Optimality Theory and the Development of *Do*-Support in Children's *Wh*-Questions

Nobuyo Fukaya

***Abstract*—This study aims to explore *where*-, *how*-, and *why*-questions produced by young L1 English-speaking children and to account for how children develop *wh*-questions within Optimality Theory (OT). For this purpose, the data have been collected from the Child Language Data Exchange System database. The analysis showed that where-questions were produced earlier than *how*- and *why*-questions and that early *where*-questions tended to fail subject-auxiliary inversion in the early stages. Another finding is that children produced *how*-questions with subject-auxiliary inversion even when *how*-questions started to be attested. To account for the developmental differences observed in the data, I propose an OT analysis. This study shows the applicability of OT to syntactic language development and demonstrates that OT provides a unified account of the development of *wh*-questions by re-ranking the same constraint set. Moreover, a developmental difference shown in *wh*-phrases can also be accounted for by assuming that the constraint, Operator in Specifier (OP-SPEC), can be divided into sub-constraints.**

***Index Terms*—*where*-questions, *how*-questions, *why*-questions, optimality theory, CHILDES**

I. Introduction

This study aims to explore L1 English-speaking children's *wh*-questions in the early stages of language development and to analyze the syntactic development of *wh*-questions within Optimality Theory (OT). Children begin to produce *wh*-questions in the early stages; however, their *wh*-questions tend to be simplified compared with those of adults in that some elements fail to appear, as the following examples show [1].

(1) a. What doing? (Crain & Lillo-Martin, 1999, p. 209)

b. Where Daddy? (Crain & Lillo-Martin, 1999, p. 209) (1a) and (1b) lack the subject and the verb *be*, respectively. Moreover, all *wh*-questions are not acquired simultaneously. According to Thornton (2016), children first produce *what*- and *where*- questions. When their *what*- and *where*-questions involve subject-auxiliary inversion (SAI), children also produce other *wh*-questions such as *how*-questions [2].

In this paper, I present a longitudinal analysis of *where*-, *how*-, and why questions. Collecting children's utterances and comparing the three *wh*-questions will provide renewed insights for *wh*-questions and language development. This study aims to answer the following research questions.

(2) a. Is there any developmental difference between *where*-questions, *how*-questions, and *why*-questions with respect to the development of SAI?

b. If each *wh*-question develops differently, how can OT explain the language development?

Manuscript received November 15, 2022, and received on, January, 5, 2023; accepted February 15, 2023.

Fukaya Nobuyo is with Niigata Agro-Food University, Japan.
E-mail: nobuyo-fukaya@ nafu.ac.jp (F.N.)

This paper is organized as follows. Section II reviews Guasti's (2000) analysis on the development of *wh*-questions [3]. Section III introduces the Child Language Data Exchange System (CHILDES) database (MacWhinney, 2000) and then presents the children's data collected from the CHILDES database [4]. Section IV provides an OT account of *wh*-questions and shows how the ranking changes during language development. Section V reviews the research questions (2) and concludes the paper.

II. Literature Review

Guasti (2000) examined children's use of *wh*-questions in the early stages of language development. The examples in (3) show that English *wh*-questions involve two operations with the exception of subject questions: *wh*-movement and SAI. *Wh*-movement requires the *wh*-phrase to move to the front (more specifically, the specifier of CP) and SAI requires the auxiliary, which is base-generated in I, to move to C. When the sentence does not contain an auxiliary such as (3b), the semantically empty auxiliary *do* is inserted in I and then moves to C.

(3) a. $[_{CP}$ What$_i$ can$_j$ $[_{IP}$ I t$_j$ $[_{VP}$ do t$_i$ for you$]]]$?

b. $[_{CP}$ Where$_i$ did$_j$ $[_{IP}$ you t$_j$ $[_{VP}$ go t$_i]]]$?

Guasti (2000) assumed that these operations are triggered by the *wh*-criterion shown in (4).

(4) a. A *wh*-operator must be in a Spec-head relation with a head carrying the *wh*-feature.

b. A head carrying the *wh*-feature must be in a Spec-head relation with a *wh*-operator.

Guasti's (2000) main goal was to address the question of whether the *wh*-criterion is satisfied even in the early stages of language development. She collected *wh*-questions produced by four English-speaking children (Adam, Eve, Sarah, and Nina) and divided the data into the following five groups.

(5) a. +SAI
b. -SAI
c. Wh S V
d. Wh S V-ing
e. Wh S Vfin

(5a) is the *wh*-question that involves SAI, while (5b) lacks SAI such as *where he is doing?*; the other three groups all lack auxiliaries. (5c), (5d), and (5e) contain the bare verb, the progressive aspect *-ing*, and the finite form of the verb, respectively. Relevant examples are provided below.

(6) a. What he like? = (5c)
b. What you doing? = (5d)
c. What he likes? = (5e)

The results are shown in Table I.

