

CROSSING BRANCHES: THE CHALLENGE OF ACQUIRING RECURSIVE NOMINAL STRUCTURES IN L2 ACQUISITION *

Natalia Docteur, Jida Jaffan, Ana T. Pérez-Leroux, and Mathusha Sanjeevan
University of Toronto

1. Introduction

Can second language learners fully acquire syntactic structures in their second language (L2) that do not exist in their first language (L1)? To which extent do L1 structures set the limits on what can be achieved in the L2? When L2 acquisition occurs after L1 constraints become entrenched, speakers face a dual burden of learning the new L2 rules while simultaneously un-learning the L1 rules (Rothman & Pascual y Cabo, 2013). Therefore, overlap in rules and structures between the L1 and L2 may facilitate L2 acquisition. Our goal is to examine the effect of L1 similarity or difference as an intervening factor in the acquisition trajectory of a L2 learner.

Several experimental studies have focused on specific transfer effects of a speaker's L1 on their L2. Luk and Shirai (2009) examined the L2 developmental trajectories of English nominal functional elements (i.e., possessives, plural -s, and articles) by L1 speakers of Japanese, Chinese and Spanish. These authors predicted that structural overlap in grammatical morphemes between the L1 and L2 would promote early L2 acquisition, and that an absence of similar syntactic representations in the L1 would hinder L2 acquisition. Their predictions were borne out; L1 speakers of Chinese & Japanese acquired the English possessive -s more readily than the plural -s, while L1 speakers of Spanish acquired plural -s and articles before possessive -s.

While structural overlap may predict positive transfer or facilitation, lack of structural overlap can predict negative transfer. Snape, Mayo and Gürel (2013) found that L1 speakers of Turkish, Spanish, and Japanese performed significantly differently on a forced elicitation task targeting English articles. Performance was predicted by the nature of the article system in the L1 (or the lack thereof). L1 Spanish speakers were target-like in their use of definite and indefinite articles as the Spanish article system mirrors that of English. L1 Turkish speakers performed well on indefinite articles, but regularly omitted definite articles, which do not exist in their L1. Similarly, L1 Japanese speakers also exhibited high rates of omission as predicted by the lack of article system in their native language.

*This work was supported by SSHRC 435-2014-2000 Development of Noun Phrase Complexity in Children grant to Ana T. Pérez-Leroux and Yves Roberge. We acknowledge the input and advice of various members of The Complexity and Recursion Group at The University of Toronto. Special thanks to Gaby Klassen and Erin Pettibone.

Language variation is not limited to functional elements. Structural depth may also vary across languages: while the Saxon genitive marker ‘s denotes possession in German *ex. Marias Haus*, its use cannot be extended to a second NP to create a recursive structure like the one existing in English *ex. Maria’s dad’s bike* (Roeper & Snyder, 2005). Limbach and Adone (2010) compared preschool-aged monolingual English speaking children with native German speakers who had acquired English in adulthood, and found that the acquisition trajectory of recursive sentences in L2 learners were qualitatively distinct from their L1 counterparts. Their results demonstrate that while second language learners (SLL) were equally successful (62%) in their comprehension of recursive structures as the native speaking children (~60%), both groups performed far below native English-speaking adults (~82%). While the monolingual English speaking children appeared to progress through developmental stages of errors to adult-like recursive phrases, German SLL’s of English displayed no consistent patterns in error types. This suggests that despite comparable success rates with the native speaking children (62% vs. ~60%), the SLL’s had a generalized comprehension deficit that could not receive a structural explanation. Nelson (2016) also documents difficulties but also establishes improvement with proficiency.

1.2 Our Study

The current paper aims to examine the acquisition of recursive embedding by L2 learners of English. Specifically, by examining two languages that contrast in branching directionality (Spanish and Mandarin) we provide a structural explanation for qualitative differences associated with the L2 acquisition of recursive embedding. Branching references which side of a head noun (in a noun phrase) modifiers attach. For example, *the cat’s toy*, is considered left branching since *toy* is the head noun, and the modifier, *cat*, appears to its left. On the other hand, *the dog under the table*, is right branching, since the modifier *under the table*, appears to the right of the head noun, *the dog*. English is considered both right and left branching, while Spanish is mainly right branching and Mandarin is left branching (Her, 2012).

