Microvariation in Multilingual Situations: The Importance of Property-by-Property Acquisition[•]

Marit Westergaard
UiT – The Arctic University of Norway & NTNU Norwegian University of Science and Technology

Address
Department of Language and Culture
UiT The Arctic University of Norway
P. O Box 6050 Langnes
9037 Tromsø
Norway

Abstract
In this paper, I argue that L1, L2 and L3 acquisition are fundamentally the same process, based on learning by parsing. Both child and adult learners are sensitive to fine linguistic distinctions and language development takes place in small steps. While the bulk of the paper focuses on crosslinguistic influence in L2/Ln acquisition, I first briefly outline the Micro-cue Model of L1 acquisition (Westergaard 2009a, 2014), arguing that children build their I-language grammars incrementally, paying attention to small distinctions in syntax and information structure from early on. They are also shown to be conservative learners, generally not producing overt elements or performing movement operations unless there is positive evidence for this in the input, thus minimizing the need for unlearning. I then ask the question how this model fares with respect to multilingual situations, more specifically L2 and L3 acquisition. Discussing both theoretical and empirical evidence, I argue that, although L2 and L3 learners are different from L1 children in that they are not always conservative learners, they are also sensitive to fine linguistic distinctions, in that transfer/crosslinguistic influence takes place on a property-by-property basis. Full Transfer is traditionally understood as wholesale transfer at the initial state of L2 acquisition. However, I argue that it is impossible to distinguish between wholesale and property-by-property transfer in L2 acquisition on empirical grounds. In L3 acquisition, on the other hand, crosslinguistic influence from both previously acquired languages would provide support for property-by-property transfer. I discuss a few such cases and argue for what I call Full Transfer Potential (FTP), rather than Full (wholesale) Transfer, within the Linguistic Proximity Model (LPM) of L3 acquisition. Thus, rather than assuming that *everything does transfer*, I argue that *anything may transfer*.

Keywords
Micro-cue, economy, incremental learning, L1A, L2A, L3A, structural similarity, typological proximity, linguistic proximity, LPM, Full Transfer Potential

1. Introduction/Overview

In this paper, I start out by showing that children are sensitive to fine distinctions in the input from early on and that their non-target production is generally affected by economy/avoidance of complexity. Outlining the Micro-cue Model (Westergaard 2009a, 2014), I argue that the genetic endowment for language acquisition makes children able to parse the input, and based on this, they build the L1 grammar incrementally in small steps. I then extend this model to multilingual situations, arguing that L2/Ln acquisition is also step-wise learning by parsing. In this process, all previously acquired languages remain active and may influence the L2/Ln, which I refer to as Full Transfer Potential. I thus argue against wholesale transfer models (Schwartz and Sprouse 1996, Rothman 2015). Within the Linguistic Proximity Model (LPM, Westergaard, Mitrofanova, Mykhaylyk and Rodina 2017), I argue that L2/Ln acquisition is incremental and that transfer/crosslinguistic influence takes place property by property, (mainly) based on structural linguistic similarity.

2. Background: Parameters and microvariation

In the highly influential Principles and Parameters framework of Chomsky (1981, 1986), variation across languages was accounted for in terms of different settings of a number of innate macro-parameters, e.g., the head parameter (responsible for basic word order) or the pro-drop parameter (differentiating between languages that allow null subjects and those that do not). It was also argued that several (sometimes seemingly unrelated) properties would cluster together in the same parameter, and this would explain the speed and ease with which children learn language: After being exposed to just a few examples of the relevant input, children would set the corresponding parameter, and many linguistic properties would automatically fall into place, sometimes even without exposure. This is formulated by Pinker (1994: 112) in the following way: "All [children] have to learn is whether their particular language has the parameter value head-first, as in English, or head-last, as in Japanese. They can do that merely by noticing whether a verb comes before or after its object … Huge chunks of grammar are then available to the child, all at once, as if the child were merely flipping a switch to one of two possible positions."

However, the parameter approach has not lived up to the expectations from the 1980s and has also met with considerable criticism. The extensive work of Newmeyer (2005) shows, convincingly to my mind, that the parameter idea does not hold up with respect to the typological clustering that it was intended to account for, and that the empirical reality of the variation found among grammars is simply too large for "the hopeful vision of UG as providing a small number of principles each admitting of a small number of parameter settings [to be] workable." He also argues against the idea that the innate endowment for language should be used to explain typological generalizations in the first place, and instead posits a UG with principles only, supplemented by additional language-particular rules which should be constrained by these principles. Despite this kind of criticism, the term parameter has continued to be used in the field, although not necessarily with the original meaning. For example, for Biberauer and Roberts (2012), parameters are no longer innate, but emergent in the acquisition process. And as a result of the need to account for microvariation, their parameters are broken down into a hierarchy of four distinct types, dependent on the size of the class of elements undergoing the relevant process (macroparameters, mesoparameters, microparameters and nanoparameters).

Furthermore, a number of studies on syntactic variation have shown that there is considerable micro-variation *within* one and the same language variety. Work coming out of the many

dialect projects carried out over the last 15-20 years, e.g., ScanDiaSyn (Scandinavian Dialect Syntax), has shown that there is considerable variation both within and across dialects. Examples from Norwegian include verb second (V2) word order in some linguistic contexts and non-V2 in others (e.g., Vangsnes 2006, Westergaard 2009b), possessives either preceding or following the noun (e.g., Lødrup 2011, Julien 2005), and subjects and objects sometimes preceding and sometimes following adverbs or negation (e.g., Anderssen, Bentzen, Rodina and Westergaard 2010), etc. We take a brief look at only two of these variation phenomena here, Subject Shift and V2.

Subject Shift refers to the word order variation illustrated in (1)-(2), where the subject may either precede or follow adverbs or negation. The word order choice is dependent on a number of factors, the most important being the category of the subject, pronouns preceding and full DPs following negation (e.g., Westergaard 2011).

(1)   *Denne boka   har **ikke Peter** lest.*
        this   book.DEF has not   Peter   read
        'This book Peter hasn't read.'

(2)   *Denne boka   har **han ikke** lest.*
        this   book.DEF has he   not   read
        'This book he hasn't read.'

There is also considerable variation across Norwegian dialects with respect to V2 word order (Lohndal, Westergaard and Vangsnes forthcoming). Again, the choice is dependent on a number of factors, e.g., clause type or subject type, but here we only consider the initial element in wh-questions: If the wh-element is long, V2 is required, if the wh-element is monosyllabic, both word orders are allowed (and the variation dependent on information structure), and finally, if the initial element is *what*, the word order is virtually always non-V2 (see e.g., Westergaard 2009b).[1] This is illustrated in (3)-(5).

