Workshop on Heritage Language Acquisition
September 19th-20th, 2016
Tromsø, Norway

Wh-Question comprehension in Heritage English: Attrition, lack of resources or processing breakdown?

Efrat Harel, Sharon Armon-Lotem and Irena Botwinik

1. Introduction

Object *wh*-questions are more difficult for children to comprehend than **subject** *wh*-questions ((1b, d) vs. (1a, c)).

The D-linked ***which*-questions** are more difficult than ***who*-questions** ((1b) vs. (1d)), resulting in children's chance performance of the former (Avrutin 2000; Freidmann, Belletti & Rizzi 2009; Goodluck 2005, 2010).[1]

(1) a. Who t scratches the dog? S-who
b. Who does the dog scratch t ? **O-who**
c. Which cat t scratches the dog? S-which
d. Which cat does the dog scratch t ? **O-which**

Focus of the study: comprehension of object *wh*-questions (1b, d) in sequential bilingual acquisition with English as a Heritage Language (HL) and Hebrew as the second language (L2).

Derivation of Hebrew object *wh*-questions
Hebrew object *wh*-questions are **derived like their English counterparts** via movement of *wh*-phrase to Spec-CP. The relevant difference between the two languages: Hebrew lacks the auxiliary *do* (2).

(2) Hebrew object *wh*-questions
a. *et mi ha-kelev soret*?
acc who the-dog scratches
'Who does the dog scratch?'
b. *et eize xatul ha-kelev soret*?
acc which cat the-dog scratches
'Which cat does the dog scratch?'

An intriguing finding of our study: Comprehension of object *which*-questions in Heritage English was significantly worse than in L2 Hebrew.

Main claim: The lack of processing resources required in computing the reference of *which*-questions affects the processing of these structures in both languages, aggravated in English due to the auxiliary *do* that has no equivalent in L2 Hebrew.

2. Methods

Participants
Two typically-developing bilingual groups: 15 **younger**, aged 4;4-5;4, and 20 **older**, aged 5;5-6;4, with length of exposure (LoE) to Hebrew of at least 1;6 years, and with age of onset (AoO) not later than 4;0.[2] There is a significant difference between the groups in chronological age and LoE but not in AoO (Table 1).

[1] There is no significant difference between subject *who*- and *which*-questions.
[2] There was also a control group of 20 Hebrew monolinguals (10 from each age group).

Table 1. The division of the 35 bilinguals according to age, LoE, and AoO (in months)

Age group	Chronological age		LoE		AoO	
	Mean (SD)	**Range**	**Mean (SD)**	**Range**	**Mean (SD)**	**Range**
4;4-5;4	58.73 (3.2)	52-64	28.2 (12.7)	18-56	30.53 (13.8)	3-46
5;5-6;4	72.2 (2.98)	67-76	41.05 (15.1)	20-74	31.15 (14.5)	0-51

Task and materials
Comprehension of *wh*-questions was tested via a selection task (pictures from Cohen 2008, Figure 1). The task included 40 questions, 20 targeting subject questions and 20 targeting object questions with both *who*- and *which*-questions. The habitual aspect was used in all of the English questions, and the present tense in the Hebrew questions:

Hebrew	English
a. *et mi ha-kelev soret*?	*Who does the dog scratch*?
acc who the-dog scratches	
b. *et eize xatul ha-kelev soret*?	*Which cat does the dog scratch*?
acc which cat the-dog scratches	

Figure 1.

3. Results
Since subject questions yielded high levels of success in both languages, we focus **only** on **object questions**.
Since no correlation was found between the bilinguals' performance and their LoE or their AoO, the results focus only on the division into two age groups.

3.1 Object *Who*-questions
Who-questions yielded a between-age-group difference both in the subjects' HL English and in Hebrew (the older participants were better than the younger participants). In English: 74% vs. 54% ($t(33)=-2.24$, $p<0.05$); in Hebrew: 81% vs. 59% ($F(1,33)=15.31$, $p<0.001$). No significant main effect was found for language for the object *who* questions ($F(1,33)= 1.96$, $p=0.171$).

