

# *A systematic review of transfer studies in third language acquisition*

Article

Accepted Version

Puig-Mayenco, E., González Alonso, J. and Rothman, J. (2020) A systematic review of transfer studies in third language acquisition. *Second Language Research*, 36 (1). pp. 31-64. ISSN 1477-0326 doi: 10.1177/0267658318809147 Available at <https://centaur.reading.ac.uk/79366/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1177/0267658318809147>

Publisher: Sage Publications

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

**CentAUR**

Central Archive at the University of Reading

Reading's research outputs online



## A Systematic Review of Transfer Studies in Third Language Acquisition

Eloi Puig-Mayenco<sup>1</sup>, Jorge González Alonso<sup>2</sup> & Jason Rothman<sup>1,2</sup>

University of Reading<sup>1</sup> & UiT, the Arctic University of Norway<sup>1</sup>

### Abstract

The present systematic review examines what factors determine when, how and to what extent previous linguistic experience (from the L1, L2 or both languages) affects the initial stages and beyond of adult L3 acquisition. In doing so, we address what a bird's eye view of the data tells us regarding competing theoretical accounts of L3 morphosyntactic transfer. Data couple together to suggest that some factors are more influential than others. As discussed, the systematic review transcends the field of adult multilingualism precisely because of what it reveals, as a *prima facie* example in behavioral research, in terms of how different types of methodological considerations impact the way data are interpreted to support or not particular claims.

**Keywords:** *Systematic review, L3/Ln acquisition, Transfer selection, Methodological considerations, Typological Proximity.*

## 1. Introduction

The study of non-native (i.e., non-primary) language acquisition and processing has long been concerned with the interplay between ‘old’ and ‘new’ linguistic knowledge (an issue already discussed in Weinreich, 1953), both in vocabulary and grammar (e.g., Jarvis, 2000; Odlin, 1989; Schwartz and Sprouse, 1996; White, 2003). Non-native language learners often speak more than one language at the onset of acquiring a new one—e.g., immigrants that arrive in Europe or the USA from India or Malaysia are likely to speak several previous languages. Accumulating evidence seems to indicate that third or more language (L3/L $n$ ) acquisition presents differently from second language acquisition (L2) (see De Angelis, 2007; Falk and Bardel, 2010; González Alonso, Rothman, Berndt, Castro, and Westergaard, 2017). While in second language acquisition, the learner can only rely on her experience with one language, in L3/L $n$  acquisition more than one system of linguistic representation is available.

With these observations in mind, it is not surprising that a substantial amount of research in L3/L $n$  acquisition has focused on determining which of the previous languages, if any, exerts a larger amount of influence on the initial representations in L3/L $n$  interlanguage grammars and thus affects the L3/L $n$  learning process. Theoretical proposals attempting to model the role of linguistic transfer in L3/L $n$  acquisition invariably contain two underlying assumptions, namely, (i) that one or more variables determine when and how transfer will take place (i.e., it is not random), and (ii) that this combination of variables is indeed weighted, such that all things being equal one variable will take precedence over the others. Thus, the models we will discuss here differ along two main dimensions. The first is what variable(s) they advocate as being ultimately explanatory for linguistic transfer in L3/L $n$  acquisition. The second is whether the model is limited in scope to one developmental stage in particular—e.g., initial, intermediate, advanced stages—or if it is meant to account for linguistic transfer at any and all points in the developmental sequence.

This paper offers a systematic review of a sizeable subset of L3 studies, focusing on morphosyntactic transfer. It is important to clarify from the outset, however, that this is not a meta-analysis in the traditional sense, for reasons that pertain to the nature of these studies and, to some extent, to our specific motivations in undertaking this task. A meta-analysis uses calculations based on individual studies' effect sizes—or some other measure of strength—to derive conclusions about the effects of a particular treatment on a specific population, targeted by all included studies (see Boulton and Cobb, 2017; Norris and Ortega, 2000; Plonsky and Oswald, 2012). Unfortunately, a majority of the studies reviewed here do not meet the requirements to conduct a meta-analysis of the type just described: effect sizes are not reported, and they often cannot be directly or indirectly estimated from the information reported in the studies (only 60.9% of the entire pool of studies provides enough information to calculate effect sizes based on, e.g. Boulton and Cobb, 2017; Plonsky and Kim, 2016; Plonsky and Oswald, 2012).

Given our main point anyway—to understand what methodological choices might confer for interpreting data in light of specific models—a systematic review is a more appropriate choice. Collective data weigh in best on debates among competing theories when they come from methodologies that fairly represent as many available theories as possible. As a whole, the group of studies we analyze in this paper have deficiencies in two related departments: they often lack the necessary detail in their description and/or reporting to replicate or re-analyze the data, and they sometimes ignore field-specific methodological considerations which directly affect their interpretation (in light of *all* available theories).

To be clear then from the start, we will employ contingencies precisely because the goal is to reveal if there are associations between method/practice and outcome. This review thus provides a bird's eye view of the field, in an attempt to evaluate how much of what we have ascribed to linguistic variables can also be explained by potential inadvertent methodological

choices. Our systematic review comprises 71 studies, where we examine methodological practices. Furthermore, since linguistic transfer—its source, its extent, its timing—feeds into the very definition of individual L3/*L<sub>n</sub>* learnability tasks and can also, especially and uniquely in the case of multilingualism, reveal insights into how the mind economizes more generally, a review of this type is non-trivial on several planes.

## 2. Setting the stage

Studying the role of transfer in the acquisition of a third or further language can contribute to our understanding of cognitive economy in ways that studying L1 or L2 acquisition cannot. This is not to say that L3/*L<sub>n</sub>* acquisition is fundamentally different, as a whole, from L1 or L2 acquisition (see Rothman, 2013, 2015). However, the fact that an L3 learner has varying amounts of previous experience with more than one language makes transfer a multidimensional factor: now the learner's brain has *choice*—however unconscious such is likely to be—for many if not most domains of grammar. Because languages (may) have different and often incompatible representations for the same structure or grammatical function, the selection of L1 over L2 representations (or vice versa) for transfer into the L3 is not a trivial issue. This is so because it might have differentially facilitative results depending on what the target L3/*L<sub>n</sub>* grammar specifies for each linguistic property—as it might resemble the L1, the L2, or neither. Crucially, however, since there is no way to know *a priori* what the most facilitative choice might be in each case, the brain is forced to make an unconscious 'best guess' as to what will most efficiently assist the creation of a linguistic representation that is able to parse the L3/*L<sub>n</sub>* input. The question thus becomes the following: what guides this informed guess? Different theories and models of morphosyntactic transfer in L3/*L<sub>n</sub>* acquisition have addressed this question by considering a substantial number of variables: type of linguistic experience, age of acquisition, similarity between the languages (overall or at the level of specific properties), among others. No model explicitly denies the simultaneous involvement

of various factors; the delineation between them, however, rests in what is ascribed as the primary factor. The list of models we present below is not exhaustive, but contains the proposals that have received the most attention for the past 15 years—and, therefore, the ones that have had a chance at the time of writing this to be systematically assessed through L3-specific empirical work. The *Scalpel Model of third Language acquisition* (Slabakova, 2017) and the *Linguistic Proximity Model* (Westergaard, Mitrofanova, Mykhaylyk, and Rodina, 2017) are not considered directly precisely because their recency translates to a dearth of studies that incorporate their predictions into the experimental design. To include them precipitously after a year of existence would thus not be fair to these new models. Many details aside, both predict that both languages can influence L3 simultaneously, in other words, they predict some level of hybridity from both sources. We have coded for hybrid transfer, which can then be used indirectly in view of these models.

### **2.1. Models of morphosyntactic transfer in L3/*L<sub>n</sub>* acquisition**

In general terms, there are two possibilities with respect to transfer at the onset of L2 acquisition: that it comes from the L1 or that there is no transfer at all—a debate with a long history in SLA studies (e.g., Epstein, Flynn, and Martohardjono, 1996; Odlin, 1989; Schwartz and Sprouse, 1996; Vainikka and Young-Scholten, 1996; see Foley and Flynn, 2013, for updated review). The picture in L3/*L<sub>n</sub>* acquisition is somewhat more complex in what pertains to potential sources of transfer, since we need to consider four logical possibilities *a priori*: (i) there is no transfer; (ii) transfer comes exclusively from the L1; (iii) transfer comes exclusively from the L2; (iv) transfer may come from either language, or from both at the same time, in whole or in parts. Some of these possibilities—notably (iii) and (iv)—have been articulated into models or hypotheses proposed within the last 15 years, which we will introduce below. No formalized model to date has been put forward in line with possibilities (i) and (ii), although

the latter—default L1 transfer—has been indirectly suggested from (at least partially) supportive data from a number of studies.

### **2.1.1. A privileged role of the L1**

Some of the work on L3 grammar acquisition seemed to support the idea of a dominant role of the native language (e.g., Hermas, 2010, 2015; Jin, 2009; Na Ranong and Leung, 2009). That is, that the default source of transfer or the only source of possible transfer is the native, first-acquired language. Even in studies which have claimed to support this with empirical data, there is no discernible explanation as to why this should be so. It is possible, for example, that the L1 is privileged for all subsequent language transfer because native L1s tend to remain the dominant language of successive bilinguals (see Lloyd-Smith, Gyllstad, and Kupisch, 2017, for phonology) and, therefore, it occupies somehow a more accessible and economic blueprint for other languages to be learned. Whatever the reason turns out to be, it runs in parallel to the main claim that the L1 trumps all other linguistic knowledge.

With the exception of Hermas' work, most studies highlighting a potential L1 default effect predate the present L3/ $L_n$  models of transfer, meaning that the data in these pre-existing studies (and even Hermas' work) could be equally accounted for by, or is compatible with, the currently available formal models—in consideration today of things not considered at the time. An L1 default in transfer source selection is indeed a strong hypothesis, precisely because it makes very clear and straightforward predictions that are amenable to testing, and thus falsifiable by evidence of transfer from the speakers' L2(s).

