

Intervention in Relative Clauses: Effects of Relativized Minimality on L2 Representation and Processing

Vera Yunxiao Xia, Lydia White, Natália Brambatti Guzzo
yunxiao.xia@mail.mcgill.ca, lydia.white@mcgill.ca, nataliaguzzo@me.com

McGill University

Introduction

The usefulness of a competence-performance distinction, such as that proposed by Chomsky (1965, and subsequently), has been much debated. Such a distinction has been adopted, implicitly or explicitly, in many different domains of language acquisition, particularly in research coming from the generative linguistic perspective. The distinction is often formulated as a representational versus processing issue, namely the unconscious linguistic knowledge that speakers attain and their access to that knowledge in real-time sentence comprehension or production. In terms of methodology, offline measures are typically considered to reflect linguistic representations, whereas online measures are taken to indicate processing. In fact, one cannot make such a rigid distinction between what offline and online measures show: sometimes online measures provide a better reflection of linguistic representation while offline measures show the effects of processing (e.g., Chondrogianni and Marinis, 2012; Hopp, 2009; Juffs, 2004; Slabakova, White and Guzzo, 2017; Sorace and Filiaci, 2006).

In the field of second language (L2) acquisition, there has been considerable discussion of the extent to which second language learners and speakers (henceforth L2ers) are similar to or different from native speakers in terms of the nature of their linguistic representations and/or their processing resources. A question has been whether or not any differences that are found reflect non-convergence as far as linguistic competence is concerned. In fact, it has been suggested by a number of researchers that, while L2ers' linguistic representations may converge with those of native speakers, their processing or computational capacity, or working memory span, can lead to performance differences (Felser and Cunnings, 2012; Hopp, 2010; Slabakova et al., 2017; Sorace, 2011; amongst many others).

In this paper, we investigate how L2ers represent and process English relative clauses in comprehension, using offline and online measures. We will revisit the well-known observation that L2ers find object relative clauses (ORCs) more difficult than subject relative clauses (SRCs), but considered in the framework of Relativized Minimality (RM) (Rizzi, 1990). We expect the

difficulties that adult L2ers have with ORCs to be similar to the difficulties shown by adult native speakers, reflecting processing problems, and different from younger child L1 acquirers, for whom relative clauses impose a greater computational burden.

Relative clauses and RM

The types of relative clauses at issue are SRCs and ORCs, the difference being the role of the relative head within the clause. In (1a), the head noun (*the king*) is interpreted as the subject of the relative clause, whereas in (1b) it is the object.

1. a. SRC: The king who _ pushed the boy.

b. ORC: The king who the boy pushed _.

ORCs are known to cause greater problems than SRCs in L1 acquisition (Adani, van der Lely, Forgiarini and Guasti, 2010, amongst others) and in adult processing (Gordon, Hendrick and Johnson, 2001, amongst others), as discussed below. To account for these differences between relative clause types, Friedmann, Belletti and Rizzi (2009) and Villata, Rizzi and Franck (2016) advance featural Relativized Minimality (see (2)), which incorporates revisions to the original Relativized Minimality (RM) proposal put forward by Rizzi (1990).

2. Featural Relativized Minimality

In a configuration X . . . Z . . . Y . . .

a local relation between X and Y is disrupted when:

a. Z structurally intervenes between X and Y (i.e., Z c-commands Y and Z does not c-command X)

b. Z matches the specification in morphosyntactic features of X

RM (including featural RM) is a syntactic locality principle, governing the behaviour of various kinds of dependencies, including the A' dependencies involved in relative clauses. A dependency relationship must be established between the relative head (=X) and the position (the gap) from which movement has taken place or where the head is interpreted (=Y), in other words, the subject or object position, depending on the type of relative clause. A local dependency can be disrupted by intervening material, the disruption being worse if the intervenor shares certain kinds of features with the elements that it intervenes between. In the case of ORCs, intervention effects occur because the NP subject of the ORC (=Z) is located between the relative head and its source, making

it harder to establish the appropriate dependency (see 3a). In the case of SRCs, there is no intervening NP, as shown in (3b).

3. a. The king who **the boy** pushed _.

X Z Y

b. The king who _ pushed the boy.

X Y

A further claim is that intervention effects are reduced if the relative head and the intervenor differ in certain features, such as number; the idea is that a mismatched intervenor is less likely to be interpreted as a potential local filler for the gap; compare (4a) and (4b). However, only features that participate in the computation of A' dependencies, such as phi-features, can have this kind of effect (e.g., Belletti, Friedmann, Brunato and Rizzi 2012; Friedmann, Rizzi and Belletti, 2016).

4. a. Feature match (sg sg): I know **the king** who **the boy** pushed.

b. Feature mismatch (sg pl): I know **the king** who **the boys** pushed.

Previous research

Investigation of relative clause acquisition and processing is of long-standing.¹ Here, we concentrate on research that is relevant to potential intervention effects, particularly the role of featural differences. A number of studies on child L1 acquisition and adult L1 processing have been used to support the claim that ORCs are harder to acquire/process than SRCs, and that acquisition/processing of ORCs is easier when there are featural differences between an intervenor and the elements that it intervenes between.

The work of Gordon et al. (2001, 2004) provides examples from adult processing, although not conducted within the RM framework. In both of these studies, processing (reading times) of English relative clauses by adult native speakers was investigated, by means of a series of self-paced reading tasks. Stimuli included SRCs and ORCs; the NP occurring in the relative clause was

¹ Much of this research looks at the effects of holding constant the role of the head noun within the main clause and the relative clause (SS, OO) or varying them (SO, OS) (following Sheldon, 1974). This issue is not under consideration here.

sometimes of the same category as the head noun and sometimes different (NP/pronoun/proper name). See (5a) (SRC) and (5b) (ORC).

5. a. The barber that admired the lawyer/you/John climbed the mountain.

b. The barber that the lawyer/you/John admired climbed the mountain.

The results show that ORCs take longer to process than SRCs (as measured by reading times at the critical region) when the same type of NP occurs as the relative head and within the relative clause (e.g., *the barber/the lawyer*). When the NPs were mismatched (e.g. *the barber/you* or *the barber/John*), there was no longer an advantage for SRCs. In other words, the effect of a mismatch was to increase the speed of processing of ORCs.

As for L1 acquisition, several recent studies have looked at relative clause acquisition in different languages from the perspective of featural differences between the head noun and potential intervenors. A number of differences have been examined, including NPs versus pronouns (similar to the adult studies discussed above), as well as features like gender and number. For these comprehension studies, a general finding is that accuracy is lower on ORCs than SRCs, with ORCs sometimes being misinterpreted as SRCs. When there is a mismatch between certain characteristics of the intervenor and the head, accuracy on ORCs improves significantly.

Friedmann et al. (2009) look at the comprehension of Hebrew relative clauses by children aged 3;7 to 5. Tasks involved matching orally presented sentences to pictures or to scenarios acted out by the experimenter. Children performed above chance on headed SRCs but at chance on headed ORCs. When the relative clause included an arbitrary pro (*the cow/someone*) instead of an NP similar to the head (*the cow/the chicken*), there was dramatic improvement in performance on ORCs, suggesting beneficial effects of a mismatched intervenor.

Belletti et al. (2012) compare effects of gender in L1 Hebrew and L1 Italian children, aged 3;9 to 5;5. Once again, the task involved picture matching. For Hebrew, while 14 out of 31 participants performed at chance on ORCs when the gender of head and intervenor matched, eight of the 14 showed significantly better performance on ORCs when the gender was mismatched. For Italian, on the other hand, it made no difference whether gender was matched or mismatched, which the authors explain in terms of syntactic differences in how gender works in the two languages; only in Hebrew is gender an inflectional feature in the relevant sense, as it shows

agreement on the verb. Adani et al. (2010) similarly found that gender mismatches in Italian had no effect.

We turn now to number, the feature we are concerned with in the present study. Several studies have shown an advantage for comprehension of ORCs when the intervenor differs in number from the head. Adani et al. (2010) look at Italian-speaking children aged 5;3 to 10;1, while Adani, Forgiarini and Guasti (2014) investigate English-speaking children with SLI (ages 9;5 to 16) compared to typically developing children matched for grammatical development (ages 6;0 to 7;5). In both studies, the task involved picture identification, administered on a computer screen. In both studies, a number mismatch between intervenor and head significantly improved performance on ORCs, for children with SLI as well as typically developing children of all ages tested. The five year olds in Adani et al. (2010) were particularly inaccurate, with 41% accuracy on matched ORCs, and 64% accuracy on mismatched ORCs. The SLI group in Adani et al. (2014) performed even less accurately, while the typically developing children were at around chance on matched ORCs; both groups nevertheless showed significantly better performance on mismatched items.

