DE GRUYTER MOUTON Zeitschrift für Sprachwissenschaft 2017; 36(1): 107–134

Open Access

Holger Hopp*

The processing of English which-questions in adult L2 learners: Effects of L1 transfer and proficiency

DOI 10.1515/zfs-2017-0006

Abstract: In a visual-world eye-tracking study, we investigate the on-line comprehension of subject and object which-questions in 60 German-English adult second-language (L2) learners at various stages of proficiency. In particular, we examine whether adult L2 learners follow the same acquisitional trajectories as children or whether L1 transfer dictates a different development. Previous research on monolingual English children shows that child learners initially commit to a strong subject preference and use inflectional cues, i.e. number marking, to overcome structural biases (e.g. Contemori and Marinis 2014b). The adult data on comprehension accuracy show strong effects of proficiency. Intermediate-proficiency L2 learners have an overwhelming structural preference for subject questions, and they are not sensitive to inflectional cues that disambiguate towards object readings. High-intermediate learners come to exploit inflectional cues for revising an initial subject-preference to the target object reading. Finally, advanced learners display native-like comprehension patterns. In contrast to the off-line accuracy data, the analysis of the eye movements shows that all groups of L2 adults make incremental use of inflectional cues in on-line comprehension.

These findings indicate that inflectional cues are integrated in the course of L2 development in similar ways as in monolingual development. We discuss the results in the context of current approaches to L2 processing and child-adult differences in language development.

Keywords: Wh-questions, L2 processing, L2 acquisition, syntax

***Corresponding author: Holger Hopp,** Institut für Anglistik und Amerikanistik, Technische Universität Braunschweig, E-mail: h.hopp@tu-bs.de

1 Introduction

Two of the longest-standing issues in research on bilingualism have been the questions whether and how the first language (L1) affects the developing L2 grammar and whether age entails differences in development and outcomes between child and adult second language acquisition.

More recently, the debate on L1 effects and maturation in different types of L2 acquisition has received a new impetus by a change in focus to the role of language processing. Comparisons between child and adult learners in language processing can shed light on differences in acquisitional trajectories that off-line measures may fail to detect and, more importantly, research on parsing can uncover the processing mechanisms that underlie language acquisition at different ages (Phillips and Ehrenhofer 2015). In particular, the use of on-line methods allows for assessing whether non-target-like performance is due to grammatical development, L1 transfer or development of the parser and parsing routines (for discussion, see Omaki and Lidz [2015]).

In this context, the present paper compares and contrasts the processing of subject and object which-questions in adult L2 learners with that in monolingual child learners and mature adult speakers. In a visual-world eye tracking experiment, we investigate whether the off-line and on-line comprehension of filler-gap dependencies proceeds along the same lines and according to the same grammatical cues in native child and non-native adult development or whether transfer of L1 word order dictates a partially divergent course of development on adult L2 learners.

In doing so, the present study, first, supplies off-line and on-line data on the role of L1 transfer and adds to previous off-line studies that suggest that L1 grammatical options are accessed in L2 comprehension of wh-questions (e.g. Rankin 2013, 2015). Second, the data speak to maturational models of L2 sentence processing which claim that adult learners differ from child learners in projecting syntactic dependencies (e.g. Clahsen and Felser 2006), proceduralizing grammatical knowledge (e.g. Ullman 2005) or integrating inflectional morphology in building a parse (e.g. Jiang 2007). Against this backdrop, the present study considers the comprehension and processing of English which-questions in German-English adult learners at different stages of proficiency. Building on a previous study with monolingual child learners and adult speakers of English (Contemori and Marinis 2014b), we assess the degree to which the off-line and on-line comprehension of wh-questions is initially affected by L1 word order options, and whether, at lower proficiency levels, L2 adult learners resemble children in their processing performance.

This paper is structured as follows. In Section 2, we review previous research on the acquisition and the processing of wh-movement in L1 and L2 acquisition, and we conclude by outlining the research questions and experimental predictions. In Section 3, we present the study, and we discuss its findings in Section 4.

2 Questions in and about language acquisition and processing

2.1 Wh-questions in L1 development and L1 processing

The grammatical development of wh-questions has been much studied in first and second language development (see de Villiers and Roeper [2011] for L1 acquisition; Pozzan and Quirk [2014] for L2 acquisition). In generative analyses, wh-questions involve the displacement of the wh-phrase from its base position to a clause-initial position, with this dependency being marked by traces or copies of the movement in the base position (e.g. Chomsky 1995). For subject wh-questions, the wh-phrase moves from the subject position (as in [1a]), while object questions are derived by moving the wh-phrase from the object position (1b). In main clauses, subject-auxiliary inversion ensues.

(1) a. *Who$_i$ t$_i$ is hitting the elephant?*
b. *Who$_i$ is the elephant hitting t$_i$?*

In sentence comprehension, movement of the wh-phrases in (1) creates a filler-gap dependency which needs to be reconstructed in parsing for accurate sentence interpretation (e.g. Bever 1970). Across languages, native speakers have been shown to have a preference for integrating the wh-phrase as a subject rather than an object (e.g. van Gompel 2013). The greater difficulties of object questions in sentence comprehension have been argued to reflect that the filler-gap dependency created by object wh-questions induces greater memory load (Gibson 1998), is structurally more complex (e.g. O'Grady 1997) or involves a more complex derivation (e.g. Jakubowicz 2011) than the construction of a subject wh-question.

Correspondingly, object wh-questions are acquired later in child first and bilingual language acquisition than subject questions (e.g. Guasti 1996). In addition, among object questions, wh-questions with lexical restrictions (e.g. 'which dog') emerge significantly later in L1 acquisition than bare wh-questions

(e.g. 'who') and they are associated with difficulty until about age 7 in monolingual development across languages (e.g. Avrutin 2000; Stromswold 1995; Tracy 1994).

From a processing perspective, it has been argued that overt which-questions are subject to greater similarity-based interference from subject wh-phrases than bare wh-phrases that lack rich referential features (e.g. Belletti et al. 2012; Omaki and Lidz 2015). According to cue-based memory models of parsing (e.g. Lewis and Vasishth 2005; McElree et al. 2003), the difficulty of retrieving a filler from memory is related to the degree of interference from other potentially available intervening elements in the string. For an object question such as *Which girl is the woman pushing?*, the intervening NP *the woman* constitutes an interfering cue, since the subject NP matches the wh-filler in number and gender and competes for reactivation at the gap site.

For monolingual English children, a recent eye-tracking study by Contemori and Marinis (2014b) with 31 children aged 5–7;10 (mean age: 6.04; *SD*: 0.1) found that children have a strong subject preference in the comprehension of object which-questions in terms of comprehension accuracy. In addition, eye movement data revealed that children were less successful and significantly slower than adults to recover from an incremental subject bias when listening to object which-questions. Importantly, recovery from subject readings was facilitated for the child learners if the subject and object which-phrase differed in number and inflectional number marking on the auxiliary disambiguated towards an object reading (e.g. *Which cow are the horses pushing?* vs *Which cow is the horse pushing?*). In the former sentences, intervention effects were reduced by virtue of number mismatches. These effects of similarity in intervention effects add to findings of the difficulties in the comprehension of object which-questions in typically-developing children and children with SLI (e.g. Friedmann and Novogrodsky 2011), and they complement research on number mismatches in the comprehension of object relative clauses (e.g. Adani et al. 2010; Contemori and Marinis 2014a). In sum, then, the acquisitional delay of which-questions is mirrored in the on-line comprehension of lexical wh-phrases.

2.2 L2 acquisition of wh-questions

In research on wh-movement in adult L2 acquisition, the focus has been on effects of maturational constraints, proficiency and effects of the first language, which we discuss in turn.

Direct comparisons between child and adult L2 learners in the acquisition of wh-questions are few and far-between. On the basis of corpus analyses of spontaneous production data, Chilla and Bonnesen (2011, 2012) argue that the

developmental trajectories, error types and outcomes in question formation differ between child and adult French-German and German-French learners. Unlike child learners, neither naturalistic nor instructed L2 learners showed correlations of morphological finiteness marking and syntactic position in the production of questions. Further, adults produced errors of (missing) subject-auxiliary inversion, non-target word orders and agreement in wh-questions in both L2 French and L2 German. In contrast, child L2 learners were target-like with respect to these properties, which suggests that they go through similar acquisitional stages as child L1 learners (for comparable evidence from comprehension, see Schulz [2013]; Tracy and Lemke [2011]). Chilla and Bonnesen (2011, 2012) interpret the differences between child and adult L2 learners as evidence that question formation in adult L2 acquisition is fundamentally different from child grammatical development. However, the adult data they analyse come from low-proficiency learners, and it remains to be seen how they develop as proficiency rises and how more advanced adult L2 leaners form and comprehend wh-questions.

