Children's *Wh*-questions across Languages: Some Preliminary Results

Kazuko Yatsushiro, Chiara Dal Farra, Aurore Gonzalez, Johannes Hein, Silvia Silleresi, Alicia Avellana, Aijun Huang, Johnson F. Ilori, Gayathri G. Krishnan, Lilla Pintér, Maria Teresa Guasti, and Uli Sauerland

1. Introduction

This paper investigates the acquisition of *wh*-questions in five different languages: German, Hungarian, Malayalam, Mandarin and Yoruba. It does so by focusing on the production of five structures, which were tested across all languages on the basis of (nearly) identical material. These included subject *who*-questions (1a) and object *who*-questions (1b), subject *which*-questions (1c) and object *which*-questions (1d), and object *what*-questions (1e). In this paper we will focus only on *who*- and *which*-questions.

(1) a. Who is pushing the ant?
     b. Who are the horses pulling?
     c. Which ladybug is waking the cats?
     d. Which duck is the rabbit pushing?
     e. What is the woman drawing?

Acquisition of *wh*-questions has been extensively investigated in the literature, but most of the studies focused on the comprehension of these structures

---

(see Friedmann & Novogrodsky 2011; Sauerland et al. 2016 among many others), rather than on their production (see Stavrakaki 2006; Guasti et al. 2012). The results of comprehension studies are quite heterogeneous in this domain. Specifically, they show that children's comprehension of *wh*-questions depends on the type of questions and on the target language, as will be discussed below. As for production studies, one of the few that has been carried out so far, Guasti et al. (2012), shows that 5-year-old Italian-speaking children are able to produce both subject and object *wh*-questions, despite the fact that the comprehension of the latter has been shown to be particularly difficult (De Vincenzi et al. 1999), i.e., production may seem more advanced than comprehension. Interestingly, what children produced was not always what adults did.

Production tasks are more telling, since they reveal (i) whether children acquired and are able to utter the grammatical form of these structures, (ii) which kind of errors are produced (in case of incorrect questions), and crucially (iii) whether children have to rely on alternative strategies (in case of correct questions, but different from the target ones: avoidance strategies). Hence, we decided to carry out a task testing production.

Previous studies on *wh*-questions showed that these structures are mastered relatively early cross-linguistically (see Guasti 2017 for an overview), but several studies detected some asymmetries, both in production and in comprehension. These include:

- **argument type asymmetry**: subject *wh*-questions are acquired earlier, and children are more accurate in their production, than object *wh*-questions (see Yoshinaga 1996 and Sauerland et al. 2016 among many others for studies on comprehension, and Stavrakaki 2006 and Guasti et al. 2012 for studies on production). This asymmetry lasts longer for some languages, see e.g. Italian as opposed to English and Hebrew (De Vincenzi et al. 1999 for Italian, Avrutin 2000 for English and Friedmann et al. 2009 for Hebrew), depending on various features involved in the formation of questions;

- **complexity of the *wh*-phrase**: *who* type questions are easier than *which N* type questions (Ervin-Tripp 1970; Friedmann et al. 2009; Guasti et al. 2012; Sauerland et al. 2016). Moreover, according to Friedmann et al. 2009, there is an interaction between the argument type of the *wh*-phrase, and the complexity of the *wh*-phrase, to the point that in some languages the argument type asymmetry is evident only in *which*-questions.

The goals of our study are twofold: on the one hand there is a purely linguistic one, and on the other hand there is a methodological one. Let us spell these out more clearly in the form of research questions.

On the linguistic side, we were interested in understanding whether previously observed asymmetries also hold for the languages we took into account and could be detected with a production task. In this paper we will focus on two of them:

- RQ1a: Is there an asymmetry between subject *wh*-questions and object *wh*-questions?

- RQ1b: Is this asymmetry found in both *who* and *which N* questions?

- RQ2: Is there an asymmetry due to the complexity of the *wh*-phrase (*who* vs. *which N*)?