Contemori and Marinis (2014) look at the processing of relative clauses by English-speaking adults and children (ages 6;0 to 8;11), looking at the effects of number mismatches on accuracy and speed in online interpretation of ORCs. The task was a self-paced listening task. Participants listened to sentences presented phrase by phrase, pressing a key after each phrase, while looking at two pictures on a computer screen. At the end of each sentence, they had to select which of the pictures matched the sentence. This provided an accuracy measure. Their response times (RTs) after each segment provided a measure of processing. Children were quite accurate overall (78% or higher) but nevertheless more accurate in the mismatched cases than the matched cases. As for RTs, these authors found an effect for number but not for match vs. mismatch. The RT results suggest an effect for plurality at different segments, with plural nouns generally faster than singular at the NP1 position (head noun) and the NP2 position (noun within the relative clause).

Researchers have also investigated other filler-gap dependencies, including wh-questions and clitic left dislocation, and have reported that a number mismatch assists children in processing (e.g. Contemori, Carlson and Marinis, 2018; Durrleman, Marinis and Franck, 2016; Manetti, Moscati, Rizzi and Belletti, 2016).

It is worth pointing out that stimuli used in these L1 studies (for both Italian and English) in fact provide two overt cues to number within the relative clause, hence introducing a possible confound (as also noted by Adani et al., 2010). That is, number is indicated not only on the NP subject of the relative clause (the intervenor) but also on the verb, as shown in (6) (test items from Adani et al., 2014).

- 6. a. The hippo that the rhino is washing...
- b. The hippo that the rhinos are washing...

One could argue that such items do not provide a pure test of the effects of featural RM, since the verb is not a potential filler for the gap but nevertheless provides an extra indication of the difference between matched and mismatched cases. Since inflectional features are implicated in relativized RM, it is not surprising that they show up overtly on the verb as well as the subject. Bentea and Durrleman (2017) controlled for this issue by looking at relative clauses in French-speaking children (ages 5 to 8). In French, verbal agreement can be overt or covert, depending on the conjugation class of the verb in question, as shown in (7) (example (6) from Bentea and Durrleman).

- 7. a. Le garçon lave le chat
The boy washes the cat
- b. Les garçons lavent le chat
The boys wash the cat

In verbs of the first conjugation, like *laver* ‘wash’, the singular and plural verb forms in the present tense, *lave/lavent*, are pronounced in exactly the same way (namely, [lav]). For verbs of the other conjugation classes, singular and plural verb forms are phonologically distinct. Note also that the plural inflection on the noun *garçons* in (7b) is not pronounced; it is the plural article *les* that indicates that this NP is plural.

These researchers found that French-speaking children, like Italian- and English-speaking children, had difficulties with ORCs but not SRCs in an aurally presented picture-matching task. As far as number match versus mismatch was concerned, mismatches did not improve the performance of the 5 year olds but did yield better performance in the older children. Crucially, the mismatch effect in the older children was observed regardless of whether inflection was audible on the verb. Bentea and Durrleman suggest that overt inflection on the verb is not crucial; rather,

covert inflection yields the same effects. Alternatively, their study could be taken to show that it is the number marking on the NP (via the article in the case of French) that is crucial and that the additional verbal inflection plays no role. We return to this issue in the discussion.

As far as L2 acquisition is concerned, there has been considerable research on L2 acquisition of relative clauses (e.g., Doughty, 1991; Eckman, Bell and Nelson, 1988; Gass, 1982), focusing on aspects that are harder or easier to learn. Much of this early research investigated whether the Noun Phrase Accessibility Hierarchy (AH) (Keenan and Comrie, 1977) predicts the difficulty of learning different kinds of relative clauses. The AH places relativization on an implicational hierarchy, such that the least marked type of relative clause is one with a relativized subject. Relativized objects are more marked, and other positions are yet more marked. In other words, on this kind of account, ORCs are expected to be harder than SRCs, though for a different reason from that advanced by the RM approach. Featural mismatches are not discussed in the context of the AH and would, presumably, not be expected to make a difference.

Previous research has not examined potential intervention effects in L2 comprehension of relative clauses. However, there has been some preliminary research on intervention effects on L2 production. A pilot study reported in Belletti and Guasti (2015) used an elicitation task to elicit relative clauses in spoken production of adult learners of Italian. Earlier studies with L1 acquirers of Italian as well as adult native speaker controls found that participants avoided producing ORCs while having no problems with SRCs (e.g., Contemori and Belletti, 2014). Instead, older children produced passives, as did adults.² In a passivized relative clause, intervention effects are avoided (compare *the king who the boy pushed* with *the king who was pushed by the boy*). Like the children acquiring Italian as L1 and the adult native speakers, advanced proficiency L2 learners of Italian were much more likely to produce a passivized relative clause than a non-passivized ORC, again avoiding intervention. However, this was not the case for intermediate proficiency learners who produced ORCs to a greater extent than passives; Belletti and Guasti suggest that this is because passives are harder to acquire.

On similar lines, Suzuki (2019) and Suzuki and Hirakawa (2019) look at relative clause production by adult Japanese-speaking learners of English, at a low proficiency level. Both studies elicited written production of relative clauses. As in the studies reported for Italian, relatively few problems were found in the production of SRCs by either group. In contrast, when ORCs were

² Costa, Lobo and Silva (2011) report the same for adult native speakers of Portuguese.

elicited, passives were produced instead of ORCs. Suzuki (2019) manipulated animacy as a feature, on the assumption that clauses in which the ORC head and the intervenor match in animacy (*the boy/the girl*) would be harder than those in which they are mismatched (*the ball/the girl*).³ A higher level of production of passives was found in the cases of animacy match than mismatch. Suzuki and Hirakawa (2019), on the other hand, manipulated number; however, they found no effects for number match versus mismatch.

The above production studies on L2 Italian and L2 English report avoidance of ORCs, suggesting an intervention effect, although not always an effect of featural differences.

To summarize, intervention effects, relating to processing times, comprehension and production, have been reported in children and adults, L1 and L2 acquirers. ORCs cause more problems than SRCs; mismatched intervenors sometimes help to reduce problems caused by the presence of an intervenor.

Representation and processing

Before going on to present our experiment, it is important to discuss whether the reported problems with relative clauses reflect problems of linguistic representation, problems in online processing, or both. We assume that the parser and the grammar are distinct and that the parser draws on the grammar in online L2 processing (e.g., Hopp, 2010; Omaki and Schultz, 2011; Phillips and Ehrenhofer, 2015; amongst others). Thus, complex computations required by the grammar may translate into processing difficulties. Adult native speakers are accurate on ORCs whether intervenors are matched or mismatched, suggesting that ORCs in particular, and RM more generally, are part of the adult grammatical system. Nevertheless, adults sometimes experience processing delays, particularly in cases where the intervenor in an ORC matches the relative head in some feature(s). For children, on the other hand, the problem is arguably not simply a processing issue, since, at least for the younger children reported in the experiments above, accuracy on ORCs in offline tasks is around 50%, especially when features match.

³ It is not clear, however, that a semantic feature like animacy is expected to have an effect according to the RM account, which restricts the relevant features to a subclass of inflectional features involved in syntactic computation (see Belletti and Rizzi, 2013).

Rizzi and colleagues (Belletti and Rizzi, 2013; Belletti et al., 2012; Friedmann et al., 2009; Rizzi 2018; Villata et al., 2016) argue that it is not RM as such that is at issue. They propose a distinctness hierarchy, relating to the relationship between the target (X), in this case the relative clause head, and the intervenor (Z), in this case the subject of an ORC. They advance four possibilities: disjunction, intersection, inclusion and identity, defined as in (8).⁴

8. a. Disjunction: X and Z share no relevant features

+R, +wh, +NP < +R, +wh >⁵

whoever the boy pushed _

b. Intersection: X and Z share some relevant features but not all

+R, +NP, -Pl +NP, +Pl < +R, +NP, -Pl >

the king who the boys pushed _

c. Inclusion: The features of Z are properly included in the features of X

+R, +NP +NP < +R, +NP >

the king who the boy pushed _

d. Identity: X and Z share all relevant features

*+R, +NP +R, +NP < +R, +NP >

who who pushed _

Feature mismatch is a case of intersection, whereas feature match is a case of inclusion. Total identity of X and Z is never possible. In adult grammars, disjunction, intersection and inclusion are all possible, since adults produce and comprehend ORCs, regardless of feature match/mismatch. However, featural RM can result in a processing disadvantage for adults in cases of inclusion. Belletti and Rizzi (2013) suggest that syntactic computations involving inclusion are more complex than those involving non-inclusion (disjunction or intersection) and that this complexity difference shows up in online processing: cases involving inclusion take longer to parse. In other words, the adult grammar allows both inclusion and non-inclusion; processing speed in cases involving inclusion can nevertheless be affected by computational complexity.

⁴ In many of the papers by these researchers, only 3 categories are described, namely disjunction, inclusion and identity.

⁵ +R is a feature marking the relative head; +NP is a feature indicating a lexical noun phrase; ±Pl indicates number; +wh indicates a free relative operator.