### **2.1.2. The L2 Status Factor Hypothesis**

The main claim of the L2 Status Factor hypothesis (henceforth L2SF; Bardel and Falk, 2007; Bardel and Sánchez, 2017; Falk and Bardel, 2011), as originally formulated, is that an L2

acquired in adulthood will have a privileged status as a source of morphosyntactic transfer. The L2SF's claim is that the L2 will be active throughout L3/L $n$  development and not only at the initial stages. In its most current instantiation, this model is conceptually aligned to Paradis' (2009) Declarative/Procedural model, which argues that the grammars of native and non-native languages acquired after puberty are sustained by different memory systems. The claim is that, while the L1 grammar is fundamentally procedural, all other grammars acquired in adulthood (plus all lexicons, including that of the L1) are mediated by declarative memory. Under this assumption, the L2SF maintains that an L2 will be more likely to influence the process of L3/L $n$  acquisition because, in Bardel and Falk's (2012) terms, the L2 and L3 are cognitively more similar (than the L1 and the L3) in their status as (adult) non-native languages.<sup>1</sup>

Recent instantiations of the model (Bardel and Sánchez, 2017; Falk et al., 2015) have begun to address certain subset situations within sequential bilingualism where the two-way distinction between implicit L1 competence and explicit L2 knowledge may not be so clear-cut, thus making it difficult to derive straightforward predictions from the initial premises of the L2SF. These situations include, most notably, the case of L3 learners who have received substantial metalinguistic training in their L1, which, may lead to the presence of L1-specific grammatical knowledge in these learners' declarative memory. Which prior language is then selected as the source of transfer largely depends, according to Bardel and Sánchez (2017), on individual differences in cognitive function such as working memory capacity and attention control, which are crucially involved in the process of evaluating and comparing the L3 input to the relevant representations from previously acquired languages. Under these premises, non-

---

<sup>1</sup> The theoretical underpinnings of the L2SF implicitly seem to confine its scope to the case of L3/L $n$  learners who have already acquired an L2 in adulthood (i.e., sequential bilinguals); most other bilingual populations, including simultaneous bilinguals, heritage speakers and child L2 learners, inevitably fall outside of the model's scope (see also Hopp, 2018).

facilitative transfer is not ascribed to a default in transfer source selection, but rather to shortcomings in cognitive capacities that lead to the selection of a non-targetlike representation.

### **2.1.3. The Cumulative Enhancement Model**

The Cumulative Enhancement Model (CEM; Berkes and Flynn, 2012; Flynn, Foley, and Vinnitskaya, 2004) proposes that both previously acquired languages are available for transfer, at any point in the process of L3 acquisition. The model is predicated on the principles of non-redundancy and maximal facilitation in successive language acquisition, which entails that transfer from previously acquired languages is only expected to obtain when such facilitates the acquisition of the target L3/*L<sub>n</sub>* property. In terms of transfer source selection, this translates into two main scenarios: (a) if one of the languages contains the target property and the other one does not (or has a non-target-like value for it), the former will transfer; and (b) if none of the languages may be of help, transfer will not obtain and the target property will be acquired in the same way it is in L1 acquisition. In short, the CEM proposes that transfer is selectively applied in L3/*L<sub>n</sub>* acquisition at the level of individual linguistic properties, if and only the creation of a target-like linguistic representation in the new grammar is facilitated. The idea of a mechanism sensitive to small, property-specific variation in the target L3/*L<sub>n</sub>* input first proposed by the original CEM paper is a valuable contribution that has been resurrected in the most recent models (e.g., Slabakova, 2017; Westergaard, Mitrofanova, Mykhaylyk, and Rodina, 2017).

### **2.1.4. The Typological Primacy Model**

The Typological Primacy Model (TPM; Rothman, 2010, 2011, 2015) proposes that, at the very beginning of L3/*L<sub>n</sub>* acquisition, all grammars of previously acquired languages are available for transfer. Paralleling Schwartz and Sprouse's (1994, 1996) Full Transfer/Full Access model of L2 acquisition, the TPM assumes that one of these grammars is transferred in its entirety, as

early in the process as possible—as soon as the linguistic parser has gathered enough information to adjudicate between the available choices.

The TPM argues that the linguistic parser selects the previously acquired language for which the highest degree of typological (structural) proximity<sup>2</sup> is detected, this being, potentially, a proxy for the largest amount of structural crossover between the L3 and the different possible sources (L1 or L2). Rothman (2015) proposes an implicational hierarchy of linguistic cues hypothesized to guide the parser in this task: language specific Lexicon → Phonology → Morphology → Syntax. The parser scans the available L3 input, assessing the degree of structural similarity between the L3 and the previously acquired languages at each of these levels, until a critical threshold of activation is reached for one of the prior languages. The fact that this is an implicational hierarchy means that, in some cases, the lower levels will not be considered, because the threshold will have already been met by a higher level in the hierarchy.

Similarly to the L2SF, the fact that only one of the prior languages is selected for transfer entails that the outcome of transfer will in some cases be non-facilitative. Unlike theories advocating transfer on a when-needed, domain-by-domain basis (e.g., Flynn et al., 2004; Slabakova, 2017; Westergaard et al., 2017), there is no need for the model to posit additional factors in order to explain a particular non-facilitative outcome of transfer, since this possibility

---

<sup>2</sup> Linguistic typology as a system of classification is based on shared formal (grammatical) features across languages, independently of genetic relationship. What this means, essentially, is that linguistic typology is complementary, but crucially not equivalent, to linguistic genealogy—i.e., the classification of languages into families according to their documented or presumed common origin. For example, both Japanese and Tamil are head-final languages (phrase heads are typically preceded by their complements) and therefore typologically similar with respect to head-directionality, but it seems unlikely that they are even remotely related. That said, genetically related languages tend to have moderately high degrees of overlap in typological terms, when their grammars are compared in whole. And so, by typology the TPM refers to underlying structural similarity (see González Alonso and Rothman, 2017; Rothman, 2011, 2015 for discussion). Throughout the paper, we use the term “typology” in strictly linguistic terms, “typological proximity/similarity” as measured over the whole grammar, and “typological transfer” to be the one predicted by the specific learner-internal mechanisms proposed in the TPM—i.e., *not* by anecdotal perceptions of overall similarities.

follows straightforwardly from the relative amount of mismatch between the transferred and target grammars.

### **3. Rationale and research questions**

Our main goal is to explore, describe and critically analyze methodological practices currently followed in studies on morphosyntactic transfer in L3/L $n$  acquisition, in an effort to shed better light on what the collective whole of the data reveal. We hope to lay the ground for more robust consensus, showing that some of the disparities in argumentation and seemingly mutual exclusivity of positions in the field are, at least in part, predicated on the interpretation of data stemming from methodological issues. We seek to uncover, to the extent they exist, potential associations between methodological choices/practices and data outcomes. If on the right track, this will then form the basis to argue for consolidating consistency in future experimental design for the purposes of reliability/replicability and maximal comparability across studies. We are guided by the following leitmotif query:

⇒ *What will examining a critical mass of studies reveal specifically for the role previous linguistic experience has for linguistic transfer in successive adult multilingual acquisition?*

To answer this question, we follow standard practices in other methodological syntheses/reviews in the field of SLA (e.g., Plonsky and Kim, 2016; Roessingh, 2004), as detailed in the following section.

In conducting this review, we do not mean to ignore the fact that certain theoretical questions demand particular methodological choices, and that the theory one subscribes to is the first and foremost factor in adopting some choices in experimental design. Having said this, however, it is important to recognize when such a conventional truth holds and when it should not. To illustrate this with a variable from our review, testing the domain of grammar in the L1 and L2 to be examined in the L3 to know for sure what each individual has as a potential source

of transfer, should be of no consequence to the theoretical debate between the models. It is a question of what potential future standard practice should be in this emerging field. If a comparison between adhering to this practice or not reveals an insightful, significant trend of differences then we might simply agree as a byproduct of showing this to be conservative in future expectations of L3 studies. Who would deny that the more conservative practice of assessing what is available from an L1 and L2 for transfer is best practice—after all if an L2er does not have a unique L2 representation or has one that is not fully developed they could only transfer what would appear to be the L1 or an L1-influenced one even if coming from the L2 grammar inventory. The question is whether such a practice yields a benefit? Besides being more precise in the obvious ways, is it actually necessary given that it represents time and resources? Beyond opinion, answering questions of this type can only be done in a quantifiable manner by a review like the present one.

#### **4. Design of the systematic review**

##### **4.1. Retrieval of studies**

Two main types of studies were included in the review: (a) studies published in peer-reviewed publications (journal articles, book chapters and conference proceedings) and (b) doctoral dissertations with a special emphasis on transfer in L3/ $L_n$  acquisition. The search, exhaustive to the extent possible, was conducted through Google Scholar, Proquest and Language and Linguistic Behavior Abstracts (LLBA). Relevant studies were located using the models' names as keywords, as well as inspecting the citing articles for each model's main publications.

After each citation was manually examined, a second filter was applied: we included only those publications which (a) included original data sets—i.e., we excluded epistemological commentaries and review articles—and (b) met one or more of the following criteria: (i) focused on transfer in L3/ $L_n$  acquisition; (ii) focused on testing specifically the models of L3/ $L_n$  acquisition discussed above; and/or (iii) focused on modelling L3/ $L_n$  acquisition.

In total, 41 independent publications/dissertations were included in the analysis. When one of the independent publications or dissertations contained more than one experiment, each experiment was coded as an individual study. In the final analysis, a total of 71 different studies were examined. The 71 different studies in the dataset were published—or defended—by 48 different researchers between 2004 and 2017 (see Appendix A for studies included in the analysis and further information on them).

#### 4.2. Coding procedure

The coding was done independently to what the authors of each study had argued from their interpretation of the results. The reason for this is that a number of the studies pre-dated the suggestion of some of the variables under consideration, and so the authors had not included them in the analysis. Even though their interpretation tended (in most studies) to coincide with our coding, we decided to apply an independent coding scheme to all studies. To do so, we examined the methodological choices and results presented in each study and we consistently coded each study following the same two-step process. In order to probe for potential compatibilities with more than one model at the same time—besides the one(s) to which each study claims to lend support—the first step was to code each experiment using a binary scheme with five macro-variables meant to capture the source (and type) of transfer: (a) *L1 transfer*; (b) *L2 transfer*; (c) *Typological transfer* (as defined in Rothman (2015), see 2.1.4 above and 4.2.1 below); (d) *Hybrid transfer* (simultaneous transfer from both languages); and (e) *Non-facilitative transfer* (See appendix B). Table 1 offers a summary of these macro-variables and the coding value associated to each level. Note that it is possible for each of the 71 studies to, in principle, get a check for several factors.

