The present study investigates the extent to which L1 versus adult L2 syntactic systems resist influence from a third language (L3). Based on the tenets of the Phonological Permeability Hypothesis (Cabrelli Amaro, 2013; Cabrelli Amaro & Rothman, 2010), we test the hypothesis that adult-acquired L2 syntactic systems are different from early-acquired systems with respect to stability. In this study, we investigate the effects of Brazilian Portuguese (BP) on Spanish in two groups of English/Spanish bilinguals that are advanced L3 BP speakers. The first group is made up of sequential L1 English/adult L2 Spanish speakers (n = 13), while the second group consists of L1 Spanish/adult L2 English speakers (n = 7). To test our hypothesis, we examine the phenomenon of raising across a dative experiencer (TPExp). Spanish and BP differ with regard to TPExp; while BP allows TPExp due to defective T, Spanish does not. Only learners who had acquired the relevant BP feature configuration (as measured by acceptance of TPExp in BP) were included in the study. Spanish data from an acceptability judgment task indicate that although both experimental groups rate TPExp higher than the Spanish control group (n = 18), L2 Spanish speakers are more accepting of TPExp than the early Spanish learners and the Spanish controls. We take these results to support our hypothesis of differential stability conditioned by age of acquisition, and address the role of regressive transfer as a correlate of L3 proficiency gains.

Key words: Third language (L3) Acquisition, Multilingualism, Morphosyntax, Regressive Transfer

1. Introduction

In the last decade, the investigation of transfer in third language (L3) acquisition has gained a lot of ground, particularly in the domain of morphosyntax. This upturn can be attributed to an interest in unique questions that are internal to L3 acquisition, as well as to the recognition that L3 acquisition can contribute to unresolved debates regarding the constitution of native versus non-native systems. Much of the research has focused on the mechanisms that drive transfer of existing linguistic systems at the initial stages of L3 acquisition (see García-Mayo & Rothman, 2012 for a review). Considerably less research concerns L3 development, a topic that informs a) L1 and L2 differences and b) how a L2 developmental path might differ from an L3/Ln developmental path. The minimal developmental research that is available largely examines
progressive transfer, that is, transfer of the L1 and/or L2 to the L3 (see e.g., Cabrelli Amaro, Amaro, & Rothman, 2015; Foote, 2009; Lozano, 2002). When it comes to regressive morphosyntactic transfer, that is, transfer of the L3 to the L1 and/or L2, there are two studies (Hui, 2010; Matthews, Cheung, & Tsang, 2014), both of which investigate whether L3→L2 transfer occurs. While these studies establish that an L3 can indeed influence an L2, there are no reports of how an L3 affects an L2 compared to an L1. As noted in Cabrelli Amaro and Rothman (2010), examination of regressive transfer in L3 acquisition can contribute to the perennial L1 vs. L2 debate by investigating potential influence on the L1 and/or L2 driven by L3 experience. That is to say, if an L2 is significantly more susceptible to L3 influence than an L1, it is possible that there is a critical difference between L1 and L2 morphosyntactic systems that has not been observable via L2 acquisition research alone. This line of investigation is the focus of the present study, in which we test the Differential Stability Hypothesis (DSH). The DSH is an extension of the Phonological Permeability Hypothesis (Cabrelli Amaro, 2013; Cabrelli Amaro & Rothman, 2010), and states that linguistic systems acquired after adolescence are more vulnerable to influence from an L3 system than those acquired in early childhood. We discuss the hypothesis in greater detail in section 2.

To test the DSH, we compare cross-linguistic effects of an L3 system on an L2 acquired in adulthood versus an L1. Specifically, we examine the influence of L3 Brazilian Portuguese (BP) on the Spanish morphosyntax of two groups of sequential bilinguals: L1 English/L2 Spanish and L1 Spanish/L2 English. We find that the Spanish system of the L2 learners reflects influence from L3 BP that is not evidenced in the L1 Spanish system to the same degree. Thus, our data support the hypothesis, and confirm that the study of L3 regressive transfer provides a novel (and needed) source for the testing of SLA theories.

The rest of this chapter is organized as follows. In section 2, we evaluate the literature on the modification of existing linguistic systems, and present the Differential Stability Hypothesis. In section 3, we present the morphosyntactic phenomenon we investigate to test the DSH, subject-to-subject raising across a dative experiencer (TPExp). While BP (and English) have a feature configuration of embedded T that allows this type of raising, Spanish does not. In section 4, we outline the predictions that the DSH makes for crosslinguistic effects on the Spanish system via influence of a system with a different feature configuration. In section 5, we present the methodology, including a description of the mirror-image participant groups, the acceptability judgment task, and the three conditions that we test. In section 6, we present results from a series of Repeated Measures ANOVA, and in section 7 we discuss our findings in
light of the DSH and provide future directions. Section 8 consists of a brief conclusion.

2. Regressive transfer

Regressive transfer is a type of crosslinguistic influence in which a more recently acquired language influences an existing linguistic system (here, the L1 and/or L2). Given the direction of regressive transfer, we associate this type of influence with attrition, and use these terms interchangeably herein. Iverson (2012) defines attrition as “the erosion of first language competence after exposure to another language” (p. 7), which we modify here to the erosion of *competence in an existing language* after exposure to another language. Employing this broad definition affords the encompassing of a variety of linguistic profiles, including the adult L2ers that are the focus of this study.