Rizzi and colleagues propose that children have a different cut-off on the hierarchy in (8), such that inclusion is not permitted in their grammars. According to Friedemann et al. (2009: 84), “children adhere to a somewhat stricter version of RM, requiring not just a *distinct* featural specification of the target with respect to the intervenor, but imposing the stronger requirement of a *disjoint* specification”. We take this to mean that the child’s representation differs from the adult’s. In the course of acquisition, children change from a more restrictive system allowing only disjunction or intersection to a less restrictive one. In fact, some of the researchers who have tested child participants have argued that, for older children at least, the problem is one of parsing rather than the underlying grammar (e.g. Adani et al., 2010). Given that older children (6+) in several of the experiments were relatively successful, they might be past the age of acquisition of structures implicating RM, their processing instead being affected in the same way as adults.

In the case of L2 acquisition, we assume that the grammars of adult learners, like adult native speakers, are constrained by RM. We hypothesize that featural differences between a relative head and an intervenor may prove advantageous in adult L2 parsing, as is the case in adult L1. There might, then, be differences in offline relative clause accuracy and online processing speed, both for SRCs versus ORCs and for mismatched versus matched items. Before turning to our experiment, we consider the extent to which featural RM effects might be derivable from the L1.

Crosslinguistic differences: Properties of relative clauses in Chinese

In the present study, we look at the acquisition of English relative clauses by native speakers of Mandarin Chinese. Chinese relative clauses differ from English in that they are prenominal rather than postnominal. This means that predictions concerning the relative ease or difficulty of SRCs versus ORCs are potentially different for the two languages, depending on whether linear or hierarchical effects are considered.

Consider the sentences in (9) and their corresponding structures in (10), from Suzuki and Hirakawa (2018). (9a) and (10a) represent SRCs, while (9b) and (10b) show ORCs. From a linear point of view, SRCs should be more difficult than ORCs, since *nuhaizi* (‘girl’) in (9a/10a) intervenes between the gap and the head noun, whereas *nanhaizi* (‘boy’) in (9b/10b) does not. This, then, is the opposite of the situation in English. However, if one considers hierarchical

structure rather than linear order, then Chinese and English are alike: *nuhaizi* ('girl') in (9a/10a) does not c-command the gap, so does not intervene, whereas *nanhaizi* ('boy') in (9b/10b) does c-command, so is a potential intervenor in terms of the definition of featural RM provided in (2). Researchers differ as to which type of relative clause is considered more difficult in Chinese. Mandarin speakers are argued to have greater difficulty with SRCs on some accounts (Chen, Ning, Bi and Dunlap, 2008; Gibson and Wu, 2013; Hsiao and Gibson, 2003) and ORCs on others (Hu, Gavarró, Vernice and Guasti, 2015; Hu, Gavarró and Guasti, 2016; Lin and Bever, 2006). Hu and colleagues argue that a hierarchical account in terms of RM is implicated.

9. a. [zhui nuhaizi de] nanhaizi (SRC)

chase girl DE boy

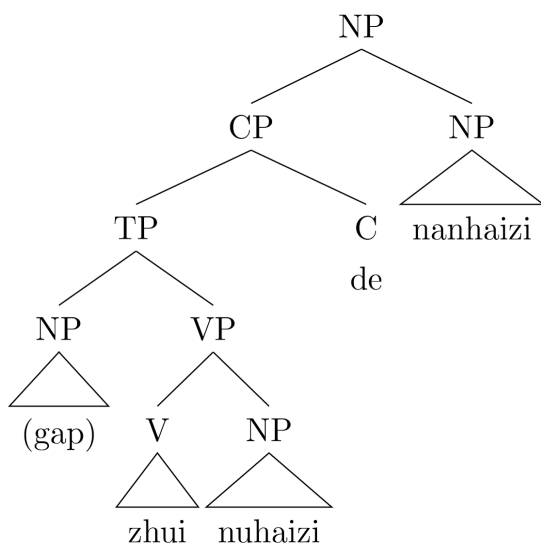
'The boy who chased the girl'

b. [nanhaizi zhui de] nuhaizi (ORC)

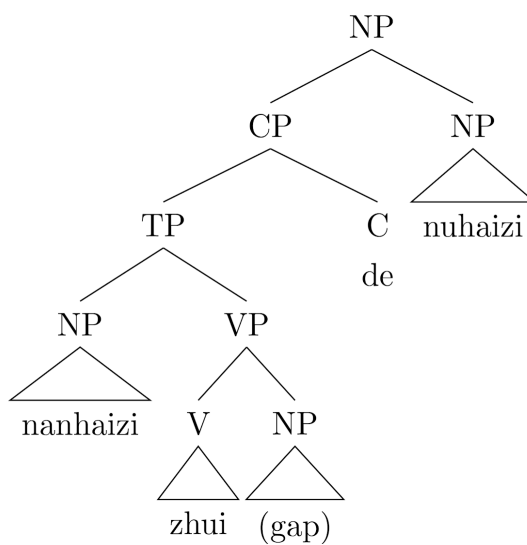
boy chase DE girl

'The girl who the boy chased'

10. a.



b.



Taking hierarchical structure into account, Chinese and English are more alike than they appear on the surface. This means that effects of L1 transfer cannot be ruled out: intervention arises in the case of ORCs in both languages. However, there is another relevant difference between the two languages, given that we will be examining the role of featural mismatches involving number. Unlike English, Chinese does not have obligatory number marking on nouns; in fact, number

marking is quite restricted. This means that there is no equivalent mismatch effect and that any potential advantage for number-mismatched NPs in the L2 could not be derived directly from the L1.

Experiment

An experiment was conducted to investigate the status of SRCs and ORCs in the comprehension of English relative clauses by Chinese speakers. We hypothesize that adult L2ers will be able to interpret SRCs and ORCs appropriately but that they will show intervention effects when parsing ORCs, particularly in cases of matching features.

Participants

Participants were tested in Montreal and Ottawa, Canada, and were compensated for their time. Proficiency was determined by means of a cloze test completed after the experiment; only people scoring at least at the intermediate level were retained for the study.⁶ All participants were Mandarin-speaking learners of English, the majority from the People's Republic of China. They were first immersed in English as (young) adults, aged 17 or older, principally by moving to an English-speaking country (usually Canada). Many participants reported non-immersive childhood exposure to English, which is not uncommon in Chinese-speaking countries.

There were three groups: L2ers of intermediate proficiency (n=13), L2ers of advanced proficiency (n=26), and adult native speaker controls (n=16). The intermediate group ranged in age from 19 to 56 (median age 23). Their age of first exposure to English ranged from 5 to 14 (median 10). Their age of immersion in English ranged from 17 to 46 (median 19). The advanced group ranged in age from 18 to 29 (median age 21). Their age of first exposure to English ranged from 3 to 13 (median 6.5). Their age of immersion in English ranged from 17 to 26 (median 18).

Pretest

A production task was included in order to determine whether plural marking was present in the interlanguage grammar, in other words, whether the L2ers would be sensitive to a number distinction between singular and plural in L2 English. As is well known, L2ers often omit

⁶ The cloze test was previously available for downloading as a self-assessment test in L2 English from the website of the Dept. of Psychology, University of Ottawa.

inflectional morphology; nevertheless, many researchers have suggested that underlying features are present (e.g., Haznedar and Schwartz, 1997; Lardiere, 1998; Prévost and White, 2000). Although absence of overt plural marking does not necessarily indicate absence of number in the underlying grammar, we assume that appropriate suppliance of overt morphology is an indication that the relevant feature is represented.

Participants were given pictures to describe containing multiple items, designed to elicit plural count nouns. Accurate production of plural marking on count nouns in plural contexts was generally high, with a mean of 90% for the advanced group, and a mean of 78% for the intermediate group. We take this to indicate that number is represented in the grammar, and hence that it is reasonable to explore the role of featural mismatches in their interpretation of relative clauses.⁷

Sentence types

The two main tasks (picture selection and self-paced reading) involved 24 test sentences and 12 fillers, presented in an order randomised for each participant, with the same wording used in the complex NPs (head nouns and relative clauses) in each task. Test items manipulated relative clause type (subject/object) and intervenor (number match/mismatch). In the picture selection task, the words *Show me* preceded the complex NP. In the self-paced reading task, the words *I know* preceded the complex NP.

All sentences were semantically reversible and contained only human NPs as heads and within the relative clause. There were 24 test sentences, which were divided into 6 sets, involving male and female characters. Within any particular sentence, the characters were either both male or both female. Note that, unlike the test items used in the L1 studies described above, there was only one cue to number, namely morphology on the noun. The verbs were in the past tense in order to avoid giving additional information on match/mismatch via verbal inflection. A full list of the test items is provided in the Appendix.