Robust evidence that proficiency moderates wh-questions comes from research on L2 sentence processing which studies the real-time comprehension of sentences. Like native speakers, adult L2 learners engage in incremental parsing, positing and computing filler-gap dependencies (Williams et al. 2001; Williams 2006). For long-distance wh-dependencies, Omaki and Schulz (2011) report that very advanced, immersed learners project abstract hierarchical structure (see also Pliatsikas and Marinis 2013), while less proficient learners may resort to surface-based strategies in establishing grammatical filler-gap dependencies (Marinis et al. 2005; Miller 2014). These findings resemble previous off-line research on constraints on long-distance wh-dependencies with adult L2 learners, where rising proficiency is associated with greater target-like performance in judgment tasks (for review, Hawkins 2001).

As for L1 effects, the initial development of wh-questions in production largely proceeds along similar stages across L1s in adult L2 acquisition (Johnston 1985; Pozzan and Quirk 2014; though see Zobl 1995). In contrast, L2 learners transfer L1 word order options in wh-movement, e.g. depending on whether the L1 licenses partial wh-movement or wh-copies (Lutz et al. 2000). For instance, German-English learners show different production and comprehension patterns of long-distance wh-movement compared to Japanese-English adults (Schulz 2006, 2011; Slavkov 2015). Hence, grammatical options in wh-movement afforded in the L1 appear to persevere into intermediate and advanced stages of adult L2 learning.

In a series of studies on the comprehension of simple wh-questions, Rankin (2013, 2015) reports that German-English intermediate to advanced adult

learners transfer V2 and OV word order of German wh-questions to English. In off-line picture selection tasks, participants were presented with English subject and object questions as in (2).

(2) a. *What chases the cat?* (subject question, inanimate wh-pronoun)
b. *What does the cat chase?* (object question, inanimate wh-pronoun)
c. *Which animal chases the cat?* (subject question, which-question)
d. *Which animal does the cat chase?* (object question, which-question)
e. *Which animal is chasing the cat?* (subject question, which-question)
f. *Which animal is the cat chasing?* (object question, which-question)

Due to differences in surface linearization, English subject and object questions are not ambiguous, with the noun phrase following the thematic verb obligatorily being the object. However, according to the V2/OV order of German, the word-by-word translation equivalents of the subject questions as in (2a) and (2c) are ambiguous between subject and object wh-questions.[1] Analogously, the word-by-word translations of the object questions (2b), (2d), and (2f) can receive a subject interpretation in German on the basis of the colloquial *tun*-support construction in German (*Was tut die Katze jagen?*). Given the underlying OV order of German, the preverbal noun phrase 'the cat' can thus be accommodated as the object in German. Although non-standard, the *tun*-construction is wide-spread in southern dialects, including Austrian German.

In a picture selection task, L1 German learners treated English subject questions as ambiguous between subject and object questions, assigning them an object reading in about a quarter of cases (Rankin 2013, 2015). For English object questions (2d) and (2f), the learners gave a sizeable proportion of subject readings, which suggests that they analysed the English strings with a V2 grammar on the basis of the German periphrastic *tun*-construction (Rankin 2015). Moreover, when presented with ungrammatical English wh-questions like *Which animals chases the cat?*, German-English advanced adult learners interpreted these sentences as object questions, seeing that agreement would license an object reading according to a German V2 grammar. In contrast, native English speakers either failed to assign an interpretation to the ungrammatical string or interpreted it as a subject question as per surface word order (Rankin 2014). Seen in conjunction, these experiments on simple wh-questions indicate that the L1 German grammar is residually active in the comprehension English of wh-questions and licenses an object interpretation for English subject wh-questions.

1 Of course, case marking can disambiguate syntactic function in German; yet, in all examples used in this paper, the translations would be ambiguous due to case syncretism with feminine and neuter nouns.

When viewed through the lens of generative approaches to L2 acquisition, these findings lend support to probabilistic models of parameter (re-)setting in variational learning models (e.g. Yang 2002). These models hold that grammatical development is not deterministic and UG-sanctioned grammatical options that were adopted in previous development may linger in developing and mature grammars. In the context of L2 acquisition, variational learning suggests that L1 word order options are never completely expunged from Interlanguage grammars and may be recruited in L2 comprehension even at advanced stages (Slabakova 2008). In representational terms, L2 speakers encode multiple grammars in Interlanguage grammars in which L1 and L2 grammatical options coexist (e.g. Amaral and Roeper 2014). In language processing, the L1 grammatical representations can probabilistically be accessed by the parser (Sharwood Smith and Truscott 2014), resulting in non-target interpretations of English wh-questions.

In this paper, we extend previous research on L1 transfer by considering the effects of L2 proficiency on activation of the L1 in L2 comprehension, and we investigate the time-course of potential L1 activation in the processing of wh-questions. The present study investigates and compares the following types of sentences (3)–(6); where "S" is for "subject", "O" for "object", "P" and "S" for "plural" and "singular", respectively).

(3) *Which cows are pushing the goat?* (S-WH-PS, number mismatch)

(4) *Which cow is pushing the goat?* (S-WH-SS, number match)

(5) *Which cow are the goats pushing?* (O-WH-SP; number mismatch)

(6) *Which cow is the goat pushing?* (O-WH-SS, number match)

The subject and object questions in (3) through (6) differ in word order, such that the element following the auxiliary reliably disambiguates between a subject and an object question. However, in (5), number marking on the auxiliary allows for an earlier disambiguation towards an object interpretation. In the same vein, in (3), plural marking on the auxiliary affords the immediate selection of the subject reading, provided the context offers only one set of plural objects. In contrast, the parser needs to wait until encountering either a verb or a noun phrase following the auxiliary in the number match conditions (4) and (6) before being able to adopt a subject or an object reading, respectively. Under serial processing accounts like the Garden-Path model (e.g. Frazier 1987), the parser initially commits to the structurally most parsimonious parse, i.e. a subject question, and subsequently needs to revise this analysis to an object

reading when encountering the auxiliary in (5) or the noun phrase *the goat* in (6). According to constraint-based models, the parser computes several analyses in parallel (e.g. Trueswell et al. 1994) and will then demote one option and commit to the target subject or object interpretation when reaching the auxiliary in the number mismatch conditions or the following segment in the number match conditions. Under either account, disambiguation will occur earlier in the number mismatch than the number match conditions. Monolingual children were indeed found to benefit from the number mismatch in that they revised more quickly and more reliably to the target object interpretation (Contemori and Marinis 2014b).

2.3 Research questions and predictions

On the assumption that L1 grammatical options partially persevere in L2 adult comprehension, German learners of English should demonstrate a strong preference for assigning a subject reading to English object questions, since English object which-questions are ambiguous between subject and object questions in German due to German V2/OV order (7).

(7) *Welche Kuh tut die Ziege schubsen?*
which cow is the goat pushing
'Which cow is pushing the goat? / Which cow is the goat pushing?'

In particular, the number match condition in the object which-question in (6) would correspond to a licit subject question in German with subject-verb agreement between the auxiliary and the first noun phrase. Hence, the sentences in (3) to (6) lend themselves well to testing effects of L1 transfer. We address the following questions:

1. Do adult L2 learners activate the L1 German V2/OV order in the comprehension of English object which-questions?
2. Does the development in the on-line comprehension of wh-questions in adult L2 acquisition resemble monolingual L1 development?
3. Do adult L2 learners reach native-like comprehension of which-questions, i.e. does proficiency moderate comprehension and processing accuracy?

These questions translate into the following predictions.

1. If L1 transfer of the V2/OV order initially informs the L2 comprehension of wh-questions, (lower-proficiency) adult L2 learners should systematically misinterpret object which-questions as subject questions.

2. If adult L2 learner's processing of wh-questions follows the same pattern as in monolingual acquisition, they should show an asymmetry in the comprehension of object which-questions depending on matching or mismatching number marking.
3. If adult L2 comprehension can reach native-like levels, then the most advanced learners should be target-like in the comprehension and processing of both subject and object which-questions, irrespective of number (mis-) matches.

3 The study

3.1 Participants

We recruited 60 adult learners of English (mean age: 20.7 yrs; 47 female) for the study. None of the learners reported any history of cognitive, visual or language impairments. At the time of testing, all participants were students of English at the University of Mannheim. Students had predominantly been raised in Southern Germany, in particular the Palatinate region, where the use of *tun*-support is widespread and productive (e.g. Post 1992). As part of the experiment, all participants completed the LexTale test of English proficiency (Lemhöfer and Broersma 2012). In addition, participants reported their proficiency estimates across various dimensions. Participant information is given in Table 1.

Of the 60 participants, three were excluded from further analysis because they had stated that they were early bilinguals with a dominant language other than German. The data from the remaining 57 participants was analysed. Even

Table 1: Participant information (all L2 learners; n = 60).