TABLE 1: NUMBER OF THE *WH*-CONSTRUCTIONS FOR THE FOUR CHILDREN

	+SAI	-SAI	Wh S V	Wh S V-ing	Wh S Vfin
Adam	816	64	469	268	207
Eve	58	2	42	44	3
Sarah	250	15	127	23	26
Nina	316	4	14	19	1

(Guasti, 2000, p. 109)

According to Guasti (2000), (5c) and (5d) should be treated in the same way as *wh*-questions with +SAI. This is because we can assume that they involve null auxiliaries in C. Moreover, (5e) should be left out when we consider children's accessibility of SAI because not all the children produce such *wh*-questions. Table II presents the comparison between (5a) and (5b) to determine whether children use the *wh*-criterion even in the early stages. Since more than 90% of the *wh*-questions produced by the children involved SAI, Guasti argued that the *wh*-criterion plays a role in the early stages of language development.

TABLE II: COMPARISON BETWEEN *WH*-QUESTIONS WITH AND WITHOUT SAI

	Adam	Eve	Sarah	Nina
+SAI	93	96.6	94.3	98.7
-SAI	7	3.3	5.6	1.2

(Guasti, 2000, p. 110)

III. DATA

A. The CHILDES Database

In this study, I collected data from the CHILDES database. The subjects were three L1 English-speaking children: Aran (Manchester corpus), Naomi [5], and Nina [6]. Table III summarizes the children's ages.

TABLE III: THREE PARTICIPANTS EXAMINED

Subjects	Age range
Aran	1;11.12 - 2;10.28[1]
Naomi	1;02.29 - 4;09.03
Nina	1;11.16 - 3;03.21

To conduct a longitudinal analysis of language development, I used Brown's stages [7]. Brown's stages consist of five stages from Stages I to V and each stage has its MLU range as shown in Table IV. In this analysis, I define the stage whose MLU is 4.50 or more as Stage V+.

TABLE IV: BROWN'S STAGES

Stage	MLU range
Stage I	1 – 1.99
Stage II	2.00 – 2.49
Stage III	2.50 – 2.99
Stage IV	3.00 – 3.74
Stage V	3.75 – 4.49
Stage V+	4.50 –

Brown (1973) assumed that children's language development is based not on their ages, but on the morphemes in their utterances. Let us examine (7) from Nina's *why*-questions.

(7) Why do you make little holes in fingers? (2;11.12)
1 1 1 1 1 2 1 2

The *why*-question in (7) contains 8 words but consists of 10 morphemes: *holes* has two morphemes, namely, *hole* and *-s*. The word *fingers* can also be broken down into *finger* and *-s*. MLU is calculated by dividing the total number of morphemes by the total number of utterances. Therefore, if there are 250 morphemes in 100 utterances that a child produced, the child's MLU is calculated by dividing 250 (the total number of morphemes) by 100 (the total number of utterances); the result is 2.5. We easily count MLU values by means of one of the commands of the Browsable Database in the CHILDES database. Table V shows the correspondence of each stage to the age ranges of the three children. The notation x; yy. zz stands for years; months. days. For example, 2;10.28 stands for 2 years, 10 months, and 28 days, respectively.

TABLE V: BROWN'S STAGES AND AGE RANGES

	Aran	Naomi	Nina
Stage I	1;11.12 – 2;00.02	1;02.29 – 1;10.28	
Stage II	2;00.09 – 2;02.25	1;11.02 – 2;02.00	1;11.16 – 2;01.22
Stage III	2;03.02 – 2;08.19	2;02.25 – 2;11.10	2;01.29 – 2;02.12
Stage IV	2;09.02 – 2;10.28	2;11.11 – 3;08.19	2;02.28 – 2;10.13
Stage V		4;07.28 – 4;09.03	2;10.21 – 3;01.06
Stage V+			3;01.07 – 3;03.21

I collected all the utterances containing *where*-, *how*-, or *why*- produced by the three children. The utterances were then divided into three types, as shown in Table VI.

TABLE VI: THREE TYPES OF *WH*-QUESTIONS INVESTIGATED

Type	Pattern
Type A	Wh Subject (Auxiliary) Verb?
Type B	Wh Auxiliary Subject Verb?
Type C	Wh's Subject (Verb)?

Type A is the question that lacks *do*-support, as in (8). This type also includes questions in which the auxiliary remains in situ, as in (9).

(8) a. Where the monkey go? (Nina 2;04.26)
b. How he goes on? (Aran 2;07.28)
c. Why you doing that? (Nina 3;01.07)

(9) a. Where Bumbo can go? (Aran 2;05.03)
b. Why I can't come out? (Aran 2;07.28)

Type B involves SAI, as in (10) and (11) with *do*-support.