**Table 1. Binary Value Assignment to Macro-Variables and Factors in the Study**

Macro-variables	Levels
L1 Transfer	Yes (+), No (-)

L2 Transfer	Yes (+), No (-)
Typological transfer	Yes (+), No (-)
Hybrid transfer	Yes (+), No (-)
Non-facilitative transfer	Yes (+), No (-)

Each study was then further coded for five different methodological factors relevant to the field of L3/*L<sub>n</sub>* acquisition, to determine whether the use of a specific methodology might correlate to the source (and type) of transfer: (a) *Proficiency* of the participants in the L3; (b) *Languages tested* (i.e., whether they were tested only in the L3, or also in one or more of the previously acquired languages); (c) type of methodology (i.e., whether the study examined production or comprehension data); (d) *Mirror-image groups*, whether mirror-image participant groups were examined (e.g., L1 Spanish, L2 English, L3 Catalan vs. L1 English, L2 Spanish, L3 Catalan) and (d) *Language combination* (i.e., whether either or both previous languages were genetically related to the L3, e.g., L1 Spanish, L2 English, L3 Catalan, where Spanish and Catalan are genetically close, versus L1 Japanese, L2 English, L3 Arabic, where none of the languages are related); (See Appendix C). These categories are explained in more detail below. Like the macro-variables, our five methodological factors were coded as binary variables. As noted above, in principle each study could check off several of these variables at a time. Table 2 contains a summary of the factors and a description of variable levels:

**Table 2 Methodological predictors/factor included in the study**

<b>Methodological factors</b>	<b>Levels</b>
Proficiency	Beginner (+), Post-beginner (-)
Languages tested	L3 only (+), L3 + L2 and/or L1(-)
Methodology	Production (+), Comprehension (-)
Mirror-image groups	Yes (+), No (-)
Language combination	High degree of relatedness (same language family) (+), Little to no overt relatedness (-)

#### 4.2.1. Macro-variables

The five macro-variables listed in table 1 are self-explanatory, in that we coded for whether a given study's results are potentially compatible with the constructs of (exclusive) *L1* or *L2* transfer, *Typological transfer*, *Hybrid transfer* or *Non-facilitative transfer*. As we alluded to above, not all of these distinctions/variables are mutually exclusive. Experimental designs/choices can inadvertently obscure the path to meaningfully testing the models against one another, by confounding predictions or due to real-world limitations concerning availability of very specific subjects with the right language pairings, at precisely the right moments in time along the L3 developmental continuum (González Alonso and Rothman, 2017). As a result, a study may receive a positive value in just one, two, or several of these macro-variables. For example, two of the groups compared in Rothman and Cabrelli Amaro (2010) are particularly relevant: L1 English→ L2 Spanish→ L3 French and L1 English→ L2 Spanish→ L3 Italian. The results suggest that transfer obtained from the L2 into the L3 for both groups (i.e., L2 Spanish into both L3 Italian and L3 French). Since Spanish, a Romance language like French and Italian, was the L2 for both groups of learners, the *L2 transfer* and *Typological transfer* variables were confounded in this case; a positive value was thus assigned to both macro-variables in our analysis. This, however, does not apply when only half of the data within the same experiment/study can be accounted for by a macro-variable. A good example are studies where there is a mirror-image methodology used specifically to test between default status transfer (the L1 or L2) versus a more nuanced situation of transfer where it would depend on some variable other than order of acquisition alone. In Rothman's (2010) study looking at word order restrictions and relative clause attachment preferences, for example, the mirror-image groups were L1 Spanish→ L2 English→ L3 Brazilian Portuguese (BP) and L1 English→ L2 Spanish→ L3 BP learners. *L1* and *L2 transfer* macro-variables were not counted as positive, since Spanish was transferred in both groups—thereby showing L1 or L2 transfer is not an

absolute default and, in this case, selection seems compatible with overall typological/structural proximity.

While three of the macro-variables (*L1* and *L2 transfer* and *Non-facilitative transfer*) are self-explanatory, it is worth highlighting what we mean by the labels *Typological Transfer* and *Hybrid transfer* here. In the first case, and since the macro-variables are meant to capture the main predictors of transfer source selection as defined in each of the models, *Typological transfer* is operationalized as that which is predicted by applying Rothman's (2015) TPM hierarchy to each case. *Hybrid Transfer* refers to those cases where influence from both languages could be observed for the same group, in either of three possible situations: combined influence on the same linguistic property (a true hybrid value); influence on different properties, that is, when in a single experiment with two conditions one is seemingly influenced by language X(L1), and the other by language Y(L2); and, finally, those situations where it was not possible to exclude a hybrid value (tease out the L1 from the L2) because both the L1 and L2 are functionally the same. For example, in an interpretation task it could be the case that participants assign an interpretation from the L1 40% of the time and 60% from the L2 to a condition in the L3. Essentially, this macro-variable operationalizes two different, but related, theoretical positions: that transfer obtains selectively on a property-by-property basis (e.g., Flynn, Foley, and Vinnitskaya, 2004; Slabakova, 2017), and that it may consist of a combined influence from both languages, even within a single linguistic property (Westergaard et al., 2017).

#### **4.2.2. Methodological factors**

##### *Proficiency in the L3*

This factor concerned whether participants were tested at the initial stages of L3/*L<sub>n</sub>* acquisition, or later in development. Our aim is twofold and grounded in theoretical as well as methodological reasons. First, as discussed, not all of the theories presented above are intended

to model transfer throughout L3/*Ln* development: the TPM, in particular, contends that the grammar of one of the learner's previous languages is transferred in whole shortly after first exposure, but has little to say about what the dynamics of cross-linguistic influence will be at various later stages of L3/*Ln* acquisition thereafter. One can derive (some) predictions, however, for intermediate and advanced proficiency learnability issues that follow from the TPM's initial stages transfer predictions (González Alonso and Rothman, 2017), making it a viable option to test with more advanced L3 development in limited contexts. The second reason is methodological in nature, and dovetails with the first. Learners make fewer errors as their proficiency increases, which means that, as we move away from the initial stages, it is less and less likely to come across errors, including those that can be attributed to transfer from previously acquired languages. In other words, the concentration of instances of our object of study (linguistic transfer) is inversely proportional to proficiency level, which makes the initial stages a more suitable testing ground. After all, failure to see an influence at an intermediate or advanced levels tells you nothing about whether or not it obtained at a lower proficiency level and has since been "worked out". Since the CEM and the L2SF make predictions that hold equally at any stage of L3/*Ln* development, data from novice learners are valid for the purpose of vetting these theories. When considering these two arguments together, it seems reasonable to assume that the stage at which participants were tested may have an impact in the way a dataset can appear to support one model over others. And so, we used two levels in our coding of this factor: *Beginners* and not beginners (i.e., *Post-beginner* learners), which for our purposes capture the necessary distinction.<sup>3</sup>

---

<sup>3</sup> The range of levels within what we have included under 'post-beginner' is wide, from intermediate (e.g., Santos, 2013) to even near-native learners (García Mayo and Slabakova, 2015; Slabakova and García Mayo, 2015). However, and besides the reasons we have just offered, we limited the levels of this variable to two for ease of comparability: measures of proficiency vary greatly across studies (ranging from self-assessment to standardized tests), and therefore it would have been difficult—if possible at all—to develop an independent taxonomy where the studies could confidently be assigned to different levels.

### *Languages tested*

Determining the source of transfer in L3/ $L_n$  acquisition is not always straightforward. In a property-by-property sense, it is not possible to test all language combinations for the purpose of this question. That is, the tripartite language pairing in juxtaposition to the grammatical property being tested, and in consideration of the research question being asked, matters a great deal. In order for the combination to be an appropriate one—in the sense of being able to address *a priori* the question of transfer source—one must first ensure that the L1 and L2 themselves, in the mind of each participant, have different values for the property tested.

Once it is established that the grammars themselves, in principle, have two different values for the target property, we indeed have a suitable combination to begin; all things being equal, relative influence from one grammar or the other can be teased apart empirically. However, the mere fact that the languages in an L1/L2 combination have, in principle—that is, at least for native monolinguals of the two languages—distinct representations for a given property does not mean that an individual L2 learner herself has (already) acquired two distinct representations. Decades of work in second language acquisition documenting differences in ultimate attainment and lingering effects of L1 transfer, even at so-called near native levels of L2 acquisition, show that such an assumption would be inappropriate (e.g., Abrahamsson and Hyltenstam, 2009; Bylund, Abrahamsson, and Hyltenstam, 2012; Clahsen and Felser, 2006; DeKeyser, 2000; Granena and Long, 2013; Hawkins and Casillas, 2008; Hawkins and Chan, 1997; Johnson and Newport, 1989; Long, 2005; Sorace, 2011; Tsimpli and Dimitrakopoulou, 2007).

Overcoming the potential confounds of not choosing appropriate language combinations, and/or appropriate subjects in terms of L2 attainment for the domain of grammar, is relatively simple. In the first place, one simply must choose a property that has distinct representations in

the grammars that constitute the contributing L1 and L2s in the triad. If testing a specific grammatical property is, for independent reasons, more important to the researcher than the combination of languages itself, then selecting the right combination of languages becomes crucial. Secondly, testing each participant's competence for the specific grammar domain of interest in all three languages, in order to know the actual state of linguistic representations available for L3 transfer, is also crucial. In an attempt to quantify the potential impact of not knowing for sure what is available for transfer in the L2, we classified studies into two types: those where participants were tested in the L3 alone (*L3 only*) and those in which minimally the L2 was also tested, if not both the L1 and L2 were also tested for the same linguistic property (*L3 + L2 or L1/L2*).

### *Methodology*

Research in related areas of language development, such as L2 acquisition and heritage language bilingualism, has frequently discussed mismatches in the outcomes of studies as a function of the type of methodology used, particularly along two axes: online (i.e., real-time) vs. offline measures, and comprehension vs. production tasks (e.g., Bialystok, 1979, 1982; Bowles, 2011; Dussias, 2003, 2004; Ellis, 2005; Godfroid et al., 2015; Jegerski, Keating, and VanPatten, 2016; Villegas, 2014; among many others). Given this record in parallel subfields, it is reasonable to consider that the type of task employed might also be an important factor in L3/ $L_n$  acquisition research, and that we might find some patterns of correlation between studies' methodologies and the general direction of their results. Owing to the dearth of relevant studies that have employed truly online measures (e.g., eye-tracking, event-related potentials) in adult L3 acquisition, there is not enough data to explore potential effects within the online-offline methodological continuum. There is, however, considerable variability as to whether

studies analyze production or comprehension data. Therefore, we coded the *Methodology* factor in two levels: *Production vs. Comprehension*.