As mentioned in Section 1, regressive influence on L1 versus L2 systems is a phenomenon that has not been examined in L3 morphosyntax. However, there is an established body of research on the effects of a late-acquired L2 on an L1 syntactic system that informs the current study and aids in making predictions. We report on these findings in the next section.

2.1. L1 morphosyntactic attrition

Attrition has been shown to present itself in various forms, ranging from minor modifications of individual language systems to what has been termed as language death (e.g., Dorian, 1981) as a result of extended contact with another language across generations. Major (2010) divides the study of L1 attrition into macro and micro levels, with the macro level investigating languages that have been in contact for more than one generation. The micro level of attrition research consists of the investigation of a smaller number of speakers (as opposed to a community of speakers) that have acquired an L2 in adulthood. Given the scope of this study, we focus on research on attrition at a micro level, specifically on structural changes, and the predictions borne out of this research.

Studies examining L1 morphosyntactic attrition within a generative framework have distinguished between attrition of syntactic structures whose realization is strictly dependent upon linguistic factors (purely syntactic or via interaction with other linguistic domains), and attrition of structures moderated by contextual information. This distinction stems from the tenets of the Interface Hypothesis (Sorace & Filiaci, 2006; Sorace 2011), which predicts that syntactic phenomena conditioned by extralinguistic factors will be more prone to attrition due to increased processing loads and the depletion of attentional resources.
However, data from Tsimpli, Sorace, Haycock, and Filiaci (2004), Tsimpli (2007), and Iverson (2012) suggest that the feature configurations linked to syntactic phenomena that are not conditioned by extralinguistic factors (such as TPExp) are in fact vulnerable to modification.\(^1\)

In spite of the aforementioned evidence of attrition, existing data from L1 Spanish speakers lead us to predict that the L1 Spanish participants in the current study will not show evidence of L3 BP influence. First, recall from section 1 that while Spanish does not allow TPExp, both BP and English do. Following from the conclusions of Tsimpli (2007) and Iverson (2012), the relevant Spanish feature configuration of an L1 Spanish/L2 English speaker could be vulnerable to influence from the English feature configuration. However, Cabrelli Amaro et al. (2015) show that L1 Spanish speakers that are near-native speakers of L2 English and have been in the United States for a minimum of five years still reject TPExp categorically. If a learner’s acceptability of TPExp does not change after five or more years of exposure to a language that allows TPExp, it is reasonable to predict that comparably limited exposure to BP will not affect L1 Spanish. On the other hand, it is certainly possible that the addition of a novel system that allows TPExp (i.e., BP) could have a compounding effect that leads to modification of the Spanish configuration (see Hui, 2010; Matthews et al., 2014), but this remains to be seen. Regardless of the outcome, we predict that L1 Spanish syntactic systems will be less vulnerable to influence from L3 BP than near-native L2 Spanish syntactic systems. We outline the basis of this prediction in the following section.

2.2. The Differential Stability Hypothesis

While we predict that L1 Spanish speakers will not experience L3 influence on acceptance of TPExp, recent research shows that an L2 Spanish system is vulnerable to L3 influence even after as little as 16 weeks of exposure. Data from Cabrelli Amaro et al.’s (2015) study of L3 BP development examines the acceptability of TPExp by L1 English/L2 advanced Spanish/L3 advanced BP speakers in Spanish and BP. The results show that, although the learners overcome initial non-facilitative transfer of Spanish and acquire the L3 BP feature configuration, acquisition comes at a cost to the L2: The learners accept ungrammatical TPExp in Spanish significantly more than the Spanish controls.

\(^1\) In light of the controversy over the definition of internal and external interfaces (see Sorace, 2011 and responses, Sorace, 2012), the linguistic vs. extra-linguistic distinction we make here is purposefully vague. Most recently, Sorace (2012) has considered Hopp’s (2012) contention that the likelihood of feature erosion is not dependent upon the type of interface, but rather the computational complexity of a given structure. We direct the reader to the aforementioned sources for further discussion of this issue.
The question then becomes, if evidence suggests that an L1 syntactic system will not undergo influence, why would a native-like L2 system do so? There is independent evidence from two recent studies (Cabrelli Amaro et al., 2015; Campos-Dintrans, Rothman, & Pires, 2014) that late L2 Spanish learners can acquire the Spanish feature configuration of embedded T. This aligns with a large body of research that shows that L2 learners can acquire L1-like mental representations, including uninterpretable features that should only be acquirable via continued access to linguistic universals (see Rothman & Pascual y Cabo, 2013, for evidence specific to L2 Spanish acquisition). Thus, if it is the case that there is not a critical period (operationalized here as > 12 years of age) for the L2 development of a native-like grammar, what would drive L3 \( \rightarrow \)L2 crosslinguistic effects that are not predicted to be found between an L2 and an L1? This is the question we attempt to answer herein.