(10) a. Where are you going? (Naomi 3;04.00)
b. How can we open it now? (Nina 2;09.21)
c. Why is he Billy? (Naomi 3;05.04)

(11) a. Where did Kimberly put that piece? (Naomi 2;08.14)
b. How do these things open? (Nina 2;09.26)
c. Why did it fall on the cow? (Aran 2;09.02)

Type C contains a contracted form of the auxiliary, as in (12). Both *'s* and other types of contracted forms are also included in this type.

(12) a. Where<u>'s</u> Mommy? (Naomi 1;11.11)
b. Where<u>'s</u> my dolly? (Nina 2;01.29)
c. Where<u>'d</u> it go? (Naomi 1;10.28)

This analysis focuses on the development of SAI; therefore, one-word *wh*-questions (13a), *wh*-questions with subject omission (13b), and *wh*-questions with verb omission (13c) were excluded from this analysis.

(13) a. Where? (Nina 2;09.26)

[1] The notation x; yy. zz stands for years; months. days. For example, 2;10.28 stands for 2 years, 10 months, and 28 days, respectively.

b. Where going on Friday? (Aran 2;06.17)
c. Where the boy? (Aran 2;00.09)

B. Results and Discussion

Fig. 1 shows the overall results of *where*-, *how*-, and *why*-questions produced by the children. As is clear from this figure, the number of *where*-questions jumped from Stage I to Stage II, while the number of *how*-questions and *why*-questions increased from Stage III to Stage IV.

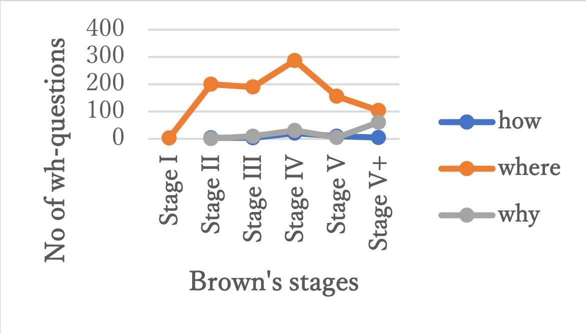

Fig. 1. Number of *wh*-questions.

Let us first examine *where*-questions. Fig. 2 shows the results of *where*-questions by type, and Table VII presents the number of utterances by child.

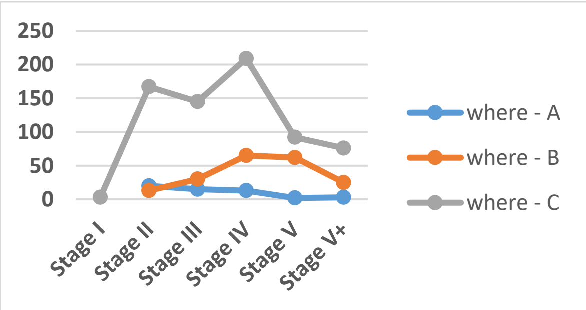

Fig. 2. Number of *where*-questions by type.

TABLE VII: RESULTS OF *WHERE*-QUESTIONS FOR THE THREE CHILDREN

	Type A	Type B	Type C	Total
Stage I			**3**	**3**
Naomi			3	3
Stage II	**20**	**13**	**167**	**200**
Aran	12		3	15
Naomi	8	11	145	164
Nina		2	19	21
Stage III	**15**	**30**	**145**	**190**
Aran	13	21	77	111
Naomi	1	9	51	61
Nina	1		17	18
Stage IV	**13**	**65**	**209**	**287**
Aran	2	21	26	49
Naomi		8	28	36
Nina	11	36	155	202
Stage V	**2**	**62**	**92**	**156**
Naomi		5	2	7
Nina	2	57	90	149
Stage V+	**3**	**25**	**76**	**104**
Nina	3	25	76	104
Total	**53**	**195**	**692**	**940**

The data revealed that Type C cases jumped from Stage I to Stage II, while those of Type B increased from Stage III.

Let us turn to *how*-questions. As is clear from Fig. 3 and Table VIII, there were no utterances without SAI.

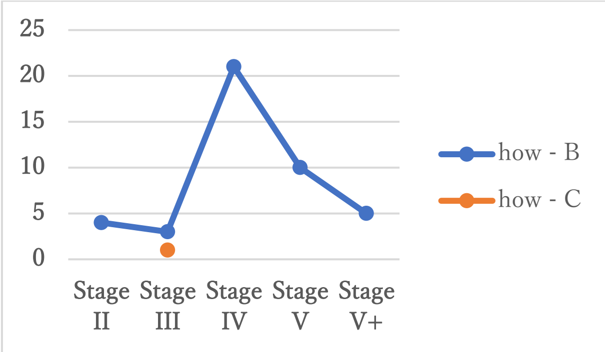

Fig. 3. Number of *how*-questions by type.