Each set of test items contained 4 sentences: an SRC match, an SRC mismatch, an ORC match and an ORC mismatch. Within each set, the 4 sentences used identical vocabulary. Every sentence was 8 words long (*Show me/I know* followed by the 6-word complex NP). All sets used

⁷ In preliminary analyses we removed those participants who supplied plural marking on count nouns less than 75% of the time; in fact this made no difference to the overall results, so all participants have been retained.

two singular NPs (sg sg) for the matched sentences (see 11a, 11c, 12a and 12c). For the mismatched sentences, 3 of the sets used a singular plural NP (sg pl) order (see 11b and 11d) while 3 used a plural singular NP (pl sg) order (see 12b and 12d).⁸ This made for a total of 6 SRC match sentences (sg sg), 6 SRC mismatch (3 sg pl and 3 pl sg), 6 ORC match (sg sg) and 6 ORC mismatch (3 sg pl and 3 pl sg).

11. a. SRC match: ...the king who pushed the boy. (sg sg)
b. SRC mismatch: ...the king who pushed the boys. (sg pl)
c. ORC match: ...the king who the boy pushed. (sg sg)
d. ORC mismatch: ...the king who the boys pushed. (sg pl)
12. a. SRC match: ...the girl who hugged the queen. (sg sg)
b. SRC mismatch: ...the girls who hugged the queen. (pl sg)
c. ORC match: ...the girl who the queen hugged. (sg sg)
d. ORC mismatch: ...the girls who the queen hugged. (pl sg)

The 12 filler sentences (6 with singular subjects, 6 with plural) involved intransitive verbs in the relative clause, as shown in (13).

13. a. Filler: the queen(s) who slept.
b. Filler: the girl(s) who walked to school.

Tasks

The two main tasks (picture identification and self-paced reading) were administered via E-Prime 2.0 (Schneider, Eschman and Zuccolotto, 2002). For the picture identification task, participants were presented with sentences, displayed on a laptop computer. Each sentence was accompanied by four pictures; participants had to press a key to indicate which picture illustrated the meaning of the sentence. This task was untimed. An example is shown in Figure 1.

⁸ One logically possible type, pl pl, was not included for either type of RC in order to reduce the number of test items.

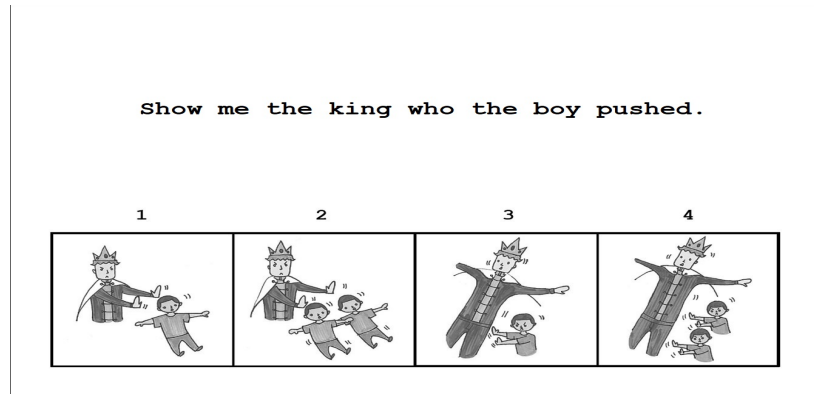


Figure 1. Sample item from the picture identification task.

For the moving window self-paced reading task, each sentence was initially presented as a series of dashes with no words or punctuation visible. Participants pressed a key to uncover the first region of a sentence while the other regions remained covered. Pressing the key a second time uncovered the next region of the sentence, with the previous region replaced by dashes. Thus, participants read the entire sentence region by region, with only one region visible at any time. The regions in the self-paced reading task are shown in (14). Each sentence was followed by a comprehension question, requiring an answer of True or False (see 15). Reading times were recorded for each region and for the entire sentence.

14.

	Region 1	Region 2	Region 3	Region 4	Region 5	
SRC	I know	the king(s)	who	pushed	the boy(s)	.
ORC	I know	the king(s)	who	the boy(s)	pushed	.
Filler	I know	the girl(s)	who	walked	to school	.

15. a. I know the king who pushed the boy.

The boy pushed the king. (True/False)

b. I know the girls who hugged the queen.

The girls hugged the queen. (True/False)

c. I know the musician who the waiter touched.

The waiter touched the musician. (True/False)

d. I know the athlete who the doctors kicked.

The athlete kicked the doctors. (True/False)

SRCs and ORCs contained 5 regions. Region 4 is the point at which it becomes clear whether an SRC or an ORC is involved (see (14)). In the case of ORCs, region 4 is the critical region, namely the point at which it can be seen whether the features match or mismatch. For SRCs, the match/mismatch issue is only resolved at region 5. For the fillers, those with PPs contained 5 regions, while those without PPs contained 4. The period was a separate region from the rest of the text.⁹

To prevent effects of task order or familiarization with test items, order of presentation of the picture identification and self-paced reading tasks was alternated, with the production pretest always taken between them.¹⁰ In total, the three tasks took around 40 minutes to complete.

Predictions

As discussed above, we assume that adult L2 learners are constrained by RM as far as their grammatical knowledge is concerned, and that they permit disjunction, intersection and inclusion. In other words, we predict accuracy in both types of relative clauses (SRCs and ORCs) on both tasks (picture identification and self-paced reading). We also hypothesize that adult L2ers may experience processing difficulties in the case of ORCs and that such effects will be reduced if the intervenor differs in number features from the relative head, in accordance with featural RM. Thus, as far as the self-paced reading task is concerned, response times to ORCs are expected to be slower than to SRCs; faster response times are expected for ORCs with mismatched intervenors than with matched intervenors. In general, we expect L2ers to behave similarly to adult native speakers, revealing possible processing effects. Our predictions are summarized in (16).

16. i. All groups will show accuracy on SRCs and ORCs in both tasks, regardless of match or mismatch.
- ii. Response times will be faster to SRCs than ORCs for all groups.

⁹ For fillers without a PP, the period was in region 5, and E-Prime was scripted to skip over the blank sixth region and proceed directly to the comprehension question.

¹⁰ Task order did not affect participants' responses. Logistic regressions show that neither accuracy (in both tasks) nor overall response times are affected by task order ($p > 0.5$).

- iii. Response times will be faster to mismatched items in ORCs than to matched items for all groups, whereas no such difference is expected for SRCs.

Results

Accuracy

We first consider the results relating to accuracy in interpreting SRCs and ORCs on the picture identification task and the self-paced reading task.¹¹ In the picture identification task, accuracy was determined by the choice of pictures. In the self-paced reading task, accuracy was determined by means of responses to the comprehension questions that followed the stimuli. On each task, participants received a score of 1 for each accurate response and 0 for inaccurate. Accuracy on the two tasks was modeled with logistic regressions (one per group) in R (R Core Team, 2019), using the `glmer()` function in the *lme4* package (Bates, Maechler, Bolker and Walker, 2015), which implements contrast coding by default.

The statistical models contained the maximal number of random effects possible. For the picture identification task, the maximal converging model for all groups includes *Type of RC* and *Match* as main effects, as well as a by-participant random intercept. For the self-paced reading task, the maximal converging models include *Type of RC* and *Match* as main effects, and by-participant and by-item random intercepts. All models were compared with equivalent models including an interaction between *Type of RC* and *Match* via ANOVAs (likelihood ratio tests). As there is no significant statistical difference between models with and without interaction,¹² we report the results for the models without interaction.

¹¹ All groups were highly accurate, indeed almost at ceiling, on the fillers in both tasks (for picture identification, intermediates: 95.5%; advanced: 96.5%; native speakers: 97.4%; for self-paced reading, intermediates: 96.7%; advanced: 98%; native speakers: 99.5%). Fillers involved intransitive verbs in the relative clause, so no choice between NPs was required as far as interpretation is concerned.

¹² Comparison between models with data from intermediate learners: $\chi^2=1.01$, $df=1$, $p=0.31$ (picture identification task); $\chi^2=1.18$, $df=1$, $p=0.27$ (self-paced reading task). Comparison between models with data from advanced learners: $\chi^2=0.33$, $df=1$, $p=0.56$ (picture identification task); $\chi^2=0.89$, $df=1$, $p=0.34$ (self-paced reading task). Comparison between models with data from native speakers: $\chi^2=2.95$, $df=1$, $p=0.08$ (picture identification task); $\chi^2=0.34$, $df=1$, $p=0.55$ (self-paced reading task).

Results from the picture identification task show that all groups interpret relative clauses appropriately, with a high level of accuracy, as shown in Figure 2. Across both the L1 and L2 groups, inaccurate responses were largely for ORC sentences, with about half the errors resulting from misinterpreting NP number, and the other half resulting from misinterpreting the ORC as an SRC. These results suggest that all groups are overall more accurate with SRCs than ORCs, and that the L2 learners are less accurate with mismatched sentences. The models confirm that the intermediate learners are significantly more accurate with SRCs than ORCs ($\hat{\beta}=1.37$, $SE=0.61$, $z=2.24$, $p=0.02$) and less accurate with mismatch than match ($\hat{\beta}=-2.24$, $SE=0.78$, $z=-2.87$, $p=0.004$). *Type of RC* is not statistically significant for the advanced group ($\hat{\beta}=0.89$, $SE=0.46$, $z=1.9$, $p=0.057$); however, like the intermediates, this group is significantly less accurate with mismatch than match ($\hat{\beta}=-1.36$, $SE=0.51$, $z=-2.63$, $p=0.008$). Native speakers are significantly more accurate with SRCs than ORCs ($\hat{\beta}=2.88$, $SE=0.83$, $z=3.44$, $p=0.0005$); *Match* is not significant for this group ($\hat{\beta}=-0.14$, $SE=0.53$, $z=-0.26$, $p=0.78$).