The second goal of the study is mainly methodological, and that is (i) to compare different types of languages, and (ii) to investigate less studied languages. As for the former point, the majority of the studies carried out so far – not only in the domain of *wh*-questions, but in any linguistic domain – focus on one language only, or compare at most two languages. There are a few exceptions, but mostly on the comprehension rather than the production side (see Sauerland et al. 2016 for *wh*-questions, Armon-Lotem et al. 2016 for passives – but see Varlokosta et al. 2016 on a study of clitic and pronoun production across 16 languages). However, a study designed to compare more than two languages allows us to identify language-specific factors and properties that help children to acquire these structures. In our case, some of the properties could involve the position of the *wh*-element (ex-situ in German, Yoruba and Hungarian, in-situ in Malayalam and Mandarin), or the presence of morphological case on the *wh*-word (in German, Hungarian and Malayalam).

For the second point, not only did we include under-studied languages, but some of the languages we tested were never investigated experimentally before in the domain of acquisition of *wh*-questions, to the best of our knowledge: Hungarian, Yoruba and Malayalam.

Even though more recent studies started to investigate acquisition with large-scale cross-linguistic experiments, these studies mostly focused on languages spoken in Europe (although typologically different ones). Here, instead, we offer a first start towards a more inclusive investigation. As will be explained below, this required development of materials as identical as we could across languages, and collaboration between researchers based in Germany and Italy (remote investigators) and local investigators working from several countries (China, Hungary, India and Nigeria).

In summary, our goals were hence to expand our knowledge on the acquisition of *wh*-question by (i) working on production, and (ii) working on less-studied languages.

The paper is organised as follows: section 2 focuses on some properties of each of the five languages we investigated. Section 3 describes the experiment itself and hence presents participants (3.1), materials (3.2) and the procedure we adopted (3.3). In section 4 we present some preliminary results, going back to the RQs presented in this section, and a first discussion is offered. Conclusions can be found in section 5.

2. Some properties of investigated languages

The languages we investigated represent five language families: Germanic (German), Uralic (Hungarian), Dravidian (Malayalam), Sino-Tibetan (Mandarin), and Niger-Congo (Yoruba). Different subsets of these languages pattern together with regard to different properties that relate to *wh*-questions. In order to give an overview let us first discuss some of the properties on which the languages we have investigated differ, focusing on the position of the *wh*-phrase and case-marking on the *wh*-phrase.

German is a verb final language (SOV word order) with the verb-second property (see e.g. Haider 2010). In the matrix clause, the finite verb appears in the second position, whether it is a declarative sentence or a *wh*-question. In a *wh*-question with a single *wh*-phrase, the *wh*-phrase must be placed in the sentence initial position (obligatory *wh*-movement).

(2)   Wen   jagt   der Elefant?
      who.acc chases nom elephant
      'Who is the elephant chasing?'

In the case of the items used in the present study, the experimental set-up was such that the verb was produced in present tense. This resulted in the main verb appearing in the second position (after the *wh*-phrase).

In German, the wh-phrase for 'who' is obligatorily case marked for Nominative (*wer*) and Accusative (*wen*) cases.[1] Thus, the case of the *wh*-phrase disambiguates which argument position it is associated with, as in (2) vs. (3).

(3)   Wer   jagt   die Giraffe?
      who.nom chases the giraffe
      'Who is chasing the giraffe?

For the *wh*-phrase for 'which', on the other hand, nominative and accusative case are overtly distinguished only for masculine singular head nouns (*welcher* vs. *welchen*). When the head noun is feminine singular, neuter singular, or plural, nominative and accusative forms are homophonous because of syncretism.

(4)   Masculine head noun

      a.   Welcher Affe   kratzt   das Gespenst?
           which.nom monkey.masc scratches the ghost
           'Which monkey is scratching the ghost?
      b.   Welchen Frosch   kratzt   die Maus?
           which.acc frog.masc scratches the mouse
           'Which frog is the mouse scratching?'