(3)   *Katti **kommer du**?*   (V2)

---

[1] This is a robust and relatively stable system that exists in the majority of northern and western dialects.

why come   you

'When are you coming?'

(4)   *Kor   du bor? / Kor   bor du?*   (V2/non-V2)

where you live / where live you

'Where do you live?'

(5)   *Ka   du   sir?*   (non-V2)

what you say

'What are you saying?'

Although V2 is traditionally assumed to constitute a parameter, the parameter cannot account for this variation that is found within a language or dialect. Furthermore, the factors affecting the choice of V2 or other word order phenomena vary across languages and dialects, e.g., clause type, information structure, category type, etc. As there seems to be little, if any, exceptionless clustering of these phenomena, this kind of variation must be *learned from the input*. In the next section, we take a look at how children acquire this kind of variation.

3. First Language Acquisition

3.1 General findings

As has been well known ever since the groundbreaking work of Brown (1973), young children's utterances are systematic and to a large extent target-like from early on, with non-target-consistent production generally constituting errors of omission rather than errors of commission (e.g., Snyder 2007). For the phenomena discussed in the previous section, one typically finds no overgeneralization across linguistic contexts with different word orders: For Subject Shift, Norwegian child data show that both word orders are attested from the earliest relevant utterances, with pronouns generally preceding and full DPs following negation; see Anderssen and Westergaard (2010). For V2 vs. non-V2 word order in wh-questions, it is also found that children pay attention to the distinction between different types of wh-elements from early on and produce the two word orders in target-appropriate contexts, with similar frequencies as in adult data (see e.g., Westergaard 2009c, 2014).

Given such early and target-consistent production for variation that must be learned from the input, an obvious question is why children should need to rely on innate parameters for

properties such as head directionality or pro-drop, for which there is both frequent and salient evidence in the primary linguistic data. In my view, such variation should also be learnable from exposure, and thus, L1 acquisition provides no argument or rationale for the existence of innate macro-parameters either.

Furthermore, parameter setting would predict (massive) overgeneralization in cases of input variation, which is generally not attested in syntax (but see below). In fact, we occasionally find the opposite, and Roeper (1999:175) states that 'there is widespread evidence for "undergeneralization" in child language.' Spontaneous corpus data typically show children producing less than what the input requires, e.g., less syntactic movement or fewer overt elements. Examples of the former are found in (6)-(8): lack of subject shift and V2 in Norwegian, and lack of subject-auxiliary inversion in English (Westergaard 2009c). Overgeneralization of such movement operations is virtually unattested, e.g., shifted DP subjects, V2 in embedded questions in Norwegian, or inversion in declaratives in English (Anderssen and Westergaard 2010, Westergaard and Bentzen 2007, Radford 1992). In previous work, I have argued that the most typical non-target production in child syntax is due to economy/avoidance of complexity; i.e., children will generally not develop a movement operation or add overt elements in their I-language grammars without positive evidence in the input for such phenomena (Westergaard 2009a, 2014).[2]

(6)   *nå   skal **ikke dem** sove.* (Ole, 2;3.15)
        now shall not   they sleep
        'Now they won't sleep.'

(7)   **der   Ann har* et.* (Ann, age 2;1.28)
         there Ann has one
        'There Ann has one,' Target: der har Ann et.

(8)   *Why **he can't** hit? (Adam, age 3;4.01)

---

[2] This does not mean that other errors do not occur in child language, e.g., verb movement in embedded declaratives (Waldmann 2008; Westergaard and Bentzen 2007) or medial wh-elements in complex wh-questions (Thornton 1990). However, such errors are possibly the result of misanalysis of the input, while non-target-consistent production that is due to economy seems to be the general case.

It may also be shown that this undergeneralization is systematic and occurs in small steps. Perhaps the best example of this comes from lack of inversion in child English, where some children have been found to produce inversion first with *be* and then with auxiliaries, and only with some wh-elements and not others (see e.g., Westergaard 2009c, where one child is found to produce inversion with *what* 96.4% and with *why* only 11.9%). Note that in all these cases, the children are making distinctions that are not found in the input, which means that their production could not be explained by simple input frequencies, as often argued in constructivist accounts of language acquisition (e.g., Tomasello 2003, Rowland and Pine 2003).

3.2 The Micro-cue Model

The Micro-cue Model was originally inspired by Fodor's (1998) idea of treelets as well as Lightfoot's (1999, 2006) cue-based approach to acquisition and change. A treelet is "a small piece of tree structure … that is made available by UG and is adopted into a learner's grammar if it proves essential for parsing input sentences" (Fodor 1998: 6). Similarly, a cue is *a piece of abstract syntactic structure*, crucially different from a trigger, which is a string in the input.[3] Lightfoot's cues are formulated in terms of major categories (e.g., V or DP), which means that they make the same predictions as macro-parameters, e.g., objects precede verbs or finite verbs are in second position, as in (9)-(10).

(9)   Cue for OV word order: $_{VP}$[DP V]

(10)   Cue for V2 word order: $_{CP}$[XP $_C$V...]

But as shown above, many languages display considerable (micro-)variation with respect to these phenomena, and children are sensitive to much finer distinctions than what is expressed in (9)-(10). Thus, in Westergaard (2009a, 2014) I argue that the cues must also specify the linguistic context in which this constellation applies, e.g., clause type, subject or verb type, etc. This means that V2 is not one big rule (or parameter), but several smaller rules, or *micro-*

---

[3] Since the term cue refers to a small piece of the grammar, it is in a way a misnomer, sometimes leading to confusion. Unfortunately, this also applies to the term micro-cue. I have nevertheless decided to keep these terms, for consistency with previous work.

cues. Micro-cues for the two phenomena discussed above, Subject Shift and V2, could be formulated as (11)-(12), the latter being only one of several rules for V2.[4]

(11) Micro-cue for Subject Shift:

$$_\text{IntP/DeclP/PolP}[(XP)\ V\ _\text{InTopP}[\ DP_{[-FOC]}\ AdvP/Neg\ ...]$$

(12) Micro-cue for V2 in questions with monosyllabic *wh*-elements:

$$_\text{IntP}[\ _\text{Int°}[wh]\ _\text{TopP}[\ _\text{Top°}[V...\ XP_{[+FOC]}...\ ]]]$$

The Micro-cue Model is a generative approach which assumes a role for a genetic endowment for language, which we may refer to as UG and which may well contain categories, features, principles, and constraints, as often assumed in the literature, but crucially no parameters. This genetic endowment enables children to *parse* the linguistic input, thereby producing representations of abstract structures and gradually building micro-cues in their I-language grammars, from which they can generate new sentences. This means that, unlike Lightfoot's cues and Fodor's treelets, micro-cues are not provided by UG; instead, they develop in the acquisition process and are part of the grammar of a *specific language*.