	Mean	Range	Standard deviation
Age	20.7	18–30	2.6
LexTale Score (in %)	71.5	19–100	13.2
Self-rating English reading (out of 10)	8.0	5–10	1.1
Self-rating English writing (out of 10)	7.5	4–10	1.3
Self-rating English speaking (out of 10)	7.4	2–10	1.3
Self-rating English comprehension (out of 10)	7.8	5–10	1.2
Length of Exposure to English (in yrs)	10.7	6–20	2.9
Length of Residence in English-speaking countries (in yrs)	0.3	0–1.5	0.4

though the participants came from the same student population, the proficiency scores varied considerably, such that effects of proficiency could be assessed.

3.2 Materials and Method

The materials were adopted from Contemori and Marinis (2014b). The factors were Sentence Type (Subject vs Object) and Matching Number (SS – singular-singular vs SP – singular-plural). Each display consisted of two pictures depicting the target and the competitor, respectively. Ten quadruplets of sentences were created as in (8)–(11). Figure 1 illustrates the display for subject and object questions in the match condition.

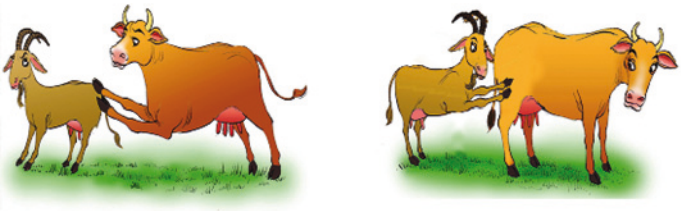

Figure 1: Display for singular match conditions in (8) and (9).

(8) S-WH-SS: *Which cow is pushing the goat?*

(9) O-WH-SS: *Which cow is the goat pushing?*

Figures 2 and 3 show the displays for the subject mismatch and the object mismatch condition, respectively.

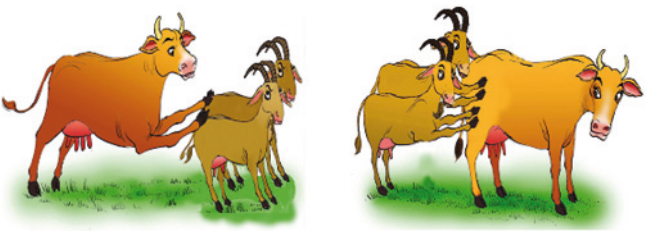

Figure 2: Display for mismatch condition in (10).

(10) O-WH-SP: *Which cow are the goats pushing?*

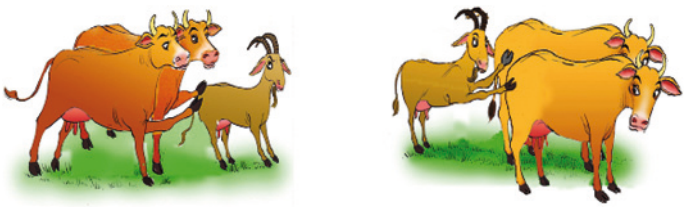

Figure 3: Display for mismatch condition in (11).

(11) S-WH-PS: *Which cows are pushing the goat?*

The materials were pseudo-randomized and presented in two lists that counterbalanced the order of pictures in the displays. Eye movements were collected using an SMI-RED tracker with a visual acuity of less than 0.4 degrees, collecting data at 60 Hz. Each participant was tested individually the lab and sat at a distance of approximately 70 cm in front of a 19-inch TFT screen on which the displays were presented. Participants saw the pictures and listened to an S-WH-question or an O-WH-question. Each display was presented for 2000 ms and participants saw a central fixation cross on the screen before the recording of the sentences started. After the recording, the participants had to press a button on a game pad (left or right) in order to match the sentence to the picture on the screen. The displays remained on the screen until after the participant had responded. Subsequently, the following trial started immediately. Three practice trials preceded the experiment. During the experiment, the participants did not receive any feedback from the experimenter. In all, the experiment took around seven minutes. Participants received course credit for participation.

3.3 Results: Comprehension Accuracy

Table 2 displays the off-line accuracy for the L2 adults and, for comparison, the data of the monolingual children and adults from Contemori and Marinis (2014b).[2]

2 Unfortunately, the complete data set from Contemori and Marinis (2014b) was not available, so that no direct statistical comparisons between the adult L2 data, on the one hand, and the child and adult L1 data, on the other hand, could be computed.

Table 2: Off-line accuracy data in per cent, absolute numbers, participants' means and standard deviation. All groups.

Conditions	Adult L2 (n = 57)		Adult L1 (n = 21)		Children L1 (n = 31)	
S-WH-SS (8)	97 %	553/570 (M = 9.7, SD = 0.6)	99 %	208/210 (M = 9.9, SD = 0.3)	95 %	296/310 (M = 9.5, SD = 0.7)
S-WH-PS (11)	97 %	551/570 (M = 9.7, SD = 0.6)	99 %	208/210 (M = 9.9, SD = 0.3)	96 %	299/310 (M = 9.6, SD = 0.6)
O-WH-SS (9)	75 %	430/570 (M = 7.5, SD = 2.9)	95 %	201/210 (M = 9.8, SD = 0.35)	62 %	195/310 (M = 6.3, SD = 3.1)
O-WH-SP (10)	84 %	476/570 (M = 8.4, SD = 2.3)	98 %	207/210 (M = 9.6, SD = 1.2)	88 %	275/310 (M = 8.9, SD = 1.5)

For the entire group of L2 adults, a Repeated-Measures ANOVA with Sentence Type (subject vs object) and Match (match vs mismatch) as within-subject factors and proficiency score as a covariate yielded significant main effects of Sentence Type ($F(1, 55) = 20.590$, $p = .001$, $\eta^2 = .272$), Match ($F(1, 55) = 4.862$, $p = .032$, $\eta^2 = .081$) as well as a significant interaction of Sentence Type and Proficiency ($F(1, 55) = 12.557$, $p = .001$, $\eta^2 = .186$). There was no significant interaction of Sentence Type and Match ($F(1, 55) = 2.488$, $p = .120$, $\eta^2 = .043$) nor a significant three-way interaction of Sentence Type, Match and Proficiency ($F(1, 55) = 0.945$, $p = .335$, $\eta^2 = .017$). Subsequent correlational analyses between the Proficiency score and off-line accuracy for the respective sentence types showed highly significant correlations between Proficiency and accuracy on O-WH-SP sentences ($r(57) = .388$, $p = .003$) and Proficiency and accuracy on O-WH-SS sentences ($r(57) = .417$, $p = .001$). In order to illustrate and analyse the differences in judgments according to proficiency further, we followed previous studies by executing a three-way split of the entire L2 group into intermediate, high-intermediate and advanced proficiency subgroups according to their LexTale scores (Table 3).[3]

For the intermediate-proficiency group, a Repeated Measures ANOVA returned a main effect of Sentence Type ($F_1(1, 18) = 19.418$, $p < .001$, $\eta^2 = .519$; $F_2(1, 9) = 149.118$, $p < .001$, $\eta^2 = .943$), a (marginal) main effect of Match

3 An anonymous reviewer asks whether there were any differences in age of onset between the three proficiency groups. A one-way ANOVA found no significant differences between the proficiency groups according to age of onset ($p > .05$).

Table 3: Off-line accuracy data in per cent, absolute numbers, participants' means and standard deviation. All L2 adult groups by proficiency group.

Conditions	Intermediate L2 (n = 19)		High-intermediate L2 (n = 19)		Advanced L2 (n = 19)	
LexTale scores	58 % (SD = 10.8)		71 % (SD = 2.7)		85 % (SD = 6.2)	
S-WH-SS (8)	96 %	183/190 (M = 9.6, SD = 0.7)	96 %	183/190 (M = 9.6, SD = 0.6)	98 %	187/190 (M = 9.8, SD = 0.4)
S-WH-PS (11)	97 %	185/190 (M = 9.7, SD = 0.5)	95 %	181/190 (M = 9.5, SD = 0.8)	97 %	185/190 (M = 9.7, SD = 0.6)
O-WH-SS (9)	58 %	110/190 (M = 5.8, SD = 3.9)	76 %	144/190 (M = 7.6, SD = 1.9)	93 %	176/190 (M = 9.3, SD = 1.1)
O-WH-SP (10)	67 %	128/190 (M = 6.7, SD = 3.2)	88 %	168/190 (M = 8.9, SD = 1.3)	95 %	180/190 (M = 9.5, SD = 0.7)

($F_1(1, 18) = 4.077$, $p = .059$, $\eta^2 = .185$; $F_2(1, 9) = 11.250$, $p = .008$, $\eta^2 = .556$) yet no interaction of Sentence Type and Match ($F_1(1, 18) = 2.155$, $p = .159$, $\eta^2 = .107$; $F_2(1, 9) = 2.729$, $p = .129$, $\eta^2 = .237$).