---

[1] Also Dative (*wem*), which is not relevant for this study.

(5) Feminine head noun
    Welche   Katze kitzelt das   Gespenst?
    which.nom/acc cat.fem tickles the.nom/acc ghost.neut
    'Which cat is tickling the ghost?' or
    'Which cat is the ghost tickling?'

Note that an example with singular nouns that are not masculine, for both subject and object, as in (5), results in ambiguity, because the definite determiner shows similar syncretism when the head noun is singular feminine or neuter. It was hence important that the question could be disambiguated in German: either one of the arguments is singular masculine so that case is morphologically expressed on either *wh*-phrase or the determiner of the in-situ argument, or there is a mismatch in number between the subject and object, so that verbal agreement disambiguates the type of *wh*-question. Target sentences like (5) were therefore excluded from the experiment (see Section 3.2).

    The canonical word order of Hungarian is SVO, although there has been some debate in the literature, given that the word order is relatively free (see Horvath 1986, Surányi 2006). As with German, Hungarian obligatorily places the *wh*-phrase to the pre-verbal position and marks the *wh*-element in object *who*-questions with the accusative case, as in (6-b). In the case of *which*-questions the marker of acc-case appears on the N as in (7-b).

(6) a. Ki kerget-i   a   zsiráf-ot?
        who chase-3SG the giraffe-acc
        'Who is chasing the giraffe?'
    b. Ki-t   kerget az elefánt?
        who-acc chase the elephant
        'Who is the elephant chasing?'

(7) a. Melyik majom vakar-ja   a szellem-et?
        which   monkey scratch-3SG the ghost-acc
        'which monkey is scratching the ghost?'
    b. Melyik béká-t vakar-ja   az egér??
        which   frog-acc scratch-3SG the mouse
        'which frog is the mouse scratching?'

Malayalam is a head-final language, and the canonical word order is SOV. In Malayalam *wh*-questions, the *wh*-phrase does not obligatorily appear in the sentence initial position, as they do in German and Hungarian (Jayaseelan 2001, Aravind 2018 and work cited therein). Instead, the *wh*-phrase remains in-situ, appearing in the subject, sentence initial position in subject *wh*-questions ((8-a) and (9-a)) and in the post-subject, pre-verbal position in object *wh*-questions ((8-b)

and (9-b)).[2] As German and Hungarian, Malayalam marks accusative case on the *wh*-element in object *wh*-questions, as in (8-b). In contrast, example (9-b) shows that in *which*-questions accusative case is marked on the noun. The *wh*-word for 'which' is only marked for case when the noun is dropped (example not shown here for the sake of space).

(8) a. aaɾ-aa   ɟiraafi-ne oot̪iykk-un̪n̪-e?
         who-cop giraffe-acc chase-prs.prog-nmlz
         'Who is chasing the giraffe?'

     b. aana   aaɾ-ey-aa   oot̪iykk-un̪n̪-e?
         elephant who-acc-cop chase-prs.prog-nmlz
         'Who is the elephant chasing?'

(9) a. eet̪i   koɾaŋŋan-aa preet̪at̪t̪i-ne coriy-un̪n̪-at̪i?
         which monkey-cop ghost-acc   scratch-prs.prog-nmlz
         'Which monkey is scratching the ghost?'

     b. ii eli   eet̪i   t̪avaʎe-ney-aaɳi pit̪iykk-un̪n̪-e?
         dem mouse which frog-acc-cop   catch-prs.prog-nmlz
         'Which frog is the mouse catching?'

In Mandarin Chinese, the base-generated word order is SVO. *Wh*-phrases in Chinese, as in Malayalam, do not appear in the sentence initial position (Huang 1982, Aoun & Li 1993). When the *wh*-phrase is the object of the predicate, it remains in-situ, appearing in the post-verbal position. Mandarin noun phrases and *wh*-phrasese are not marked with case.