TABLE VIII: RESULTS OF *HOW*-QUESTIONS FOR THE THREE CHILDREN

	Type A	Type B	Type C	Total
Stage I			**3**	**3**
Naomi			3	3
Stage II	**20**	**13**	**167**	**200**
Aran	12		3	15
Naomi	8	11	145	164
Nina		2	19	21
Stage III	**15**	**30**	**145**	**190**
Aran	13	21	77	111
Naomi	1	9	51	61
Nina	1		17	18
Stage IV	**13**	**65**	**209**	**287**
Aran	2	21	26	49
Naomi		8	28	36
Nina	11	36	155	202
Stage V	**2**	**62**	**92**	**156**
Naomi		5	2	7
Nina	2	57	90	149
Stage V+	**3**	**25**	**76**	**104**
Nina	3	25	76	104
Total	**53**	**195**	**692**	**940**

The results of *why*-questions are shown below in Fig. 4 and Table IX.

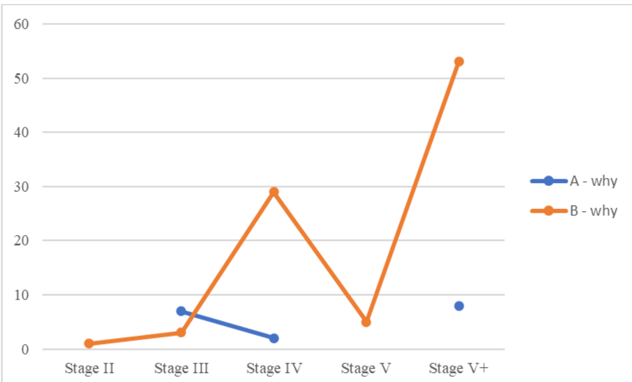

Fig. 4. Number of *why*-questions by type.

TABLE IX: RESULTS OF *WHY*-QUESTIONS FOR THE THREE CHILDREN

	Type A	Type B	Total
Stage II		**1**	**1**
Aran		1	1
Stage III	**7**	**3**	**10**
Aran	5	3	8
Naomi	2		2
Stage IV	**2**	**29**	**31**
Aran	1	12	13
Naomi		14	14
Nina	1	3	4
Stage V		**5**	**5**
Nina		5	5
Stage V+	**8**	**53**	**61**
Nina	8	53	61
Total	**17**	**91**	**108**

The data offer important findings for the development of *wh*-questions. First, Type C with *where* was produced in the early stages, when *how*- and *why*-questions were hardly observed. Looking at Type C, we may say that the children acquired *where*-questions earlier than *how*- and *why*-questions. However, the data also show that at around Stage IV, *wh*-questions with SAI (Type B) began to appear through the three *wh*-types. Another finding is that Type A, was observed in *where*-questions even at Stage IV, whereas this type was not observed in *how*-questions. In Section IV, I address research question (2b) and show that OT provides a unified account of the development of *wh*-questions by re-ranking the constraint set.

IV. OPTIMALITY THEORY

A. The Mechanism of Optimality Theory

OT was proposed by Prince and Smolensky (1993) in the field of phonology, and it has been applied to other fields such as syntax and morphology [8]. OT assumes the main mechanism where Generator (GEN) and Evaluator (EVAL) are mediated between the input and output. GEN produces an infinite number of candidates based on the input. The candidate set is then handed down to EVAL to determine an optimal candidate. When choosing optimality, OT assumes that every language has a language-particular ranking consisting of constraints. OT constraints have two important features: universality and violability. Optimality is determined not by the number of violations of constraints, but by the number of the least serious violations of universal constraints. A violation of a high-ranking constraint leads to a serious violation. Let us examine the following schematic Table X.

TABLE X: $C_1>>C_2>>C_3$

	C_1	C_2	C_3
☞Candidate (a)		*	*
Candidate (b)	*!		
Candidate (c)		**!	

In Table X, there are three constraints, C_1, C_2, and C_3. Under the ranking of $C_1>>C_2>>C_3$, C_1 is ranked the highest, C_2 is in the middle, and C_3 is the lowest. The optimal candidate is (a) with the pointing finger. Candidates (b) and (c) cannot be optimal under this ranking. Candidate (b) violates the highest ranked constraint C_1 and candidate (c) has C_2 twice. If the constraints are ranked as in $C_3>>C_2>>C_1$, candidate (b) is chosen as optimal. This is shown in Table XI.

TABLE XI: $C_3>>C_2>>C_1$

	C_3	C_2	C_1
Candidate (a)	*!	*	*
☞Candidate (b)			*
Candidate (c)		*!*	

B. An OT Analysis

In this section, I first introduce five constraints that are considered relevant in the development of wh-questions. The first constrain is OP-SPEC [9].

(14) Syntactic operators must be in specifier position. (Grimshaw, 1997, p. 374)

OP-SPEC requires that the *wh*-phrase move to a specifier position of the clause. What is important is that this constraint does not specify any position such as Spec CP. Therefore, OP-SPEC is satisfied when the *wh*-phrase is positioned not only in Spec CP, but also in other specifier positions such as Spec IP.