#### *Use of mirror-image groups*

One of the many ways to classify the current models of morphosyntactic transfer in L3/Ln acquisition is by whether or not they contend that the order of acquisition crucially determines the default, or at least predominant, source of transfer. While the L2SF and L1 default proposals assign a prominent role to the L2(s) and the L1(s), respectively, historically established models such as the CEM and the TPM as well as the two newest models, the Linguistic Proximity Model (LPM; Westergaard, Mitrofanova, Mykhaylyk, and Rodina, 2017) and the Scalpel Model (Slabakova, 2017) predict the source of transfer on the basis of factors that hold irrespective of whether the selected language is the learner's L1 or her L2. This can lead to overlapping predictions by various theories depending on several factors, for example, the specific property being tested, as described in detail above using the Rothman and Cabrelli Amaro (2010) study as an example where L2 status and typological proximity were confounded.

Since the most powerful dataset is one that is able to consider as many theories as possible within the same experimental design, some authors (e.g., Falk and Bardel, 2010; Rothman, 2010b; Rothman and Cabrelli Amaro, 2010) have encouraged the use of a specific method that helps researchers to tease apart predictions. Although getting such groups is not always possible, this involves the use of 'mirror-image' participant groups, for whom the L3 is shared and the L1 and L2 are the same languages but in reversed order of acquisition. For example, in a study examining the acquisition of Catalan as the L3 of Spanish-English learners, the mirror-image groups would be L1 English, L2 Spanish, L3 Catalan, and L1 Spanish, L2 English, L3 Catalan. With this type of design, models such as the L2SF predict, at least in principle, a

difference between the groups, since transfer will obtain from different languages. The CEM, for example, would expect both groups to behave similarly, because they predict the source of transfer to be determined by factors that are independent of chronological order of acquisition. This methodological factor had a straightforward binary coding: *Use or No use of mirror-image* groups.

### *Language combination*

As we discussed in previous sections, linguistic typology in a “genetic” sense has featured prominently in models of L3/*L<sub>n</sub>* morphosyntactic transfer, although it has invariably been alluded to as a (learner-external) proxy for the actual variables considered by these theories, which are cognitive in nature and thus internal to the learner. In other words, the fact that two languages are genetically related—or have a long history of more direct(ly relevant) contact—guarantees some degree of crossover in at least lexis and perhaps, especially in the case of languages belonging to the same family, phonology, syntax, morphology, information structure and beyond. To be clear, we used language family in the subset sense (Germanic, Romance, Slavonic) as opposed to the superset sense (e.g. Indo-European). If, as models such as the TPM or the LPM propose, structural similarity between the L3 and previously acquired languages is an extremely important, if not the most deterministic variable in the selection of a transfer source, genetic relatedness might be a broad-brushstroke pointer to the likely predominant linguistic influence. There is, of course, no actual guarantee that this will be the case, since typology (in both its diachronic and synchronic senses) is merely a learner-external factor that tends to correlate more or less strongly with variables the linguistic parser is indeed able to evaluate. Nevertheless, and in order to vet our theories beyond their most immediate scenarios (i.e., those in which they originated), research on language combinations where genealogical relatedness is present as well as those where it is absent is equally advisable. For this variable,

we coded studies depending on whether a genetic relation existed between the L3 and the L1 or the L2 (e.g., our previous case of L1 Spanish, L2 English, L3 Catalan, where the L1 and the L3 are closely related). Studies where neither the L1 nor the L2 were straightforwardly related to the L3 (an extreme case would be, for example, L1 Basque, L2 Spanish, L3 Swahili) were coded as *Not related*. Note that, as explained in our description of the macro-variables above, this methodological factor is *not* operationalized or calculated in the same way as the *Typological transfer* macro-variable.

## **5. Results**

### **5.1. Reporting and analysis**

In order to better navigate the results of this systematic review, we present them broken down by the macro-variables explained in section 4. Also, note that this section presents the results without evaluative assessment or other type of interpretation—discussion and unpacking of what the results reveal follow in section 6. As we discuss each macro-variable in turn, we provide an overview of how the methodological factors presented in section 4 distribute across the subset of the total studies whose outcome can be ascribed to the macro-variable in focus. Note that the tables summarizing by-methodological factor distributions in 5.2. through 5.6 necessarily reflect only the subset of studies pertinent to each macro-variable, and so percentages should be read with both these subset totals and the grand superset total of 71 studies in mind. This means that the methodological factors should be interpreted *within* as well as *across* the macro-variable distribution. For example, if it happens to be the case that a majority of the studies pointing to the L2 transfer macro-variable are, say, production studies, this does not necessarily mean that production methodologies reliably predict L2 transfer. What it means is that, for these studies available in the literature, such an association exists, implications of which are left open for discussion. In order to see if production itself truly

correlates with the outcome of L2 transfer, one would need to consider the distribution of the *Methodology* factor across the superset: it might be that a majority of all available studies employing production methodologies support other macro-variables as well, or better.

In consideration of a battery of Fisher’s exact tests—recall that each methodological factor is coded in a binary fashion—we report, for each subsection, whether any significant associations are observed between methodological factors and the specific outcome captured by the macro-variable. The choice of this statistical test over the more common Pearson chi-square was motivated by the fact that some of the cells did not meet the minimum raw count requirements of a chi-square test. Since we are limited by availability from the literature itself, Fisher’s exact test is the more appropriate method to explore the associations in 2x2 contingency tables when some of the cells have lower numbers (e.g., Wong, 2011).

## 5.2. L1 Transfer

Out of the 71 studies considered, 10 studies, 14.1% of the total, show transfer coming exclusively from the L1. Table 3 includes raw counts and percentages relative to the same distributions of each methodological variable over the whole sample of 71 studies.

**Table 3.** Distribution of studies by methodological factor within the L1 Transfer subset (n=10), and *p* values for Fisher exact tests on the associations between distribution and outcome. Bolded values indicate a significant result ( $p < .05$ ).

Variable	Level	n in L1T (%)	n in Other (%)	Sig. ( <i>p</i> )
<b>Proficiency</b>	Beginner (n=30)	3 (10%)	27 (90%)	.51
	Post-beginner (n=41)	7 (17.1%)	34 (82.9%)	
<b>Languages tested</b>	L3 only (n=55)	9 (16.4%)	46 (83.6%)	.43
	L3 (+L2/L1) (n=16)	1 (6.2%)	15 (93.8%)	
<b>Methodology</b>	Comprehension (n=45)	9 (20%)	36 (80%)	.08
	Production (n=26)	1 (3.8%)	25 (96.2%)	
<b>Mirror-Image</b>	No Use (n=47)	10 (21.3%)	37 (78.7%)	<b>.01</b>
	Use (n=24)	0 (0%)	24 (100%)	
<b>Combination</b>	Related (n=34)	3 (8.8%)	31 (91.2%)	.19
	Not related (n=37)	7 (18.9%)	30 (81.1%)	

As this is the first such chart, it is worth breaking down how to read it and thus the ones in the next sections. *Proficiency*, binarily coded as *Beginner* or *Post-beginner*, has a distribution of 3 (studies) and 7 (studies), respectively, over the relevant 10 studies for this variable (first two cells in the column “n(umber) in L1T”). For the same methodological factor, the following column (“n in Other”) reports the number of studies where *Beginners* or *Post-beginners* are used, respectively, within the remaining 61 studies (out of the 71 superset): 27 beginners and 34 post-beginners. The numbers in these two columns, the quadrant highlighted in grey, will always add up to 71, the total number of studies in the analysis. In both columns, percentages are relative to the total number of studies from the 71 broken down in the “Level” column, so whatever percentage of the 30 Beginner studies (3 out of 30) or the Post-beginner studies (7 out of 41) these 10 relate to across the whole. Incidentally, the two numbers in the “Level” column will also always equal the total number of studies, or 70. And so, 10% relates to 3 studies showing L1 transfer exclusively out of 30 studies that use beginners, and 17.1% to 7 studies showing L1 transfer out of the 41 where post-beginner learners were examined.

Fisher’s exact tests conducted to detect potential associations between the distribution of each factor and the *L1 Transfer* outcome revealed only one significant case: reporting L1 transfer effects is significantly associated to only one methodological factor, the absence of mirror-image groups (10 vs. 0) in these studies’ experimental designs ( $p=.01$ ).

### 5.3. L2 Transfer

Of the total 71 studies, 20 (28.2%) suggest that transfer comes exclusively from the L2. Table 4 below shows how the methodological factors we coded for distribute across this subset of 20 studies.

**Table 4.** Distribution of studies by methodological factor within the L2 Transfer subset (n=20), and *p* values for Fisher exact tests on the associations between distribution and outcome. Bolded values indicate a significant result ( $p<.05$ ).

Variable	Level	n in L2 (%)	n in Other (%)	Sig. ( <i>p</i> )
----------	-------	-------------	----------------	-------------------

<b>Proficiency</b>	Beginner (n=30)	8 (26.7%)	22 (73.3%)	.51
	Post-beginner (n=41)	12 (29.3%)	29 (70.7%)	
<b>Languages tested</b>	L3 only (n=55)	17 (30.9%)	38 (69.1%)	.52
	L3 (+L2/L1) (n=16)	3 (18.8%)	13 (81.3%)	
<b>Methodology</b>	Comprehension (n=45)	8 (17.8%)	37 (82.2%)	<b>.02</b>
	Production (n=26)	12 (46.2%)	14 (53.8%)	
<b>Mirror-Image</b>	No Use (n=47)	19 (40.4%)	28 (59.6%)	<b>.01</b>
	Use (n=24)	1 (4.2%)	23 (95.8%)	
<b>Combination</b>	Related (n=34)	10 (29.4%)	24 (70.6%)	.51
	Not related (n=37)	10 (27%)	27 (73%)	

As is shown in Table 4 above, two methodological factors are significantly associated with an L2 Transfer outcome. The first is *Methodology* (12 vs. 8 studies,  $p=.02$ ) in the favor of production methodologies. In other words, having chosen a production experiment seems to correlate with observing L2 Transfer effects. The second association, as in the *L1 Transfer* macro-variable above, is the correlation to *L2 Transfer* when a mirror-image design was not employed (19 vs. 1;  $p<.01$ ).

#### 5.4. Typological Transfer

Out of the 71 studies, the results of 43 of them (60.1%) can be ascribed to transfer that is typologically determined (see section 4.2.1 and footnote 2). Table 5 below shows the distribution of the methodological factors across these 44 studies, and the respective statistical results of Fisher's exact tests.