For the high-intermediate group, there were main effects of Sentence Type ($F_1(1, 18) = 20.382$, $p < .001$, $\eta^2 = .531$; $F_2(1, 9) = 33.065$, $p < .001$, $\eta^2 = .786$), Match ($F_1(1, 18) = 9.855$, $p = .006$, $\eta^2 = .354$; $F_2(1, 9) = 7.835$, $p = .021$, $\eta^2 = .465$) as well as an interaction of Sentence Type and Match ($F_1(1, 18) = 14.417$, $p = .001$, $\eta^2 = .445$; $F_2(1, 9) = 7.835$, $p = .022$, $\eta^2 = .457$). Subsequent pairwise comparisons yielded a significant difference between the O-WH-SS and the O-WH-SP conditions ($t_1(18) = 3.911$, $p = .001$; $t_2(9) = 3.032$, $p = .014$), yet not for the subject questions (all t's < 1).

Finally, the advanced group displayed a main effect of Sentence Type ($F_1(1, 18) = 4.800$, $p = .042$, $\eta^2 = .211$; $F_2(1, 9) = 7.579$, $p = .022$, $\eta^2 = .457$), yet no effect of Match ($F_1(1, 18) = 0.136$, $p = .716$, $\eta^2 = .008$; $F_2(1, 9) = 0.130$, $p = .726$, $\eta^2 = .014$) nor an interaction of the two ($F_1(1, 18) = 1.543$, $p = .230$, $\eta^2 = .079$; $F_2(1, 9) = 0.448$, $p = .520$, $\eta^2 = .047$).

In sum, the accuracy data attested clear differences according to proficiency. The intermediate-proficient L2 adults show higher accuracy in comprehension for subject over object which-questions, and the number mismatch did not help them to arrive at the target object reading. This pattern is different both from the child L1 and adult L1 speakers in Contemori and Marinis (2014b). In contrast, the high-intermediate-proficiency L2 group displayed sensitivity to

number marking, with mismatches leading to higher target comprehension of object which-questions. This pattern corresponds to the performance shown by child L1 learners. Finally, the highest-proficiency group had high accuracy overall, irrespective of number matches, as did the monolingual English adults in Contemori and Marinis (2014b).

3.4 Eye Tracking Results

In the following, we analyse the fixation proportions to Target and Competitors as participants listened to the wh-questions.

Figures 4 and 5 show the proportion of looks to Target and Competitor in the subject which-questions and the object which-questions for the L2 group. We only plot data for the items in which the comprehension question was answered accurately. The graphs are time-locked to the onset of the auxiliary verb (*is*/*are*) at 0 ms.

Figure 4 shows that the two types of subject questions did not differ in that looks to the target rise continuously from the onset of the auxiliary. For object questions in Figure 5, however, the L2 group first demonstrated higher looks to the competitor corresponding to a subject interpretation and then revised their interpretation to the target. This revision was delayed for the sentences with matched singular referents.

For the analysis of the eye movement data, we followed the procedure in Contemori and Marinis (2014b) and divided the continuous eye-movements into five time windows of 400 ms each, starting at the onset of the auxiliary.[4] We analysed differences in the proportion of eye movements to Competitor and Target as within-subject factor and Proficiency Group as a between-subject factor using Repeated Measures ANOVAs. We present the findings for each of the time windows (TW1 (0–400 ms), TW2 (420–800 ms), TW3 (820–1200 ms), TW4 (1220–1600 ms), TW5 (1620–2000 ms)) separately.

For TW1, we found a main effect of Sentence Type ($F_1(1, 54) = 15.069$, $p < .001$, $\eta^2 = .218$; $F_2(1, 27) = 8.542$, $p = .007$, $\eta^2 = .240$), with more looks to the target items in subject than in object which-questions. There were no further main effects or interactions. In TW2, there were main effects of Sentence Type ($F_1(1, 54) = 41.765$, $p < .001$, $\eta^2 = .436$; $F_2(1, 27) = 33.246$, $p < .001$, $\eta^2 = .552$) and

4 An anonymous reviewer wonders whether the decision to follow Contemori and Marinis (2014) in using 400 ms bins for analysing the data is "the right way of analysing these data". We agree that there are alternative ways of analysing continuous eye-tracking data, but we chose the analysis for sake of comparability with the child and adult L1 results by Contemori and Marinis (2014b).

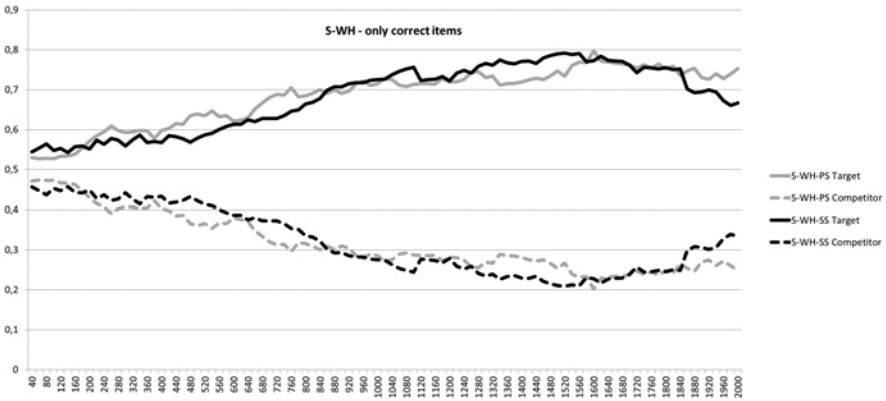

Figure 4: Proportions of looks to Target and Competitor in subject which-questions for correctly comprehended questions in the L2 group (n = 57).

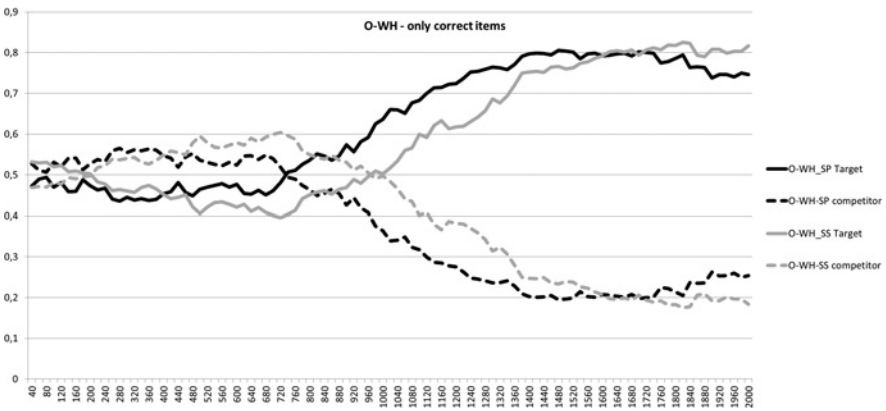

Figure 5: Proportions of looks to Target and Competitor in object which-questions for correctly comprehended questions in the L2 group (n = 57).

Match ($F_1(1, 54) = 13.372$, $p < .001$, $\eta^2 = .198$; $F_2(1, 27) = 10.546$, $p = .003$, $\eta^2 = .281$) as well as marginal interactions of Match and Proficiency Group in the analysis by subjects ($F_1(2, 54) = 2.531$, $p = .089$, $\eta^2 = .086$; $F_2(2, 27) = 1.996$, $p = .155$, $\eta^2 = .129$) and Sentence, Match and Proficiency Group ($F_1(2, 54) = 2.937$, $p = .060$, $\eta^2 = .099$; $F_2(2, 27) = 3.050$, $p = .064$, $\eta^2 = .184$). For TW3, the analysis returned main effects of Sentence Type ($F_1(1, 54) = 40.154$, $p < .001$, $\eta^2 = .426$; $F_2(1, 27) = 28.525$, $p < .001$, $\eta^2 = .514$) and Match ($F_1(1, 54) = 12.791$, $p < .001$, $\eta^2 = .192$; $F_2(1, 27) = 5.469$, $p = .027$, $\eta^2 = .168$), yet no further significant interactions. For TW4, we found no main effect of Sentence Type ($F_1(1, 54) = 1.370$,

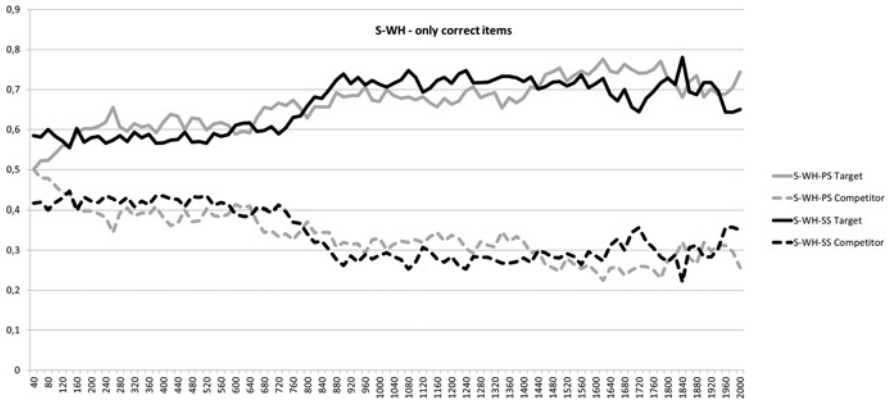

Figure 6: Proportions of looks to Target and Competitor in subject which-questions for correctly comprehended questions in intermediate-proficiency group (n = 19).