The second constraint is OP-SCOPE, proposed by Baković (1998) [10].

(15) Syntactic operators must c-command the extended projection over which their scope is interpreted. (Baković, 1998, p. 39)

OP-SPEC and OP-SCOPE play an important role in the production of *wh*-questions. Let us examine the following Table XII and see how they work.

TABLE XII: OP-SPEC AND OP-SCOPE

	OP-SPEC	OP-SCOPE
a. $[_{XP}$ wh do $[_{VP}$ S V]]		
b. $[_{VP}$ wh $[_{VP}$ S V]]	*	

Candidate (a) places the *wh*-phrase in Spec XP; therefore, OP-SPEC and OP-SCOPE are both satisfied. In candidate (b), the *wh*-phrase is adjoined to VP. This candidate satisfies OP-SCOPE, but violates OP-SPEC.

The third constraint is STAY.

(16) Trace is not allowed. (Grimshaw, 1997, p. 374)

STAY prohibits any movement. Therefore, if constraints requiring movement are ranked higher than STAY, the candidate with the movement is chosen as optimal. Under the opposite ranking, the candidate without movement should be optimal.

The fourth constraint is CASE.

(17) DPs must be case marked. (Grimshaw, 1997, p. 374)

The fifth constraint is FULL-INT.

(18) Lexical conceptual structure is parsed. (Grimshaw, 1997, p. 374)

According to Grimshaw (1997), the syntactic input contains the argument structure and some information such as topic and tense. The semantically empty auxiliary *do* does not appear in the input. Candidates with *do*-support generated by GEN have a violation of FULL-INT.

Before presenting an OT analysis, let us focus on OP-SPEC. Baković (1998) assumed that OP-SPEC is divided into the following four sub-constraints:

(19) a. ARGOP-SPEC: Argument operators must be in specifier position.
 b. LOCOP-SPEC: Location operators must be in specifier position.
 c. MANOP-SPEC: Manner operators must be in specifier position.
 d. REASOP-SPEC: Reason operators must be in specifier position.

When all the constraints in (19) are ranked above STAY, it is akin to saying that the language has the ranking of OP-SPEC>>STAY. English is supposed to have this ranking. However, there is the possibility that some sub-constraints are ranked higher than STAY, while others are ranked lower. In some Spanish dialects, matrix *wh*-questions behave differently according to types of *wh*-phrases: argument *wh*-questions require that the verb move to the second position of the clause, while adjunct *wh*-questions do not. Baković (1998) proposed that the dialects should have the ranking of ARGOP- SPEC>>STAY>>ADJOP-SPEC (LOCOP-SPEC, MANOP-SPEC, REASOP-SPEC). Tables XIII and XIV show how optimality is created.

TABLE XIII: ARGOP- SPEC>>STAY

	ARGOP-SPEC	STAY
a. ☞ [$_{VP}$ Arg[$_{V'}$ V [$_{VP}$ Subj [V t$_v$ t$_{wh}$]]]]		**
b. [$_{VP}$ Arg [$_{VP}$ Subj [$_{v'}$ V t$_{wh}$]]]	*!	*

(Baković, 1998, p. 44)

TABLE XIV: STAY>>ADJOP-SPEC

	STAY	ADJOP-SPEC
a. [$_{VP}$ Adj[$_{V'}$ V [$_{VP}$ Subj [$_{V'}$ t$_v$ t$_{wh}$]]]]	**!	
b. ☞ [$_{VP}$ Adj [$_{VP}$ Subj [$_{V'}$ V t$_{wh}$]]]	*	*

(Baković, 1998, p. 44)

In this study, I also assume that OP-SPEC can be divided into sub-constraints. LOCOP-SPEC, MANOP-SPEC, and REASOP-SPEC are relevant here.

1) Where-questions

I first address the question of why the children initially produced Type A in *where*-questions. I assume that the children started with ranking (20).

(20) OP-SCOPE, FULL-INT>>LOCOP-SPEC, STAY>>CASE

OP-SCOPE and FULL-INT are equally ranked, being ranked above LOCOP-SPEC and STAY. CASE is the lowest ranked constraint. Under this ranking, it is better to move the *wh*-phrase to a position in which the *wh*-phrase can c-command the clause, as shown in Table XV. Note that in this analysis, constraints placed in the same stratum are assumed to be equally ranked.