Thus, the Micro-cue Model argues that children start out with small pieces of structure and build up their I-language grammars incrementally. Nevertheless, adult language is clearly more than accumulated knowledge of specific rules and constructions, which means that generalization must take place in the developmental process, based on positive evidence in the input. According to the Micro-cue Model, this generalization is not an across-the-board process, but takes place as step-by-step development. Importantly, the steps are small, typically involving the addition of a sub-category, a lexical item, or feature to an already existing micro-cue (see Westergaard 2014 for examples). That is, children do not immediately extend a process or rule to a major category (e.g., all nouns, all verbs, or all *wh*-elements), but only within a small class or subcategory. Nevertheless, syntactic *over*-generalizations are

---

[4] The (somewhat complex) notation is not important here, but rather the principle that fine distinctions must beincluded in the micro-cue, since they are part of speakers' knowledge of the language. Note that the micro-cues have categories and structure, which means that they are different from chunks and frames in usage-based accounts of L1 acquisition. The micro-cue in (11) specifies that DP subjects that are given information (specified as [-FOC] here), typically pronouns, move to a position in front of adverbs or negation in non-subject-initial clauses, while the micro-cue in (12) specifies that in wh-questions with monosyllabic wh-elements (heads), the verb moves to a head position called Top°, but only when the subject is new information, i.e., [+FOC].

occasionally found in child language, e.g., well-known overextensions of causative constructions, dative alternation, or locative alternation (Bowerman 1988), see also footnote 2. But an important fact about these overgeneralizations is that they do not affect a major category such as all verbs or all nouns, but only take place across a class or subcategory, sometimes only one lexical element. The obvious advantage of development in small steps is that the need to backtrack in the learning process will be minimal.

The L1 acquisition process is illustrated in Figure 1: The child is equipped with a genetic endowment for language acquisition (UG), and this enables the child to parse the input and build micro-cues (pieces of syntactic structure, referred to as M here) in the L1 grammar, which is initially quite small (the first image). Throughout the language acquisition process, more micro-cues are learned and the L1 grammar is simply expanded, as shown in the second image.
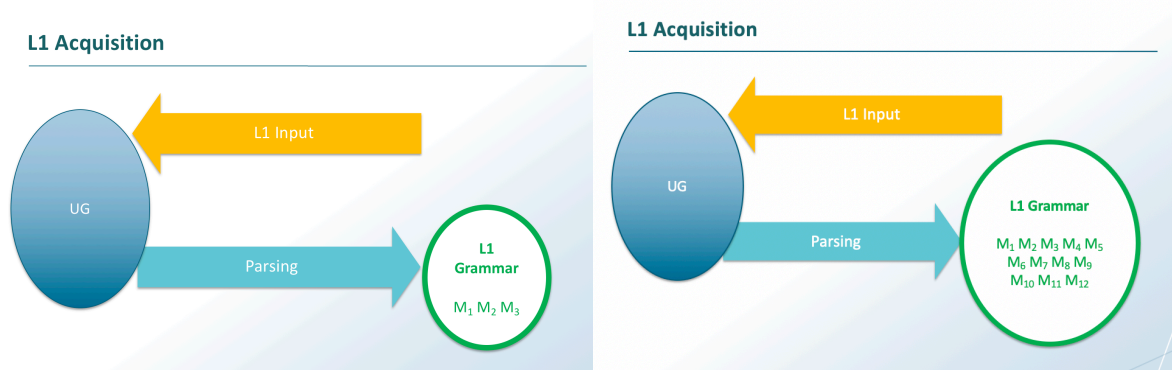

Figure 1: The process of L1 acquisition according to the Micro-cue Model: Parsing and building a grammar in small steps.

To summarize, in this section I have outlined the Micro-cue Model, which is a non-parametric generative approach to language acquisition arguing that children are sensitive to fine distinctions in syntax and information structure from early on. According to this model, L1 acquisition is learning by parsing, not setting (macro-)parameters, and children build their I-language grammars incrementally. Furthermore, children are generally conservative learners, not making massive overgeneralizations, and the generalizations they do make are across small categories, thus minimizing the need for unlearning.

The main question addressed in this article is whether the Micro-cue Model may be extended to multilingual situations. That is, are bi- and multilingual children and adults conservative learners, like L1 children? Are they also sensitive to fine syntactic distinctions, learning in a step-wise fashion? And when there is transfer/crosslinguistic influence in L2 and L3/Ln acquisition, does this apply to the full grammar in one fell swoop or does it take place property by property?[5] In the next two sections, I discuss these issues in L2 and L3/Ln acquisition respectively.

4. L2 Acquisition

4.1 Conservative learning, microvariation and transfer/crosslinguistic influence

A brief look at studies of L2 acquisition will reveal that, unlike L1 children, L2 learners are generally not conservative. While monolingual children typically do not overgeneralize movement operations, this has often been found in L2 learner data, e.g., for verb movement in embedded clauses in German (Clahsen, Meisel and Pienemann 1983, Clahsen and Muysken 1986). A recent example is found in Anderssen, Bentzen, Busterud, Dahl, Lundquist and Westergaard (2018), which is an L2/Ln study of the acquisition of Subject Shift in Norwegian. Recall that Subject Shift entails pronominal subjects moving across adverbs/negation, while DP subjects remain in situ; cf. examples (1)-(2) above. The Anderssen et al. (2018) study is an Acceptability Judgement Task, testing pronominal and DP subjects in both word orders (S-Neg and Neg-S). The results show that Norwegian natives clearly prefer pronominal subjects preceding and DP subjects following negation, as expected. The L2/Ln learners (who had a variety of L1s/previously acquired languages), on the other hand, prefer *all* subjects in the high position (thus displaying a behavior that is the opposite of L1 children, who had an early preference for the low position). That is, the adult L2/Ln learners over-accept movement, indicating that they are not conservative learners.

Does this mean that L2 acquisition is *fundamentally* different from L1 acquisition (see Bley-Vroman 2009 for an overview)? Not necessarily. Much of the difference between L1 and L2 acquisition can be explained by the obvious fact that L2 learners already know another language, which may (and typically will) interfere in the process. Furthermore, while L1 language development is mainly affected by economy/avoidance of complexity (cf. section

---

[5] While I adopt the original definition of crosslinguistic influence as an overarching concept which includes a number of different processes (Sharwood Smith 1983), I nevertheless use the terms transfer and crosslinguistic influence interchangeably in this paper; cf. section 5.4 for further discussion.

3.1), older learners, who presumably more often find themselves in situations where the communication requirements extend their current knowledge of the L2, may be more sensitive to frequency in the input (and for Norwegian Subject Shift, the L2 learners' preference represents the more frequent word order).

