	11,742 (4.07)	3,649 (3.33)	5,179 (3.42)	1,983 (5.15)	6,365 (5.27)	37,031 (4.53)
3;06-3;11	6,873 (6.02)	6,597 (4.48)	2,261 (3.57)	312 (3.30)	2,217 (5.16)	3,835 (5.75)	22,095 (5.13)
4;00-4;05	3,279 (6.30)	5,600 (5.24)	819 (3.39)		1,947 (5.66)	3,160 (5.84)	14,805 (5.55)
4;06-4;11	3,168 (6.44)	4,655 (5.26)	1,221 (3.42)		865 (3.78)	2,495 (6.56)	12,404 (5.54)
5;00-5;05						2,192 (7.23)	2,192 (7.23)
5;06-5;11			112 (4.28)			1,897 (5.75)	2,009 (5.67)
6;00-6;05			352 (4.52)			2,703 (5.12)	3,055 (5.05)
6;06-6;11						2,468 (4.74)	2,468 (4.74)
7;00-7;05			170 (3.42)			1,946 (4.76)	2,116 (4.65)
7;06-7;11						565 (5.58)	565 (5.58)
Total	31,485 (5.61)	44,484 (3.90)	19,420 (2.80)	12,865 (3.19)	10,807 (4.57)	34,726 (5.17)	153,787 (4.38)

**Appendix B.**

Parents' detailed information of the subcorpora from the CHILDES database used for the current study based on childrens' age periods, utterance numbers, and MLUs

Age period	Abe	Adam	Laura	Nathaniel	Roman	Ross	Total
1;00-1;05			1,527			1,671	3,198

			(4.07)			(4.78)	(4.44)
1;06-1;11			6,517 (4.25)			460 (5.45)	6,977 (4.33)
2;00-2;05	1,066 (5.87)	2,488 (4.87)	5,950 (4.28)	2,892 (5.15)	330 (5.94)	638 (5.47)	13,364 (4.80)
2;06-2;11	8,225 (5.81)	5,171 (4.95)	1,726 (4.55)	7,607 (5.38)	1,227 (5.61)	6,275 (5.11)	30,231 (5.33)
3;00-3;05	6,575 (5.60)	5,203 (5.25)	2,931 (4.40)	8,257 (5.78)	1,102 (5.82)	6,573 (5.47)	30,641 (5.46)
3;06-3;11	5,005 (4.86)	3,905 (5.32)	1,361 (5.11)	598 (4.51)	1,370 (6.27)	3,848 (5.44)	16,087 (5.24)
4;00-4;05	2,037 (4.63)	2,042 (5.48)	1,466 (4.78)		1,454 (7.24)	4,008 (5.96)	11,007 (5.64)
4;06-4;11	2,237 (4.97)	1,136 (5.76)	1,696 (5.03)		558 (5.95)	2,861 (6.44)	8,488 (5.65)
5;00-5;05						2,578 (6.15)	2,578 (6.15)
5;06-5;11			7 (6.14)			5,689 (5.75)	5,696 (5.75)
6;00-6;05			484 (5.60)			10,581 (5.49)	11,065 (5.50)
6;06-6;11						7,601 (5.49)	7,601 (5.49)
7;00-7;05			187 (3.86)			6,038 (5.24)	6,225 (5.20)
7;06-7;11						2,374 (5.68)	2,374 (5.68)
Total	25,145 (5.40)	19,945 (5.19)	23,852 (4.45)	19,354 (5.49)	6,041 (6.24)	61,195 (5.54)	155,532 (5.32)

**Jungsoo Kim**

Assistant Professor

Department of English Language Education

Incheon National University

Michuhol Campus, 12 Gaetbeol-ro, Yeonsu-gu

Incheon 21999, Korea

E-mail: jungsookim@inu.ac.kr

Received: 2024. 04. 02.

Revised: 2024. 05. 31.

Accepted: 2024. 05. 31.