Following Luk and Shirai (2009) and Snape, Mayo and Gürel (2013) we predict that syntactic overlap may aid L2 acquisition, and that SLL’s may prefer to use their native structures in the L2 when faced with lexical ambiguity. If overlapping structures in the L1 and L2 benefit SLL’s, because the L1 parameters are entrenched and early encounters with the L2 can only be computed through the lens of the L1, L2 learners may initially rely on pre-existing L1 structures to anchor L2 learning and then gradually shift to nascent L2 syntax over time. Below we report on the results of a production task targeting English recursive possessives, comitatives, locatives and relational nouns. This study sought to answer the following questions:

- Do L2 English speakers and monolingual English speakers produce similar response patterns?
- Are experienced adult L2 speakers of Spanish and Mandarin comparable to native speakers and to each other in their ability to produce recursive structures?

- To what extent does similarity from the L1 facilitate acquisition of recursive structures?

2. Methods

A total of 49 adults (ages 18 and over) participated in the study. Thirty-six were L2 English speakers whose L1 was either Mandarin (n=20) or Spanish (n=16) with thirteen monolingual controls. Participants were recruited through personal connections within the university setting, where they were tested in the Language and Complexity Lab at The University of Toronto. All participants completed the tasks in English.

2.1 Tasks

Participants completed a Production Task designed to elicit recursive structures of four types of noun phrases: possessives, comitatives, locatives and relational nouns. In the Production Task participants were shown a PowerPoint presentation which contained a written narrative accompanied by images, as well as an audio recorded description. The descriptive preamble highlighted the multiple referential competitors present in the picture scenario, and introduced all the relevant vocabulary items needed to provide a complete response. Importantly, participants needed to produce a recursively modified NP in order to provide an adequate response to the referential question (*Which x is...?*). Participants were asked to state their answer out loud to ensure quality of recording. In total, the production task consisted of 42 items. These included six trials each of our four target conditions: possessives, comitatives, locatives, and relational nouns, plus additional distractor sentences.

Possessive Condition:

Description: Dora brought a friend, and Diego brought his friend too. They each have a backpack.

Prompt: Which backpack is blue?

Target Response: Dora's friend's backpack.

Comitative Condition:

Description: These two women are taking their babies out. This woman is carrying her bag. This woman has flowers. But one of the babies started to cry.

Prompt: Who's crying?

Target Response: The baby with the woman with the flowers.

Locative Condition:

Description: In the yard of this house, there are two trees. One is far from the house, one is near the house. They let the dogs out in the yard, but one of the dogs started to bark.

Prompt: Which dog is barking?

Target Response: The dog next to the tree next to the house.

Relational Noun Condition:

Description: These little mice live in a doll house. Some of the mice are wearing a tie, some are not.

Prompt: Which mouse is happy?

Target Response: The mouse with the tie inside the house.

In addition, participants completed a Language Questionnaire, a Cloze Task, and a Vocabulary Assessment. The questionnaire prompted for language history, current rates of usage and self-rated proficiency. Each blank in the Cloze task had four different words as options for participants to select. For the vocabulary assessment, participants were given 50 definitions and were asked to select the word that best matched the definition.

2.2 Coding

Participants' productions were transcribed and coded. The coding system included a) the level of embedding of the structure produced, b) a referential coding that described whether and how referential success was achieved and c) the types of links (-s, PP, relative clauses, etc.) used to formulate these structures. Levels of embedding include:

Single NP: *the alligator*

Level 1: *the alligator [in the water]*

2Level1: *the girl [with the bow] [with the bike]*

Level 2: *the bird [on the alligator [in the water]]*

The referential analysis simultaneously considered the syntax and semantics of structures elicited from the Production Task. Referential context was coded as:

Alternative- A different syntactic structure from the target structure that was able to achieve referential success.

Target Response: *Batman's baby's lollipop.*

Participant's Response: *The one that batman's baby is carrying now is bigger.*

Incomplete- Did not make reference to all target referents and thus the description was not uniquely referential.