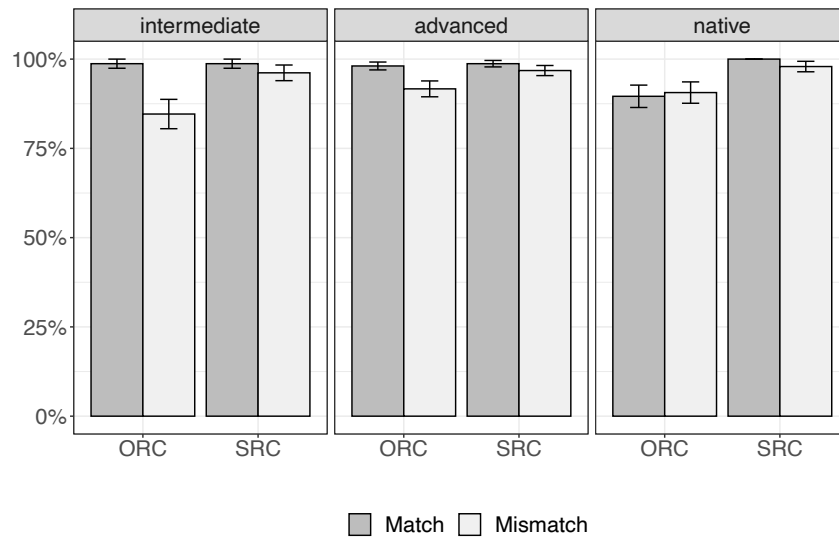


Figure 2. Accuracy on the picture identification task.

On the self-paced reading task, overall accuracy was also high, as shown in Figure 3. Inaccurate responses for both native speakers and L2ers again largely consisted of interpreting ORCs as if they were SRCs, rather than interpreting SRCs as ORCs. Examination of the proportions in Figure 3 suggests that both intermediates and native speakers are more accurate

with SRCs than ORCs, while advanced speakers perform similarly with both. The models confirm that intermediates are significantly more accurate on SRCs ($\hat{\beta}=1.19$, $SE=0.54$, $z=2.18$, $p=0.02$); no effect of *Match* was obtained for this group ($\hat{\beta}=-0.7$, $SE=0.53$, $z=-1.31$, $p=0.18$). For the advanced group, neither *Type of RC* ($\hat{\beta}=0.2$, $SE=0.39$, $z=0.51$, $p=0.6$) nor *Match* ($\hat{\beta}=-0.28$, $SE=0.4$, $z=-0.71$, $p=0.47$) is statistically significant. Native speakers are also significantly more accurate with SRCs than ORCs ($\hat{\beta}=1.85$, $SE=0.44$, $z=4.12$, $p<0.0001$); no effect of *Match* was found for this group either ($\hat{\beta}=0.1$, $SE=0.36$, $z=0.27$, $p=0.78$).¹³

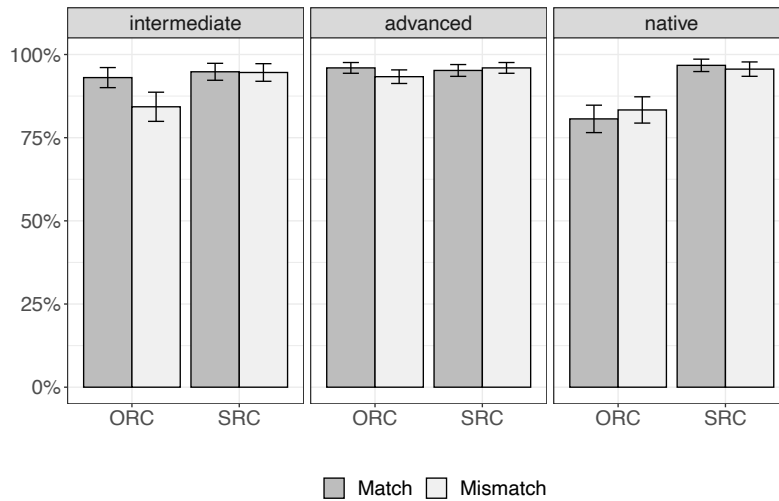


Figure 3. Accuracy on the self-paced reading task.

Reading times

We turn now to reading times, tested only on the self-paced reading task. Items are retained for the analysis only if the answers to the comprehension questions were accurate; outliers (more than 2 SDs above/below the mean response times for any individual participant) have also been removed. Out of a total of 1980 responses, 191 have been removed (86 outliers and 105 wrong answers), which corresponds to 9.6% of the data. Response times (RTs) were modelled via linear

¹³ The panel for intermediates in both Figures 2 and 3 suggests that there might be an interaction between Type of RC and Match for this group. Examination of the results for the logistic regressions including interaction for this group indicates no significant interaction (picture identification task: $\hat{\beta}=1.64$, $SE=1.59$, $z=1.02$, $p=0.3$; self-paced reading task: $\hat{\beta}=1.18$, $SE=1.09$, $z=1.8$, $p=0.27$). Recall that models containing interaction have a worse fit than models without interaction, which is why we do not report them in the main text.

regressions (one per group) in R using the `lmer()` function in the *lme4* package; *p* values were obtained with the *lmerTest* package (Kuznetsova, Brockhoff and Christensen, 2014). Response time data were log-transformed in the statistical models.

Total RTs (as measured at the end of the sentence, excluding the period) and RTs at specific regions in the target sentences were modelled separately. As in the accuracy models, the RT models contained the maximal number of random effects possible. The maximal converging models for Total RTs included *Type of RC* and *Match* as main effects, as well as by-participant and by-item random intercepts. Like the accuracy results, the Total RT models were also run with an interaction between *Type of RC* and *Match*. ANOVAs (likelihood ratio tests) comparing models with and without interaction indicate no statistical difference between them.¹⁴ We thus report the results for the models without interaction in this section.

Regarding RTs at specific regions in the sentence, we examined the effect of *Match* for ORC and SRC in regions where mismatch is observed: region 4 for ORC and region 5 for SRC (see (14)). We also examined the effect of *Match* for ORC at region 5, the spillover region. The maximal converging models for regions 4 and 5 included *Match* as main effect and a by-participant random intercept. (Since the material in regions 4 and 5 differs for SRCs versus ORCs (see (14)), it is not possible to compare these RC types at the same region.)

Figure 4 shows the results for Total RTs, which suggest that responses overall are faster with SRCs than ORCs, with no effect of *Match*. The models for Total RTs indicate that neither *Type of RC* ($\hat{\beta}=-0.07$, $SE=0.04$, $t=-1.77$, $p=0.09$) nor *Match* ($\hat{\beta}=0.01$, $SE=0.04$, $t=0.28$, $p=0.77$) is, in fact, statistically significant for intermediates. On the other hand, the advanced group is overall significantly faster with SRCs than ORCs ($\hat{\beta}=-0.09$, $SE=0.03$, $t=-2.43$, $p=0.02$); no effect for *Match* was found in this group ($\hat{\beta}=0.05$, $SE=0.03$, $t=1.35$, $p=0.19$). Native speakers are also significantly faster with SRCs than ORCs ($\hat{\beta}=-0.09$, $SE=0.03$, $t=-2.61$, $p=0.01$); no effect of *Match* was observed for this group either ($\hat{\beta}=-0.03$, $SE=0.03$, $t=-1.12$, $p=0.27$).

¹⁴ Comparison between models with data from intermediate learners: $\chi^2=0.11$, $df=1$, $p=0.73$. Comparison between models with data from advanced learners: $\chi^2=1.28$, $df=1$, $p=0.25$. Comparison between models with data from native speakers: $\chi^2=1.13$, $df=1$, $p<0.0001$.

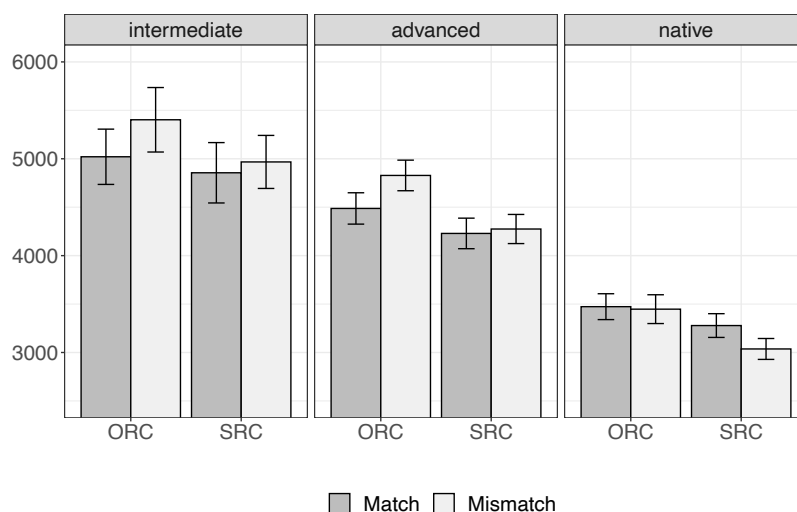


Figure 4. Total response times in the self-paced reading task.