(10) a. shuí zài zhuī chǎng-jǐng-lù?
          who ASP chase giraffe
          'Who is chasing the giraffe?'

      b. dà xiàng   zài zhuī shuí?
          big elephant ASP chase who
          'Who is the elephant chasing?'

(11) a. nǎ-zhī hóu-zǐ zài náo   yōu-líng?
          which monkey ASP scratch ghost
          'Which monkey is scratching the ghost?'

      b. lǎo shǔ   zài náo   nǎzhī qīngwā?
          old mouse ASP scratch which frog
          'Which frog is the mouse scratching?'

---

[2] The examples provided in (8) and (9) involve clefting, as this is the preferred way of asking *wh*-questions in Malayalam (see Jayaseelan 2001 for instance).

Yoruba, too, is a SVO language. As in German and Hungarian, the *wh*-phrase obligatorily appears in the sentence initial position, as shown in (12) and (13) (Sonaiya 1989).

(12) a. Ta ló ń lé àgùnfọn?
who foc prog chase giraffe
'Who is chasing that giraffe?'

b. Ta ni erin ń lé?
Who foc elephant prog chase
'Who is the elephant chasing?'

(13) a. Ọ̀bọ wo ló ń yún òjùjú lára?
monkey which foc prog scratch ghost loc.body
'Which monkey is scratching the ghost?'

b. Ọ̀pọ̀lọ́ wo ni eku ń yún lára?
toad which foc mouse prog scratch loc.body
'Which toad is the mouse scratching?'.

Although case is not marked on the *wh*-phrase itself, the form of the directly adjacent focus marker may disambiguate between a subject and an object question. While its general form is *ni* (cf. (12-b), (13-b)) there is an alternative form *ló* which only occurs in subject questions ((12-a), (13-a)). It is a fusion of the focus marker *ni* and a tense-marking high tone syllable (HTS) *ó* that usually appears between the subject and the verb (14) (cf. Awobuluyi 1975).

(14) Ta ni {erin-ín/erin-ó/erín} ń lé?
who foc elephant-HTS/elephant-HTS/elephant prog chase
'Who is the elephant chasing?'

As this fusion under adjacency is only possible once the subject is displaced the focus marker *ló* is only available in subject questions. Despite its frequent occurrence the fused form is however not obligatory (cf. (15) which is a version of (12-a) without a fused focus marker).

(15) Ta ni ó ń lé àgùnfọn?
who foc HTS prog chase giraffe
'Who is chasing the giraffe.'

3. Experiment
3.1. Participants

A total of 97 children aged between 3;10–6;7 were recruited for the study. Consent forms in accordance with the respective ethics requirements for each language were obtained from the parents or guardians of all children prior to test-

ing. For German 21 monolingual children were recruited from a kindergarten in Berlin. For Hungarian 23 children were recruited from a kindergarten in Budapest, of which 3 had to be excluded for not completing the task. Of the remaining 20 children, one had a bilingual exposure. For Malayalam 17 monolingual children were recruited from a kindergarten in Alappuzha, India, of which one was excluded for not completing the task. For Mandarin 24 children were recruited from a kindergarten in Suzhou, China, of which 2 had to be excluded for not completing the task. Of the remaining 22 children 8 had a bilingual exposure. For Yoruba 12 children were recruited from schools and from their homes in Ibadan, Nigeria. As it is common in this region, all of them were Yoruba-English bilingual with Yoruba being their dominant language in day-to-day communication at home and in school. More information on the participants whose performance was factored into the results is given in Table 1.