Target Response: *Dora's friend's backpack.*

Participant's Response: *Um Dora's friend.*

Sequential- The target referent was identified but not in an integrated utterance. This often resulted from the use of follow up questions.

Target Response: *The boy's cat's tail is fluffy.*

Participant's Response: *Uh ... the- the cat is held by the ... uh ... by the boy is uh has a fluffy tail.*

Target Embedded- Target response achieved with recursively embedded NPs.

Target/ Participant's Response: *The clown's monkey's balloon.*

Nonembedded- Forms such as coordination, compounding, apposition, etc. were used to relate the relevant elements.

Target Response: *The hotdog with ketchup on the table*

Participant's Response: *Um ... is the hotdog in the blue table and ... with mustard.*

In addition, target structures were further coded along additional parameters. Responses were identified as ‘strict targets’ only if the specific type of recursive structure representing the condition was produced uniformly (i.e., for the possessive condition only two –s possessives were counted as strict target). Target responses were also coded by what combination of embedding strategies were used in the structure (i.e., either two possessives, a mix of possessive and other strategy, only PPs, only relative clauses, a mix of PPs and relative clauses). This was used to gage the overall branching directionality of the recursive structures produced: right-branching, left-branching or mixed.

Transcriptions were completed by two student researchers which were then checked for discrepancies. Coding of the Production Task was completed by three student researchers and was subsequently checked by an independent coder for accuracy.

3. Results

Participants in the study were two groups of L2 English speakers with different first languages: Mandarin (n=20), and Spanish (n=16). One group of monolingual native English speakers (n=13) served as controls. Participants were recruited from a university setting. As Table 1 shows, the L2 groups started learning English as children and adolescents, and had lived in Canada for an average of 3-4 years. There were no significant differences in age of onset of L2 exposure ($W = 136$, $p = 0.459$). On average, the Spanish speakers lived about one year less in an English-speaking environment ($W = 247$, p -value = 0.005). They reported higher levels of current use, but this difference did not achieve significance ($t = -1.819$, $df = 26.253$, $p = 0.08$).

Table 1. Experience profiles for L2 participants.

Language Group	Age of Onset	SD	% of L2 Use	SD	Years in country	SD
Mandarin	10.50	(3.76)	37.30	(17.15)	3.65	(1.53)
Spanish	11.63	(5.41)	50.28	(24.06)	2.85	(3.66)

Table 2 compares English language proficiency measures between the Spanish and Mandarin SLL. The two L2 groups were close in overall English language ability, as measured by the Cloze task ($W = 146$, $p = 0.671$). Spanish-speakers rated their own

English abilities higher than the Mandarin speakers did ($W = 96$, $p = 0.042$), and indeed had higher vocabulary scores ($W = 41$, $p < .001$).

Table 2. English language ability for L2 participants.

Group	Self-rating	SD	Cloze Test	SD	Vocabulary	SD
Mandarin	6.63	(1.55)	13.30	(2.05)	28.55	(7.14)
Spanish	7.37	(1.18)	12.94	(3.23)	38.69	(4.63)

Are the three language groups different in their ability to produce recursive NPs? As expected, native English speakers produced substantively more recursive NPs than the two L2 groups. Table 3 reports recursive responses in two ways: all types included, or only the expected target type (strict target). We calculated, for each speaker, the proportion of total recursive responses. As these data were normally distributed, we conducted an ANOVA to evaluate the differences between groups. The effect of group was marginally significant ($F_{2,46} = 3.0847$, $p = 0.055$). Both L2 groups were significantly different from the English monolinguals ($t = -2.184$, $p = 0.034$; for Spanish L1, $t = 2.228$, and for Mandarin L1, $p = 0.031$), but no different from each other ($t = 0.057$, $p = 0.954$).

Table 3. Mean proportion of recursive responses per group.

