

## The acquisition of *wh*-ex-situ and in-situ questions in Mandarin

Our aim is to investigate the production of subject and object *wh*-questions in Mandarin, an in-situ language, and explain children's production of (pseudo-)cleft questions, when a simpler structure with *wh*-in-situ is available and the specific pattern of the subject/object (S/O) asymmetry.

In languages displaying overt *wh*-movement, subject questions are easier to comprehend and produce than object ones (De Vincenzi et al. 1999; Guasti et al. 2012), with this S/O asymmetry lasting longer time in which- than who-questions (Friedmann et al. 2009). Different explanations have been offered (Friedmann et al. 2009; Jakubowicz 2004; Guasti et al. 2012), which all have in common that *wh*-movement is costly and thus predict (1) that children should prefer *wh*-in-situ, if this is an option, and (2) display no S/O asymmetry in *wh*-in-situ questions. Evidence from French production (a.o., Haman 2000; Zuckerman & Hulk 2001) confirms the first prediction, while from Akan production (Omane & Höhle 2021) displays a more complex picture with younger children producing more ex-situ questions (like adults) than older ones. However, these studies offer a partial picture, because they do not include which-questions and do not address prediction (2). In addition, they do not explain why children produce ex-situ questions at all. In this paper, we address all these issues, based on the Mandarin production of *wh*-questions.

Mandarin features *wh*-in-situ and, according to theoretical work, allows *wh*-elements to be in sentence initial position in clefts sentences (Cheung 2008, 2014). We conducted a production experiment using a one covered character task (Guasti et al. 2012) and involved 100 typically developing Mandarin-speaking children ranging in age from 3;0 to 6;11 dividing them in 4 age groups (age 3, 4, 5 and 6) plus 32 adults. We used a 2 (type of *wh*-type: who vs which) by 2 (type of argument: S vs O) by 5 (age groups) design (with 6 items per condition) and analyzed data with a general linear mixed-effects model (with subject and items as random factors).

Descriptive results are reported in Table 1. Here, we focus on simple and complex *wh*-questions. As it can be seen, children as well as adults produced *wh*-in-situ more frequently than complex *wh*-questions (cleft questions, as in (1) or pseudo cleft, as in (2)). These increased with age, with 6-year-olds performing not differently from adults. *Wh*-in-situ questions were more frequent in object than subject position for 5, 6, and adults regardless of the *wh*-type, while complex *wh*-questions were more frequent for subject than for object. Participants also produced other grammatical modified questions (passive, questions with a change of the *wh*-type or a pronoun rather than a noun), and errors, which are conflated together in Table 1. These results are supported by the statistical analyses (see Table 2).

These findings lend support to the Derivational theory of complexity, whereby external merge is more complex than internal merge (Jakubowicz 2004). But, at the same time, they raise the questions of why ex-situ questions are produced at all when a simpler strategy is available and why object-questions favor *wh*-in-situ, while subject-questions favor *wh*-ex-situ.

We explain these facts as follow. Cross-linguistically, (pseudo-)clefts interpret exhaustive focus (Kiss 1998), and this is also confirmed for Mandarin (Hole 2011; Pan 2019). When the reference of the inquired argument is known to be unique and exhaustive (as in our design), the speaker can choose to overtly focalize it, independently from its being a +Q(uestion) element. The overt focus marking by using (pseudo-)cleft would increase the transparency of the question, facilitating the interlocutor to narrow down the range of the answers, but at the same time also increase the cost of production for the speaker, since the formation of (pseudo-)clefts requires more steps of movement. The participants thus kept balancing between pragmatic transparency and syntactic simplicity and for children the aim of transparency become more accessible as they get access to more complex syntax or more computational resources. As for the S/O asymmetry, following Belletti (2008) and Paul & Whitman (2008), we hypothesize that subject and object (pseudo-)clefts in Mandarin have different derivations, with object ones more costly than subject ones either because they give rise to intervention effects or require pied-piping. Speakers thus value simplicity more than transparency in object-question production and use fewer complex questions.

- (1) Shì nǎ liǎng gè nánhái zài dǎ zhè gè nǚhái? (Cleft question)  
BE which two CL boy ASP hit this CL girl  
“Which two boys is it that are hitting this girl?”
- (2) Zhuī zhè liǎng gè nánshēng dē (rén) shì shuí yā? (Pseudo-cleft question)  
chase this two CL boy DE person BE who Q  
“Who is the person that is chasing these two boys?”

**Table 1.** Percentages (%) and raw scores of *wh*-questions featuring *wh*-in-situ (Simple), (pseudo)cleft (Complex), and other structures or errors (Other R).