The concept of differential influence of an L3 on an L1 versus L2 was first proposed in Cabrelli Amaro and Rothman (2010). The Phonological Permeability Hypothesis (PPH) states that late-acquired phonological systems, even when seemingly native-like, are fundamentally different than early-acquired systems with regards to stability. That is, late-acquired systems are more permeable to influence, even when the learners have converged on a native-like L2 target. The PPH was developed to explore a new way of determining potential differences (or lack thereof) in L1 and L2 phonological systems. It assumes that adult acquisition of a native-like representation is possible, that is, there is no critical period for phonological acquisition. Instead, there is a critical period after which the stability of a phonological grammar is less than that of a native grammar. Cabrelli Amaro (submitted) presents data that support a weak version of the hypothesis. She compared Spanish perception and production of word-final vowels by L1 English/L2 Spanish and L1 Spanish/L2 English learners of BP, predicting that evidence of vowel reduction in Spanish would indicate BP influence. While there was no evidence of perceptual vulnerability in a forced-choice preference task, BP-like changes in vowel height were noted in the L2 Spanish speakers’ production that were not found in the L1 Spanish speakers’ production. These data led to a reformulation of the PPH, which states that the addition of a novel phonological system can affect aspects of speech production to a larger degree in late-acquired systems, at least when the languages under observation are typologically related. The author hypothesizes that cognitive control mechanisms such as failure to inhibit non-target forms in speech output could be responsible for the asymmetry in perception and production. However, Cabrelli Amaro et al.’s (2015) finding of L3 influence on L2 Spanish TPExp leads us to consider the original tenets of the PPH as they apply to the domain of syntax. We refer to this modification as the Differential Stability
Hypothesis (DSH): There is a maturationally conditioned fundamental difference in stability between linguistic systems that are acquired before versus after the critical period. We test the DSH here via investigation of L1 English/L2 Spanish and L1 Spanish/L2 English learners of L3 BP that have acquired the BP feature configuration associated with TPExp. By confirming acquisition of the L3 configuration, we can be certain that they have something available to influence the L1/L2 Spanish system. In the following section, we review the phenomenon under investigation (TPExp in Spanish and BP), outline the learning and modification tasks, and formalize our research questions and predictions.

3. Subject-to-subject raising across a dative experiencer in Spanish and Brazilian Portuguese

To test the DSH, we examine subject-to-subject raising across a dative experiencer (TPExp). This property is ideal to test our hypothesis because TPExp is permitted in BP, but crucially not in Spanish, due to the feature specification of embedded T in each language. In Spanish, the dative clitic of the verb parece ‘to seem’ cannot co-occur with raising across a dative experiencer, as in (1). Conversely, this co-occurrence is acceptable in BP (2). This contrast reflects different featural specifications of embedded T (Ausín & Depiante, 2000, for Spanish; A. Pires, personal communication, 2014, for BP): In BP, embedded T is [-finite] and cannot assign nominative Case to the experiencer. Therefore, the experiencer can overtly raise, and case will be checked in the matrix clause. However, embedded T in Spanish is [+finite], and the Case feature is assigned to the embedded subject. This Case assignment blocks the subject from raising across the experiencer to the matrix clause.

(1) *Lorenzo me parece [T[non-def] t; adorar a Irene]
   Lorenzo me.exp seems to adore to Irene
   ‘Lorenzo seems to me to adore Irene.’

(2) Lourenço me parece [T[def] t; adorar (a) Irene]
   Lourenço me.exp seems to adore (the) Irene
   ‘Lourenço seems to me to adore Irene.’

As in BP, embedded T is [-finite] in English (3).

(3) Lawrence seems to me [T[def] t; to adore Irene]

4. Research questions and predictions
Given what we have presented about TPExp in Section 3, we can assume that English/Spanish bilinguals with native-like grammars in both languages have the English feature configuration available for (facilitative) transfer to L3 BP. In fact, this was confirmed for L1 English/L2 Spanish and L1 Spanish/L2 English learners at the initial stages of L3 BP acquisition in Cabrelli Amaro et al. (2015). In spite of access to this facilitative resource, however, both learner groups transferred the non-facilitative [+finite] Spanish configuration to L3 BP. This result aligns with the pattern found in the existing research on English/Spanish bilinguals acquiring L3 BP. That is, these bilinguals transfer Spanish at the L3 BP initial stages regardless of whether a) Spanish is the L1 or L2 and b) transfer is facilitative (e.g., Giancaspro, Halloran, & Iverson, 2015; Montrul, Dias, & Santos, 2011). These studies substantiate the claim that initial stages transfer is driven by structural similarity and that a single linguistic system transfers to the L3, formalized in the Typological Primacy Model (TPM, e.g. Rothman, 2015). Given the evidence that Spanish transfers to L3 BP, we test only Spanish and BP in this study, and not English.

Once Cabrelli Amaro et al. (2015) established that both groups of English/Spanish bilinguals had transferred Spanish to L3 BP, they compared the L1 English/L2 Spanish initial stages learners with L1 English/L2 Spanish advanced L3 BP learners. The authors determined that although the learners had converged on the L3 BP target, doing so appeared to come at a cost to the learners’ Spanish. That is, the advanced L3 group assigned a significantly higher acceptability rating to TPExp in Spanish than both the Spanish control group and initial stages L1 English/L2 Spanish/L3 BP group. This finding has led us to question how this particular learner group compares with its mirror image, i.e., a group of L1 Spanish/L2 English learners. Thus, our main research question, which has two parts, is as follows: First, does the reconfiguration of embedded T in BP influence the Spanish system? In other words, do L1 and L2 Spanish speakers that accept TPExp in BP, accept TPExp in Spanish more than a native Spanish control group? Second, does the degree of influence depend upon the age of acquisition of Spanish? That is, do L2 Spanish speakers assign a significantly higher acceptability rating to TPExp in Spanish than L1 Spanish speakers? Since Cabrelli Amaro et al.’s data do not indicate any influence of L1 English TPExp on L1 Spanish/L2 English speakers that had been an English immersion environment for at least five years, we predict that L3 BP TPExp will not influence Spanish TPExp ratings. As such, we expect that the degree of BP influence on the two groups’ acceptance of TPExp in Spanish will not be equal. Rather, the L2 Spanish group will accept TPExp at a higher rate than the L1 Spanish group will.
5. Methodology