Group	All recursive (Mean)	SD	Strict Targets (Mean)	SD
English	0.526	(0.13)	0.342	(0.14)
Mandarin	0.401	(0.13)	0.124	(0.10)
Spanish	0.398	(0.21)	0.135	(0.13)

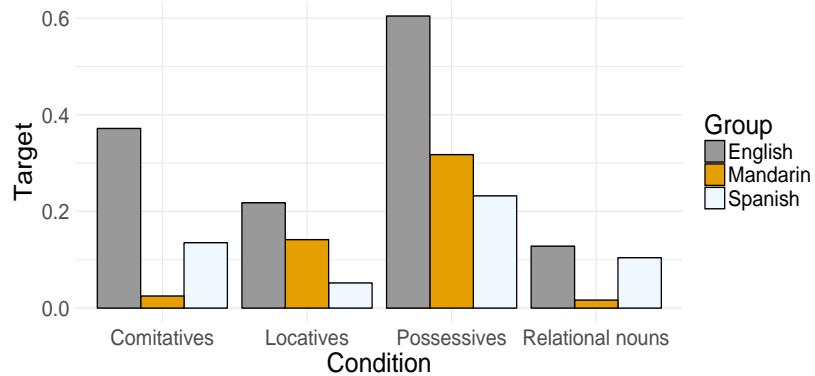
The external variables were not associated with the overall ability to produce recursive structures. Only general language ability, as measured by the Cloze test, showed a positive correlation with proportion of all recursive structures produced. The correlation was weak but significant ($r = 0.356$, $p = 0.033$).

Do L2 speakers and monolingual English speakers produce similar response patterns? To explore the effect of directionality contrasts, we analysed the effect of group for each condition separately. Figure 1 displays an analysis of strict targets by condition. This analysis concentrates only on responses where the expected type of embedding was considered ('s for the possessive, PPs for the other conditions), i.e., the strict target. The data was entered into a linear mixed effect (logit) model with Target as the dependent variable, Group as a fixed effect and Participant and Item as random effects. The formula was as in (x)

$$(x) \text{Target} \sim \text{Group} * (1 | \text{Participant}) + (1 | \text{Item})$$

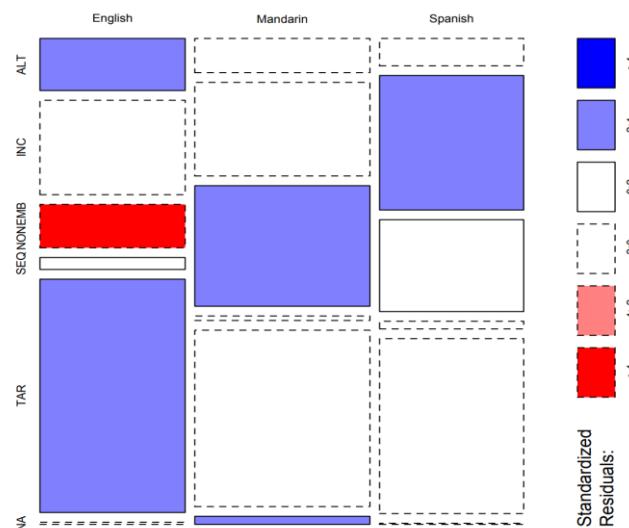
For possessives, the Mandarin speakers were marginally different from the monolingual group (Estimate=-0.8283, $z=1.714$, $p= 0.086$), and the Spanish speakers were also significantly different (Estimate=-1.369, $z=-2.745$, $p=0.006$). For comitative PPs, only Mandarin speakers differed significantly from controls (Estimate, -1.706, $z=-3.556$, $p< 0.001$). The group effect for relational nouns or locatives was not significant.

Figure 1. Mean proportion of target response (strict types) per condition per group.



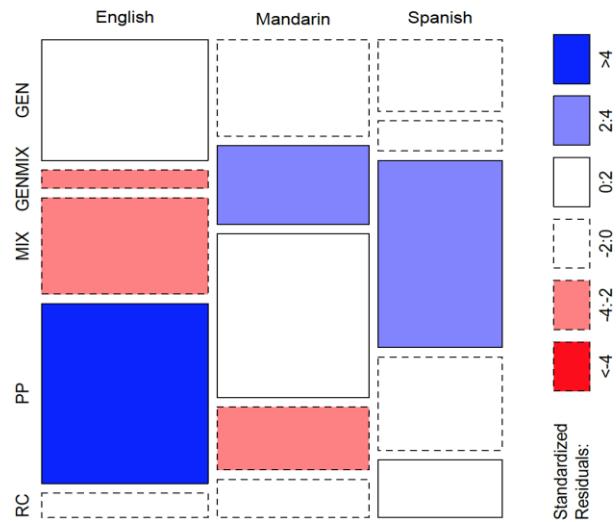
As shown in Figure 2, there are differences in the distribution of all response types across groups ($\chi^2=66.623$, $df=8$, $p<.001$). Overall, the analysis of residuals shows that English speakers produce significantly more target responses and less non-embedded responses than the L2 groups. L2 speakers generally produced more non-target responses than the English speakers, but of different kinds. Mandarin speakers produced significantly more non-embedded responses, while the Spanish speakers produced significantly more incomplete responses.