Table 2. Mean levels of success on *who*-questions in both languages (percentage correct out of all responses**)**

	OBJECT *WHO*	
	English (SD)	**Hebrew (SD)**
4;4-5;4	***54* (31.12)**	***59.06* (13.37)**
5;5-6;4	***74.38* (22.79)***	***81* (18.32)***
Total	*65.65* (28.17)	*71.6* (19.55)

Note: We assume the chance level to be at 33% for random guessing between three possible agents: the target character, the theta-role reversal character, and the wrong patient character. This provided us with a chance level of 6 or fewer correct responses out of 10. That is only 7 and above is above chance. Tables 3 and 4 present the distribution of chance levels within the two groups.

Table 3. Distribution of chance levels within the YOUNGER group

	ENGLISH		HEBREW	
	N	**percent**	**N**	**percent**
Below chance (0-32%)	**4**	26.7	**0**	0
At chance (33-69%)	**5**	33.3	**12**	80
Above chance (70-100%)	**6**	40	**3**	20

In order to check whether the distribution is different from the one of expected frequencies, we performed chi-square tests. The tests for the younger group yielded a significant difference only in Hebrew ($\chi 2$ (1)=5.4, $p<0.05$). **In Hebrew, most children are at chance level while in English, they are evenly distributed across the levels.**

Table 4. Distribution of chance levels within the OLDER group

	ENGLISH		HEBREW	
	N	**percent**	**N**	**percent**
Below chance (0-32%)	**2**	10	**0**	0
At chance (33-69%)	**4**	20	**3**	15
Above chance (70-100%)	**14**	70	**17**	85

Chi-square tests for the older groups showed a significant difference in both English ($\chi 2$ (2)=12.4, $p<0.01$) and Hebrew ($\chi 2$ (1)=9.8, $p<0.01$). **In both languages most of the participants (17/20 in Hebrew and 14/20 in English) are above chance on object *who*-questions.**

Error analysis reveals that the most common error is the theta-reversal error, and it goes down with age (Table 5).

Table 5. Reversal errors in *who*-questions in both languages (percentage out of all responses)

	ENGLISH		HEBREW	
	Mean	**SD**	**Mean**	**SD**
4;4-5;4	36.66	29.92	24.13	14.51
5;5-6;4	18.55	18.69	13	15.92
Total	26.31	25.42	17.77	16.11

Table 6. 'Other' errors (choice of the middle figure in the picture) (percentage out of all responses)

	ENGLISH		HEBREW	
	Mean	**SD**	**Mean**	**SD**
4;4-5;4	9.33	9.61	16.8	14.01
5;5-6;4	7.05	8.03	6	6.80
Total	8.03	8.68	10.62	11.66

3.2 Object *which*-questions

Unlike ***who***-questions, ***which***-questions yielded no between-age-group difference (Table 7).[3]

Table 7. Mean levels of success on ***which***-questions in both languages (percentage correct out of all responses)

	OBJECT *WHICH*	
	English (SD)	**Hebrew (SD)**
4;4-5;4	***26.66* (22.25)**	***49.33* (19.44)***
5;5-6;4	***31.55* (25.39)**	***56* (30.67)***
Total	***29.46*** (23.88)	***53.14*** (26.32)*

A between-language difference was observed for each age group and for the full sample, $F(1,33)=26.95$, $p<0.001$).

[3] An intriguing difference was attested on **Hebrew** object ***who*-questions** between young bilinguals and their **monolingual** peers (59% vs.77% correct). Among bilinguals there was a between-age-group difference both in Hebrew and in English, while there was no difference between the two age groups among the monolinguals: (77% success for younger, 87% for older monolinguals).

Tables 8 and 9 present chance level distribution for object ***which***-questions.