TABLE XV: *WHERE*-QUESTIONS (1)

		OP-SCOPE	FULL-INT	LOCOP-SPEC	STAY	CASE
a.	☞ [$_{VP}$ where$_i$ [$_{VP}$ S V t$_i$]]			*	*	*
b.	[$_{IP}$ where$_i$ do[$_{VP}$ S V t$_i$]]		*!		*	*
c.	[$_{IP}$ where$_i$ [$_{IP}$ S$_j$ [$_{VP}$ t$_j$ V t$_i$]]]			*	**!	
d.	[$_{CP}$ where$_i$ do$_k$ [$_{IP}$ t$_j$ t$_k$ [$_{VP}$ S$_j$ V t$_i$]]]		*!		***	
e.	[$_{VP}$ S V where]	*!		*		*

In candidate (a), the *wh*-phrase is adjoined to VP and the subject remains in situ. Therefore, this candidate violates LOCOP-SPEC, STAY, and CASE. In candidate (b), the semantically empty auxiliary *do* is inserted into I, which causes a serious violation of FULL-INT. Candidate (c) adjoins the *wh*-phrase to IP and moves the subject to Spec IP, violating LOCOP-SPEC and STAY twice. In candidate (d), the *wh*-phrase moves to Spec CP, the semantically empty auxiliary to C, and the subject to Spec IP. Such movements cause three violations of STAY. Candidate (e) represents the situation in which every element remains in situ. Under this ranking, candidate (a) is chosen as optimal.

One question arises at this point: why were a huge number of Type C-cases observed in the early stages? I assume that Type C in these stages was not generated under the relevant ranking, but can be taken as a schema because *where's* always appeared irrespective of the subject.

(21) a. Where's my pictures? (Nina 2;01.06)
 b. Where's our cookies? (Naomi 2;08.14)
 c. Where's those go? (Aran 2;04.27)

I assume that at around Stage IV, ranking (20) was re-ranked as in (22).

(22) OP-SCOPE>>LOCOP-SPEC, FULL-INT, STAY, CASE

Under this ranking, LOCOP-SPEC, FULL-INT, STAY, and CASE are tied in the same stratum. The correctness of this ranking is illustrated in Table XVI.

TABLE XVI: *WHERE*-QUESTIONS (2)

		OP-SCOPE	LOCOP-SPEC	FULL-INT	STAY	CASE
a.	☞ [$_{VP}$ where$_i$ [$_{VP}$ S V t$_i$]]		*		*	*
b.	☞ [$_{IP}$ where$_i$ do[$_{VP}$ S V t$_i$]]			*	*	*
c.	☞ [$_{IP}$ where$_i$ [$_{IP}$ S$_j$ [$_{VP}$ t$_j$ V t$_i$]]]		*		**	
d.	[$_{CP}$ where$_i$ do$_k$ [$_{IP}$ t$_j$ t$_k$ [$_{VP}$ S$_j$ V t$_i$]]]			*	***!	
e.	[$_{VP}$ S V where]	*!	*			*

Under this ranking, candidates (a), (b), and (c) are selected as optimal. As Fig. 2 illustrates, *where*-questions with *do*-support begin to appear. A representative example is shown in (23a). Moreover, candidates (a) and (c) may look similar, but candidate (c) is more complicated in that it has the IP structure. This is supported by (23b) in which the verb agrees with the subject.

(23) a. Where does this go in here? (Naomi 2;11.12)
 b. Where it goes? (Aran 2;07.21)

Ranking (22) is then supposed to be re-ranked as in (24) and reaches adult grammar.

(24) OP-SCOPE, LOCOP-SPEC>>FULL-INT, CASE>>STAY

In ranking (24), OP-SCOPE and LOCOP-SPEC are top-ranked, dominating FULL-INT and CASE, which are in turn ranked above STAY. What is important here is that LOCOP-SPEC is ranked the highest; therefore, the *wh*-phrase is expected to move to a specifier position, and not to be adjoined to the clause.

TABLE XVII: *WHERE*-QUESTIONS (3)

	OP-SCOPE	LOCOP-SPEC	FULL-INT	CASE	STAY
a. [$_{VP}$ where$_i$ [$_{VP}$ S V t_i]]		*!		*	*
b. [$_{IP}$ where$_i$ do[$_{VP}$ S V t_i]]			*	*!	*
c. [$_{IP}$ where$_i$ [$_{IP}$ S_j [$_{VP}$ t_j V t_i]]]		*!			**
d. ☞ [$_{CP}$ where$_i$ do$_k$ [$_{IP}$ S_j t_k [$_{VP}$ t_j V t_i]]]			*		***
e. [$_{VP}$ S V where]	*!	*			*

As Table XVII shows, the optimal candidate (d) has *do*-support, resulting in the violation of FULL-INT.

2) How- and why-questions

Let us move on to *how*- and *why*-questions. Importantly, constraint-rankings are not construction-specific; therefore, the rankings proposed in the previous section imply that the children were supposed to have the core rankings, as in (25). Under (25a), LOCOP-SPEC, which plays a role in the decision of *where*-questions, was placed in the same stratum as STAY.