Table 1: Participants

Language	*N*	Age range	*M*	Location
German	21	3;10–5;9	5;2	Berlin
Hungarian	20	4;2–4;9	4;6	Budapest
Malayalam	16	4;2–5;0	4;7	Alappuzha
Mandarin	22	5;3–6;2	5;9	Suzhou
Yoruba	12	3;10–6;7	5;1	Ibadan

3.2. Materials

The experiment was adapted from the elicitation task in Guasti et al. (2012). In order to investigate children's production of *wh*-questions we used 30 pictures showing one or more protagonists (agents) carrying out an action affecting one or more other protagonists (patients). Of these, 12 were used to elicit a subject question (6 *who*, 6 *which*) and 18 were used to elicit an object question (6 *who*, 6 *what*, 6 *which*). The protagonists' characters were such that they are easily identifiable by children (animals and well-known human characters like witches, boys, girls, etc.). We also tried to ensure that the depicted action could be described with simple verbs, i.e. no particle verbs or multi-verb constructions. An example target question (in English) for each of the five question types was given in (1), and reported here as (16), We also used six additional pictures for a warm-up phase preceding the actual experiment.

(16) Target questions of each of the five question types

    a.   Who is pushing the ant?   (Who S; 6 items)
    b.   Who are the horses pulling?   (Who O; 6 items)
    c.   Which ladybug is waking the cats?   (Which S; 6 items)
    d.   Which duck is the rabbit pushing?   (Which O; 6 items)
   e.   What is the woman drawing?   (What O; 6 items)

The experiment was set up and piloted in German. It was then distributed to the local investigators for each additional language so that it could be made linguistically and culturally appropriate. This involved checking pictures and target sentences for valency and complexity of verbal expressions, ambiguities, uncommon or inappropriate protagonists and uncommon or inappropriate actions or pictures (for example, a picture with a cup, rather than what looks like a wine glass, was used in some languages).

3.3. Procedure

Children were tested in a quiet room, usually at the institution where they were recruited, with an experimenter and a second person present. The experimenter led the procedure and interacted with the child, while the second person manipulated a puppet and responded to the questions. The pictures were displayed using a presentation software on a computer. For the questions with a simple *wh*-word participants were presented with a picture part of which—either the agent or the patient of the verb—was occluded (see Fig. 1). For the questions containing a complex *wh*-word, i.e. '*which* N', the occluded picture was preceded by a non-occluded one showing two tokens of the to-be-questioned protagonist in order to make a *which* N question felicitous (see Fig. 2).
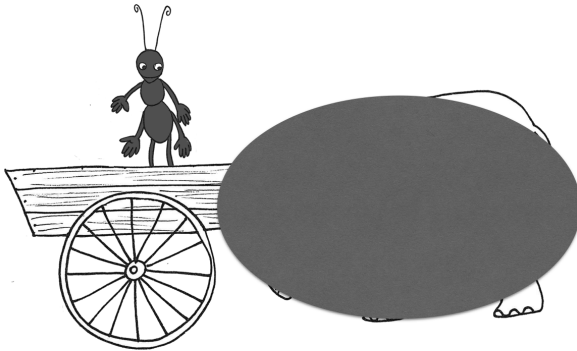
Figure 1: Stimulus picture for 'Who is pushing the ant?'
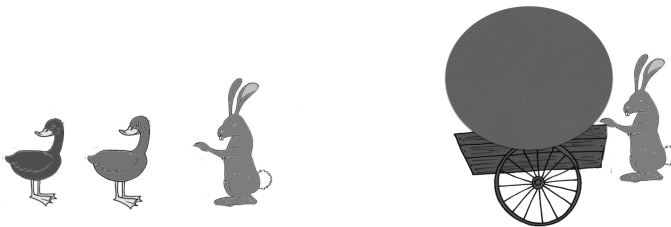
O7   O7

Figure 2: Picture sequence for 'Which duck is the rabbit pushing?'

Children were told that the puppet knew the identity of the protagonist behind the occlusion, but could not hear adults' voices. Children's task was to find out the identity by asking a question to the puppet. They were encouraged to use a full question including a verb (Thornton (1990), and others). Each picture was first described by the experimenter, replacing the occluded protagonist with an indefinite (*someone*, *something*). The child was then reminded that the puppet knew the identity of the indefinite and was prompted to ask a question. An example of a description and a prompt is given in (17-a) for a simple *wh*-word question and in (17-b) for a complex *wh*-word question.