Figure 5 shows participants' RTs for *Match* at region 4, for ORC sentences only. Region 4 is the critical region to examine intervention effects in ORCs, as this is the region where NP2 is encountered. Examination of the bars in the figure suggests that learners, particularly those in the intermediate group, are faster with sentences where the two NPs match in number. This is contrary to expectation, as an advantage for the mismatched conditions was predicted (see (16iii)). The statistical models for *Match* with ORCs at region 4 confirm this: intermediates are significantly slower with mismatched sentences ($\hat{\beta}=0.27$, $SE=0.1$, $t=2.65$, $p=0.009$), as are the advanced learners ($\hat{\beta}=0.12$, $SE=0.05$, $t=2.26$, $p=0.02$). No statistical effect was obtained for native speakers ($\hat{\beta}=-0.05$, $SE=0.06$, $t=-0.81$, $p=0.41$). In Figure 5, the lines indicate the behaviour of each participant. It can be seen that the majority of intermediate and advanced learners exhibit the pattern that the models capture (i.e., an increase in RT in the mismatched condition). This is not the case for the native speakers.

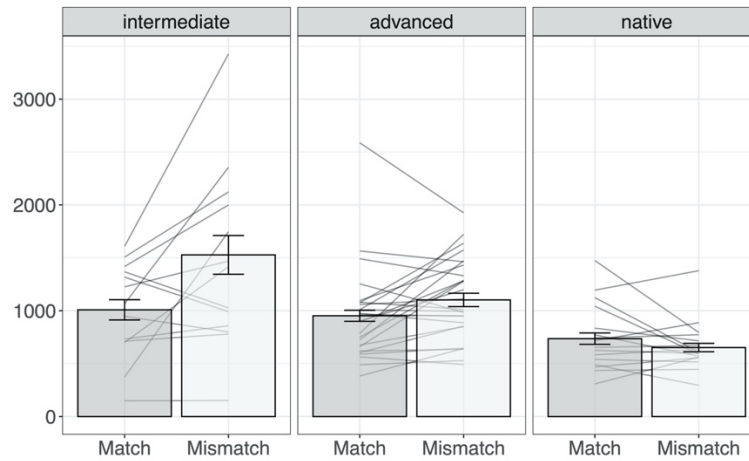


Figure 5. Response times for ORCs at region 4.

Participants' RTs at region 5 for ORCs were also examined, to check whether the effect found for *Match* in the critical region is also observed in the spillover region. The statistical models indicate that *Match* is not significant at region 5 for any of the groups (intermediates: $\hat{\beta}=-0.08$, $SE=0.1$, $t=-0.79$, $p=0.43$; advanced: $\hat{\beta}=0.06$, $SE=0.06$, $t=1.01$, $p=0.31$; native speakers: $\hat{\beta}=-0.04$, $SE=0.06$, $t=-0.6$, $p=0.53$).

We also probed whether the same effect of *Match* is found for SRCs at the region where NP2 appears, namely, region 5. *Match* is not statistically significant for any of the groups (intermediates: $\hat{\beta}=0.01$, $SE=0.12$, $t=0.15$, $p=0.87$; advanced: $\hat{\beta}=0.11$, $SE=0.07$, $t=1.6$, $p=0.11$; native speakers: $\hat{\beta}=-0.02$, $SE=0.06$, $t=-0.35$, $p=0.72$). These results are consistent with the idea that only ORCs will exhibit an effect of *Match*, as there is no NP intervening between the head of the relative clause and the gap in SRCs (see (3)).

The possibility arises that the (unexpected) results found for *Match* in ORCs at region 4 might not be the result of intervention effects, but, rather, a more general effect of number. Previous studies have obtained different responses for singular and plural NPs, regardless of the match/mismatch issue. For example, Contemori and Marinis (2014) found that the head noun of the relative clause (NP1) was processed faster if it was plural than if it was singular.

To determine whether this might be the case for our participants, we looked at RTs at the point where NP1 is encountered (region 2). We examined all sentence types together (ORCs, SRCs and fillers), as all of the sentences have the same profile at this point (*I know the N who ...*). Figure

6 shows RTs for singular and plural NPs at region 2. Each group's responses were modelled with a linear regression including *Number* as main effect as well as a by-participant random intercept. *Number* at region 2 is not statistically significant for any of the groups (intermediates: $\hat{\beta}=0.06$, $SE=0.05$, $t=1.21$, $p=0.22$; advanced: $\hat{\beta}=-0.06$, $SE=0.03$, $t=-1.92$, $p=0.55$, native speakers: $\hat{\beta}=-0.007$, $SE=0.03$, $t=-0.21$, $p=0.82$).

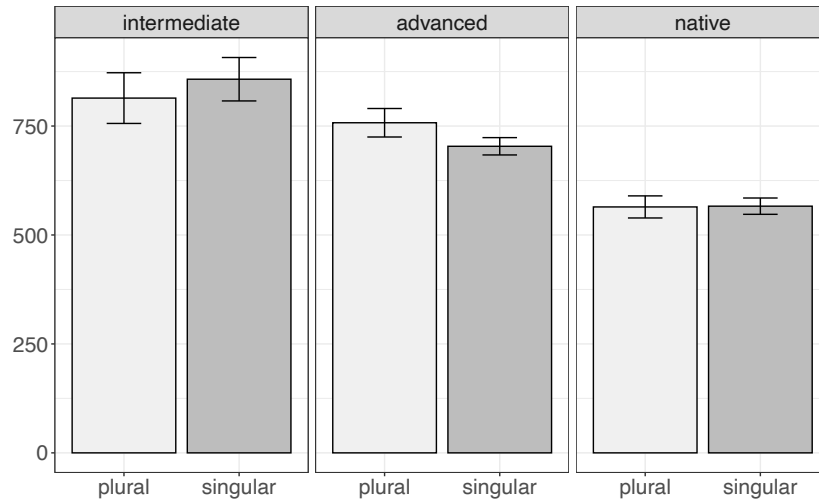


Figure 6. Response times for singular vs. plural NPs at region 2.

In summary, the results concerning SRCs and ORCs largely confirm the hypotheses, in that participants were generally accurate on both relative clause types, with the intermediate group and the native speakers being significantly more accurate on SRCs in both tasks. Similarly, SRCs were processed faster than ORCs by the advanced group and the native speakers. The results relating to performance on matched versus mismatched ORCs, on the other hand, are contrary to what was predicted, with matched items responded to faster than mismatched by both learner groups, at least at region 4, an effect which disappeared at region 5. The implications of the results are discussed in the next section.

Discussion

To summarize the results, as far as differences between SRCs and ORCs are concerned, we have found evidence for intervention effects in both tasks, as indicated by accuracy results in the picture

identification and self-paced reading tasks, as well as by response times on the self-paced reading task, all of which show an advantage for SRCs over ORCs. Consistent with prediction (16i), all groups were very accurate in interpreting relative clauses of both types, suggesting that SRCs and ORCs are appropriately represented and that adult L2ers observe all aspects of the distinctness hierarchy in (8) in the same way as adult native speakers, in other words permitting both types of relative clauses. Recall that in many of the studies on child L1 acquirers, discussed above, children performed at only around 50% accuracy on ORCs involving inclusion. It is noteworthy that the adult L2ers did not show any such difficulty. Nevertheless, accuracy on SRCs was significantly higher than ORCs in almost all cases, indicating a processing effect: parsing of ORCs is computationally more complex than SRCs, in that the relative head has to be retained in working memory longer before resolution is possible (Gordon et al., 2001; Belletti and Rizzi, 2013; amongst many others). This additional complexity leads to occasional inaccuracies in interpreting ORCs, for native speakers and L2ers alike.¹⁵ Consistent with prediction (16ii), response times to ORCs were significantly slower than to SRCs for the advanced L2ers and the native speakers, again suggesting a processing effect when intervention is involved. However, the effect was not significant for the intermediate group.