Table 8. Distribution of chance levels in English and Hebrew *which*-questions within the YOUNGER group

	ENGLISH		HEBREW	
	N	**percent**	**N**	**percent**
Below chance (0-32%)	**9**	66	**5**	33.3
At chance (33-69%)	**5**	33.3	**6**	40
Above chance (70-100%)	**1**	6.7	**4**	26.7

Table 9. Distribution of chance levels in English and Hebrew *which*-questions within the OLDER group

	ENGLISH		HEBREW	
	N	**percent**	**N**	**percent**
Below chance (0-32%)	**12**	60	**6**	30
At chance (33-69%)	**6**	30	**6**	30
Above chance (70-100%)	**2**	10	**8**	40

Chi-square tests showed a significant difference only in English, for both younger ($\chi 2$ (2)=6.4, $p<0.05$) and older ($\chi 2$ (2)=7.6, $p<0.05$). **In English, almost all younger children are either below or at chance level, while in Hebrew they are evenly distributed across the levels**.[4]

Error analysis: In both languages, the most common error was the reversal error, but in English, choosing the wrong (identical) figure took place significantly more than in Hebrew ($t(34)=-6.13$, $p<0.001$) (Table 10).

Table 10. Reversal errors in *which*-questions in both languages (percentage out of all responses)

	ENGLISH		HEBREW	
	Mean	**SD**	**Mean**	**SD**
4;4-5;4	*63.33*	22.25	*34.66*	19.22
5;5-6;4	*58.83*	24.21	*32*	25.87
Total	**60.76**	23.16	**33.14***	22.97

Table 11. 'Other' errors in *which*-questions in both languages (percentage out of all responses)

	ENGLISH		HEBREW	
	Mean	**SD**	**Mean**	**SD**
4;4-5;4	*10*	11.33	*16*	15.49
5;5-6;4	*9.61*	9.58	*12*	15.76
Total	**9.77**	10.21	**13.71**	15.54

To summarize, the performance on object ***who***-questions suggests that the children in the study have the syntactic knowledge necessary to comprehend object questions but are still in the process of developing the capacity to handle the unique aspects of the interpretation of ***which***-questions, a difficulty which is enhanced in their HL, English.

4. Discussion

The comprehension of object ***who***-questions improves with age (around 56% by younger; above 70% correct by older) → children's performance on it is related to the need to hold the *wh*-phrase longer (de Vincenzi et al. 1991).

What is the source of difficulty underlying the comprehension of *which*-questions? Although adults, unlike children do not err on object *which*-questions, studies have shown that even for them the *wh*-dependency is more demanding for *which*-questions than for *who*-questions (e.g. Kaan et al. 2000).

[4] In addition, chi-square tests comparing between the age groups within each language showed a significant difference in favor of the older subjects only for Hebrew object ***who***-questions.

4.1 Set restriction as the source of difficulty
The difficulty of the *which*-questions might have two sources:
(i) the need to establish the link to the previous discourse (The Discourse (D)-linking Hypothesis, Avroutin 2000; Pesetsky 1987, 2000; Shapiro 2000).
or
(ii) it might be caused by the mechanism of "set-restriction" inherent to *which*-questions (Donkers et al. 2013, Friedmann et al. 2009, Goodluck 2010).

(ii) seems more promising, based on the following.
Processing of **generic** *which*-questions (e.g. *which person/animal*) by children and adults is on a par with *who*-questions (3a, b), rather than with the non-generic *which*-questions (3c) (Donkers et al. 2013 for adults, Goodluck 2005 for children).

(3) a. Who does the dog scratch t ? *who*-question
b. Which animal does the dog scratch t ? generic *which*-question
c. Which cat does the dog scratch t ? non-generic *which*-question

If D-linking was the crucial part of the computation of these questions, both generic and non-generic *which*-questions (3b, c) should have patterned together, contrary to fact.