(17) a. Look! There is an ant, and someone is pushing her. Puppet knows who. Ask puppet who.
       b. [first picture] Look! There is a brown duck, a green duck, and a rabbit. [second picture] The rabbit is pushing one of the ducks. Puppet knows which duck. Ask puppet which duck.

A warm-up phase consisting of six items allowing the child to understand and get used to the task was followed by the experimental phase consisting of 30 target items. The target items were presented in two blocks. The first block contained all 18 items with a simple *wh*-word (*who* S, *who* O, *what* O) in randomized order. The second block contained all 12 items with a complex *wh*-word (*which* S, *which* O) in randomized order. The order within both blocks was kept constant across children and languages as was the order of the blocks themselves. Each session was recorded with an audio-recorder for later transcription and analysis. The experimental sessions were conducted and the recordings transcribed, glossed, and entered into a spreadsheet by the local investigators (and their affiliates). Data were then coded as a joint effort between the local and remote researchers. We coded each response for whether it was the target response or not target but nonetheless correct, i.e. grammatical with a target, or close to target, interpretation. In addition, the type(s) of error(s) were coded. In this paper, we will focus exclusively on target responses, leaving correct responses for future investigation.

4. Preliminary results and discussion

Let us go back to the two research questions presented in the introduction:

- RQ1a: Is there an asymmetry between subject *wh*-questions and object *wh*-questions?

- RQ1b: Is this asymmetry found in both *who* and *which N* questions?

- RQ2: Is there an asymmetry due to the complexity of the *wh*-phrase (*who* vs. *which N*)?

Here we report some preliminary results based only on target answers.

Concerning the asymmetry between argument types [RQ1a], we constructed models for each language (using the lme4 package with R, Bates et al. 2015), with argument type (subject vs. object) as a predictor, taking data only from the *who*-question and *which*-question items. We found a subject advantage in German (t = 6.619, p < .01), Malayalam (t = 2.831, p < .01), and Yoruba (t = 2.831, p < .01), an object advantage in Mandarin (t = −5.907, p < .01), and no difference in Hungarian (t = 0.39, p = .696).

Next, we separated *wh*-questions by argument-type of the *wh*-phrase[RQ1b], and constructed models with argument type as a predictor for each language. A subject advantage emerged with *which*-questions but not with *who*-questions in Hungarian (which: t = 2.665, p < .01; who: t = −0.173, p = .863) and Malayalam (which: t = 2.310, p < .05; who: t = 1.000, p = .318); a subject advantage in both conditions was found in German (which: t = 4.219, p < .01; who: t = 6.109, p < .01) and Yoruba (which: t = 2.702, p < .01; who: t = 3.249, p < .01); while an object advantage emerged in Mandarin for both types of questions (which: t = −3.770, p < .01; who: t = −2.865, p < .01. An overview of the asymmetries on argument type is presented in Table 2.

Table 2: RQ1: Asymmetries between subject and object in *who/which*-questions across languages

	who-questions	*which*-questions
German	subj > obj	subj > obj
Hungarian	subj = obj	subj > obj
Malayalam	subj = obj	subj > obj
Mandarin	obj > subj	obj > subj
Yoruba	subj > obj	subj > obj

Moving to the asymmetry between *who*-questions and *which*-questions [RQ2], there seems to be an advantage of *who*-questions in German, and an advantage of *which*-questions in Yoruba, while no difference emerged for the other three languages. Table 3 summarizes the main findings on *who* vs. *which N* asymmetries.[3]

Table 3: RQ2: Asymmetries between *who* and *which N*

German	who > which (t = 3.166, p < .01)
Hungarian	who = which (t = −1.268, p = .205)
Malayalam	who = which (t = −0.845, p = .399)
Mandarin	who = which (t = −0.466, p = .642)
Yoruba	who < which (t = −2.235, p < .05)

---

[3] For each language, we constructed models with the type of *wh*-phrase (*who* vs. *which N*) as the predictor for this analysis.