Figure 2. Analysis of all response types by group.



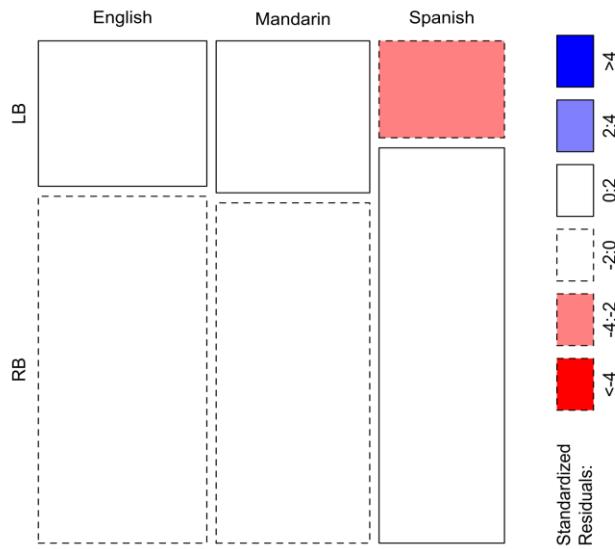
Further exploring the data, we compared the distribution of various linkers used in each condition (possessives, locatives etc.) across language groups. Shown in Figure 3, the analysis indicates that where English speakers produced more PP's and less mixed types, L2 speakers produce more mixed types (Mandarin with mixed structures that include a genitive, and Spanish speakers by mixing RC and PP). The differences in overall distribution were significant ($\chi^2 = 76.17$, $df = 8$, $p < .001$).

Figure 3. Distribution of types of target responses across languages.



Abstracting from condition, and including all forms of response types, not just the strict target, we see that there is an overall effect of branching ($\chi^2 = 12.956$, $df = 2$, $p = 0.002$), primarily driven by the Spanish speakers, who used left branching less than other groups. This can be seen in Figure 4.

Figure 4. Distribution of left branching (LB) and right branching (RB) links across the recursive responses (all types) across languages.



4. Discussion

The analysis of the measures of general language ability and the language history questionnaire suggest that both groups of participants were very similar with one exception, Spanish speakers had higher vocabulary rates. We believe that the higher scores on the vocabulary task from Spanish speakers may be attributed to the greater similarity between the Spanish and English lexicons. However, the design of this study ensured that having a larger vocabulary would not affect one's ability to produce different syntactic structures since all referents in the display were named prior to sentence production. Furthermore, the participants were equal on all other tasks and as such our results could not be attributed to differences in proficiency. We can therefore attribute the main difference between the two sets of participants to their L1 language.

We observed no clear effects of age of onset of exposure, self-rating, or current levels of language use in L2 speakers' ability to produce recursive targets. Only global language ability, as reflected by performance in the Cloze test, was moderately correlated with recursive ability.

The analysis of response types (Figures 1 & 2) shows that native speakers succeed more often than non-natives, and produce less non-embedded responses. This is what we would expect since L2 speakers are mature adults, aware of the referential requirements of the task. While Spanish speakers produce more incomplete responses than Mandarin L2 speakers, it is not clear that a linguistic explanation can be attributed to this finding.

The analysis of strict target per condition (Figure 1) provides clear evidence of an effect of directionality, at least for the two most productive constructions, possessives and comitatives. Mandarin speakers produced significantly less strict targets than controls for the (right-branching) comitative condition, and Spanish speakers produced

significantly fewer (left-branching) possessives than the native controls. The Mandarin speakers produced less recursive possessives than controls, but the trend did not reach significance. Other effects showed no significance.