Furthermore, such findings do not mean that L2 learners are not sensitive to microvariation, which can be shown in the way the L1 influences the L2. While many early studies of generative L2 acquisition focused on the transfer of L1 parameter settings, results often showed more selective transfer (or unlearning), e.g., White (1985) on the null subject parameter, finding that Spanish learners of L2 English treated the various properties of this parameter differently, or White (1990) on the verb movement parameter, showing that French learners of L2 English make a distinction between the word order of declaratives and questions. That is, these studies failed to show the expected clustering effects; instead, learners seemed to be transferring certain fine distinctions from the L1 into the L2.[6] Such results have also been attested in Norwegian learners of English, where unlearning V2 word order has been found to be harder in subject-initial than non-subject-initial clauses (Westergaard 2003, Jensen 2017). Furthermore, Westergaard (2003) found that the distinction between V2 and non-V2 in Norwegian dialects was carried over into L2 English, as young learners were significantly more likely to accept sentences such as (13), which are grammatical in Norwegian, than sentences such (14), which are ungrammatical in the L1; cf. examples (3)-(5) above. That is, the learners were not transferring a parameter setting for V2, but the fine-grained variation found in the L1.

(13)   *Where the ball is? (grammatical in the L1)

(14)   *What color the ball is? (ungrammatical in the L1)

During the 1990s, L2 acquisition research was heavily focused on defining the initial state, and several models of transfer/crosslinguistic influence were developed: Full Transfer Full Access (Schwartz and Sprouse 1996), Minimal Trees (Vainikka and Young-Scholten 1996), Weak Transfer (Eubank 1993/94), and the Initial Hypothesis of Syntax (Platzack 1996),

---

[6] Others have of course also argued against parameters in L2 acquisition, perhaps most notably Lardiere (2009) in work on the Feature Reassembly Hypothesis.

arguing for different positions with respect to what transfers from the L1 into the L2, ranging from transfer of the complete L1 grammar to no transfer at all. Since then, numerous studies have provided considerable and convincing evidence for full transfer and generally shown that partial transfer models do not hold up against empirical evidence (e.g., Grüter 2006). It is nevertheless a pertinent question what full transfer actually means. Does it mean that the complete L1 grammar transfers in one fell swoop, or does it mean that the whole grammar may transfer in a step-wise fashion, one property at a time?

4.2 Wholesale or property-by-property transfer

In the seminal article by Schwartz and Sprouse (1996: 40-41), the Full Transfer Full Access (FT/FA) model is defined as in (15a), with the extra specification in (15b):

(15) a. '[T]he initial state of L2 acquisition is the final state of L1 acquisition.

b. '[T]he entirety of the Ll grammar (excluding the phonetic matrices of lexical/morphological items) … immediately carr[ies] over as the initial state of a new grammatical system on first exposure to input from the target language (TL). This initial state of the L2 system will have to change in light of TL input that cannot be generated by this grammar; that is, failure to assign a representation to input data will force some sort of restructuring of the system …'

According to FT/FA, L2 acquisition must be a two-step process, consisting of (i) copying of the L1 grammar onto the L2, and (ii) restructuring of the transferred grammar due to parsing failure of further L2 input. I have illustrated this in Figure 2 (where the grammar is seen as a collection of micro-cues): The first image shows the initial state of an L2 learner (equipped with UG and an L1 grammar) who is exposed to L2 input and immediately makes a copy of the complete L1 grammar, labeling it L2.[7] At the subsequent stage (image 2), the L2 learner has two identical grammars (illustrated here by the same color for all the morphosyntactic rules). With further exposure to L2 input that deviates from the L1, the L2 grammar will need to be restructured, illustrated by a change of color for some micro-cues in the second image.

---

[7] Ideally, UG and the L1 should also be visible on the right-hand side of each image. However, due to space limitations, the right-hand side only shows the change. This also applies to the remaining figures in this paper.
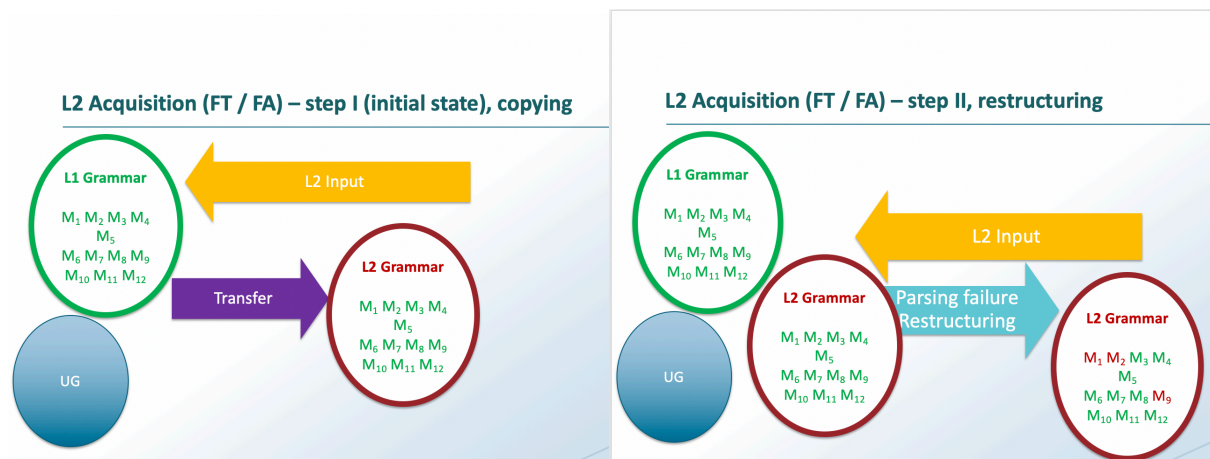
Figure 2: L2 acquisition according to FT/FA: Copying of the L1 at the initial state (image 1); subsequent restructuring based on parsing failure (image 2).

It seems unclear whether the idea of a copy has been intended and interpreted as an abstraction or an actual neurobiological process. As far as I am aware, there is no description in the literature detailing what neuroanatomical mechanism would be responsible for this copying process and what would trigger it. Nevertheless, the intention of Schwartz and Sprouse (1996) seems to have been that the brain does make a real copy of the L1. This is most recently claimed in Rothman et al. (2019: 151), which refers to personal communication with Rex Sprouse (March 3, 2018), stating "that the authors are clear about what full transfer means, namely a full copy of L1 grammatical representations."[8]

However, it does not seem to be the case that full transfer has always been *interpreted* as a literal copy, as the following quote from White (2003: 68) indicates an interpretation of copying more as an abstraction (boldface added by me).