An interrogative phrase like *which person/which animal* is **semantically identical** to the interrogative phrase *who*, both mean: (for) which x, x (human/animate); *who* and *which animal* refer to the same non-specific, set of entities.

→ the computation involved in **set restriction** underlies the difficulty in processing non-generic *which*-questions.

But why?

4.2 Reference-set computation all over again
Inspiration
With restricted sets, listeners keep all possible alternatives active in memory, leading to some overload in the **working memory** resources (Cowles 2003, chapter 3).[5] We take Cowles' suggestion to be about computational complexity involved in set restriction. This, in turn, is reminiscent of another kind of computational complexity, the **reference-set computation** that is argued to be at play in the computation of meaning in a variety of structures: quantifier scope ambiguities (4), reference resolution of ordinary pronouns (5), scalar implicatures (6) (Grodzinsky & Reinhart 1993, Reinhart 1999, Chierchia et al. 2001, Gualmini et al. 2001, respectively):

(4) A teacher will interview every student.
Meaning 1 (easily accessible): there is **some teacher** and s/he will interview **every student**
Meaning 2 (much more difficult to access): for **every student**, there is **some teacher** that will interview him/her.

(5) Mama Bear is touching her.
Children accept the reading whereby *Mama Bear* and *her* are coreferential (Chien & Wexler, 1990).

(6) Every boy chose a skate-board or a bike.
Children do not accept the inclusive-*or* reading (Chierchia et al. 2001).

Reference set computation involved in (one of the meanings of) (4)
Meaning 1 is easy to access, it can be computed directly from the S-Structure (spell-out representation) of (4), as the scope relations are on a par with c-command relations between the two quantifiers (*some* c-commands *every*).
To get meaning 2, an operation of QR (Quantifier Raising) should be performed (after spell-out/at LF): *every student* should be moved from its surface/spell-out position to a position where it c-commands the existential quantifier, *some teacher* (Reinhart 1993/1997).

[5] It is nowadays generally assumed that the major function of the working memory system is to temporarily store the outcomes of intermediate computations when problem solving, and to perform further computations on these temporary outcomes (e.g. Baddeley 1986). There are no claims that children's general memory resources are limited. The huge amount of information that children manage to learn suggests, in fact, the opposite.

QR is **not** a derivationally **necessary** operation → it will be performed only if there is no other way to achieve this meaning. How can we know that there is no other way? We have to construct a reference-set including all the possible derivations and to compare them. This is difficult, causing a conscious effort for adults, and might cause a processing breakdown for children (Grodzinsky & Reinhart 1993, Reinhart 1999).[6]

Proposal: The process of set restriction is akin to reference-set computation. Both types of computation involve holding several alternatives and choosing between them. This causes a great burden on the working memory, presenting a processing challenge for adults, and probably exceeding preschool children's processing resources.

Failure to execute some computation → some strategy enabling to bypass the difficulty
What is the strategy in our case?

4.3 The strategy employed by children and its consequences

Children ignore the *which* element, viewing an interrogative *which*-DP as a non-interrogative, lexical $DP_{[-wh]}$.
This has consequences for the processing of the questions including such a DP.

4.3.1 Theta-role reversal in Hebrew and English

Working assumption: Processing is done according to predicates rather than according to words (Pritchett 1992; Siloni 2014); the syntactic structure is formed only upon the arrival of the verb.

(7) *et eize xatul ha-kelev soret*?
acc which cat the-dog scratches
'Which cat does the dog scratch?'