The status of featural differences, on the other hand, is less clear. The feature we have manipulated is number. Recall the claim that ORCs where the intervenor and head mismatch in features, as in (4b), will be easier to acquire/process than cases where there is a match, as in (4a). However, we obtained the opposite result, with match showing an advantage over mismatch, or no difference between them. As far as accuracy is concerned, on the picture identification task, L2 participants were significantly less accurate on stimuli involving mismatched features, although there was no effect on accuracy in the self-paced reading task. As far as processing is concerned, prediction (16iii) was that mismatched features would give a processing advantage; in other words, that participants would respond faster to ORCs where the intervenor mismatched the head in number (sg pl or pl sg versus sg sg). In fact, RTs at the critical region (region 4) showed that mismatched items were responded to more slowly by both L2 groups. It is well known that L2 participants are often delayed in giving expected responses. At region 5, the spillover region, the

¹⁵ See Hopp (2010) and Slabakova et al. (2017), who argue that computational complexity in offline tasks results in difficulties for L2ers that reflect processing rather than representation.

mismatched cases no longer resulted in longer RTs. At best this provides weak support for the idea that the mismatch was in fact having a beneficial effect. (Gordon et al. (2004) make a similar argument involving the disappearance of an effect to motivate advantages for mismatched NPs.) However, we acknowledge a problem with our stimuli in that the sentences involving SRCs did not include a spillover region. In future research, it would be desirable to add a phrase (such as *last week*) to both SRC and ORC stimuli, in order to control more precisely for potential spillover effects.

As for the advantage for matched items in ORCs (accuracy in the picture identification task, RTs at region 4 in the self-paced reading task), it is possible that the L2ers pay more attention to – and take more time over – a number mismatch, as their L1 does not have obligatory number marking on nouns. In other words, the L1 is indirectly implicated after all, not in the sense that L2ers fail to represent number, but rather in the sense that overt number stands out for them, leading to minor inaccuracies or delays. Consequently, we might expect to see greater inaccuracies or delays in pl pl sentences due to overt number on both the NPs. However, such sentences were not included in our stimuli. In future research, it would be prudent to use all four possible combinations of sg sg, sg pl, pl sg and pl pl.

As far as RTs are concerned, since no effect of *Number* on NP1 was found, the significant difference in learners' RTs for match and mismatch at region 4 does not seem to be a consequence of taking longer to process singular or plural as such. It is possible that, once presented with a mismatch, L2ers scan the sentences again in their minds (since the earlier parts of the sentence are not visible on the screen), which results in longer RTs at the region where the mismatch is observed. In order to determine whether absence of inflection in the L1 is playing a role here, it would be advisable to add L2ers whose L1 has richer inflection; such L2ers might be expected to show the predicted benefit for mismatched cases.¹⁶

¹⁶ A reviewer suggests that, because of absence of inflection in the L1, these L2ers might not compute the number feature in the way required for featural RM, that is, as a feature involved in computing A' dependencies. This reviewer points out that, in that case, one might expect difficulties with agreement as well. However, there is well-established L2 research that has shown that L2ers (including L2ers with L1 Chinese) have underlying knowledge of number and agreement and their syntactic effects even when inflection is omitted (e.g., Haznedar and Schwartz, 1997; Lardiere, 1998; Prévost and White, 2000).

The absence of a mismatch effect is not in principle problematic. According to the hierarchy in (8), disjunction, intersection and inclusion are all permitted in adult grammars. In other words, it is not the case that effects of featural differences will inevitably be found. Rather, if featural differences are found, they are expected to favour the mismatched cases. No difference between matched and mismatched items was in fact observed for the native speakers on all measures, for the advanced L2ers on most measures, and for the intermediates on some measures. An issue, then, is why we obtained no effect for certain groups on certain measures whereas others have reported mismatch effects. We note, however, that there are other studies which have failed to find a mismatch effect for number. Using an online listening task, Contemori and Marinis (2014) found an effect for number (in that plural marking speeded up processing) but not for match vs. mismatch in the response times of English-speaking children and adults. As far as L2 learners are concerned, Suzuki and Hirakawa (2019) found no effects for number match versus mismatch in elicited production of relative clauses, which were in fact replaced by passives regardless of match/mismatch, thus circumventing the intervention problem.

Several studies on L1 acquisition, on the other hand, do find an effect for number (e.g., Adani et al., 2010; Adani et al., 2014). As noted above, in these studies, there were two overt cues to the match or mismatch, namely number marking on the noun as well as verbal agreement, as shown in (6). It is therefore not possible to determine whether a mismatch on the intervenor alone would have had an effect or whether what was observed was not solely an RM effect. The only study to address this issue experimentally is Bentea and Durrleman (2017), who manipulated presence or absence of inflection on the verb in French. They found that children aged 6-8 showed an advantage in mismatched ORCs in French, regardless of presence or absence of overtly pronounced verbal inflection, whereas 5 year olds showed no effect for mismatch. Number marking on NPs in French is pronounced only on the article; it could be that the difference between *le* and *les* is more salient (at least for the older children) than the difference between singular and plural inflection on a noun, as is the case for English.¹⁷

In order to determine the extent to which overt marking of inflection on both noun and verb explains the results appearing to support featural RM in L1 acquisition, it is important to replicate

¹⁷ It is also possible that the older children were familiar with the fact that in the written language, singular and plural verbal agreement are spelled differently, so that effectively there is indeed double marking of inflection on all verbs in French if written forms are taken into consideration.

these studies using verb forms that do not indicate number, such as past tense verbs in English, as used in our study. Additionally, our experiment could be rerun with stimuli in the present tense, so that number distinctions are realized on the verb as well as the NPs involved. It remains, then, an open question as to whether the effects of featural RM have in fact been truly tested in many previous studies, since the verb is not a potential filler for the gap but nevertheless provides an extra indication of differences between matched and mismatched items. We suspect that double marking of inflection is in fact necessary (but not sufficient; see Contemori and Marinis (2014)) for the mismatch effect to show up as far as language learners (L1 or L2) are concerned but this remains to be investigated.

In conclusion, we have found differences between SRCs and ORCs in native speaker and L2 learner performance, and we have suggested that these are consistent with featural RM and are attributable to processing rather than reflecting a representational problem. Differences between SRCs and ORCs would also be consistent with other theories, such as the Noun Phrase Accessibility Hierarchy (Keenan and Comrie, 1977) or the similarity based interference approach of Gordon and colleagues (Gordon et al., 2001, 2004), amongst others. Many L2 researchers have proposed that working memory capacity impinges on L2 performance (e.g., Cunnings, 2017; Hopp 2010). An account in terms of working memory would also expect better performance on SRCs than ORCs.

Belletti and Rizzi (2013) point out that the featural RM account makes different predictions from more general accounts attributing problems to working memory. In particular, the RM account expects difficulties under very specific circumstances, involving certain kinds of morphosyntactic features. An account in terms of working memory would not, presumably, expect differences between disjunction, intersection and inclusion; that is, it would not predict performance differences depending on the properties of certain morphosyntactic features.

In other words, the more crucial test case for featural RM involves match versus mismatch of features and here our results are contrary (match better than mismatch) or neutral (no effect for mismatch). Even though we did not get the expected mismatch effect, we did get an effect (match preferred) in certain cases, which is also presumably not consistent with a general working memory account. We reiterate the point that featural RM does not in fact predict that featural differences must result in performance differences; rather it anticipates that if there are differences they will be in one direction rather than the other. We have suggested that some of the earlier research that

has reported effects in the predicted direction for number in L1 acquirers may be attributable to the kinds of sentences that were used as stimuli, where number marking was not restricted to the intervenor. This is an issue that clearly requires additional research on Italian and English, along the lines initiated by Bentea and Durrleman (2017) for French.

Acknowledgements

We would like to thank the SLR reviewers for their helpful comments, as well as the audiences at EUROSLA 2018, GALANA 2018 and SLRF 2018. Thanks to Erica Brandao for assistance with recruitment and testing and to Guiherme Garcia for advice on the statistical analyses.

Funding

This research was supported by research grant #2016-SE-188196 from le Fonds de recherche du Québec – société et culture (to Lydia White, PI) and by the ARIA program of the Faculty of Arts, McGill University (to Vera Xia).

References

- Adani A, Forgiarini M and Guasti MT (2014) Number dissimilarities facilitate the comprehension of relative clauses in children with (grammatical) specific language impairment. *Journal of Child Language* 41: 811-841.
- Adani F, van der Lely HJK, Forgiarini M and Guasti MT (2010) Grammatical feature dissimilarities make relative clauses easier: A comprehension study with Italian children. *Lingua* 120: 2148–2166.
- Bates D, Maechler M, Bolker B and Walker S (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67: 1–48.
- Belletti A and Guasti MT (2015) *The acquisition of Italian*. Amsterdam: John Benjamins.
- Belletti A and Rizzi L (2013) Intervention in grammar and processing. In: Caponigro I and Cecchetto C (eds) *From grammar to meaning: The spontaneous logicity of language* Cambridge UK: Cambridge University Press, pp. 294-311.