**Table 5.** Distribution of studies by methodological factor within the Typological Transfer subset (n=43), and  $p$  values for Fisher exact tests on the associations between distribution and outcome. Bolded values indicate a significant result ( $p<.05$ ).

<b>Variable</b>	<b>Level</b>	<b>n in TT (%)</b>	<b>n in Other (%)</b>	<b>Sig. (p)</b>
<b>Proficiency</b>	Beginner (n=30)	17 (56.7%)	13 (43.3%)	.62
	Post-beginner (n=41)	26 (63.4%)	15 (36.6%)	
<b>Languages tested</b>	L3 only (n=55)	28 (50.9%)	27 (49.1%)	<b>.01</b>
	L3 (+L2/L1) (n=16)	15 (93.8%)	1 (6.2%)	
<b>Methodology</b>	Comprehension (n=45)	28 (62.2%)	17 (37.8%)	.80
	Production (n=26)	15 (57.7%)	11 (42.3%)	
<b>Mirror-Image</b>	No Use (n=47)	22 (46.8.3%)	25 (53.2%)	<b>.01</b>
	Use (n=24)	21 (87.5%)	3 (12.5%)	
<b>Combination</b>	Related (n=34)	27 (79.4%)	7 (20.6%)	<b>.01</b>

Not related (n=37)	16 (43.2%)	21 (56.8%)
--------------------	------------	------------

The distributions of three of the methodological factor coded for are significantly associated to a *Typological Transfer* outcome. The first is related to the *Combination* of languages. 27 of these 43 studies were conducted with combinations where at least one of the previous languages was genetically related to the L3, versus 16 studies where all languages were genetically unrelated ( $p < .01$ ). The second association is with use of a *Mirror-image* methodology. Contrary to the L1 and L2 Transfer macro-variables where *Mirror image* also turned out to correlate, the significant association here is found in the opposite direction; using a mirror imagine methodology was done by more studies in the relevant subset (22 vs. 21,  $p = .01$ ). Finally, a significant association is found between Typological Transfer and the *Languages tested* factor ( $p = .01$ ), which, as you will recall, relates to whether a study tested only the L3 or if it indeed also tested knowledge of the target domain in at least the L2 (if not the L2 and L1).

### 5.5. Hybrid transfer

So far, we have examined macro-variables relating to transfer from one linguistic system, be it the L1 or the L2—for reasons of order of acquisition or structural similarity. The macro-variable we have labeled *Hybrid transfer* considers those cases in which a study reported evidence of transfer from both the L1 and the L2 within the same subjects. 17 of the 71 studies (23.9%) found some evidence of transfer from both languages.

**Table 6.** Distribution of studies by methodological factor within the Hybrid Transfer subset (n=17), and  $p$  values for Fisher exact tests on the associations between distribution and outcome. Bolded values indicate a significant result ( $p < .05$ ).

Variable	Level	n in HT (%)	n in Other (%)	Sig. (p)
<b>Proficiency</b>	Beginner (N=30)	7 (23.3%)	23 (75.6%)	.57
	Post-beginner (N=41)	10 (24.4%)	31 (75.6%)	
<b>Languages tested</b>	L3 only (N=55)	16 (29.1%)	39 (70.9%)	.07
	L3 (+L2/L1) (N=16)	1 (6.2%)	15 (93.8%)	
<b>Methodology</b>	Comprehension (N=45)	7 (15.6%)	38 (84.4%)	<b>.04</b>
	Production (N=26)	10 (38.5%)	16 (61.5%)	
<b>Mirror-Image</b>	No Use (N=47)	15 (31.9%)	32 (68.1%)	<b>.03</b>
	Use (N=24)	2 (8.3%)	22 (91.7%)	

<b>Combination</b>	Related (N=34)	4 (11.8%)	30 (88.2%)	.06
	Not related (N=37)	13 (35.1%)	24 (64.9%)	

The statistical tests reveal that two methodological factors (*Methodology* and *Mirror-image*) are significantly associated with an outcome of *Hybrid transfer*. Considering whether a particular study showing hybrid transfer (n=17) employed a production versus a comprehension type of method seems to matter whereby L3 production correlates to transfer hybridity (10 vs. 7,  $p = .04$ ). Moreover, of the relevant subset, studies not using a mirror-image methodology are associated with studies that reveal *Hybrid transfer* (15 vs. 2,  $p = .03$ ).

### 5.6. Non-facilitative transfer

Recall that this last macro-variable refers to the apparent transfer of a linguistic property into the L3 from a previously acquired language that does not facilitate grammar building towards the target.

**Table 7.** Distribution of studies by methodological factor within the Non-facilitative Transfer subset (n=62), and  $p$  values for Fisher exact tests on the associations between distribution and outcome. Bolded values indicate a significant result ( $p < .05$ ).

<b>Variable</b>	<b>Level</b>	<b>n in NT (%)</b>	<b>n in Other (%)</b>	<b>Sig. (p)</b>
<b>Proficiency</b>	Beginner (n=30)	29 (96.7%)	1 (3.3%)	.17
	Post-beginner (n=37)	33 (89.2%)	4 (10.8%)	
<b>Languages tested</b>	L3 only (n=51)	47 (92.1%)	4 (7.9%)	.81
	L3 (+L2/L1) (n=16)	15 (93.8%)	1 (6.2%)	
<b>Methodology</b>	Comprehension (n=42)	39 (92.8%)	3 (7.1%)	.21
	Production (n=25)	23 (92%)	2 (9%)	
<b>Mirror-Image</b>	No Use (n=43)	41 (95.3%)	2 (4.7%)	.21
	Use (n=24)	21 (87.5%)	3 (12.5%)	
<b>Combination</b>	Related (n=33)	30 (90.9%)	3 (9.1%)	1
	Not related (n=34)	32 (94.1%)	2 (5.9%)	

As can be seen in Table 7, the general picture clearly suggests it is possible and indeed quite likely to experience non-facilitative transfer in L3/ $L_n$  acquisition: in fact, 62 out of the 67 studies (92.5%) show evidence of non-facilitative transfer, as opposed to the 5 studies (7.5%) where all prior language influence seems to be facilitative. It is worth mentioning that, within

the 71 studies included in the review, 4 of them were coded as not applicable for this variable because these could probe for the possibility of non-facilitative transfer from both languages—i.e., the linguistic property or properties they test could only provide facilitative transfer or simply not obtain at all—and so it is impossible to determine if non-facilitative transfer could obtain for the same learners and the same languages testing different properties. The statistics reported above show that no significant associations were found; that is, irrespective of all potential methodological choices non-facilitative transfer is found equally robustly.

## **6. General discussion**

Several trends can be observed in the results, which we endeavor to unpack now. Recall that we did not take at face value support or lack thereof for any particular theory claimed by the authors of included studies. Instead, we coded each study for all the same variables and essentially reduced the models themselves to a particular combination of positive and negative values for those variables, namely, *L1 transfer*, *L2 transfer*, *Typological transfer*, *Hybrid transfer* and *Non-facilitative transfer*. To start, such an approach attempts to avoid overt and implicit biases on several levels, not the least could be our own implicit biases. In doing so, we were able to capture most neutrally what the data support irrespective of what is claimed in any particular study and to entertain all models for each data set, even if the study itself was limited to a subset of theories considered. Furthermore, since the models' predictions are not always entirely incompatible with each other our approach allowed us to capture when a given data set is compatible with more than one theory. Additionally, other factors related to methodological choices were encoded—e.g., *Proficiency* in the L3, whether all three languages were tested, whether the task examined production or comprehension, among others—to test the hypothesis that datasets in seeming discord in terms of what they reveal about multilingual transfer might be better explained as a byproduct of high order interactions. Before unpacking things, it is prudent to point out that the overall snapshot reveals significant variation across the studies and

across all relevant areas, that is, differences exist related to the backgrounds of the subjects tested, the languages in the trilingual pairings, the domains of grammar tested and several non-trivial distinctions in type, creation and administration of the testing methodology. As we saw in the previous section, the systematic review shows that some of the methodological factors we coded for were, indeed, significantly associated with the outcomes/claims of the studies.

Why should methodology matter? All methods employed contain some level of implicit biases towards particular outcomes—and this is not necessarily a bad thing, just one we need to be mindful of. The challenge becomes one of choosing the methods that convey the least or are best fit-for-purpose in line with our research goals. The first step in choosing the best cohort of methodological practices is to consider, upon the achieving of a critical mass of studies in a given field, the (inadvertent) effects of them. If inevitable effects are neutral as they pertain to our research questions, we can acknowledge them and put them aside. If they possibly obscure; however, we can and should consider what alternatives are more neutral and less entangling. We turn to this task now.

As pertains to the type of methodology used, significant associations were found between either production or comprehension-based methodologies and two of the macro-variables: *L2 only transfer* (e.g., Bardel and Falk, 2007; Tavakol and Jabbari, 2014) and *Hybrid transfer* (e.g., Angelovska, 2017; Fallah and Akbar Jabbari, 2016). Research in other populations has typically found a divide between production and comprehension data, as reported for child L1 acquisition (e.g., Hendriks, 2014), child L2 acquisition (e.g., Unsworth, 2007) and adult L2 acquisition (e.g., Gershkoff-Stowe, L. and Hahn, 2013). It is, thus, not entirely surprising that in L3 acquisition this divide is also apparent.

In order to understand language, the mind must in some ways reverse-engineer input received juxtaposed against whatever system is able to decode language (-specific) information. This is not to suggest that production does not require the same (in the opposite order of course);

we simply wish to point out that it requires much more, and this can add complexity to the task and thus extraneous noise to the proverbial signal we are trying to disentangle. Comprehension principally requires decoding, whereas production has further and more complex requirements (e.g., selecting words from the mental lexicon, assigning syntactic representations, passing from the mental computational representation to the phonological form for articulation, etc.). It might be the case, then, that production itself, especially at lower levels of proficiency, introduces variables that make the L2 more likely to be accessed for production—above and beyond when other co-occurring factors are at play.

As discussed in Falk and Bardel (2011) and Bardel and Falk (2012), the L2 might be more accessible for production because of its non-native status (potentially represented and stored differently). If on the right track, this could account for the association revealed within the subset of studies that show L2 only transfer—12 of 20 or 60%—but it would leave unexplained the overall results when considering the superset of 71 studies from which 26 were production methodologies (12 of 26 or 46.2%). However, one must also concede that insofar as production is more susceptible to influences beyond grammatical representation, studies showing seemingly default L2-based influence in production might capture processing based influence at a more superficial level than being truly reflective of underlying representations in the emerging L3 system—the latter being what all theories claim to be focusing on.