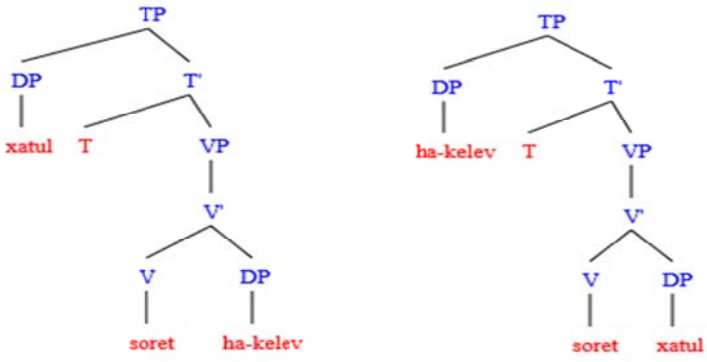
Processing of (7)

As the main verb enters the processor, there are two lexical $DPs_{[-wh]}$ waiting to be incorporated in the structure (*which*-DP, and "subject" DP). Any of them can be assigned θ1 (Agent) or θ2 (Theme) randomly, resulting sometimes in a reversal error.

Store — Structure formation

$[_{DP}$ eize-xatul] (the cat)

$[_{DP}$ ha-kelev] (the dog)



$[_{V}$ soret] θ1, θ2 (scratches)

The same might occur in the processing of the **English** *which*-questions. However, in English an additional ambiguity might arise, leading to more instances of theta-role reversals.

[6] Grodzinsky & Reinhart (1993) resort to this hypothesis to account for the delay of Principle B of the Binding Theory in child's language ((5) in the text). For a detailed discussion of the hypothesis that reference-set computation exceeds young children's abilities, see Reinhart (1999), where this hypothesis is extended to account for other cases where children perform at chance.

4.3.2 Even more theta-role reversal (only in English)

The additional ambiguity involved in the processing of the English *which*-questions (8); *do* can be either functional, auxiliary verb, not assigning any thematic roles, or it can be lexical, theta-assigning verb.

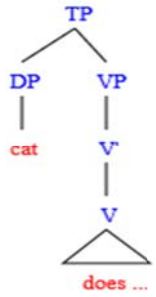
(8) Which cat does the dog scratch?

If *do* is processed as functional, the processing is done exactly as explained above for the Hebrew example.
However, when English is acquired as a HL in the context of Hebrew that has no equivalent for the functional 'do', its lexical use becomes more salient. As a result, it might be processed as such, changing the course of processing as follows: Upon the arrival of *do,* there is just one DP, *which*-$DP_{[-wh]}$ in the store. It is immediately assigned the Agent, and the incoming, "subject" DP, is assigned a Theme. This results in a reversal of thematic roles error.

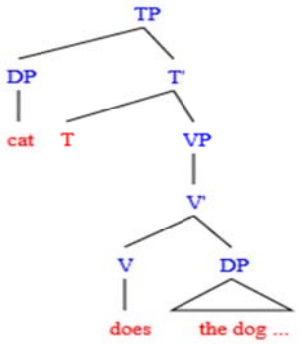
Processing of (8)

Store — Structure formation

[which-cat]



[$_V$ does] θ1, θ2



[$_{DP}$ the dog]

5. Conclusion

We have reported the comprehension of object *wh*-questions in Heritage English and L2 Hebrew by preschool bilinguals, focusing mainly on the comprehension of object *which*-questions.
We have argued that the (general) difficulty attested in the comprehension of these structures is due to the set restriction computation which, like the reference set computation, exceeds the working memory resources of preschoolers. This, we proposed, causes the children to ignore the *which* element altogether, treating the interrogative *which*-DP as a non-interrogative lexical DP. As a result, when the main verb enters the processor, its thematic roles can be linked randomly to the two lexical DPs waiting in the store, giving rise to a certain amount of reversal errors. We have attributed the higher proportion of reversal errors in English to the ambiguity of the verb *do*, and to the fact that Hebrew does not have a comparable auxiliary. This combination, we suggested, has an adverse bias towards the lexical identification of *do*, which results in an increase of the reversal of the theta-roles.
Our findings and their explanation lend further support to the claim that the source of difficulty in these structures is the computation of set restriction, showing that this computation is as demanding for bilingual preschoolers as for their monolingual peers.