(16)   The initial state in L2 acquisition is the L1 steady state grammar in its entirety. One
         needs to think of this as **in some sense** a copy (or clone) of the L1 grammar, a copy
         which can be modified without affecting the original.

The quote from White also indicates what the original rationale behind full transfer was: The copy was proposed in order to make sure that the restructuring, which at the time was

---

[8] Further communication with both authors (e-mail September 16, 2019) reveals that by copies they mean "*models* of mentally represented *knowledge*" and that they "remain strictly agnostic about the neuro-chemical structures and/or mechanisms" involved in wholesale transfer.

considered mainly to be parameter resetting, would not affect the original parameter settings of the L1.

In my view, copying is unnecessary: First, a substantial amount of research has shown that simultaneous bilinguals are generally able to keep their two languages apart in the acquisition process, with only minimal leakage between them (e.g., Meisel 1989, De Houwer 2005; see also Hulk and Müller 2000 on contexts for crosslinguistic influence in bilingual acquisition). Changes in one grammar generally do not affect the other grammar. If there is a mechanism making this possible in simultaneous bilingual acquisition, this same mechanism should also be available in sequential acquisition (L2): That is, acquisition of one language should generally proceed without affecting the other, whether the languages are acquired simultaneously or sequentially, and it should therefore be unnecessary to make a copy of one of them to make sure that it remains stable. Second, there is increasing evidence that the L2 *may* in fact influence the L1 (see e.g., the articles in Köpke, Schmid, Keijzer and Dostert 2007). Furthermore, over the last two decades, a considerable body of work on bilingualism and executive control has shown that both languages of a (simultaneous or sequential) bilingual are always active and that the language not in use must to some extent be inhibited (see e.g., Kroll, Dussias, Bogulski and Valdes Kroff 2012, Blumenfeld and Marian 2013). Recent psycholinguistic studies also indicate that the L1 and the L2 share the same networks for language processing (see e.g., Cunnings 2017 or Del Maschio and Abutalebi 2019 for overviews). This means that in order to account for L1 influence on the L2, it is no longer necessary to posit that a copy of the L1 is made, since the L1 can always be accessed directly.

An unclear aspect of the FT/FA model is the reference to "initial state" and "on first exposure." It is difficult to imagine that what was meant was that the brain would make copies of the L1 for every language that one is ever (briefly) exposed to (see Sharwood Smith and Truscott 2006, 2008, who question what would trigger such copying). For this reason, I find it more plausible to interpret full transfer as a metaphor: That is, when a learner is exposed to an L2, *the entirety of the L1 grammar remains available for parsing the L2*, from initial exposure and throughout. But on this view, the influence from the L1 should take place in a step-wise fashion, only when there is a need for a particular structure.

It is also relevant to make a comparison to L1 acquisition here: If our genetic endowment for language learning makes us able to make fine distinctions in L1 acquisition (cf. section 3),

why should we not be able to do so also in L2 acquisition? It seems to me that a view of L2 acquisition as copying and restructuring would mean that it is fundamentally different from L1 acquisition. This could of course be the case, but much recent work typically argues that L1 and L2 acquisition are *not* fundamentally different, in that there is no linguistic property that seems to be unlearnable in an L2, even when the age of onset is beyond puberty (e.g., Montrul and Slabakova 2003 on the preterite-imperfective contrast in Spanish or Donaldson 2011 on left dislocation in French). Related to this last point is the question why the brain would create a context for massive unlearning in L2 acquisition, when this is avoided in L1 acquisition (cf. section 3.2).

4.3 Full Transfer Potential: Learning by parsing

Based on the argumentation in the previous section, I would like to propose what I refer to as Full Transfer Potential, meaning that *anything may transfer*, not that *everything does transfer*. This is different from the partial transfer models of the 1990s in that there is no restriction on the kind of linguistic property that may transfer. However, it also differs from FT/FA in that there is no wholesale transfer (copying) at the initial state; instead, *transfer takes place as a result of parsing*, property by property. Full Transfer Potential can be defined as in (16):

(16)   **Full Transfer Potential (FTP)**
- The initial state of the L2 is the L1 plus the genetic endowment for language acquisition (as in FT/FA)
- The complete grammar of the L1 remains active (so no need to make a copy of it)
- L2A (like L1A) is *learning by parsing*:[9]
    - If there is an identical or similar micro-cue available in the L1, this will be used to parse the L2
    - If there is no similar structure in the L1 grammar, the learner resorts to UG
- Any property of the L1 may transfer
- Crosslinguistic influence takes place property by property

---

[9] In this respect, this model is very similar to the MOGUL (Modular Online Growth and Use of Language) framework of Truscott and Sharwood Smith (2004), Sharwood Smith and Truscott (2014), which is a processing-based model of language acquisition allowing for gradual development and optionality in terms of different activation of representations from the L1 and L2.

Learning by parsing (in comprehension as well as production) works in the following way: When exposed to L2 input, learners will try to parse it, first scanning the L1. If an identical structure (micro-cue) exists in the L1, parsing will be facilitated, and the result of this is positive/facilitative transfer. Non-facilitative transfer, on the other hand, occurs when the learner mis-analyzes L2 input that bears some resemblance to structures in the L1, assuming it is identical (when in fact it is not) and uses the L1 structure to parse it. Parsing produces a representation of the L2 structure, which is used to build the new grammar incrementally. At the beginning of the learning process, the representations are unstable, but with increased input and use, they will stabilize (if the input continues to confirm the similarity between the L1 and the L2); alternatively, they will be replaced by representations that correspond more closely to the L2 structures (if continued input reveals dissimilarities between the two languages). If there is no identical or similar structure in the L1, the learner will resort to UG for parsing. Depending on the nature of the property, this may be easy or hard, and may consist of several steps, but should eventually lead to learning the L2 structure (without transfer), e.g., learning grammatical gender in German if the L1 is English (of course, German gender is not in UG, but the learner uses the innate endowment to parse the relevant structures, as in L1 acquisition). In production, non-facilitative transfer should be more frequent than in comprehension. In cases where the target representation is lacking (or too weak) in the current version of the L2, learners will typically use a structure from their L1 when creating a representation for production. This may result in facilitative transfer if the two languages happen to be identical with respect to the property in question, but will often result in non-facilitative transfer. Again, if positive evidence against a non-facilitatively transferred structure is frequent and salient in the input, the transferred representation may be short-lived in the L2 grammar, whereas if the evidence is infrequent or perhaps even non-existent, the transferred L1 structure may persist for an extended period of time and perhaps never be target-consistently acquired (by some learners).