Recall that evidence in favor of these two asymmetries have been provided in several previous acquisition studies, both in comprehension and in production (see Yoshinaga (1996); Stavrakaki (2006); Guasti et al. (2012); Sauerland et al. (2016), among others, on the argument type asymmetry, and Ervin-Tripp (1970); Friedmann et al. (2009); Guasti et al. (2012); Sauerland et al. (2016) on the asymmetry due to the complexity of the *wh*-phrase). However, most of the previous studies carried out so far focused on one (European) language. One of the goal of the current cross-linguistic study was to find out whether these asymmetries could be found in a production task in several unrelated and less-studied languages.

The preliminary results show that the argument type asymmetry and the asymmetry due to the complexity of the *wh*-phrase were not detected in all of the languages. In particular, we found three patterns for the former. First, subject questions were easier than object questions in both German and Yoruba, and such asymmetry was found in both *who* and *which* questions. Second, for Hungarian and Malayalam subject *wh*-questions were not produced more accurately than object *wh*-questions in case of *who* questions, but the asymmetry was found in *which* questions. Third, Mandarin, in contrast to previous findings on other languages, showed a 'reversed' asymmetry: object *wh*-questions were produced more accurately than subject *wh*-questions, in both *who* and *which* questions.

As for the second type of asymmetry, the well-known 'who' advantage emerged in German, while Yoruba showed the opposite pattern. In addition, *who*-questions were not produced more accurately than *which*-questions in Hungarian, Mandarin and Malayalam. Why these asymmetries were not observed in all of the languages investigated in this study and why these five languages behaved so differently are questions we plan to pursue in future research.

As the reader may remember, German, Hungarian, Malayalam, Mandarin and Yoruba vary in the way that they form *wh*-questions. For instance, the position of the *wh*-element is not the same across these languages (ex-situ in German, Yoruba and Hungarian, in-situ in Malayalam and Mandarin). In addition, in some but not all of these languages morphological case is present on at least one *wh*-word (German, Hungarian and Malayalam). Some of these properties relevant for the formation of *wh*-questions are summarized in Table 4.

Table 4: Some properties relevant for the formation of *wh*-questions

Language	case on *wh*-phrase	*wh*-fronting	verb-final	1-*wh*
German	yes/no	yes	yes	no
Hungarian	yes	yes	no	no
Malayalam	yes	no	yes	no
Mandarin	no	no	no	no
Yoruba	no	yes	no	no

One of our goals in future work is to investigate whether these linguistic factors affect the production of target *wh*-questions by children across languages. In

fact, in a recent crosslinguistic comprehension study, Sauerland et al. (2016) found that two of these facilitated 5-year-old children's comprehension of *wh*-questions: (i) overt case morphology on the *wh*-phrase and (ii) a single lexical item for 'who' and 'which' (abbreviated as 1-*wh* in the above table). An additional feature that has been found to facilitate the comprehension of *wh*-questions was the use of synthetic verbal forms.

Future research will also be devoted to the investigation of an additional research question spelled out below:

- RQ3: Are there alternative strategies used by children in order to produce a *wh*-question?

Investigating the type of errors children produced as well as the alternative strategies they used is a crucial information, in that it may reveal which type of structure is more complex. Moreover, it is possible that errors or alternative strategies will mainly be found with some type of questions (object but not subject, for instance), further underlining the fact that in some languages, these are more complex.

5. Conclusion

This paper provides the first glimpse of the results from a *wh*-question production study, conducted in five languages. The experiment was designed to examine whether some of the well-documented asymmetries in comprehension (namely, the argument asymmetry and the complexity asymmetry) are observed in production as well. Our preliminary findings show that languages exhibit different patterns with respect to the effect of argument and complexity.

The present production study will provide a rich set of cross-linguistic data from less-studied language, some of which were never investigated experimentally before in the domain of acquisition of *wh*-questions (Hungarian, Malayalam and Yoruba). Systematic investigation of less studied languages with different grammatical properties should shed a new light on understanding what influences children's production of *wh*-questions. More detailed analyses of the types of errors children make – as well as the alternative strategies they use – is our next step.