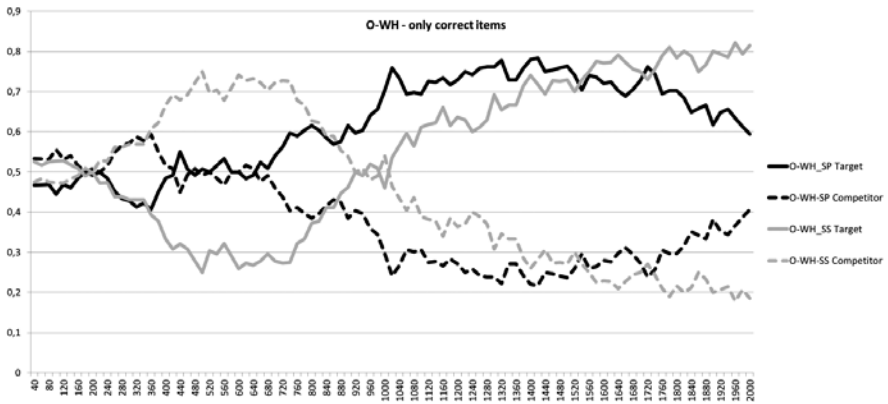

Figure 7: Proportions of looks to Target and Competitor in object which-questions for correctly comprehended questions in intermediate-proficiency group (n = 19).

p = .241, η^2 = .025; $F_2(1, 27)$ = 1.203, p = .282, η^2 = .043), yet an effect of Match that became marginally significant in the analysis by subjects ($F_1(1, 54)$ = 3.829, p = .056, η^2 = .198; $F_2(1, 27)$ = 1.142, p = .295, η^2 = .041) as well as marginal interactions of Sentence and Proficiency Group ($F_1(2, 54)$ = 2.670, p = .078, η^2 = .090; $F_2(2, 27)$ = 2.344, p = .115, η^2 = .148). In TW5, there were no significant effects.

Following up on the interactions with Proficiency Group in TW 2 and TW4, we conducted group-wise analyses for time windows from TW2 through TW4.

First, Figures 6 and 7 display the fixation proportions for subject and object sentences in the intermediate-proficiency group.

For the intermediate-proficiency group, the Repeated Measures ANOVA for TW2 returned main effects of Sentence Type ($F_1(1, 18) = 9.247$, $p = .007$, $\eta^2 = .339$; $F_2(1, 9) = 13.040$, $p = .006$, $\eta^2 = .592$) and Match ($F_1(1, 18) = 19.967$, $p < .001$, $\eta^2 = .526$; $F_2(1, 9) = 13.424$, $p = .005$, $\eta^2 = .599$), yet no interaction of Sentence and Match in the analysis ($F_1(1, 18) = 1.642$, $p = .216$, $\eta^2 = .084$; $F_2(1, 9) = 4.511$, $p = .063$, $\eta^2 = .334$). Subsequent pairwise analyses show that looks to the target were higher for O-WH-SP than for O-WH-SS sentences ($t_1(18) = 3.196$, $p = .005$; $t_2(9) = 4.189$, $p = .002$). For TW3, we found a main effect of Sentence Type ($F_1(1, 18) = 10.996$, $p = .004$, $\eta^2 = .379$; $F_2(1, 9) = 26.227$, $p = .001$, $\eta^2 = .745$), which reflects higher looks to the target for subject than for object questions. On TW4, there was but a marginal effect of Sentence Type ($F_1(1, 18) = 3.974$, $p = .062$, $\eta^2 = .181$; $F_2(1, 9) = 4.000$, $p = .077$, $\eta^2 = .308$).

For the high-intermediate group, Figures 8 and 9 show the fixation proportions.

For the high-intermediate group, the Repeated Measures ANOVA for TW2 showed a main effect of Sentence Type ($F_1(1, 18) = 25.539$, $p < .001$, $\eta^2 = .587$; $F_2(1, 9) = 13.749$, $p = .005$, $\eta^2 = .604$), yet no effect of Match ($F_1(1, 18) = 0.560$, $p = .464$, $\eta^2 = .030$; $F_2(1, 9) = 0.320$, $p = .585$, $\eta^2 = .034$), but a significant interaction of Sentence and Match in the analysis by subjects ($F_1(1, 18) = 5.669$, $p = .029$, $\eta^2 = .240$; $F_2(1, 9) = 2.313$, $p = .163$, $\eta^2 = .204$). This interaction reflects the higher looks to the competitor in the Mismatch condition compared to the Match condition. For TW3, the analysis yielded main effects of Sentence Type ($F_1(1, 18) = 36.858$, $p < .001$, $\eta^2 = .672$; $F_2(1, 9) = 7.574$, $p = .022$, $\eta^2 = .457$) and Match in the analysis by subjects ($F_1(1, 18) = 6.939$, $p = .017$, $\eta^2 = .278$; $F_2(1, 9) = 2.378$, $p = .157$, $\eta^2 = .209$). For TW4, we found a main effect of Match in the analysis by subjects ($F_1(1, 18) = 5.327$, $p = .033$, $\eta^2 = .228$; $F_2(1, 9) = 0.844$, $p = .382$, $\eta^2 = .086$).

For the advanced-proficiency group, Figures 10 and 11 show the fixation proportions.

For the advanced group, the Repeated Measures ANOVA for TW2 returned a main effect of Sentence Type ($F_1(1, 18) = 9.987$, $p = .005$, $\eta^2 = .357$; $F_2(1, 9) = 7.736$, $p = .021$, $\eta^2 = .462$). For TW3, the analysis yielded main effects of Sentence Type ($F_1(1, 18) = 7.080$, $p = .016$, $\eta^2 = .282$; $F_2(1, 9) = 7.556$, $p = .023$, $\eta^2 = .456$) and Match in the analysis by subjects ($F_1(1, 18) = 4.519$, $p = .048$, $\eta^2 = .201$; $F_2(1, 9) = 2.124$, $p = .179$, $\eta^2 = .191$). For TW4, no effects reached significance.

Summarizing, the analysis of the time curves yielded qualitatively identical patterns for the match and the mismatch condition in subject and object ques-

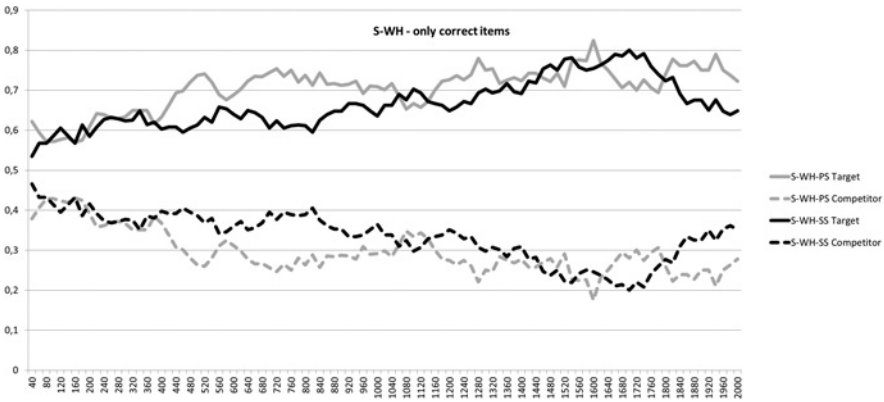

Figure 8: Proportions of looks to Target and Competitor in subject which-questions for correctly comprehended questions in the high-intermediate group (n = 19).