(25) a. OP-SCOPE, FULL-INT>>STAY>>CASE
b. OP-SCOPE>>FULL-INT, STAY, CASE
c. OP-SCOPE>>FULL-INT, CASE>>STAY

Few cases of *how*-questions and *why*-questions were available in the early stages. This provides evidence that MANOP-SPEC and REASOP-SPEC were still not operative; therefore, they did not have a specific stratum to sit in. I argue that these two constraints came to be operative when the children had the ranking in (25b). Moreover, the developmental difference between *how*- and *why*-questions suggests that MANOP-SPEC and REASOP-SPEC were ranked differently with respect to FULL-INT, STAY, and CASE. I assume that when the children started to produce *how*-questions, they had ranking (26). The initial ranking for *how*-questions is based on ranking (25b) and MANOP-SPEC is equally-ranked with OP-SCOPE. The correctness of this ranking is supported by the fact that the children produced *how*-questions with SAI.

(26) OP-SCOPE, MANOP-SPEC>>FULL-INT, STAY, CASE

Ranking (26) is illustrated in Table XVIII.

TABLE XVIII: *HOW*-QUESTIONS (1)

	OP-SCOPE	MANOP-SPEC	FULL-INT	STAY	CASE
a. [$_{VP}$ how$_i$ [$_{VP}$ S V t_i]]		*!		*	*
b. ☞ [$_{IP}$ how$_i$ do [$_{VP}$ S V t_i]]			*	*	*
c. [$_{IP}$ how$_i$ [$_{IP}$ S_j [$_{VP}$ t_j V t_i]]]		*!		**	
d. [$_{CP}$ how$_i$ do$_k$ [$_{IP}$ t_j t_k [$_{VP}$ S_j V t_i]]]			*	***!	
e. [$_{VP}$ S V how]	*!	*			*

Candidates such as (a) and (c), which are adjoined to the clause, are suboptimal under this ranking. Candidate (b), with *do*-support in IP, is chosen as optimal. Ranking (26) is re-ranked as in (27), which is illustrated in Table XIX. As with *where*-questions, candidate (d) emerges as optimal.

(27) OP-SCOPE, MANOP-SPEC >FULL-INT, CASE>>STAY

TABLE XIX: *HOW*-QUESTIONS (2)

	OP-SCOPE	MANOP-SPEC	FULL-INT	CASE	STAY
a. [$_{VP}$ how$_i$ [$_{VP}$ S V t_i]]		*!		*	*
b. [$_{IP}$ how$_i$ do [$_{VP}$ S V t_i]]			*	*!	*
c. [$_{IP}$ how$_i$ [$_{IP}$ S_j [$_{VP}$ t_j V t_i]]]		*!			**
d. ☞ [$_{CP}$ how$_i$ do$_k$ [$_{IP}$ S_j t_k [$_{VP}$ t_j V t_i]]]			*		***
e. [$_{VP}$ S V how]	*!	*			*

Let us turn to why-questions in. I assume that REASOP-SPEC was equally ranked with FULL-INT and STAY when this constraint became operative, as in (28).

(28) OP-SCOPE>> REASOP-SPEC, FULL-INT, STAY, CASE

This ranking shows similar effects as ranking (22) for *where*-questions, as show Table XX shows:

TABLE XX: *WHY*-QUESTIONS (1)

	OP-SCOPE	REASOP-SPEC	FULL-INT	STAY	CASE
a. ☞ [$_{VP}$ why$_i$ [$_{VP}$ S V t_i]]		*		*	*
b. ☞ [$_{IP}$ why$_i$ do[$_{VP}$ S V t_i]]			*	*	*
c. ☞ [$_{IP}$ why$_i$ [$_{IP}$ S_j [$_{VP}$ t_j V t_i]]]		*		**	
d. [$_{CP}$ why$_i$ do$_k$ [$_{IP}$ t_j t_k [$_{VP}$ S_j V t_i]]]			*	***!	
e. [$_{VP}$ S V why]	*!	*			*

Ranking (28) is supposed to be re-ranked, as in (29). The correctness of this ranking is illustrated by Table XXI.

(29) OP-SCOPE, REASOP-SPEC>>FULL-INT, CASE>>STAY

TABLE XXI: *WHY*-QUESTIONS (2)

	OP-SCOPE	REASOP-SPEC	FULL-INT	CASE	STAY
a. [$_{VP}$ why$_i$ [$_{VP}$ S V t_i]]		*!		*	*
b. [$_{IP}$ why$_i$ do[$_{VP}$ S V t_i]]			*	*!	*
c. [$_{IP}$ why$_i$ [$_{IP}$ S_j [$_{VP}$ t_j V t_i]]]		*!			**
d. ☞ [$_{CP}$ why$_i$ do$_k$ [$_{IP}$ S_j t_k [$_{VP}$ t_j V t_i]]]			*		***
e. [$_{VP}$ S V why]	*!	*			*

V. CONCLUSION

This study investigated *where*-, *how*-, and *why*-questions produced by L1 English-speaking children. One of the findings that emerged from the children's data is that *how*- and *why*-questions were observed later than where-questions, which supports Rowland *et al.* (2003) [11]. Second, the data offered evidence that in *why*-questions the failure of SAI (Type A) was observed, while in *how*-questions it was not [12]. Last but not the least, the development of *wh*-questions varied according to a *wh*-phrase; however, the three *wh*-phrases were common in that the first appearance of *do*-support was at around Stage IV. Therefore, looking back on the first research question, repeated here as (30), we can say that the answer is *yes*.