5.1 Participants

The experimental data we present here come from two groups that differ in their age of acquisition of Spanish. The first group is composed of L1 English/L2 Spanish speakers (n = 13) who were raised in the US and acquired Spanish after the age of 12. The second group is composed of L1 Spanish/L2 English speakers (n = 7) who were raised in a monolingual Spanish environment and acquired English after the age of 12. All participants are undergraduate or graduate students at US post-secondary institutions. Potential participants with fluency in additional languages (as assessed via a language background questionnaire) were excluded from the study. While we started with 48 participants between the two groups, a two-part screening process reduced our numbers to 20 participants (this is unfortunately a common hazard of L3 research). The first part of the screening consisted of written proficiency assessments in Spanish (for the L2 Spanish group) and L3 BP to determine minimally advanced proficiency in both late-acquired languages under investigation.2 For the Spanish assessment, participants had to score 40 out of 50 on a 50-item Spanish proficiency cloze test used extensively in generative L2 research (e.g., Montrul, 2011; Montrul et al., 2008). For the BP assessment, participants had to score 80 out of 100 on a written test implemented by the Brazil-US Cultural Association (ACBEU), previously used in Rothman and Iverson (2009, 2011). Of the 48 participants that completed these assessments, 16 did not minimum proficiency requirements either in L2 Spanish (n = 7) or in BP (L1 Spanish n = 8, L2 Spanish n = 1). The second part of the screening process confirmed target-like acceptance of TPExp in L3 BP. We administered an acceptability judgment task to the remaining participant pool in BP and Spanish, which we describe in detail in section 5. To be included in the analysis of TPExp in Spanish, learners’ mean acceptability score for TPExp had to fall within or above the confidence interval of the BP control (see next paragraph for details). Of the 32 learners, seven of 17 L1 Spanish speakers and 13 of 15 L2 Spanish speakers had a mean BP TPExp rating within the BP control confidence interval.

As mentioned, the two experimental groups were compared with a BP control group to confirm acceptability of TPExp in BP. BP control participants (n = 34) had been raised in a monolingual environment in

---

2 We did not administer an English proficiency test to the L1 Spanish learners. All of the participants had lived in the US a minimum of two years and were enrolled in US universities that require an English proficiency exam as part of their application process.
Brazil and were self-reported advanced speakers of English. All participants were university students whose age of onset of L2 acquisition was at least 12 years. The learners whose mean TPExp rating fell within the BP control confidence interval were then compared with a Spanish control group of L1 Spanish/L2 English speakers (n = 21; 18 were from Cabrelli Amaro et al., 2015; data from the remaining three controls were collected separately). These speakers did not have fluency in any additional languages and had not been exposed to BP.

5.2. Acceptability judgment task

The experimental learners completed an 80-item acceptability judgment task in Spanish and BP, and the order in which each language was tested was counterbalanced across participants. We tested three conditions, with eight items for each condition (four with proper names, four with other subject DPs). These items were first used in Campos-Dintrans et al. (2014). The remaining items consisted of 44 items used for a separate study on differential object marking, and 12 fillers (6 grammatical, 6 ungrammatical). None of the other items included raising verbs, experiencers, or infinitival complements. Sentences were judged on a 4-point Likert scale, whereby 1 was odd/strange and 4 was completely natural. Participants were instructed to leave the question blank if they were not sure.

Items in Condition 1 (vPnoExp) consist of the raising verb parecer ‘to seem’, and lack an overt experiencer. In these test items, parecer is a modal verb and does not project an embedded TP. This structure is grammatical in BP (4a) and Spanish (4b) and we thus predict that all participants will accept them.

(4) a. Jorge parece [vP precisar de ajuda]
   b. Jorge parece [vP necesitar ayuda]
   ‘Jorge seems to need help.’

We include this condition to ascertain the role of an overt experiencer in the rejection of Condition 2, which we explain in turn.

The items in Condition 2 (TPExp) are composed of the verb parecer ‘to seem’ with an embedded TP complement, projecting an experiencer. The subject of embedded TP raises to the matrix Spec, TP. Since BP projects an embedded defective TP and Spanish does not, TPExp structures are grammatical in BP (5a) and not in Spanish (5b).

(5) a. A mulher me parece [TP saber tudo]
   b. *La mujer me parece [TP saber todo]
   ‘The women seems to me to know everything.’

As expected based on the analysis of TPExp in Section 3, L1 and L2 Spanish speakers clearly reject this condition in Spanish (Cabrelli Amaro et al., 2015). Therefore, if the advanced L3 BP speakers rate this structure
in Spanish significantly higher than the advanced L2 Spanish speakers do, we can make the case for L3 BP influence on Spanish.

Similarly to Condition 1, items in Condition 3 (APExp) do not project an embedded TP, but, like Condition 2, they involve subject-to-subject raising. However, the difference between Conditions 2 and 3 lies in the origin of the raised subject: The subject in Condition 2 raises from the TP domain, while the subject in Condition 3 raises from the AP domain. Unlike with an embedded TP, Spanish does not block raising from an AP. Therefore, this Condition is grammatical in both BP (6a) and Spanish (6b).