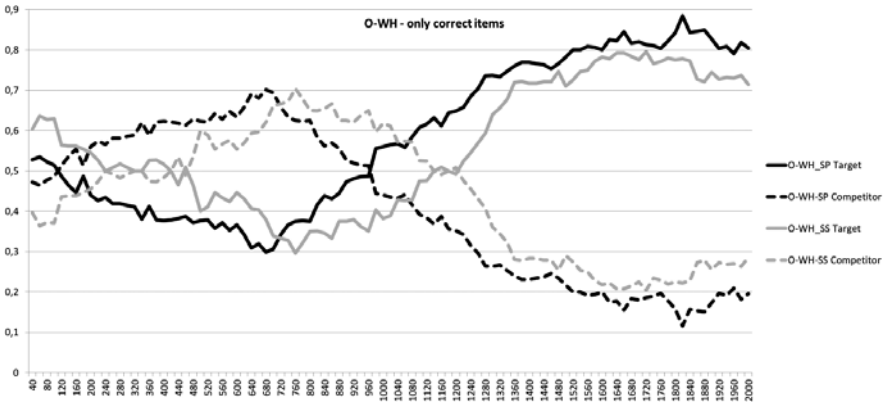

Figure 9: Proportions of looks to Target and Competitor in object which-questions for correctly comprehended questions in the high-intermediate group (n = 19).

tions in the high-intermediate and the advanced groups, with a temporal delay in the disambiguation of object which-questions in the number match condition. This delay reflects the later disambiguation by the word following the auxiliary in the number match conditions. In addition, the high-intermediate group had a strong initial subject preference, whereas the advanced group assigned an object parse to object questions throughout. In contrast, the intermediate-proficiency group showed distinct patterns for object questions, with a pronounced initial subject preference for the match condition and an initial null preference for the mismatch condition.

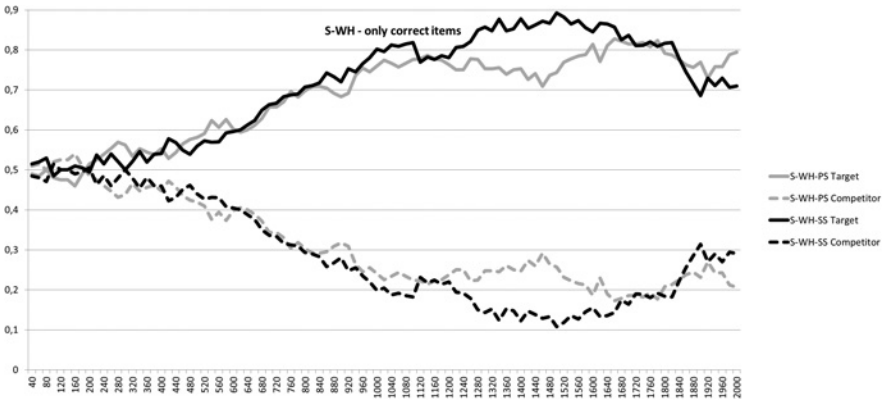

Figure 10: Proportions of looks to Target and Competitor in subject which-questions for correctly comprehended questions in the advanced group (*n* = 19).

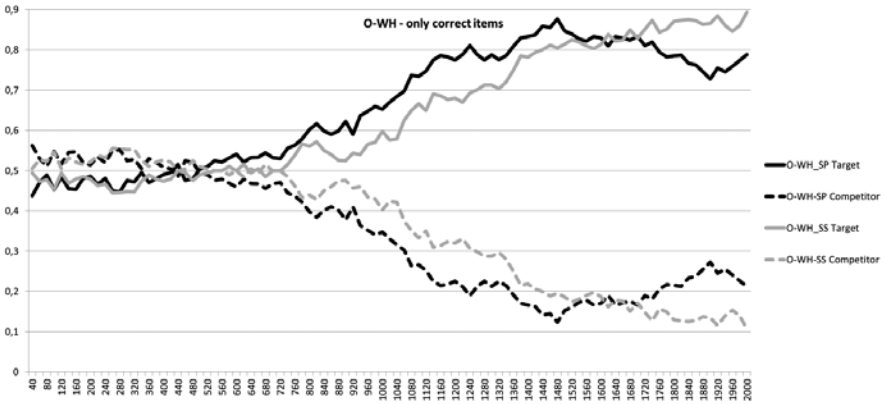

Figure 11: Proportions of looks to Target and Competitor in object which-questions for correctly comprehended questions in the advanced group (*n* = 19).

For the intermediate-proficiency group, we probed the time course of fixations further by analysing the fixation proportions of the object which-questions with incorrectly answered comprehension questions (Figure 12).[5]

Figure 12 shows that the group's gaze pattern for the incorrectly answered comprehension questions for the number match condition (grey lines) was

5 For the other groups, there were not enough items to afford an analysis of the items with incorrect comprehension accuracy.

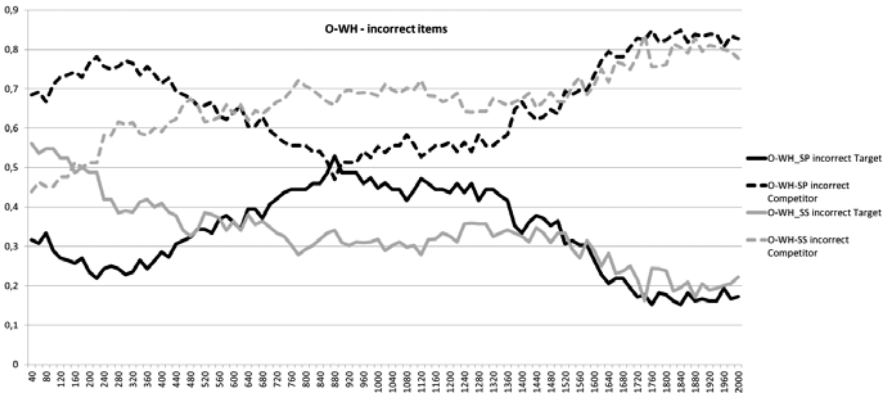

Figure 12: Proportions of looks to Target and Competitor in object which-questions for incorrectly comprehended questions in the intermediate-proficiency group (n = 19).

highly similar to their fixation proportions for subject-which questions (Figure 6). In other words, despite the word order of object which-questions being incompatible with an English subject reading, the lower-proficiency group did not retract from the initial subject preference for 42 % of the object which-questions. In contrast, for the 33 % of incorrectly answered object which-questions with number mismatches (black lines), the development of fixation proportions indicates that number marking on the auxiliary initiated a reanalysis process in TW3; however, the reanalysis attempt ultimately proved to be unsuccessful. Statistically, the differences in fixation proportions between match and mismatch object questions result in a significant main effect of Match ($F_1(1, 7) = 12.003$, $p = .010$, $\eta^2 = .632$; $F_2(1, 9) = 7.821$, $p = .021$, $\eta^2 = .465$) and a marginally significant interaction of Sentence and Match ($F_1(1, 7) = 4.855$, $p = .063$, $\eta^2 = .410$; $F_2(1, 9) = 4.625$, $p = .060$, $\eta^2 = .339$).

4 Discussion

Using off-line judgments and on-line eye-tracking measures, the present study tested whether the adult L2 comprehension of which-questions is affected by proficiency, L1 transfer and inflectional markings.

Both the accuracy data and the gaze patterns for object-wh-questions differed according to proficiency. The advanced group showed native-like performance in that it quickly revised a marginal subject preference in TW1 to the target object reading; differences in number marking did not substantially af-

fect the time course or the accuracy of comprehension in this group. In contrast, the high-intermediate group showed pronounced looks towards the subject competitor in object which-questions, and revision of the subject analysis started only in TW3. In this group, the initial subject preference obtained regardless of number marking; however, reanalysis occurred earlier for the mismatch than the match condition as shown in the significant interaction on TW2. As mentioned before, this delay is likely due to the fact that disambiguation towards an object reading is possible on the auxiliary in the mismatch condition, yet can occur only on the following segment in the match condition. In displaying delays according to number (mis-)matches, the high-intermediate proficiency group attests evidence of sensitivity to number marking in adult L2 processing. Finally, the intermediate-proficiency group showed a strong subject preference in the accuracy data, which was not modulated by number mismatches. In the on-line data, the intermediate-level learners manifested an initial subject preference only in the match condition, yet not in the mismatch condition, in which participants delayed any parsing commitment until disambiguation became possible (Figure 7). This asymmetry suggests that these learners used number information on-line in the trials that were answered correctly. As the analysis of the inaccurate trials showed, however, the intermediate group stuck to the subject interpretation in the match condition even when word order information became clearly incompatible with such a reading (Figure 12). In contrast, for the number mismatch condition, attempts at revising the subject reading were made, albeit unsuccessfully. Hence simple wh-questions present adult L2 learners with protracted comprehension difficulties, since even learners with a minimum of six years of exposure (mean for the intermediate-proficiency group: 10.4 yrs; *SD* = 2.6) do not converge on target comprehension of simple wh-questions in English.

Table 4 summarizes the sensitivity to number mismatches in off-line judgments and on-line performance in the three proficiency groups.

Next, we asked whether non-target comprehension of wh-questions reflects L1 transfer. As argued above, L1 transfer of German V2 ordering would become visible if learners adopted a subject question reading of object which-questions.

Table 4: Sensitivity to number mismatches in the L2 groups.