(30) Is there any developmental difference between *where*-questions, *how*-questions, and *why*-questions with respect to the development of SAI?

At this point, there is a question of why the children successfully produced *how*-questions with SAI. The answer to this question is related to the second research question, repeated here as (31).

(31) If each *wh*-question develops differently, how can OT explain the language development?

Although OT supposes that universal constraints are ranked in a language-specific way, not in a construction-specific way, I showed, following Baković (1998), that the developmental difference in language can be accounted for by assuming that OP-SPEC is divided into several types according to a *wh*-type. To put it precisely, LOCOP-SPEC, MANOP-SPEC, and REASOP-SPEC were first ranked differently with respect to FULL-INT, as in (32). This led to the developmental difference in *do*-support presented in Section III.

(32) a. *where*-questions: OP-SCOPE>>LOCOP-SPEC, FULL-INT, STAY, CASE
b. *how*-questions: OP-SCOPE, MANOP-SPEC >>FULL-INT, CASE, STAY
c. *why*-questions: OP-SCOPE>> REASOP-SPEC, FULL-INT, STAY, CASE

Then, the rankings were supposed to be re-ranked in the following way.

(33) a. *where*-questions: OP-SCOPE, LOCOP-SPEC>>FULL-INT, CASE>>STAY
b. *how*-questions: OP-SCOPE, MANOP-SPEC >>FULL-INT, CASE>>STAY
c. *why*-questions: OP-SCOPE>> REASOP-SPEC, FULL-INT, CASE>>STAY

In (33), all the sub-constraints of OP-SPEC are in the same stratum; therefore, these rankings were equal to (34) in which SAI is required due to OP-SPEC>>STAY.

(34) OP-SCOPE, OP-SPEC>>FULL-INT, CASE>>STAY

This OT analysis presents the following advantages. First, it has shown that language development can be accounted for by re-ranking the same constraint set. Second, differences in language development can be attributed to sub-constraints of OP-SPEC. Further research is required to determine whether OT can provide a unified account of the development of other constructions.

Conflict of Interest

The author declares no conflict of interest.

References

[1] S. Crain and D. Lillo-Martin, *An Introduction to Linguistic Theory and Language Acquisition*, Malden, MA: Blackwell Publishing, 1999.

[2] R. Thornton, "Acquisition of questions," in *The Oxford Handbook of Developmental Linguistics*, J. Lidz, W. Snyder, and J. Pater, Eds. Oxford: Oxford University Press, 2016, pp. 310–340.

[3] M. T. Guasti, "An excursion into interrogatives in early English and Italian," in *The Acquisition of Syntax*, M. Friedemann and L. Rizzi, Eds. Essex: Pearson Education Limited, 2000, pp. 105–128.

[4] B. MacWhinney, *The CHILDES Project: Tools for Analysing Talk*, vol. 2, 3rd ed. New Jersey: Psychology Press, 2000.

[5] J. Sachs, "Talking about the there and then: The emergence of displaced reference in parent-child discourse," in *Children's Language*, vol. 4, K. E. Nelson, Ed. New Jersey: Psychology Press, 1983, pp. 1–28.

[6] P. Suppes, "The semantics of children's language," *American Psychologist*, vol. 29, pp. 103–114, 1974.

[7] R. Brown, *A First Language*, Cambridge, MA: Harvard University Press, 1973.

[8] A. Prince and P. Smolensky, *Optimality Theory: Constraint Interaction in Generative Grammar, RuCCs Technical Report #2*, Piscataway, NJ: Rutgers University Center for Cognitive Science, 1983.

[9] J. Grimshaw, "Projection, heads, and optimality," *Linguistic Inquiry*, vol. 28, pp. 373–422, 1997.

[10] E. Baković, "Optimality and inversion in Spanish," in *Is the Best Good Enough? Optimality and Competition in Syntax*, P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis, and D. Pesetsky, Eds. Cambridge, MA: MIT Press, 1998, pp. 35–58.

[11] C. F. Rowland, J. M. Pine, E. V. M. Lieven, and A. L. Theakston, "Determinants of acquisition order in *wh*-questions: Re-evaluating the role of caregiver speech," *Journal of Child Language*, vol. 30, pp. 609–635, 2003.

[12] R. Thornton, "Why continuity," *Natural Language and Linguistic Theory*, vol. 26, no. 1, pp.107–146, 2008.