It makes sense that the surface output effects of production would reflect an L2 bias due to metalinguistic and/or recency effects of having learned an L2 in a similar way as an L3 (both different from an L1). Alternatively, a hybrid effect is also likely especially if production taxes the attentional/processing resource allocation. If the goal is specifically to determine the underlying representation used to parse L3 sentences, we might conclude that comprehension has a privileged status to be used and that it is thus a more appropriate methodology, especially for beginning learners. This is not to suggest that production is unimportant, quite the contrary.

We simply intend to suggest it would be more useful for other questions within L3 development and ultimate attainment, for example.

The (lack of) use of mirror-image groups showed significant associations with three of the macro-variables: *L1 transfer*, *L2 transfer* and *Typological transfer*. For the two macro-variables targeting order of acquisition as a determining factor (L1 vs. L2), their associations with the (lack of) use of this design showed that most of these studies do not employ the mirror-image design (e.g., Foote, 2009; Hermas, 2010; Na Ranong and Leung, 2009). The association with the *Typological transfer* macro-variable shows that studies with evidence for this type of transfer tend to use the design (Giancaspro, Halloran, and Iverson, 2015; Rothman, 2010). The fact that the mirror-image design is not employed in, at least, some of the studies from the former two groups is unfortunate. Recall that this design was explicitly devised and advocated for by authors of opposing theories to tease apart order of acquisition (either L1 or L2) from other potentially explanatory variables for transfer source selection (Falk and Bardel, 2010; García-Mayo and Rothman, 2012; Rothman and Cabrelli Amaro, 2010). Thus, if one study shows L1 transfer or L2 transfer but has not used a mirror-image design, we cannot rule out the possibility that the source of transfer was based on other factors rather than order of acquisition. We understand it is not always practical to find mirror-image groups. We also realize that if this were a requirement, it would severely reduce the language pairing we would be able to study for obvious practical reasons. Nevertheless, showing L1 or L2 transfer alone and using such to support a L1 or L2 privileged/default model of transfer is vacuous if one cannot rule out other possibilities the mirror-image design affords. In such cases, data are merely compatible with a given theory, not necessarily supportive of it. A reasonable alternative could be to compare L2 and L3 acquisition of the same target language when the L1 is held constant, but this too is not without potential confounds (see Cabrelli Amaro & Rothman, 2010).

With respect to the studies showing *L1 transfer* that are also compatible with other macro-variables, 4 out of 10 studies can just as well explained by *Typological transfer*. Perhaps the 8.5% of remaining studies (6/71) showing *L1 transfer* not otherwise accounted for is low enough to be taken as relative noise in an otherwise clearer signal. However, we cannot escape the fact that other variables might actually account for even this relatively low number overall. Almost none of these studies control for what the systematic review has revealed as important factors, such as using a *Mirror-image* approach and testing the status of the domain of grammar in the L2 to know for sure that a distinct L2 representation was actually available for transfer. Of the 20 studies showing *L2 transfer*, 16 also had a positive value for *Typological transfer*. Thus, the percentage of studies with unambiguous evidence for *L2 transfer* is reduced to 5.6% of the total (4 of 71 studies).

These results have two clear implications for the study of adult successive multilingualism. The first one is that order of acquisition, as postulated by original formulations of the L2 Status Factor or the group of studies advocating default L1 transfer, can hardly be considered the main factor in the selection of the source of transfer in (the initial stages of) L3/Ln acquisition. With ever larger bodies of evidence suggesting that transfer can come from an L1 or an L2 depending on other variables, L3/Ln transfer models incorporating order of acquisition defaults at the top of their hierarchy of factors will inevitably suffer to accommodate all presently available data. The second implication is that using the bi-directional mirror-image design is crucial to reveal the dynamic nature of multilingual transfer.

If a model wants to argue that strict order of acquisition (L1 or L2 as a default) is the most deterministic variable for transfer selection, then, not only does it need to provide a good explanation of what happens when this is not the case, but it also needs to be able to have accurate predictions for when order of acquisition will not be deterministic in transfer selection. The latest papers associated with the L2 Status Factor take this most seriously (Bardel and

Sánchez, 2017; Falk et al., 2015). They attempt to explain when and why L1 transfer might occur, arguing that high degrees of L1 metalinguistic knowledge trigger transfer from the L1 and/or individual differences such as working memory capacity conspire to explain unexpected outcomes. However, these are fairly new claims. Promising as they are and despite the fact that they make clear testable predictions, the methodological designs used up to now in the vast majority of studies—virtually all of the 71 reviewed here, including the ones conducted by these authors in previous years—do not allow for testing such claims.

The case for specifically testing knowledge of the grammatical domain under investigation in all three languages of each participant was made above and was pointed out in the analysis to be a key factor correlating to outcomes. Recall that to determine what the source of transfer is, we need to be confident that an individual has two distinct representations available for transfer (one clearly aligning with the L1 and the other different from the L1, if not exactly like the target L2). Given what we know about L2 acquisition from decades of research (see for reviews, Ortega, 2011; Slabakova, 2017; VanPatten and Williams, 2015), we simply cannot take for granted that all L3 learners have acquired all domains of the L2 and thus actually have multiple sources from which transfer selection can obtain. Yet, when we examine the associations between this methodological factor and the research outcomes, the only one that comes out as significant is its association with the macro-variable of *Typological transfer*. This reveals that a good portion of the studies showing *Typological transfer* have tested the L1 and L2 as well as the L3 of the same speakers with respect to the specific linguistic domain under investigation (e.g., Na Ranong and Leung, 2009; Santos, 2013). Equally, given the nature of the association itself, it reveals that when one knows for sure—because this is objectively tested—that there are two representations available from which transfer can obtain, the outcome almost always aligns with typological proximity (15 out of 16 relevant studies or 93.8%). It is true that most of these experiments, 13 of the 16 in fact, stem from papers from

Rothman's lab or former members of it, which might lead some to attribute this association more to preferences of a research group than anything else. However, any bias one might be inclined to attribute should not lead one astray from what is revealed and/or reduce the logical prudence of what is being advocated. The fact remains that these happen to be the only papers that control for L1 and L2 knowledge of the domain under investigation and the data clearly reveal that when this is done the trend is unmistakable. Who would argue, alternatively, that it is not good practice or that there is an implicit bias/confound to ensuring L3 learners have access to distinct L1 and L2 representations before investing in attempts to tease apart the source (L1 or L2) of L3 transfer. We submit that doing so should be a pre-requisite moving forward. Doing so a priori might reduce some of the variation in data we have, eliminating potential false positive of seemingly L1 based transfer.

As should not be overly surprising, *Language Combination* is significantly associated with *Typological Transfer* and, in fact, is the only one to show this. The results of the statistical test suggest that the degree of relatedness between the L1 or L2 and the L3 is a strong predictor for transfer selection. Note that out of the 34 studies that use a linguistic triad with high degree of relatedness, 27 of these studies find evidence for transfer from the language that is genetically closer to the L3. However, this is not to say that the only studies showing structurally-based typological transfer are those that test languages that are overtly, genetically related. In fact, 16 studies that use languages which are not genetically related provide data captured by comparative typological proximity when applying the TPM's implicational hierarchy (see Rothman, 2015). These results make it fair to establish and assume that the degree of similarity between languages, be it obvious or not, is crucial for transfer selection in L3/L<sub>n</sub> acquisition. Thus, it is important for any theory attempting to model the initial stages of L3/L<sub>n</sub> acquisition, and indeed trace its developmental trajectories, to factor in similarity between languages as a strong variable.

The fact that the non-facilitative claim of the CEM is refuted by over 92.5% of available datasets is quite convincing. One might ponder then, if it is time to discard this theory from further consideration moving forward. After all, the systematic review has made it clear that any adequate theory of morphosyntactic transfer in L3/L $n$  acquisition must minimally be able to accommodate instances of non-facilitative transfer from previously acquired languages. This renders a strong version of the CEM overwhelmingly unsupported. Models which follow similar principles to the CEM yet allow for the possibility of non-facilitative transfer, such as the Scalpel Model (Slabakova, 2017) and the LPM (Westergaard et al., 2017), are in many ways better suited to pursue the general idea that transfer is not wholesale in the beginning, but rather obtains on a property-by-property basis and indeed could reflect transfer/influence from both languages at the same time.

## **7. Conclusion**

Of course, no single variable, not even the one our analysis reveals as being overall the most explanatory—typological proximity—accounts for all the data. Recall that the macro-variables, which relate most closely to claims of the existing models revealed that L1 transfer is compatible with 14.1% of the results, L2 transfer is compatible with 28.2%, Typological transfer with 60.5%, and the CEM is compatible with only 5.9% of the results. And thus, it is fair to conclude that no current theory is proven correct by the analysis herein, even if some are more questioned and/or on a better track than others. This should come as no surprise. Indeed, it would be highly unlikely that any of the models, at least in their present form, would be correct in absolute terms; the field is likely too young for this to have obtained. This is also good news. It means that there is significant room for refinement to present models and space for new ones that build on the insights of its predecessors and the coverage (or lack thereof) they have of the data. The systematic review reveals that transfer/influence at multiple stages

in L3 development seems to be more dynamic than any one or any interactional combination of several variables—at least the ones considered so far—could capture. In this sense, the Scalpel Model and the LPM, especially since both take typological proximity to be an important variable, are welcome, very recent additions to this nascent field. However, it is not clear (yet) how either of these approaches predict a priori *when* non-facilitative transfer will obtain (other than assuming that it *can* obtain), nor do they seem to have defined in precise terms the mechanisms that give rise to this. They are especially promising additions because they embody both initial stages and developmental theories in one, therefore, we look forward to newer instantiations that further develop the predictive value and ecological validity of these approaches.

As is true of any review and/or synthesis of behavioral research, the findings emerging from the very exercise of doing a systematic review are relevant well beyond the field of inquiry itself. In fact, the present review can be used as a proxy to remind us of what we all know, yet due to multifarious reasons cannot always control for in all studies: (a) methodology matters a great deal, and (b) we need to triangulate various types of methodologies as well as variables considered in our analyses to tease apart co-varying factors affecting our conclusions. It takes a cohort of studies to reveal methodological implicational patterns and to use these patterns to hammer home important points. We hope to have shown just how this can effectively be done, for the benefit of scholars interested in questions of transfer in additive multilingual acquisition. We can conclude with the following recommendations to increase the comparative value of studies in this emerging field: (a) L3 studies should employ, where possible, a mirror image design (b) test the specific knowledge of the L3 domain of inquiry in the previously acquired languages, (c) use comprehension or production plus comprehension methods, especially for beginning L3 learners. To the extent that we all gravitate towards common practices and

general design in L3 studies, the more meaningful comparisons will be, and the clearer generalizations can be from the superset of L3 data.