(6)  
a. O professor me parece \[AP inteligente\]  
b. El profesor me parece \[TP inteligente\]  
‘The professor seems intelligent to me.’

6. Results

We begin this section with a report of the baseline data from the Spanish and BP controls. After establishing the difference in acceptance of TPExp in each language and the CI for BP TPExp ratings, we present a comparison of the data from the two experimental groups and the Spanish control. In addition to providing descriptive statistics (\(M, SD, CI\), and effect size as measured by Hedges’ \(g\)), we evaluate the statistical significance of intergroup and intragroup comparisons via repeated measures (RM) ANOVA and Bonferroni post hoc analyses. Before presenting our results, it should be noted that some of our control data (first reported in Cabrelli Amaro et al., 2015) was collected using a 5-point scale. To compare findings between the 4-point and 5-point scale data sets, we used a proportional transformation to convert 4-point scale data to a 5-point scale. It has been suggested that a proportional transformation to compare data from different Likert scales is not an empirically derived mathematical solution (Colman, Morris, & Preston, 1997). However, the proposed alternative of z-score transformation is incompatible with a RM ANOVA since this transformation would yield group means of zero and result in p-values of 1.00.

6.1. Control data in Spanish and BP

Acceptability judgment data from the BP (n = 34) and Spanish (n = 18) controls are presented in Table 1 and illustrated in Figure 1.

<table>
<thead>
<tr>
<th>Table 1. Mean acceptability ratings from Spanish and BP controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>vPnoExp</td>
</tr>
</tbody>
</table>

Similar to Condition 1, items in Condition 3 (APExp) do not project an embedded TP, but, like Condition 2, they involve subject-to-subject raising. However, the difference between Conditions 2 and 3 lies in the origin of the raised subject: The subject in Condition 2 raises from the TP domain, while the subject in Condition 3 raises from the AP domain. Unlike with an embedded TP, Spanish does not block raising from an AP. Therefore, this Condition is grammatical in both BP (6a) and Spanish (6b).

(6)  
a. O professor me parece \[AP inteligente\]  
b. El profesor me parece \[TP inteligente\]  
‘The professor seems intelligent to me.’
Regression transfer in L3 syntactic development

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
<th>CI</th>
<th>M</th>
<th>SD</th>
<th>CI</th>
<th>M</th>
<th>SD</th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>BP</td>
<td>4.10</td>
<td>.57</td>
<td>3.90-4.30</td>
<td>3.59</td>
<td>1.05</td>
<td>3.22-3.96</td>
<td>4.39</td>
<td>.64</td>
<td>4.16-4.62</td>
</tr>
<tr>
<td>Spanish</td>
<td>4.29</td>
<td>.66</td>
<td>3.96-4.61</td>
<td>1.64</td>
<td>.53</td>
<td>1.37-1.90</td>
<td>4.6</td>
<td>.34</td>
<td>4.43-4.76</td>
</tr>
</tbody>
</table>

Figure 1. Mean acceptability ratings for BP and Spanish controls

The results of a two-way RM ANOVA with the between-subjects factor “Language” and within-subjects factor “Condition” showed a significant Language*Condition interaction ($F(1.589, 77.838) = 47.352, p < 0.001)$. The assumption of sphericity was violated for the within-subjects effect of Condition ($X^2(2) = 14.390, p < .001$); therefore, degrees of freedom were corrected using Greenhouse-Geisser estimates ($\epsilon = .794$). While neither group accepts TPExp at the same rate as APExp and vPnoExp (all comparisons $p < .05$), the BP group’s acceptability rating of TPExp is significantly higher than the Spanish group’s rating ($p < .001, g = 2.00$). Moreover, the effect size (as measured by bias-corrected Hedges’ $g$) is categorized as large by Oswald and Plonsky (2014). This finding supports the analysis that TPExp is acceptable in BP but not in Spanish.

6.2. Spanish control and learner comparison
Establishing baseline ratings for TPExp in BP allowed us to determine which L3 learners’ ratings of TPExp fell within or above the BP CI (3.22-3.96). We then took this reduced group of learners (n = 13 L1 English/L2 Spanish, n = 7 L1 Spanish/L2 English), and compared their Spanish ratings with each other and with the Spanish control. Means, SD, and CIs for these three groups are in Table 2.

Table 2. Mean acceptability ratings in Spanish from the L1 and L2 Spanish groups.

<table>
<thead>
<tr>
<th></th>
<th>vPnoExp</th>
<th>TPExp</th>
<th>APExp</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>CI</td>
</tr>
<tr>
<td>L1 Spanish</td>
<td>4.31</td>
<td>.59</td>
<td>3.86-4.77</td>
</tr>
<tr>
<td>L2 Spanish</td>
<td>3.74</td>
<td>.81</td>
<td>3.26-4.23</td>
</tr>
</tbody>
</table>

Figure 2. Mean acceptability ratings in Spanish from the experimental and control groups.