	Off-line accuracy	**On-line performance**
Intermediate proficiency	No	Emerging
High-intermediate proficiency	Yes	Yes
Advanced proficiency	Target	Target

Since the comprehension pattern attested in the intermediate-proficiency adult L2 group differs from the performance shown in L1 development (Contemori and Marinis 2014b; see also Adani et al. 2010; Contemori and Marinis 2014a for Italian), the specific performance pattern of the L2 group can be assumed to reflect transfer of L1 properties. As shown in Table 4, the intermediate L2 adults did not differ significantly in comprehension accuracy according to mismatches in number marking. The number match condition in object questions is compatible with a German V2/OV subject question parse based on periphrastic *tun*-constructions in German. This finding resonates with previous off-line studies on the comprehension of simple wh- and which-questions by Rankin (2013, 2015) and suggests that the non-target performance by the intermediate-proficiency group reflects recruitment of L1 grammatical options in parsing.

Unlike in the study by Rankin (2015), in which the selection of subject interpretations for object questions remained below 10 %, subject readings among the present intermediate-proficiency learners were high, accounting for almost half of the trials. Importantly, the analysis of the fixation proportions of the incorrect trials in the number match condition (Figure 12) bears out that misinterpretations did not arise from performance slips or a lack of incremental analysis of the input; rather, intermediate-proficiency learners interpreted object-which questions systematically as subject questions to a large extent.

The samples in both the present study and the one by Rankin (2015) came from similar populations and did not differ much in their mean lengths of exposure; however, Rankin (2015) did not use an independent measure of proficiency, so that the two samples may have differed in their command of English. As the present study bears out, misinterpretations in object which-questions subside as proficiency rises.

These findings are compatible with approaches to L2 grammatical development (a) which identify the L1 grammar as the initial state of adult L2 acquisition with subsequent restructuring of the Interlanguage grammar as exposure and proficiency rise (e.g. Schwartz and Sprouse 1996) and (b) which model the concurrent activation of (competing) L1 and L2 grammatical representations in L2 comprehension (e.g. Amaral and Roeper 2014; Slabakova 2008). For instance, Amaral and Roeper (2014) propose that the competence of the L2 learner comprises a number of different UG-licensed grammatical representations that compete for selection. Especially at less advanced proficiency levels, the L1 grammar has a privileged status in that it will be accessed to parse input unless there is formal incompatibility between the L2 input and the parse assigned by the L1 grammar. For the intermediate-proficiency and high-intermediate learners, object-which questions can receive an interpretation according to the L1 grammar and thus they are systematically accommodated by the L1 repre-

sentations available in the Interlanguage grammar. In the course of L2 development, rises in proficiency and exposure lead to restructuring of the Interlanguage grammar or greater facility in inhibiting L1 representations or their activation levels in L2 sentence comprehension (for discussion, see Sharwood Smith and Truscott [2014]). However, conclusive evidence that the magnitude of subject interpretations of object which questions is due to L1 transfer of OV could only come from testing an additional group of L2 adults whose L1 is SVO. This way, one could definitively adjudicate whether the strong subject preference attested in the lower-proficiency group constitutes a general feature in adult L2 acquisition or whether it is due to or augmented by the recruitment of L1 word order options.

Finally, the present paper investigated the extent to which L2 development resembles monolingual development in parsing in the incremental use of inflectional morphology for reanalysis. While the intermediate-proficiency group demonstrated no sensitivity to number mismatches in comprehension accuracy, as borne out by the lack of an interaction of Sentence Type and Match, it displayed some sensitivity to number marking in the on-line comprehension data. Moreover, the analysis of the incorrectly answered trials suggests that number marking is accessible to the L2 parser for reanalysis; however, learners cannot reliably complete the reanalysis process and retain the initial subject reading (Figure 12). In contrast, the high-intermediate-proficiency group showed clear interactions of Sentence Type and Match both in the off-line and in the on-line data. This pattern bears out that adult L2 learners integrate inflectional morphology for revising initial parsing commitments and restructuring hierarchical argument relations in on-line sentence comprehension (see also Hopp 2006, 2010; Hoshino et al. 2010). When viewed developmentally, the comprehension patterns of the high-intermediate-proficiency group *vis à vis* the intermediate-proficiency group suggest that adults can make use of facilitative inflectional marking in order to overcome L1-induced structural biases in L2 processing and move towards the target processing of wh-questions in the L2.

Further, descriptive cross-study comparisons with the eye movement data in Contemori and Marinis (2014b) allow for a comparative assessment of the timing of reanalysis processes in adult L2 and child and adult L1 processing. In Contemori and Marinis (2014b), English mature native speakers showed revision of a subject to an object parse in TW3, whereas children did not revise their subject analysis until TW4. Across proficiency groups, the adult L2 learners in this study successfully revised their subject preference in TW3, which suggests that, first, the commitment to a subject interpretation is less strong in adult L2 processing than child L1 processing and, second, that reanalysis in L2 adult processing is quicker and less difficult than in child monolingual processing.

This finding is compatible with the hypothesis that the greater degree of difficulty associated with reanalysis in child processing is related to maturational limitations in working memory and/or executive control that force children to commit to a single analysis and that reduce reanalysis ability (e.g. Kidd et al. 2011; Trueswell et al. 1999). In contrast, L2 adults recover from the subject interpretation in object questions in the same time window as native speakers. Such congruence in the timing of reanalysis between non-native and native adults further suggests that L2 parsing is equivalent in the processing mechanisms involved in reanalysis, and the present evidence runs counter to suggestions that reanalysis abilities are generally compromised in adult L2 processing (e.g. Roberts and Felser 2011). Further, the congruence of the processing patterns in adult L2 with child and adult L1 patterns in the high-intermediate and advanced-proficiency groups is consistent with approaches to adult L2 processing that stress continuity between L1 and L2 processing (e.g. Hopp 2010; McDonald 2006). Finally, it does not suggest that the integration of inflectional information or structure building and revision processes are impaired in adult L1 *vis à vis* monolingual processing (e.g. Clahsen and Felser 2006; Jiang 2007).

At the same time, the study is limited in scope to testing one type of which-questions in one population of adult L2 learners. Future studies should investigate the extent to which the findings generalize to adult L2 learners with L1s other than German and present direct statistical comparisons of adult and child learners. In ongoing research, we are exploring how child L2 and L3 learners of English comprehend which-questions. Comparative studies along these lines will further delineate whether and how bilingual parsing is bounded by age and informed by L1 properties.

In sum, the present study shows that even relatively advanced L2 learners appear to recruit L1 grammatical options in L2 comprehension. In the course of L2 development, they refer to inflectional morphology in L2 parsing to overcome L1-informed structural biases and ultimately attain native-like comprehension of English which-questions.

Acknowledgments: I would like to thank Carla Contemori and Theo Marinis for allowing me to use their stimuli and two anonymous reviewers for helpful comments. Thanks also go to Carmen Lukoschek, Nico Lindheimer and Marko Mijatovic for assistance with data collection. All remaining errors are my responsibility.

References

Adani, Flavia, Heather K. J. van der Lely, Maria Forgiarini & Maria T. Guasti. 2010. Grammatical feature dissimilarities make relative clauses easier: A comprehension study with Italian children. *Lingua* 120(9). 2148–2166.

Amaral, Luis & Tom Roeper. 2014. Multiple grammars and second language representations. *Second Language Research* 30(1). 3–36.

Avrutin, Sergey. 2000. Comprehension of discourse-linked and non-discourse-linked questions by children and Broca's aphasics. In Yosef Grodzinsky, Lew Shapiro & David Swinney (eds.), *Language and the brain: Representation and processing,* 295–313. San Diego, CA: Academic Press.

Belletti, Adriana, Naama Friedmann, Dominique Brunato & Luigi Rizzi. 2012. Does gender make a difference? Comparing the effect of gender on children's comprehension of relative clauses in Hebrew and Italian. *Lingua* 122(10). 1053–1069.

Bever, Thomas G. 1970. The cognitive basis for linguistic structure. In John R. Hayes (ed.), *Cognition and the development of language*, 279–362. New York: Wiley.

Chilla, Solveig & Matthias Bonnesen. 2011. A cross-linguistic perspective on child SLA: The acquisition of questions in German and French. *Linguistische Berichte* 228, 413–424.

Chilla, Solveig & Matthias Bonnesen. 2012. A cross-linguistic perspective on questions in German and French adult second language acquisition. *Linguistik Online* 57(7). DOI: 10.13092/lo.57.248.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Clahsen, Harald & Claudia Felser. 2006. Grammatical processing in language learners. *Applied Psycholinguistics* 27(1). 3–42.

Contemori, Carla & Theodoris Marinis. 2014a. The impact of number mismatch and passives on the real-time processing of relative clauses. *Journal of Child Language* 41(3). 658–689. DOI: 10.1017/S0305000913000172.