Of course, many other scholars assume learning by parsing, also in L2 acquisition; see e.g., Dekydtspotter and Renaud (2014). FT/FA also assumes a role for parsing, as restructuring is argued to take place as a result of parsing failure (cf. Figure 2). This means that structures that are identical in the L1and the L2 do not have to be learned, since they are copied onto the L2 at the initial state. However, under FTP, parsing is the only mechanism for language learning, applying to all properties of the L2, whether they are identical to the L1, different from the L1, or non-existent in the L1 (and thus not transferred). In cases where the L2 grammar has

developed non-target-consistent structures (either due to non-facilitative transfer or imperfect learning), restructuring may take place. This is also a step-wise process, based – not on parsing failure – but actually on the learner being able to parse L2 input in a more target-consistent way. Since native speakers have been found to be able to produce representations for ungrammatical sentences (Ivanova, Pickering, Branigan, McLean and Costa 2012), L2 learners must also (at least to some extent and at some point in the acquisition process) be able to parse input that is not licenced by their L1 and/or current version of their L2. This licencing failure results in the new representation gradually becoming more productive and taking over the previous one in the L2 grammar of the learner, as in the multiple grammars approach of Amaral and Roeper (2014). Thus, this is a gradual process, reflecting the variability often found in L2 grammars.

The images in Figure 3 illustrate the process of L2 acquisition under FTP: The learner is exposed to L2 input and uses available resources to parse it. Since there is only one previously acquired language, the L1 is the only possible source of crosslinguistic influence, and whenever there is some similarity between the L2 and the L1, the learner will use existing structures (micro-cues) from the L1 to parse the incoming signal and start building the new grammar. But crucially, there is *no copy* made of the entirety of the L1 grammar. When there are clear and salient differences between the L1 and the L2, this should be easy to learn, meaning that even very early in the acquisition process, there may be differences between the L1 and the L2, illustrated by two different colors of the micro-cues in the first image. Note that the L2 grammar at this early stage is small, and as L2 acquisition proceeds, there is no fundamentally different process taking over. All that happens is that the L2 grammar is gradually expanded, as we see in the second image. Thus, L2 acquisition is learning by parsing, just like L1 acquisition (cf. Figure 1).
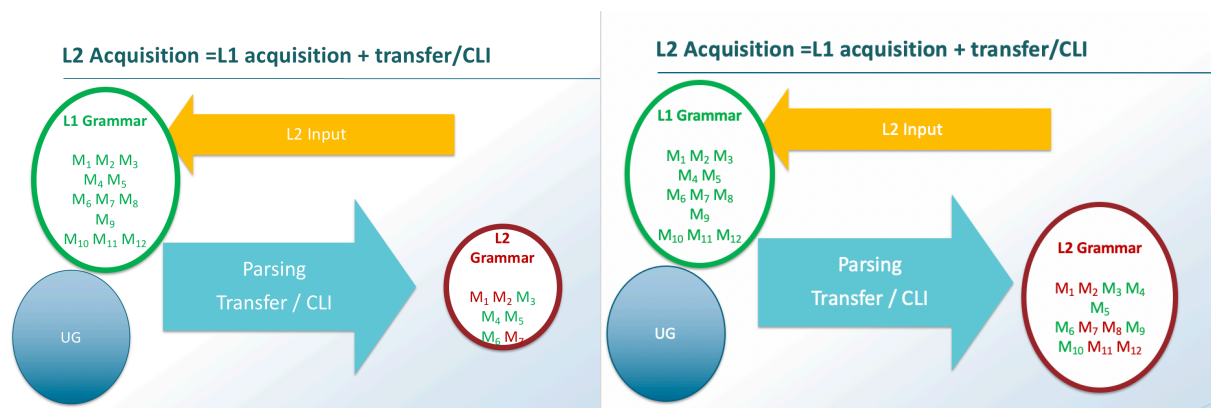
Figure 3: L2 acquisition according to Full Transfer Potential: Parsing and property-by-property transfer; building a grammar in small steps.

An important question is whether it is possible to distinguish between wholesale and property-by-property transfer in L2 acquisition: Whenever we find evidence for complete transfer, we cannot know whether everything was transferred at the initial state or whether transfer occurred as the need arose. This means that it should be virtually impossible to distinguish between the two, using linguistic data. However, it could be possible in L3 acquisition, where there are two possible sources of crosslinguistic influence. That is, if we find an effect of *both* previously acquired languages at an early stage, this would support the claim that transfer takes place property by property. We turn to L3 acquisition in the next section.

5. L3 Acquisition
5.1 Brief overview
The field of formal approaches to L3 acquisition has sparked considerable theoretical interest over the last two decades. I will only briefly mention a few relevant issues here, as overviews of this relatively new field may be found in numerous publications, most recently in Rothman et al. (2019). In addition to studies finding transfer from the L1 on the L3 (e.g., Jin 2009, Na Ranong and Leung 2009, Hermas 2010, 2014), several models have been developed, arguing for a variety of factors responsible for crosslinguistic influence in L3 acquisition. These include the Cumulative Enhancement Model (CEM), which contends that L3 acquisition is cumulative and that both previously acquired languages may positively affect the L3 (Flynn, Foley and Vinnitskaya 2001, Berkes and Flynn 2012), and the L2 Status Factor (L2SF), which claims that the L2 will be the primary source of influence in L3 acquisition, especially at early stages (e.g., Bardel and Falk 2007, 2012, Falk, Lindquist and Bardel 2015, Bardel and Sanchez 2017), generally due to similarities between the L2 and the L3 with respect to metalinguistic knowledge and storage in declarative memory (Ullmann 2001, Paradis 2009).

Neither the CEM nor the L2SF argues for wholesale transfer. The idea of making a copy of the complete grammar of a previously acquired language has re-surfaced in the Typological Primacy Model (TPM, Rothman 2011, 2013, 2015, González Alonso and Rothman 2017), formulated in Rothman et al. (2019: 24) as "reduplication of a representation from previously acquired linguistic representations, as an initial hypothesis for a given domain (literally, a

copy)." Two newer models of L3A, the Scalpel Model (Slabakova 2017) and the Linguistic Proximity Model (LPM, Mykhaylyk et al. 2016, Westergaard et al. 2017) argue explicitly against wholesale transfer. This is most clearly expressed by Slabakova (2017: 3) in the following way: "… the grammars already acquired act with a scalpel-like precision to extract the L1 or L2 options relevant to the acquisition task at hand. Therefore transfer is selective and works property by property."