Another two-way RM ANOVA was constructed, this time with “Group” as the between-subjects factor. The assumption of sphericity was not violated ($\chi^2 (2) = .060$, $p = .970$). The analysis yielded a significant Group*Condition interaction ($F(4, 74) = 13.678$, $p < .001$). Post hoc
comparisons reveal that the only significant pairwise comparisons are from the TPExp condition. The L2 Spanish group rates TPExp significantly higher than the control group does, with a very large effect size ($p < .001, g = 1.99$). Although the difference between the L1 Spanish group and the control group only approaches significance ($p = .064, g = 1.01$), the L1 Spanish mean does not fall within the control’s CI and the effect size of the between-group difference is large. That said, when we compare the L1 Spanish-control and L2 Spanish-control differences, the L1 Spanish-control effect size is approximately one standard deviation smaller than that of the L2 Spanish-control effect size. In a direct comparison of the TPExp condition between the experimental groups, we find that while the groups’ ratings do not differ significantly, the difference yields a medium effect size ($p = .106, g = -.71$). Thus, while both groups rate TPExp higher than the Spanish control, the degree of difference between the L2 Spanish group and the control group is higher.\(^3\) With that said, the large SDs in the TPExp condition indicate considerable individual variation, which we discuss in section 7.3.

7. Discussion

In this section, we begin with a discussion of the results as they relate to the Differential Stability Hypothesis. We then continue with two issues that our analysis yields: The L2 Spanish learners’ representation at the onset of L3 acquisition, and individual variation. We conclude the section with a brief overview of the next steps we plan to take in this line of inquiry.

7.1. The DSH

The DSH states that there is a fundamental difference in stability between early- and late-acquired linguistic systems. While the hypothesis assumes that late learners have continued access to linguistic universals, it posits that their systems are more vulnerable to L3 influence than an early-acquired system. L1 Spanish speakers categorically are shown to reject

\(^3\) We recognize that the experimental participants’ acceptance of all three experimental conditions could potentially indicate acquiescence bias. However, within-subjects comparisons among the experimental conditions (Bonferroni post hoc analyses) and between the grammatical and ungrammatical fillers (t-tests) suggest that this is not the case. With respect to the experimental conditions, the L1 Spanish group significantly differs between TPExp and vPnoExp ($p < .001$), and TPExp and APExp ($p < .001$). The L2 Spanish group does not differentiate between TPExp and vPnoExp ($p = .299$), but accepts APExp at a higher rate than TPExp ($p = .018$). Comparing the grammatical and ungrammatical fillers, we find that a significant difference with large effect sizes for the L1 Spanish group ($\eta(8.639), p < .001, g=3.71$) and the L2 Spanish group ($\eta(2.396), p = .034, g=1.78$).
TPExp in Spanish even after five years of L2 English immersion (Cabrelli Amaro et al., 2015), and English and BP share a featural configuration for embedded T which permits TPExp. Since these data showed that the English configuration had not yet influenced the speakers’ L1 Spanish, we predicted that similar early-acquired Spanish systems would be invulnerable to influence from the BP feature configuration. Conversely, we predicted that the late-acquired Spanish system of the L1 English/L2 Spanish group would be (more) vulnerable to L3 BP influence. However, our comparison of Spanish TPExp ratings by L1 and L2 Spanish speakers points to BP influence on both groups (albeit to different degrees, see discussion below). Evidently, it would seem that the acquisition of L3 BP has come at a price. While we observe changes to the acceptability of TPExp, it is important to note that we do not assume that all is lost when it comes to the learners’ Spanish representation. The L1 Spanish group still rates TPExp lower in Spanish than in BP \((t(-3.648), p = .011, g = 2.00)\), as does the L2 Spanish group, albeit expectedly to a lesser degree \((t(-3.151, p = .008, g = -.96)\).

It is possible that the L3 influence on Spanish, which was unexpected in the L1 Spanish group for this particular property, is the result of a cumulative effect. As mentioned in section 2, we have seen cases of cumulative influence from Hui (2010) and Matthews, Cheung, and Tsang (2014). In both studies, L1 Cantonese/L2 English speakers were compared with L1 Cantonese/L2 English/L3 French (Hui, 2010) or L3 German learners (Matthews et al., 2014). Properties were tested that were similar in the L2 and L3, but different in the L1 (specifically, English and French subject relative clauses and English and German morphological tense marking, respectively). While the bilinguals did not have target-like knowledge of these properties in L2 English, the trilinguals did. This knowledge was attributed to the cumulative effect of the L2 in conjunction with the L3; it would seem that the L3 input was the extra push the L2 needed.