Our second analysis collapsed the various types of target responses, classified as per the link types, across condition (Figure 3). This allowed us to examine the overall configuration of responses, independent of whether a given link was employed. This resulted in 5 categories of recursive responses:

a) Two linked PPs

Example: *[the mouse is sitting [on the box [with the tomato cans]]]*

b) Two linked RCs

Example: *[the bucket [that contains the ... crayons [that are bundled together]]]*

c) Mixed PP and RC links

Example: *[the mouse is sitting [on the box [that has uh ... tomato cans in- inside it]]]*

d) Two linked possessive genitives

Example: *[Elmo['s sister['s ball is flat]]]*

e) Mixed genitive and either RC or PP.

Example: *[the backpack [of Dora ['s friend]]]*

Types a-c) are uniformly right branching, type d) is left branching, and e) is of mixed directionality. Our analysis shows that L2 speakers prefer relative clauses over PPs, when compared to native speakers. This pattern also originates from the L1, as Mandarin and similar languages do not allow DP-internal PP modification, relying on relative clauses instead. Spanish allows a limited of types of PPs as DP-internal modifiers (Pérez-Leroux, 2016), and unrestricted use of relative clauses. Both L2 speakers preferred mixed structures, but whereas the Mandarin L1 group produced significantly more of the e) type above which included one instance of left-branching, Spanish speakers opted for the uniformly right branching option c).

L2 speakers are able to produce complex recursive NP structures, but not at the same rates as native speakers. Our study shows that overall, L2 speakers prefer to mix strategies over producing iteration of embedding of the same type. It seems clear that the branching typology of the L1 type clearly shapes speakers' preferences, even when alternative constructions have been acquired. One possible mechanism for transfer effects is cross-language syntactic priming. Syntactic priming refers to certain sentence constructions being preferred over others due to previous exposure. Constructs that are used in one's L1 may prime the same construct in an L2 (Mercan, 2016). This may be an explanation for the results we found, and prove a fruitful venue for future explorations.

References

- Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition*, 75(2), B13-B25.
- Branigan, H. P., Pickering, M. J., Liversedge, S. P., Stewart, A. J., & Urbach, T. P. (1995). Syntactic priming: Investigating the mental representation of language. *Journal of Psycholinguistic Research*, 24(6), 489-506.
- Her, O. S. (2012). Structure of classifiers and measure words: A lexical functional Account. *Language and Linguistics*, 13(6), 1211.
- Limbach, M. & Adone, D. (2010) Language acquisition of recursive possessives in English. *Proceedings of the 34th Annual Boston University Conference on Language Development (BUCLD34)*, ed. by Katie Franich, Kate M. Iserman and Lauren L. Keil, 281-290. Somerville, MA: Cascadilla Press.
- Luk, Z. P. S., & Shirai, Y. (2009). Is the acquisition order of grammatical morphemes impervious to L1 knowledge? Evidence from the acquisition of plural- s, articles, and possessive's. *Language Learning*, 59(4), 721-754.
- Mercan, G. (2016). Structural priming in L2 Turkish. *Second Language Acquisition of Turkish*, 59, 313.
- Nelson, Jon S. (2016) First and second language acquisition of recursive operations: Two studies. Doctoral dissertation, University of Massachusetts at Amherst.
- Pérez-Leroux, A. T. (2016). Turtles all the way down: How children harness the power of recursion, Plenary talk at the 2016 meeting of the Canadian Linguistic Association, Calgary, May 30, 2016.
- Roeper, T., & Snyder, W. (2005). Language learnability and the forms of recursion. *UG and external systems: Language, brain and computation*, 155-169.
- Rothman, J., & Pascual y Cabo, D. (2013). Generative approaches to Spanish second language acquisition. *The handbook of Spanish second language acquisition*, 46-63.
- Snape, N., Mayo, M. D. P. G., & Gürel, A. (2013). L1 transfer in article selection for generic reference by Spanish, Turkish and Japanese L2 learners. *International Journal of English Studies*, 13(1), 1-28.