5.2 Wholesale transfer in L3 acquisition

To my knowledge, the first formal model of L3 acquisition was the Interlanguage Transfer Hypothesis (ITH), first proposed by Leung (1998, 2003), although the term was coined by Jin (2009). Relying on the FT/FA hypothesis of Schwartz and Sprouse (1996), Leung (2003: 199) argues for a typology-based model with full transfer at the initial state of L3 acquisition, formulated in the following way: "the Ln initial state is the steady state of a previously acquired (inter)language which is typologically closest to Ln." However, the idea of full transfer is abandoned already in Leung (2005), where evidence from a study of L3 French by Cantonese-English bilinguals is used to argue for partial transfer in L3 acquisition (but full transfer in L2 acquisition).

Rothman (2011: 112) makes a proposal for L3 acquisition which is similar to the claim in Leung (2003), i.e., that "[s]yntactic properties of the closest (psycho)typological language, either the L1 or L2, constitute the initial state hypotheses in multilingualism, whether or not such transfer constitutes the most economical option." In later work, the TPM is refined in several ways: First, the initial state is replaced by the concept of *initial stages*, a somewhat prolonged period allowing the parser to evaluate which of the two previously acquired languages to choose for copying (Rothman 2013, 2015). Second, typological primacy is replaced by a four-level hierarchy of linguistic cues for the parser to evaluate in terms of similarity to the previously acquired languages: lexicon, phonology, morphology, and syntax. That is, the parser will start out evaluating the lexicon, and if there is no indication that the L3 is more similar to one of the previously acquired languages at this level, the parser will move to the next level, and so on, until the parser decides which of the two languages to make a copy of. This means that in those cases where there is no obvious typological similarity between the L3 and the previously acquired languages (at the lexical level), the TPM predicts that some linguistic/structural similarity will be the decisive factor. According to the TPM, the L3 acquisition process is very similar to L2 acquisition within the FT/FA model (cf.

Figure 2). I illustrate this in Figure 4, where L3 acquisition is shown as a two-step process, wholesale transfer as copying of one of the previously acquired grammars at the initial stages (first image) and subsequent restructuring based on continued L3 input as a result of parsing failure (second image). Thus, at the initial stages, the L3 is identical to one of the previously acquired grammars (which happens to be the L2 in Figure 4), while later stages see a gradual restructuring of this grammar.
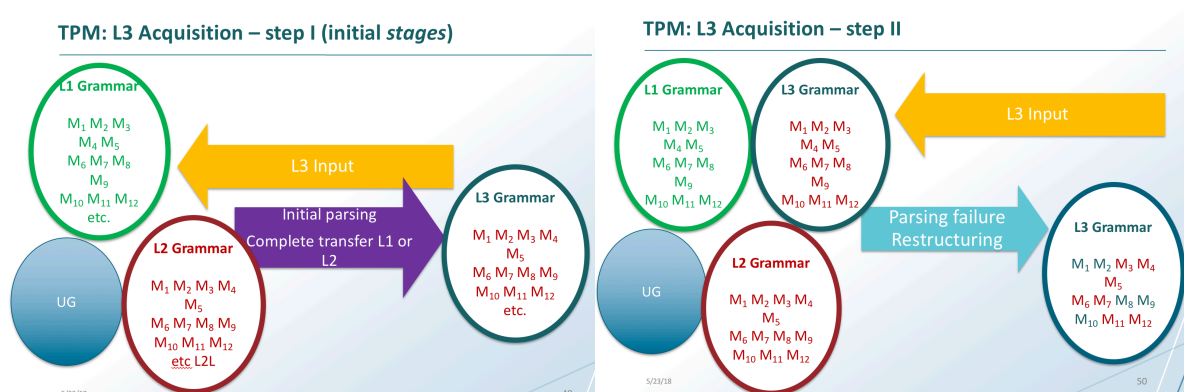

Figure 4: L3 acquisition according to the TPM: Initial parsing, copying of one of the previously acquired languages (image 1); subsequent restructuring of the L3 based on parsing failure (image 2).

Recent work has seen further refinements of the model. Rothman et al. (2019) allow for initial surface influence as well as secondary transfer taking place property by property (before and after wholesale transfer), and finally, they speculate that L4 acquisition may proceed property by property, as "L3 experiences of non-facilitation might very well mean that full transfer will be disregarded as a viable option when the mind is an experienced multilingual one, meaning L4 acquisition and beyond" (Rothman et al. 2019: 157). Thus, the TPM must assume a number of principled distinctions that seem to be motivated by the idea of wholesale transfer: distinctions between the initial state, the initial stages and later stages (the initial stages singled out as the stage disallowing property-by-property transfer), between what is referred to as representational transfer (copying) and crosslinguistic effects (CLE), the latter referring to temporary bleeding of one language into another in processing (more on this below), and possibly also between L2/3 acquisition on the one hand and L4/Ln acquisition on the other. Moreover, the TPM has been reduced in scope to a model of the initial stages, not a model of L3 acquisition, which does not make predictions for later stages, except developmental stages that derive directly from wholesale transfer, cf. González Alonso and Rothman (2017).

The motivation for wholesale transfer in L3 acquisition is different from the original rationale in L2 acquisition, which was the need to prevent restructuring from affecting the L1 (cf. section 4.2). For the TPM, wholesale transfer is explained by an economy consideration, where transfer in one fell swoop is argued to be less costly than property-by-property transfer. While recognizing that a substantial body of psycholinguistic research has shown that all languages of a bi- or multilingual are activated simultaneously, Rothman et al. (2019: 158) refer to the fact that "the same research indicates that there are differences in relative activation whereby grammars that are not needed contextually are suppressed." Thus, it is claimed that "[i]f L3 transfer were to occur on a property-by-property basis, maintaining activation of the L1 and L2 at a sufficiently high level for comparison over the extended period of acquisition would be too costly for a cognitive system that is hardwired to conserve and maximize resources."

However, cognitive economy is an elusive concept which can be (and has been) used in various ways in linguistics. Thus, it is not at all clear that transferring a lot once is cognitively more economical than transferring a little many times. In first language acquisition, children seem to follow an economy principle that is in line with the latter type of development; i.e., they do not (over-)generalize a learned pattern across-the-board to major categories, but extend their grammars in small steps, across classes or subcategories (cf. section 3.2). Moreover, higher activation of one of the previously acquired languages (as proposed in the above quote), which is of course highly plausible given what is currently known about the brain, cannot be responsible for wholesale transfer understood as copying. Moreover, it remains a mystery why the alleged principle of cognitive economy is only active at a particular stage of L3 acquisition, when the rationale for its existence is argued to be that it will ensure a less costly process "over the extended period of acquisition" (cf. the above quote).