Although the data point to a difference between the Spanish control and both of the experimental groups’ acceptance of TPExp, an analysis of effect size indicates that the L2 Spanish group accepted TPExp at a higher rate than the L1 Spanish group. The group difference is further illustrated via the individual data in Tables 3 and 4. While 11 of 13 (85%) L2 Spanish participants’ TPExp ratings fall outside of the upper limit of the Spanish control’s CI, four of seven (57%) L1 Spanish learners’ ratings do. Again, 57% was an unexpected finding for the L1 Spanish group, but one that we believe reflects the dynamic nature of these multilingual systems. We take the inter-group difference as a possible indication that the L2 Spanish system demonstrates a higher degree of instability than the L1 Spanish system, which provides tentative support for the DSH.
Recall that asymmetry in the perception and production data in Cabrelli Amaro (submitted) led to the formulation of a weak version of the PPH. She suggests that perception (which she equates with competence) is more resistant to influence than production based on data that show that L1 and L2 Spanish speakers did not differ significantly from the Spanish control on a perception task. Conversely, a difference between L1 Spanish and L2 Spanish production led to the conclusion that L2 Spanish production is more vulnerable to L3 influence, and it was posited that this difference is due to issues in cognitive control during speech production. Contrary to the results from the phonological data, the apparent L3 influence on the acceptance of TPExp by L2 Spanish speakers in Cabrelli Amaro et al. (2015) led us to take a second look at L3 effects on the morphosyntactic competence of L1 and L2 Spanish speakers. Our comparison of the L2 findings with L1 Spanish TPExp yields a result that is inconsistent with the phonological data. That is, our data support the notion that L2 morphosyntactic competence as measured by acceptance or rejection of TPExp is indeed more vulnerable to crosslinguistic influence than L1 TPExp. While we recognize that we cannot make a one-to-one comparison between morphosyntax and phonology, we believe that this finding could potentially hold for phonology as well. Specifically, we contend that the null result from Cabrelli Amaro’s perception data could stem from the nature of the task used. The author used a forced-choice preference task, in which the listener heard a BP-like stimulus and a Spanish-like stimulus and selected which stimulus sounded more natural in the language being tested. We would argue that such a task is easier than a task in which the learner is presented with a single stimulus and is required to judge its naturalness without a point of comparison. It is possible that task demands similar to those of the AJT used here could yield phonological data that is comparable with our results.

7.2. L2 Spanish acquisition

The notion that the L2 Spanish speakers are more vulnerable to L3 influence brings up the issue of the L2 Spanish representation at the time of testing. When the L2 Spanish data were presented in Cabrelli Amaro et al. (2015) without L1 Spanish data as a point of comparison, the authors questioned whether the L2 Spanish learners had ever acquired the Spanish feature configuration. It was suggested that perhaps the advanced learners’ target-like BP acceptance of TPExp was a reflection of English transfer. That is, the learners had never acquired the Spanish configuration, and had transferred the only feature configuration that they had available (i.e., the English configuration). For the purpose of the present study, we could follow this line of thought, and propose that the acceptance of TPExp in Spanish is not due to L3 BP influence; rather, it is a reflection of the fact
that the learners never acquired the configuration in Spanish. However, as Cabrelli Amaro et al. (2015) note, previous research supports the acquisition of the feature configuration by late L2 learners with similar learner profiles and proficiency criteria (Dintrans et al., 2014; the L1 English/L2 Spanish/L3 BP initial stages data from Cabrelli Amaro et al., 2015). Finally, although the majority of the L2 Spanish participants’ TPExp ratings fall above the upper limit of the Spanish control CI, there are two learners that maintain a Spanish-like rating despite native-like TPExp ratings in BP. Taken together, then, we follow the assumption that the L2 Spanish learners had acquired the feature configuration of Spanish embedded T before acquiring L3 BP.

7.3. Individual variation

Another aspect of the data that deserves attention is the individual variation in both experimental groups (Tables 3 and 4) and potential effects of variables other than AoA. As shown in Table 2, there is a large SD in both groups. The L1 Spanish participants’ mean ratings tend to pool at either end of the scale, while the L2 Spanish ratings are more variable. Within each group, we question why some of the learners’ judgments are vulnerable to L3 BP influence while others are not (as indicated by shaded cells), and why there is so much variation among the L2 Spanish learners that do not pattern with the control.

In reviewing the data from the language background questionnaire, we did not find any patterns of language use or L2 Spanish immersion that indicate a strong relationship between frequency and context of Spanish use and acceptance of TPExp. There is one L2 Spanish learner (1052) who reports using Spanish only a few times a year, and displays nearly categorical acceptance of TPExp in both languages. However, all of the other learners in both groups reportedly used Spanish on at least a weekly basis at the time of testing, and thus the learners’ Spanish Language Maintenance Effort (LME, Herdina & Jessner, 2002) is not varied enough to explicitly address its effects. Moreover, it has been found that continued language use does not necessarily prevent crosslinguistic influence (e.g., Mehotcheva, 2010; Xu, 2010, but cf. Gurel, 2007). There is also a lack of a clear pattern when connecting BP use to Spanish TPExp acceptance. While it is true that the majority of the learners that rate Spanish TPExp higher than the controls use BP at least several times a week, there are two instances in each group in which the learners acceptance TPExp in Spanish in spite of a low rate of BP use.

With respect to environment, it is important to note that seven of the 11 L2 Spanish learners were tested in a BP immersion setting, and six of these seven learners rate TPExp higher than the control. Conversely, all of the L1 Spanish learners were tested in the US. We recognize that this
difference between groups could be a confounding variable, although the data clearly demonstrate that Spanish TPExp acceptance occurs in both contexts for both groups.

Finally, we considered the possibility of a priming effect for the learners that completed the task in BP first (recall that the order was counterbalanced across participants). However, in the case of the L1 Spanish group, two of the four participants that fall above the Spanish CI completed the Spanish task first, and two of the three that fall within the Spanish CI completed the BP task first. We find a comparable pattern for the L2 Spanish group: Four of the 11 that fall above the Spanish CI completed the Spanish task first, and one of the two that fall within the Spanish CI completed the Spanish task first.

Table 3. L2 Spanish group’s TPExp ratings and language use (in days per week) in Spanish and BP.