## References

- Abrahamsson N and Hyltenstam K (2009) Age of Onset and Nativelikeness in a Second Language: Listener Perception Versus Linguistic Scrutiny. *Language Learning* 59(2): 249–306.
- Angelovska T (2017) (When) do L3 English learners transfer from L2 German? Evidence from spoken and written data by L1 Russian speakers. In: Angelovska T and Hahn A (eds) *L3 Syntactic Transfer*. Amsterdam, NL: John Benjamins.
- Bardel C and Falk Y (2007) The role of the second language in third language acquisition: the case of Germanic syntax. *Second Language Research* 23(4): 459–484. DOI: 10.1177/0267658307077080557.
- Bardel C and Falk Y (2012) The L2 status factor and the declarative/procedural distinction. In: Cabrelli Amaro J, Flynn S, and Rothman J (eds) *Third language acquisition in adulthood*. Amsterdam, NL: John Benjamins, pp. 61–78.
- Bardel C and Sánchez L (2017) The L2 status factor hypothesis revisited: The role of metalinguistic knowledge, working memory, attention and noticing in third language learning. In: Angelovska T and Hahn A (eds) *L3 Syntactic Transfer: Models, new developments and implications*. Amsterdam, NL: John Benjamins, pp. 85–102.
- Bayona P (2009) The Acquisition of Spanish Middle and Impersonal Passives. In: Leung YI (ed.) *Third Language Acquisition and Universal Grammar*. Bristol, UK: Multilingual Matters, pp. 1–29.
- Ben Abbes K (2017) *The Acquisition of French morpho-syntactic properties: Cross-linguistic Influence in the Learning of L3 French by Turkish/Spanish speakers who learned English*

as an L2. University of Essex.

- Berends S, Schaeffer J and Sleeman P (2017) Cross-Linguistic Influence in Adult Second Language Learners: Dutch Quantitative Pronoun Constructions. In: *Proceedings of the 41st annual Boston University Conference on Language Development* (eds M LaMendola and J Scott), Boston, US, 2017, pp. 74–87. Cascadilla Proceedings Project.
- Berkes É and Flynn S (2012) Multilingualism: New Perspectives on Syntactic Development. In: *The Handbook of Bilingualism and Multilingualism, Second Edition*. 2nd Edition. Chichester, UK: : John Wiley and Sons, Ltd.
- Bialystok E (1979) Explicit and implicit judgments of L2 grammaticality. *Language Learning* 29(1): 81–103. DOI: 10.1111/j.1467-1770.1979.tb01053.x.
- Bialystok E (1982) On the relationship between knowing and using linguistic forms. *Applied Linguistics* 3(3): 181–206. DOI: 10.1093/applin/III.3.181.
- Bohnacker U (2006) When Swedes begin to learn German: From V2 to V2. *Second Language Research* 22(4): 443–486.
- Borg K (2013) The Acquisition of Future of Probability in L3 Spanish. In: *Proceedings of the 12th Generative Approaches to Second Language Acquisition (GASLA 2013)*, Sommerville, MA, 2013. Cascadilla Proceedings Project.
- Boulton A and Cobb T (2017) Corpus Use in Language Learning: A Meta-Analysis. *Language Learning* 67(2): 345–393.
- Bowles M (2011) Meaning implicit and explicit linguistic knowledge. *Studies in Second Language Acquisition* 33: 247–271.
- Bruhn de Garavito J and Perpiñán S (2014) Subject pronouns and clitics in the Spanish interlanguage of French L1 speakers. In: *Canadian Linguistic Association Conference Proceedings*, 2014.
- Bylund E, Abrahamsson N and Hyltenstam K (2012) Does first language maintenance hamper

- nativelikeness in a second language? A study of ultimate attainment in early bilinguals. In: Abrahamsson Niclas and Hyltenstam K (eds) *High-level L2 Acquisition, Learning and Use*. Cambridge, UK: Cambridge University Press.
- Cabrelli Amaro J (2013) Raising across experiencers in L3 Portuguese: Further evidence for psychotypological transfer. In: Stavrakaki S, Konstantinopoulou P, and Lalioti M (eds) *Advances in language acquisition*. Newcastle upon Tyne, UK: Cambridge Scholars Press.
- Cabrelli Amaro J, Amaro F and Rothman J (2015) The relationship between L3 transfer and structural similarity across development. In: Peukert H (ed.) *Transfer effects in multilingual language development*. Amsterdam, NL: John Benjamins.
- Chin D (2009) Language Transfer in the Acquisition of the Semantic Contrast in L3 Spanish. In: Leung YI (ed.) *Third Language Acquisition and Universal Grammar*. Bristol, UK: Multilingual Matters, pp. 30–54.
- Clahsen H and Felser C (2006) Grammatical processing in language learners. *Applied Psycholinguistics* 27: 3–42.
- De Angelis G (2007) *Third or Additional Language Acquisition*. Clevedon, UK: Multilingual Matters.
- DeKeyser R (2000) The Robustness of Critical Period Effects in Second Language Acquisition. *Studies in Second Language Acquisition* 22(4): 499–533. DOI: 0272-2631/00.
- Dussias P (2003) Syntactic ambiguity resolution in second language learners: Some effects of bilinguality on L1 and L2 processing strategies. *Studies in Second Language Acquisition* 25: 529–557.
- Dussias P (2004) Parsing a first language like a second: The erosion of L1 parsing strategies in Spanish-English Bilinguals. *International Journal of Bilingualism* 8(3): 355–371.
- Ellis R (2005) Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition* 27: 111–139.

- Epstein S, Flynn S and Martohardjono G (1996) Second language acquisition: Theoretical and experimental issues in contemporary research. *Brain and Behavioral Sciences* 19: 677–758.
- Falk Y (2017) On pronouns that drop (out of German). In: Angelovska T and Hahn A (eds) *L3 Syntactic Transfer: Models, new developments and implications*. Amsterdam, NL: John Benjamins Publishing Company, pp. 127–142.
- Falk Y and Bardel C (2010) The study of role of the background languages in third language acquisition. *International Review of Applied Linguistics in Language Teaching (IRAL)* 48(2): 185–220.
- Falk Y and Bardel C (2011) Object pronouns in German L3 syntax: Evidence for the L2 status factor. *Second Language Research* 27(1): 59–82.
- Falk Y, Lindqvist C and Bardel C (2015) The role of L1 explicit metalinguistic knowledge in L3 oral production at the initial state. *Bilingualism: Language and Cognition* 18(2): 227–235. DOI: 10.1017/S1366728913000552.
- Fallah N and Akbar Jabbari A (2016) L3 acquisition of English attributive adjectives Dominant language of communication matters for syntactic cross-linguistic influence. *Linguistic Approaches to Bilingualism*. DOI: DOI: 10.1075/lab.16003.fal.
- Fallah N, Jabbari A and Fazilatfar A (2016) Source(s) of syntactic CLI: The case of L3 acquisition of English possessives by Mazandarani–Persian bilinguals. *Second Language Research* 32(7): 825–843.
- Fessi I (2014) Adquisición del contraste aspectual entre los pretéritos imperfecto e indefinido en español L3 por arabófonos tunecinos. Universidad Nebrija.
- Flynn S, Foley C and Vinnitskaya I (2004) The cumulative-enhancement model for language acquisition: Comparing adults' and children's patterns of development in first, second and third language acquisition. *International Journal of Multilingualism* 1(1): 3–16.

- Foley C and Flynn S (2013) The role of the native language. In: Herschensohn JR and Young-Scholten M (eds) *The Cambridge handbook of second language acquisition*. Cambridge, UK.
- Foote R (2009) Transfer in L3 acquisition: The role of typology. In: Leung, Y-K. I. (ed.) *Third Language Acquisition and Universal Grammar*. Bristol, UK: Multilingual matters.
- García-Mayo MDP and Rothman J (2012) L3 morphosyntax in the generative tradition. In: *Third Language Acquisition in Adulthood*. DOI: 10.1075/sibil.46.04pil.
- García Mayo M del P and Slabakova R (2015) Object drop in L3 acquisition. *International Journal of Bilingualism* 19(5). DOI: 10.1177/1367006914524643.
- Gershkoff-Stowe, L. and Hahn E (2013) Word comprehension and production asymmetries in children and adults. *Journal of Experimental Psychology* 114: 489–509.
- Giancaspro D, Halloran B and Iverson M (2015) Transfer at the initial stages of L3 Brazilian Portuguese: A look at three groups of English/Spanish bilinguals. *Bilingualism: Language and Cognition* 18(2). DOI: 10.1017/S1366728914000339.
- Godfroid A, Loewen S, Jung S, et al. (2015) Timed and untimed grammaticality judgments measure distinct types of knowledge. *Studies in Second Language Acquisition* 37: 269–297.
- González Alonso J and Rothman J (2017) Coming of age in L3 initial stages transfer models: Deriving developmental predictions and looking towards the future. *International Journal of Bilingualism* 21(6): 683–697. DOI: 10.1177/1367006916649265.
- González Alonso J, Rothman J, Berndt D, et al. (2017) Broad scope and narrow focus: On the contemporary linguistic and psycholinguistic study of third language acquisition. *International Journal of Bilingualism* 21(6): 639–650. DOI: DOI: 10.1177/1367006916653685.
- Granena G and Long M (2013) Age of onset, length of residence, language aptitude, and