Group	Who		Which	
	Subject	Object	Subject	Object
<b>Age 3</b>				
Simple	20.5 (32/156)	21.1 (33/156)	7.1 (11/156)	7.7 (12/156)
Complex	6.4 (10/156)	0 (0/156)	4.5 (7/156)	0 (0/156)
Other R	73.1 (114/156)	78.9 (123/156)	88.4 (138/156)	92.3 (144/156)
<b>Age 4</b>				
Simple	28.2 (44/156)	30.1 (47/156)	7.1 (11/156)	19.2 (30/156)
Complex	25.1 (39/156)	5.8 (9/156)	17.9 (28/156)	2.6 (4/156)
Other R	46.7 (73/156)	64.1 (100/156)	75.0 (117/156)	78.2 (122/156)
<b>Age 5</b>				
Simple	41.0 (59/144)	52.8 (76/144)	29.9 (43/144)	61.1 (88/144)
Complex	22.2 (32/144)	8.3 (12/144)	16.7 (24/144)	3.5 (5/144)
Other R	36.8 (53/144)	38.8 (56/144)	53.4 (77/144)	35.4 (51/144)
<b>Age 6</b>				
Simple	16.0 (23/144)	55.6 (80/144)	23.6 (34/144)	67.4 (97/144)
Complex	56.2 (81/144)	22.9 (33/144)	48.6 (70/144)	17.4 (25/144)
Other R	27.8 (40/144)	21.5 (31/144)	27.7 (40/144)	15.2 (22/144)
<b>Adults</b>				
Simple	50.0 (98/196)	63.8 (125/196)	39.3 (77/196)	62.3 (122/196)
Complex	45.4 (89/196)	27.1 (53/196)	38.8 (76/196)	15.8 (31/196)
Other R	2.6 (5/196)	7.1 (14/196)	19.9 (39/196)	19.9 (39/196)

**Table 2.** Summary of selected statistical analysis with linear mixed-effects models.

**Wh-in situ questions:** main effects of Group (all  $p < .001$ ), *wh*-element ( $p < .001$ ) and the interaction between Group and *wh*-element, and Group and Argument Type (all  $p < .01$ )

	Estimate $\beta$	SE	Wald Z	p
(Intercept)	-3.39	0.49	-6.94	= .000**
Group (Age 3 vs 5) $\times$ <i>wh</i> -element	-1.87	0.47	-3.97	= .000***
Group (Age 3 vs 6) $\times$ <i>wh</i> -element	-2.22	0.49	-4.56	= .000***
Group (Age 3 vs Adult) $\times$ <i>wh</i> -element	-1.30	0.47	-2.75	= .006**
Group (Age 3 vs 4) $\times$ Argument Type	-1.30	0.60	-2.17	= .03*
Group (Age 3 vs 5) $\times$ Argument Type	-1.64	0.54	-3.07	= .002**
Group (Age 3 vs 6) $\times$ Argument Type	-2.64	0.56	-4.71	= .000***
Group (Age 3 vs Adult) $\times$ Argument Type	-1.66	0.54	-3.10	= .002**

Reference level for Group = Age 3, *wh*-element = *which*, Argument Type = subject; \*\*\* $p < .001$ , \*\* $p < .01$ . \* $p < .05$ .

**Complex wh-questions:** main effect of Group (all  $p < .001$ ), *wh*-element ( $\beta = 0.54$ , SE = 0.11, Wald Z = 4.84,  $p < .001$ ) and Argument Type ( $\beta = 2.02$ , SE = 0.14, Wald Z = 14.68,  $p < .001$ ).

#### Selected references

- Cheung, C. C. H. (2008). *Wh-fronting in Chinese*. Doctoral Dissertation. Los Angeles (CA): USC.
- Cheung, C. C.-H. (2014). *Wh-fronting and the left periphery in Mandarin*. *Journal of East Asian Linguistics*, 23 (4), 393–431.
- Friedmann, N., Belletti, A., & Rizzi, L. (2009). Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua*, 119, 67–88.
- Guasti, M. T., Branchini, C., & Arosio, F. (2012). Interference in the production of Italian subject and object *wh*-questions. *Applied Psycholinguistics*, 33:185–223.
- Hole, D. (2011). The deconstruction of Chinese shì...de clefts revisited. *Lingua*, 121(11), 1707–1733.
- Jakubowicz, C. (2004). Is Movement costly? The grammar and the processor in language acquisition. *Presentation JEL (Journé'e d'Etudes Linguistiques)*, University of Nantes.
- Omane, P. O., & Höhle, B. (2021). Acquiring syntactic variability: The production of *wh*-questions in children and adults speaking Akan. *Frontiers in Communication*, 6, 604951.
- Pan, V. J. (2019). *Architecture of The Periphery in Chinese: Cartography and Minimalism*. London and New York: Routledge.
- Paul, W., & Whitman, J. (2008). *Shì ... de* focus clefts in Mandarin Chinese. *The Linguistic Review*, 25(3-4), 413–451.
- Zuckerman S., & Hulk A. (2001). Acquiring optionality in French *wh*-questions: An experimental study. *Revue Québécoise de Linguistique*, 30(2), 71–97.