References

Aoun, Joseph & Audrey Li. 1993. *The syntax of scope*. Cambridge, MA: MIT Press.

Aravind, Athulya. 2018. Licensing long-distance wh-in-situ in Malayalam. *Natural Language & Linguistic Theory* 36(1). 1–43.

Armon-Lotem, Sharon et al. 2016. A large-scale crosslinguistic investigation of the acquisition of passive. *Language Acquisition* 23(1). 27–56.

Avrutin, Sergey. 2000. Comprehension of discourse-linked and non-discourse-linked questions by children and Broca's aphasics. In *Language and the brain: Representations and processing*, 295–313. Academic Press.

Awobuluyi, Oladele. 1975. On "The Subject Concord Prefix" in Yoruba. *Studies in African Linguistics* 6(3). 215–238.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48.

De Vincenzi, Marica, Lisa Arduino, Laura Ciccarelli & Remo Job. 1999. Parsing strategies in children comprehension of interrogative sentences. In Sebastiano Bagnara (ed.), *Proceedings of the European Conference on Cognitive Science*, 301–308.

Ervin-Tripp, Susan. 1970. Discourse agreement: How children answer questions. In *Cognition and the development of language*, 79–107. New York: Wiley.

Friedmann, Naama, Adriana Belletti & Luigi Rizzi. 2009. Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua* 119. 67–88.

Friedmann, Naama & Rama Novogrodsky. 2011. Which questions are most difficult to understand? The comprehension of wh questions in three subtypes of SLI. *Lingua* 121. 367–382.

Guasti, Maria Teresa. 2017. *Language acquisition. The growth of grammar.* MIT Press.

Guasti, Maria Teresa, Chiara Branchini & Fabrizio Arosio. 2012. Interference in the production of Italian subject and object wh-questions. *Applied Psycholinguistics* 33(1). 185.

Haider, Hubert. 2010. *The syntax of German*. Cambridge: Cambridge University Press.

Horvath, Julie. 1986. *Focus in the theory of grammar and the syntax of Hungarian*. Foris.

Huang, Cheng-Teh James. 1982. *Logical relations in Chinese and the theory of grammar*. MIT dissertation.

Jayaseelan, Karattuparambil A. 2001. Questions and question-word incorporating quantifiers in Malayalam. *Syntax* 4(2). 63–93.

Sauerland, Uli et al. 2016. How do 5-year-olds understand questions? Differences in languages across Europe. *First Language*. 1–34.

Sonaiya, Remi. 1989. Wh-Movement and Proper Government in Yoruba. In Paul Newman & Robert D. Botne (eds.), *Current Approaches to African Linguistics (vol. 5)*, 109–126. Dordrecht: Foris.

Stavrakaki, Stavroula. 2006. Developmental perspectives on Specific Language Impairment: Evidence from the production of wh-questions by Greek SLI children over time. *Advances in Speech Language Pathology* 8(4). 384–396.

Surányi, Balázs. 2006. Scrambling in Hungarian. *Acta Linguistica Hungarica (Since 2017 Acta Linguistica Academica)* 53(4). 393–432.

Thornton, Rosalind. 1990. *Adventures in long-distance moving: The acquisition of complex Wh-questions*. University of Connecticut dissertation.

Varlokosta, Spyridoula et al. 2016. A cross-linguistic study of the acquisition of clitic and pronoun production. *Language Acquisition* 23(1). 1–26.

Yoshinaga, Naoko. 1996. *Wh-questions: A comparative study of their form and acquisition in English and Japanese*. University of Hawai'i at Manoa.

Proceedings of the 46th annual Boston University Conference on Language Development

edited by Ying Gong
and Felix Kpogo

Cascadilla Press   Somerville, MA   2022

Ordering information

To order a copy of the proceedings or to place a standing order, contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, sales@cascadilla.com, www.cascadilla.com