- ultimate L2 attainment in three linguistic domains. *Second Language Research* 29(3): 31–343.
- Hawkins R and Casillas G (2008) Explaining frequency of verb morphology in early L2 speech. *Lingua* 118: 595–612.
- Hawkins R and Chan C (1997) The partial availability of Universal Grammar in second language acquisition: The ‘failed functional features hypothesis. *Second Language Research* 13: 187–226.
- Hendriks P (2014) *Asymmetries between Language Production and Comprehension*. Dordrecht: Springer. DOI: 10.1007/978-94-007-6901-4.
- Hermas A (2010) Language acquisition as computational resetting: verb movement in L3 initial state. *International Journal of Multilingualism* 7(4): 202–224.
- Hermas A (2015) The categorization of the relative complementizer phrase in third-language English: A feature re-assembly account. *International Journal of Bilingualism* 19(5): 587–607.
- Hopp H (2018) Cross-linguistic influence in the child third language acquisition of grammar: Sentence comprehension and production among Turkish-German and German learners of English. *International Journal of Bilingualism*: 136700691775252. DOI: 10.1177/1367006917752523.
- Imaz Aguirre A and García Mayo M del P (2017) Transfer effects in the acquisition of double object constructions in English as an L3. In: Angelovska T and Hahn A (eds) *L3 Syntactic Transfer*. Amsterdam, NL: John Benjamins.
- Ionin T, Grolla E, Santos H, et al. (2015) Interpretation of NPs in generic and existential contexts in L3 Brazilian Portuguese. *Linguistic Approaches to Bilingualism* 5: 215–251.
- Iverson M (2009) Competing SLA hypotheses assessed: Comparing heritage and successive Spanish bilinguals of L3 Brazilian Portuguese. In: Pires A and Rothman J (eds) *Minimalist*

- inquiries into child and adult language acquisition: Case studies across Portuguese.*  
Berlin, DE: Mouton de Gruyter, pp. 221–243.
- Jaensch C (2012) L3 Acquisition of German: Do some learners have it easier?. In: *Third Language Acquisition in Adulthood*. Amsterdam, NL: John Benjamins, pp. 165–193.
- Jarvis S (2000) Methodological rigor in the study of transfer: Identifying L1 influence in the interlanguage lexicon. *Language Learning* 50(2): 245–309.
- Jegerski J, Keating G and VanPatten B (2016) On-line relative clause attachment strategy in heritage speakers of Spanish. *International Journal of Bilingualism* 20(3): 254–268.
- Jin F (2009) Third Language Acquisition of Norwegian Objects: Interlanguage Transfer or L1 Influence? In: *Third Language Acquisition and Universal Grammar*. Bristol, UK: Multilingual Matters.
- Johnson J and Newport E (1989) Critical Period Effects in Second Language Learning: The Influence of Maturational State on the Acquisition of English as a second Language. *Cognitive Psychology* 21: 60–99.
- Kulundary V and Gabriele A (2012) Examining the role of L2 syntactic development in L3 acquisition. In: Cabrelli Amaro J, Flynn S, and Rothman J (eds) *Third Language Acquisition in Adulthood*. Amsterdam: John Benjamins, pp. 195–222.
- Leung YI (2005) L2 vs. L3 initial state: A comparative study of the acquisition of French DPs by Vietnamese monolinguals and Cantonese–English bilinguals. *Bilingualism: Language and Cognition* 8(1): 39–61.
- Leung YI (2008) The verbal functional domain in L2A and L3A. Tense and agreement in Cantonese-English-French Interlanguage. In: Liceras JM, Zobl H, and Goodluck H (eds) *The Role of Formal Features in Second Language Acquisition*. Mahwah: Erlbaum, pp. 379–403.
- Lindqvist C and Falk Y (2014) When Germans begin to learn Swedish - Which is the transfer

- source for function words, content words and syntax? In: *EUROSLA Yearbook 14*, Amsterdam, NL, 2014, pp. 225–239. John Benjamins. DOI: 10.1075/eurosla.14.09lin.
- Lloyd-Smith A, Gyllstad H and Kupisch T (2017) Transfer into L3 English: Global accent in German-dominant heritage speakers of Turkish. *Linguistic Approaches to Bilingualism* 7(2): 131–163.
- Long M (2005) Problems with supposed counter-evidence to the Critical Period Hypothesis. *International Review of Applied Linguistics* 43(4): 287–317.
- Mollaie A, Jabbari A and Javad Rezaie M (2016) The Acquisition of French (L3) Wh-question by Persian (L1) Learners of English (L2) as a Foreign Language: Optimality Theory. *International Journal of English Linguistics* 6(7): 36–47.
- Montrul S, Dias R and Santos H (2011) Clitics and object expression in the L3 acquisition of Brazilian Portuguese. Structural similarity matters for transfer. *Second Language Research* 27(1): 21–58.
- Na Ranong S and Leung Y-KI (2009) Null objects in L1 Thai-L2 English-L3 Chinese: An empirical take on a theoretical problem. In: Leung Y-KI (ed.) *Third Language Acquisition and Universal Grammar*. Bristol, UK: Multilingual Matters, pp. 162–191.
- Norris J and Ortega L (2000) Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning* 50: 417–528.
- Odlin T (1989) *Language transfer: Cross-linguistic influence in language learning*. Cambridge University Press. New York, NY: Cambridge University Press. DOI: 10.1075/jpcl.8.1.13byr.
- Ortega L (2011) *Second language acquisition*. London, UK: Routledge.
- Paradis M (2009) *Declarative and Procedural Determinants of Second Languages [Studies in Bilingualism 40]*. Amsterdam, NL: John Benjamins.
- Park M (2016) *Third Language Acquisition among Early Bilinguals*. National University of

Singapore.

- Plonsky L and Kim Y (2016) Task-based learner production: A substantive and methodological review. *Annual Review of Applied Linguistics* 36: 73–97.
- Plonsky L and Oswald F (2012) How to do a meta-analysis. In: *Research methods in second language acquisition: A practical guide*. London, UK: Basil Blackwell, pp. 275–295.
- Puig-Mayenco E, Miller D and Rothman J (2018) Language Dominance and Transfer Selection in L3 acquisition: Evidence from Sentential Negation and Negative Quantifiers in L3 English. In: Cho J, Iverson M, Judy T, et al. (eds) *Meaning and structure in second language acquisition: In honor of Roumyana Slabakova*. Amsterdam, NL: John Benjamins Publishing Company.
- Roessingh H (2004) Effective high school ESL programs: A synthesis and meta-analysis. *Canadian Modern Language Review* 60: 611–636.
- Rothman J (2010a) On the typological economy of syntactic transfer: Word order and relative clause high/low attachment preference in L3 Brazilian Portuguese. *International Review of Applied Linguistics in Language Teaching (IRAL)* 48: 245–273.
- Rothman J (2010b) On the typological economy of syntactic transfer: Word order and relative clause high/low attachment preference in L3 Brazilian Portuguese. In: *IRAL - International Review of Applied Linguistics in Language Teaching*, 2010. DOI: 10.1515/iral.2010.011.
- Rothman J (2011) L3 syntactic transfer selectivity and typological determinacy: The typological primacy model. *Second Language Research*. DOI: 10.1177/0267658310386439.
- Rothman J (2013) Cognitive economy, non-redundancy and typological primacy in L3 acquisition: Evidence from initial stages of L3 Romance. In: Baauw S, Dirjkoningen F, and Pinto M (eds) *Romance Languages and Linguistic Theory 2011*. Amsterdam, NL:

- John Benjamins, pp. 217–247.
- Rothman J (2015) Linguistic and cognitive motivations for the Typological Primacy Model (TPM) of third language (L3) transfer: Timing of acquisition and proficiency considered. *Bilingualism* 18(2): 179–190. DOI: 10.1017/S136672891300059X.
- Rothman J and Cabrelli Amaro J (2010) What Variables Condition Syntactic Transfer? A Look at the L3 Initial State. *Second Language Research* 26: 189–218.
- Sánchez L and Bardel C (2017) Transfer from the L2 in third language learning: A study on L2 proficiency. In: Angelovska T and Hahn A (eds) *L3 Syntactic Transfer*. Amsterdam, NL: John Benjamins.
- Santos H (2013) *Cross-Linguistic Influence in the Acquisition of Portuguese as a Third Language*. University of Illinois at Urbana Champaign.
- Schwartz B and Sprouse R (1994) Word order and nominative case in nonnative language acquisition: a longitudinal study of (L1 Turkish) German interlanguage. In: Hoekstra T and Schwartz B (eds) *Language acquisition studies in generative grammar*. Amsterdam, NL: John Benjamins.
- Schwartz B and Sprouse R (1996) L2 cognitive states and the full transfer/full access model. *Second Language Research* 12: 40–72.
- Slabakova R (2017) The scalpel model of third language acquisition. *International Journal of Bilingualism* 21(6): 651–666. DOI: 10.1177/1367006916655413.
- Slabakova R and García Mayo M (2015) The L3 syntax–discourse interface. *Bilingualism: Language and Cognition* 18(2): 208–226. DOI: 10.1017/S1366728913000369.
- Sorace A (2011) Pinning down the concept of interface in bilingualism. *Linguistic Approaches to Bilingualism* 1: 1–33.
- Tavakol M and Jabbari M (2014) Cross-linguistic influence in third Language (L3) and fourth Language (L4) acquisition of the syntactic licensing of subject pronouns and object verb

- property: A case study. *International Journal of Research Studies in Language Learning* 3(7): 29–42.
- Tsang WL (2009) The L3 acquisition of Cantonese Reflexives. In: Leung YI (ed.) *Third Language Acquisition and Universal Grammar*. Bristol, UK: Multilingual Matters, pp. 192–219.
- Tsimpli I and Dimitrakopoulou M (2007) The Interpretability Hypothesis: Evidence from wh-interrogatives in second language acquisition. *Second Language Research* 12(1): 7–39.
- Unsworth S (2007) L1 and L2 acquisition between sentence and discourse: Comparing production and comprehension in child Dutch. *Lingua* 117: 1930–1958.
- Vainikka A and Young-Scholten M (1996) The early stages in adult L2 syntax: additional evidence from Romance speakers. *Second Language Research* 12(2): 140–176.
- VanPatten B and Williams J (2015) *Theories in Second Language Acquisition*. Routledge: New York.
- Villegas Á (2014) *The role of L2 English immersion in the processing of L1 Spanish sentence complement/relative clause ambiguities*. Pennsylvania State University.
- Weinreich U (1953) *Languages in contact, findings and problems*. New York, NY: Linguistic Circle of New York.
- Westergaard M, Mitrofanova N, Mykhaylyk R, et al. (2017) Crosslinguistic influence in the acquisition of a third language: The Linguistic Proximity Model. *International Journal of Bilingualism* 21(6): 666–682. DOI: DOI: 10.1177/1367006916648859.
- White L (2003) *Second language acquisition and Universal Grammar*. New York, NT: Cambridge University Press.
- Wong KC (2011) Chi squared test versus Fisher's exact test. *Hong Kong Med J* 17(5): 427.