<table>
<thead>
<tr>
<th>Participant</th>
<th>L3 BP</th>
<th>L2 Spanish</th>
<th>Spanish use per week</th>
<th>BP use per week</th>
<th>Spanish immersion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1069</td>
<td>3.67</td>
<td>1.83</td>
<td>3+</td>
<td>immersion 1 semester</td>
<td></td>
</tr>
<tr>
<td>1083</td>
<td>3.67</td>
<td>3.84</td>
<td>1+</td>
<td>immersion none</td>
<td></td>
</tr>
<tr>
<td>1061</td>
<td>3.83</td>
<td>2.71</td>
<td>7</td>
<td>immersion 1.5 years</td>
<td></td>
</tr>
<tr>
<td>1018</td>
<td>4.05</td>
<td>3.83</td>
<td>7</td>
<td>1 month</td>
<td></td>
</tr>
<tr>
<td>1055</td>
<td>4.50</td>
<td>3.84</td>
<td>3+</td>
<td>immersion 1 semester</td>
<td></td>
</tr>
<tr>
<td>1054</td>
<td>4.50</td>
<td>4.33</td>
<td>7</td>
<td>immersion 2 years</td>
<td></td>
</tr>
<tr>
<td>1068</td>
<td>4.50</td>
<td>4.33</td>
<td>3+</td>
<td>immersion ---</td>
<td></td>
</tr>
<tr>
<td>1037</td>
<td>4.67</td>
<td>4.33</td>
<td>7</td>
<td>6 months</td>
<td></td>
</tr>
<tr>
<td>1011</td>
<td>4.67</td>
<td>2.83</td>
<td>7</td>
<td>1 semester</td>
<td></td>
</tr>
<tr>
<td>1052</td>
<td>4.84</td>
<td>4.83</td>
<td>&lt;1</td>
<td>1+</td>
<td></td>
</tr>
<tr>
<td>1035</td>
<td>5.00</td>
<td>1.50</td>
<td>3+</td>
<td>1+ 1 semester</td>
<td></td>
</tr>
<tr>
<td>1066</td>
<td>5.00</td>
<td>4.33</td>
<td>7</td>
<td>immersion 1 year</td>
<td></td>
</tr>
</tbody>
</table>

*Means within the Spanish confidence interval are shaded.

Table 4. L1 Spanish group’s TPExp ratings and language use (in days per week) in Spanish and BP.

<table>
<thead>
<tr>
<th>Participant</th>
<th>L3 BP</th>
<th>L1 Spanish</th>
<th>Spanish use per week</th>
<th>BP use per week</th>
</tr>
</thead>
<tbody>
<tr>
<td>1103</td>
<td>3.67</td>
<td>3.67</td>
<td>7</td>
<td>1+</td>
</tr>
<tr>
<td>1114</td>
<td>3.83</td>
<td>1.83</td>
<td>7</td>
<td>3+</td>
</tr>
<tr>
<td>1108</td>
<td>3.83</td>
<td>3.67</td>
<td>7</td>
<td>3+</td>
</tr>
<tr>
<td>1115</td>
<td>4.00</td>
<td>1.67</td>
<td>7</td>
<td>3+</td>
</tr>
<tr>
<td>1038</td>
<td>4.17</td>
<td>1.33</td>
<td>7</td>
<td>&lt;1</td>
</tr>
<tr>
<td>1117</td>
<td>4.17</td>
<td>3.00</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>1109</td>
<td>4.50</td>
<td>2.33</td>
<td>7</td>
<td>1+</td>
</tr>
</tbody>
</table>

*Means within the Spanish confidence interval are shaded.
7.4. Future directions

The questions that our data have generated can be addressed via longitudinal investigation that controls for language use and environment. Following learners from the initial stages of L3 acquisition provides a baseline measurement for each learner and allows us to confirm L2 acquisition of the properties under investigation (e.g., Schmid & Mehotcheva, 2012). A longitudinal investigation of morphosyntactic and phonological development that follows learners over the course of a year of classroom-based L3 BP instruction is currently underway. In this way, we can control the context of BP acquisition context and language use. Our goal in carrying out this longitudinal study is to corroborate (or not) existing evidence of differential phonological and morphosyntactic stability in L1 versus L2 Spanish learners of L3 BP. We investigate English in addition to BP and Spanish, and examine properties that present differently in each of the three languages. In doing so, we can determine the role of structural similarity in regressive transfer. Having found differential influence on L1 and L2 systems between similar languages, we question whether this phenomenon is more likely to occur between these related languages (as suggested in Schmid & Mehotcheva, 2012, p. 16) than in a language pairing such as English and BP, due to the increased difficulty in inhibition of similar languages that can eventually lead to erosion.

8. Conclusion

The goal of this study has been to assess potential differences between L1 and L2 morphosyntactic systems via effects of L3 influence. We tested the Differential Stability Hypothesis (DSH), which posits that L2 systems differ from L1 systems in terms of stability and predicts that systems acquired after adolescence will be more vulnerable to influence from a subsequently acquired language. Our investigation of English/Spanish bilinguals acquiring L3 Brazilian Portuguese yield two findings of note. First, our data show that L1 and L2 Spanish speakers’ Spanish ratings of raising across a dative experiencer (grammatical in BP only) reflect BP influence, and thus it is possible that regressive influence is a correlate of proficiency gains. Second, and more importantly, the L2 group’s ratings are less Spanish-like than the L1 group’s ratings. We take these findings as preliminary support for the DSH, and recognize a number of confounding factors that can be dealt with via the implementation of longitudinal methodologies.

References


Sorace, A., & Serratrice, L. (2009). Internal and external interfaces in bilingual