- Belletti A, Friedmann N, Brunato D and Rizzi L (2012) Does gender make a difference? Comparing the effect of gender on children's comprehension of relative clauses in Hebrew and Italian. *Lingua* 122: 1053–69.
- Bentea A and Durrleman S (2017) Now you hear it, now you don't: Number mismatch in the comprehension of relative clauses. In: LaMendola M and Scott J (eds) *Proceedings of the 41st annual conference of the Boston University Conference on Language Development*. Somerville MA: Cascadilla Press.
- Chen B, Ning A, Bi H and Dunlap S (2008) Chinese subject-relative clauses are more difficult to process than the object-relative clauses. *Acta Psychologica* 129: 61-65.
- Chomsky N (1965) *Aspects of the theory of syntax*. Cambridge MA: MIT press.
- Chondrogianni V and Marinis T (2012) Production and processing asymmetries in the acquisition of tense morphology by sequential bilingual children. *Bilingualism: Language and Cognition* 15: 5-21
- Contemori C and Belletti A (2014) Relatives and passive object relatives in Italian-speaking children and adults: Intervention in production and comprehension. *Applied Psycholinguistics* 35: 1021-1053.
- Contemori C and Marinis T (2014) The impact of number mismatch and passives on the real-time processing of relative clauses. *Journal of Child Language* 41: 658-689.
- Contemori C, Carlson M and Marinis T (2018) On-line processing of English *which*-questions by children and adults: a visual world paradigm study. *Journal of Child Language* 45: 415-441.
- Costa J, Lobo M and Silva C (2011) Subject-object asymmetries in the acquisition of Portuguese relative clauses: Adults vs. children. *Lingua* 121: 1083-1100.
- Doughty C (1991) Second language instruction does make a difference: Evidence from an empirical study of SL relativization. *Studies in Second Language Acquisition* 13: 431-470.
- Durrleman S, Marinis T and Franck J (2016) Syntactic complexity in the comprehension of *wh*-questions and relative clauses in typical language development and autism. *Applied Psycholinguistics* 37: 1501-1527.
- Eckman F, Bell L and Nelson D (1988) On the generalization of relative clause instruction in the acquisition of English as a second language. *Applied Linguistics* 9: 1-20.

- Felser C and Cunnings I (2012) Processing reflexives in a second language: The timing of structural and discourse level constraints. *Applied Psycholinguistics* 33: 571-603.
- Friedmann N, Belletti A and Rizzi L (2009) Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua* 119: 67-88.
- Gass S (1982) From theory to practice. In: Hines M and Rutherford W (eds) *On TESOL*. Washington DC, pp. 129-139.
- Gibson E and Wu, H-H I (2013) Processing Chinese relative clauses in context. *Language and Cognitive Processes* 28: 125-155.
- Gordon PC, Hendrick R and Johnson M (2001) Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26: 1411-1423.
- Gordon PC, Hendrick R and Johnson M (2004) Effects of noun phrase type on sentence complexity. *Journal of Memory and Language* 51: 97-114.
- Haznedar B and Schwartz BD (1997) Are there optional infinitives in child L2 acquisition? In: Hughes E, Hughes M and Greenhill A (eds) *Proceedings of the 21st Annual Boston University Conference on Language Development*. Somerville MA: Cascadilla Press, pp. 257-268.
- Hsiao F and Gibson E (2003) Processing relative clauses in Chinese. *Cognition* 90: 3-27.
- Hopp H (2009) The syntax-discourse interface in near-native acquisition: Off-line and on-line performance. *Bilingualism: Language and Cognition* 12: 463-483.
- Hopp H (2010) Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua* 120: 901-931.
- Hu S, Gavarró A and Guasti MT (2016) Children's production of head-final relative clauses: The case of Mandarin. *Applied Psycholinguistics* 37: 323-346.
- Hu S, Gavarró A, Vernice M and Guasti MT (2016) The acquisition of Chinese relative clauses: Contrasting two theoretical approaches. *Journal of Child Language* 43: 1-21.
- Juffs A (2004) Representation, processing and working memory in a second language. *Transactions of the Philological Society* 102: 199-225.
- Keenan E and Comrie B (1977) Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8: 63-99.
- Kuznetsova A, Brockhoff PB and Christensen RHB (2014) lmerTest: Tests for random and fixed effects for linear mixed effect models. R package.

- Lardiere D (1998) Dissociating syntax from morphology in a divergent end-state grammar. *Second Language Research* 14: 359-375.
- Lin C-JC and Bever TG (2006) Subject preference in the processing of relative clauses in Chinese. In: Baumer D, Montero D and Scanlon M (eds) *Proceedings of the 25th West Coast Conference on Formal Linguistics*. Somerville MA: Cascadilla Proceedings Project, pp. 254-260.
- Manetti C, Moscati V, Rizzi L and Belletti A (2016) The role of number and gender features in the comprehension of Italian clitic left dislocations. In: Scott J and Waughtal D (eds) *Proceedings of the 40th Annual Boston University Conference on Language Development*. Somerville MA: Cascadilla Press, pp. 229-240.
- Omaki A and Schulz B (2011) Filler-gap dependencies and island constraints in second-language sentence processing. *Studies in Second Language Acquisition* 33: 563-588.
- Phillips C and Ehrenhofer L (2015) The role of language processing in language acquisition. *Linguistic Approaches to Bilingualism* 5: 409-453.
- Prévost P and White L (2000) Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research* 16: 103-133.
- R Development Core Team (2019) *R: A language and environment for statistical computing*. Vienna, Austria.
- Rizzi L (1990) *Relativized minimality*. Cambridge MA: MIT Press.
- Rizzi L (2018) Intervention effects in grammar and language acquisition. *Probus* 30: 339-367.
- Schneider W, Eschman A and Zuccolotto A (2002) E-Prime user's guide. Pittsburgh: Psychology Software Tools Inc.
- Sheldon A (1974) The role of parallel function in the acquisition of relative clauses in English. *Journal of Verbal Learning and Verbal Behavior* 13: 272-281.
- Slabakova R, White L and Guzzo NB (2017) Pronoun interpretation in the second language: the role of computational complexity. *Frontiers in Psychology* 8:1236. doi: 10.3389/fpsyg.2017.01236.
- Sorace A (2011) Pinning down the concept of “interface” in bilingualism. *Linguistic Approaches to Bilingualism* 1: 1-33.
- Sorace A and Filiaci F (2006) Anaphora resolution in near-native speakers of Italian. *Second Language Research* 22: 339-368.

- Suzuki K (2019) Intervention effects in relative clause production in L2 English. Poster presented at Generative Approaches to Second Language Acquisition (GASLA XV), Reno, March 2019.
- Suzuki K and Hirakawa M (2018) Relative clauses in heritage Chinese: Explanation by the Relativized Minimality approach. Poster presented at the European Second Language Association (EUROSLA 28), Münster, September 2018.
- Suzuki K and Hirakawa Y (2019) Intervention effects in relative clause production: An L2 English study. Poster presented at the European Second Language Association (EUROSLA 29), Lund, August 2019.
- Villata S, Rizzi L and Franck J (2016) Intervention effects and relativized minimality: New experimental evidence from graded judgments. *Lingua* 179: 76-96.

Appendix I. Stimuli

Test items

RC type	NP1	NP2	Sentence
SRC	sg	sg	Show me/I know the musician who touched the waiter.
	sg	sg	Show me/I know the king who pushed the boy.
	sg	sg	Show me/I know the athlete who kicked the doctor.
	sg	sg	Show me/I know the chef who watched the soldier.
	sg	sg	Show me/I know the girl who hugged the queen.
	sg	sg	Show me/I know the scientist who kissed the artist.
	sg	pl	Show me/I know the musician who touched the waiters.
	sg	pl	Show me/I know the king who pushed the boys.
	sg	pl	Show me/I know the athlete who kicked the doctors.
	pl	sg	Show me/I know the chefs who watched the soldier.
	pl	sg	Show me/I know the girls who hugged the queen.
	pl	sg	Show me/I know the scientists who kissed the artist.
ORC	sg	sg	Show me/I know the musician who the waiter touched.
	sg	sg	Show me/I know the king who the boy pushed.
	sg	sg	Show me/I know the athlete who the doctor kicked.
	sg	sg	Show me/I know the chef who the soldier watched.
	sg	sg	Show me/I know the girl who the queen hugged.
	sg	sg	Show me/I know the scientist who the artist kissed.
	sg	pl	Show me/I know the musician who the waiters touched.
	sg	pl	Show me/I know the king who the boys pushed.
	sg	pl	Show me/I know the athlete who the doctors kicked.

pl	sg	Show me/I know the chefs who the soldier watched.
pl	sg	Show me/I know the girls who the queen hugged.
pl	sg	Show me/I know the scientists who the artist kissed.

Fillers

RC type	NP1	Sentence
Intransitive	sg	Show me/I know the chef who laughed.
		Show me/I know the queen who slept.
		Show me/I know the soldier who cried.
		Show me/I know the girl who walked to school.
		Show me/I know the boy who ran from the hospital.
		Show me/I know the athlete who climbed up the stairs.
	pl	Show me/I know the chefs who laughed.
		Show me/I know the queens who slept.
		Show me/I know the soldiers who cried.
		Show me/I know the girls who walked to school.
		Show me/I know the boys who ran from the hospital.
		Show me/I know the athletes who climbed up the stairs.