Contemori, Carla & Theodoros Marinis. 2014b. Online processing of English Which-questions by children and adults: A visual world paradigm study. In Will Orman & Matthew J. Valleau (eds.), *BUCLD 38 Online Proceedings Supplement*.

Frazier, Lyn. 1987. Sentence processing: A tutorial review. In Max Coltheart (ed.), *Attention and performance XII: The psychology of reading,* 559–586. Hillsdale, NJ: Lawrence Erlbaum Associates.

Friedman, Naama & Rama Novogrodsky. 2011. Which questions are most difficult to understand? The comprehension of wh-questions in three subtypes of SLI. *Lingua* 121(3). 367–382.

Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68(1). 1–76.

van Gompel, Roger P. G. 2013. *Sentence processing*. Hove: Psychology Press.

Guasti, Maria T. 1996. The acquisition of Italian interrogatives. In Harald Clahsen (ed.), *Generative perspectives on language acquisition*, 241–270. Amsterdam: Benjamins.

Hawkins, Roger. 2001. *Second language syntax: A generative introduction*. Oxford: Blackwell.

Hopp, Holger. 2006. Syntactic features and reanalysis in near-native processing. *Second Language Research* 22(3). 369–397.

Hopp, Holger. 2010. Ultimate attainment in L2 inflectional morphology: Performance similarities between non-native and native speakers. *Lingua* 120(4). 901–931.

Hoshino, Noriako, Paola. E. Dussias & Judith F. Kroll. 2010. Processing subject-verb agreement in a second language depends on proficiency. *Bilingualism: Language and Cognition* 13(2). 87–98.

Jakubowicz, Celia. 2011. Measuring derivational complexity: New evidence from typically developing and SLI learners of L1 French. *Lingua* 121(3). 339–351.

Jiang, Nan. 2007. Selective integration of linguistic knowledge in adult second language learning. *Language Learning* 57(1). 1–33.

Johnston, Malcom. 1985. *Syntactic and morphological progression in learner English*. Canberra: Department of Immigration and Ethnic Affairs.

Kidd, Evan, Andrew J. Stewart & Ludovica Serratrice. 2011. Children do not overcome lexical biases where adults do: The role of the referential scene in garden-path recovery. *Journal of Child Language* 38(1). 222–234.

Lemhöfer, Kristin & Miriam Broersma. 2012. Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods* 44(2). 325–343.

Lewis, Richard L. & Shravan Vasishth. 2005. An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science* 29(3). 375–419.

Lutz, Uli, Gereon Müller & Arnim von Stechow (eds.). 2000. *Wh-scope marking*. Amsterdam: Benjamins.

Marinis, Theodoris, Leah Roberts, Claudia Felser & Harald Clahsen. 2005. Gaps in second language sentence processing. *Studies in Second Language Acquisition* 27(1). 53–78.

McDonald, Janet L. 2006. Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language* 55(3). 381–401.

McElree, Brian, Stephani Foraker & Lisbeth Dyer. 2003. Memory structures that subserve sentence comprehension. *Journal of Memory and Language* 48(1). 67–91.

Miller, Kate A. 2014. Accessing and maintaining referents in L2 processing of wh-dependencies. *Linguistic Approaches to Bilingualism* 4(2). 167–191.

O'Grady, William. 1997. *Syntactic development*. Chicago, IL: University of Chicago Press.

Omaki, Akira & Jeffrey Lidz. 2015. Linking parser development to acquisition of linguistic knowledge. *Language Acquisition* 22(2). 158–192.

Omaki, Akira & Barbara Schulz. 2011. Filler-gap dependencies and island constraints in second language sentence processing. *Studies in Second Language Acquisition* 33(4). 563–588.

Phillips, Colin & Lara Ehrenhofer. 2015. The role of language processing in language acquisition. *Linguistic Approaches to Bilingualism* 5(4). 409–453.

Pliatsikas, Christos & Theodoris Marinis. 2013. Processing empty categories in a second language: When naturalistic exposure fills the (intermediate) gap. *Bilingualism: Language and Cognition* 16(1). 167–182.

Post, Rudolf. 1992. *Pfälzisch – Einführung in eine Sprachlandschaft*. Landau: Pfälzische Verlagsanstalt.

Pozzan, Lucia & Erin Quirk. 2014. Second language acquisition of English questions: An elicited production study. *Applied Psycholinguistics* 35(6). 1055–1086.

Rankin, Thomas. 2013. Competing grammars in the comprehension of questions and relative clauses in L1 German-L2 English. In Jennifer Cabrelli Amaro et al. (eds.), *Proceedings of the 12th Generative Approaches to Second Language Acquisition Conference (GASLA 2013),* 170–179. Somerville, MA: Cascadilla Proceedings Project.

Rankin, Thomas. 2014. Syntactic transfer and agreement cues in L1 German-speakers' comprehension of English wh-questions. Paper presented at the *23rd European Second Language Acquisition Conference (EUROSLA)*, York, 7 September 2014.

Rankin, Thomas. 2015. Variational learning in L2: The transfer of L1 syntax and parsing strategies in the interpretation of wh-questions by L1 German learners of L2 English. *Linguistic Approaches to Bilingualism* 4(4). 432–461.

Roberts, Leah & Claudia Felser. 2011. Plausibility and recovery from garden paths in second language sentence processing. *Applied Psycholinguistics* 32(2). 299–321.

Schulz, Barbara. 2006. *Wh-scope marking in English interlanguage grammars: Transfer and processing effects on the second language acquisition of complex wh-questions*. Manoa, HI: University of Hawaii dissertation.

Schulz, Barbara. 2011. Syntactic creativity in second language English: *Wh*-scope marking in Japanese-English interlanguage. *Second Language Research* 27(3). 313–41.

Schulz, Petra. 2013. Wer versteht was wann? Sprachverstehen im frühen Zweitspracherwerb des Deutschen am Beispiel der w-Fragen. In Arnulf Deppermann (ed.), *Das Deutsch der Migranten*, 313–338. Berlin & New York: de Gruyter.

Schwartz, Bonnie D. & Rex A. Sprouse. 1996. L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research* 1(1). 40–72.

Sharwood Smith, Michael & John Truscott. 2014. *The multilingual mind: A modular processing perspective*. Cambridge: Cambridge University Press.

Slabakova, Roumyana. 2008. *Meaning in the second language*. Berlin & New York: Mouton de Gruyter.

Slavkov, Nikolai. 2015. Long-distance wh-movement and long-distance wh-movement avoidance in L2 English: Evidence from French and Bulgarian speakers. *Second Language Research* 31(2). 179–210.

Stromswold, Karen. 1995b. The cognitive and neural bases of language acquisition. In Michael S. Gazzaniga (ed.), *The cognitive neurosciences*, 855–870. Cambridge, MA: MIT Press.

Tracy, Rosemarie. 1994. Raising questions: formal and functional aspects of the acquisition of wh-questions in German. In Rosemarie Tracy & Elsa Lattey (eds.), *How tolerant is Universal Grammar: Essays on language learnability and language variation*, 1–34. Tübingen: Niemeyer.

Tracy, Rosemarie & Vytautas Lemke. 2011. Young L2 and L1 learners: More alike than different. In Marzena Watorek, Sandra Benazzo & Maya Hickmann (eds.), *Comparative perspectives on language acquisition – A tribute to Clive Perdue*, 303–323. Bristol: Multilingual Matters.

Trueswell, John C., Michael K. Tanenhaus & Susan M. Garnsey. 1994. Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language* 33(3). 285–318.

Trueswell, John C., Irina Sekerina, Nicole M. Hill & Marian L. Logrip. 1999. The kindergarden-path effect: Studying on-line sentence processing in young children. *Cognition* 73(1). 89–134.

Ullman, Michael T. 2005. A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In Cristina Sanz (ed.), *Mind and context in adult second language acquisition: Methods, theory and practice*, 141–178. Washington, DC: Georgetown University Press.

de Villiers, Jill & Tom Roeper. 2011. The acquisition path for wh-questions. In Jill de Villiers & Tom Roeper (eds.), *Handbook of generative approaches to language acquisition*, 189–246. Dordrecht: Springer.

Williams, John N. 2006. Incremental interpretation in second language sentence processing. *Bilingualism: Language and Cognition* 9(1). 71–88.

Williams, John N., Peter Möbius & Choonkyong Kim. 2001. Native and non-native processing of English wh-questions: Parsing strategies and plausibility constraints. *Applied Psycholinguistics* 22(4). 509–540.

Yang, Charles D. 2002. *Knowledge and learning in natural Language*. Oxford: Oxford University Press.

Zobl, Helmut. 1995. Converging evidence for the “acquisition – learning” distinction. *Applied Linguistics* 16(1). 35–56.