Furthermore, the timing of this complete copying continues to be a thorny issue, even with the modification that wholesale transfer takes place at what is referred to as the initial stages. It is unclear how the initial stages should be identified, as the definition offered in González Alonso and Rothman (2017: 688) is simply that the initial stages are "the period in which structurally driven wholesale transfer from the L1 or the L2 takes place." This also means that it becomes virtually impossible to test the wholesale transfer claim empirically, also in L3

acquisition: Since both initial surface influence and residual transfer may take place property by property, any finding that shows an effect of both previously acquired languages in an L3 setting could be explained as having taken place either before or after wholesale transfer.

5.3 The Linguistic Proximity Model

Like the TPM, the Linguistic Proximity Model (LPM) emphasizes the structural similarity between the languages involved in L3 acquisition. This means that the order of acquisition of the previously acquired languages should be unimportant, and as a consequence, relevant studies could include simultaneous bilinguals who do not have a clear L1 and L2. FTP applies (cf. section 4.3), meaning that *the entirety of both previously acquired languages remain available for parsing the L3.* In Westergaard et al. (2017: 670), the model defines L3/Ln acquisition in the following way:

(17)   L*n* acquisition involves incremental property-by-property learning and allows for both facilitative and non-facilitative influence from one or both previously acquired languages. Crosslinguistic influence occurs when a particular linguistic property in the L*n* input reveals abstract structural similarity with linguistic properties of the previously learned languages.

The LPM bears strong resemblance to the Scalpel Model (Slabakova 2017), in that it rejects the notion of wholesale transfer. Crosslinguistic influence is argued to take place property by property, reflecting the ability of L3 learners to make fine morphosyntactic distinctions, similar to L1 and L2 learners. Thus, transfer could be from either or both of the previously acquired languages. This means that it is perfectly possible for one language to be the major source of transfer, in some cases perhaps even the only source at an early stage, if the L3 is very similar to one of the previously acquired languages.[10] This means that the model does not recognize any special status of the initial stages, except that learners may naturally rely

---

[10] Rothman et al. (2019: 108) argue that the collective results of nine studies of Spanish-English learners of Brazilian Portuguese (BP) as an L3 indicate that transfer must be wholesale, as influence has always been shown to come from Spanish. However, given the obvious similarities between Spanish and BP, it is not surprising that these early L3 learners would activate their knowledge of Spanish when asked to provide grammaticality judgements in BP (as in e.g., Rothman 2010). According to the LPM, this would simply mean that they are parsing BP input using their Spanish grammar, not *a copy of* the Spanish grammar.

more on previously acquired languages early in the L3 acquisition process. The LPM thus attempts to account for any stage of L3 acquisition.

As stated in (17), the main factor responsible for crosslinguistic influence is abstract structural similarity. However, Westergaard et al. (2017) also acknowledge that L3 learners may often be influenced by surface typological (lexical) similarity at early stages, typically before they are able to process the sometimes very complex syntactic structure of the incoming L3 input. This should especially be the case in situations where one of the previously acquired languages bears close surface resemblance to the L3, e.g., where English and a Romance language are the previously acquired languages and another Romance language is the L3. This modification of the LPM should make the model able to account for the many studies that show surface typological transfer at an early stage, also in situations where transferring from the other language would have been felicitous, e.g., Rothman and Cabrelli Amaro (2010). Note that this modification does some of the same job as the hierarchy of the TPM (lexicon, phonology, morphology, syntax), without being tied to a strict order. Some syntax may be complex, but other syntactic properties may be quite simple and salient and therefore noticed by learners much earlier, e.g., basic word order such as SVO vs. SOV. Despite this modification for early stages of L3 acquisition, the LPM "maintain[s] that as the exposure to an L3 grows, the role of overall typological proximity should decrease, while the role of more abstract structural similarities should increase" (Westergaard et al. 2017: 677).

According to the LPM, L3 learners have access to all previously acquired linguistic knowledge, at all stages of acquisition, as shown in much recent psycholinguistic research (see references above). Furthermore, just like L1 and L2 acquisition, L3 acquisition is learning by parsing (cf. section 4.3). Thus, when a learner is exposed to L3 input, s/he uses the abstract I-language grammars of the previously acquired languages to parse the incoming signal. As a result of this parsing, the L3 learner gradually builds a new grammar. This is illustrated in Figure 5, where the first image shows how parsing L3 input may activate both previously acquired languages (called L_A and L_B here), and how this results in building a grammar incrementally. There is no second step – the process simply continues and the result is a gradually larger grammar (with transferred as well as learned structures), as seen in the second image. This means that L3 acquisition is in principle the same process as L1 and L2 acquisition (cf. Figures 1 and 3), except that the L3 learner has more resources available at the outset of learning.
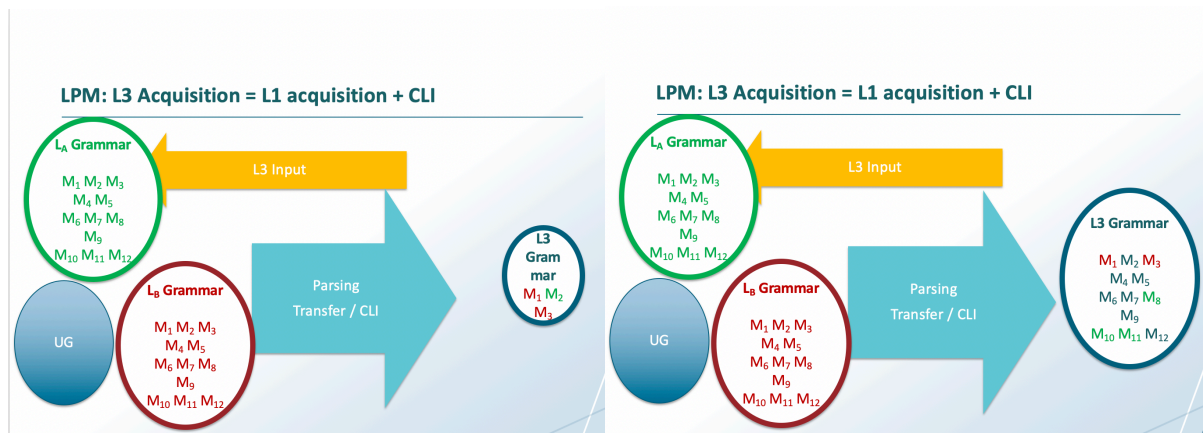
Figure 5: L3 acquisition according to the LPM: Parsing and property-by-property transfer from one or both previously acquired languages; building a grammar in small steps.

This view of acquisition means that the LPM does not make a fundamental distinction between transfer and crosslinguistic influence/effects, the way the TPM does, the former referring to copying and the latter to transient processing effects. Rothman et al. (2019) base this distinction on a traditional definition of competence and performance in generative theory (Chomsky 1965: 4) and argue that transfer takes place at the level of competence, while crosslinguistic effects are performance phenomena. However, the competence/performance distinction has been used in various ways in linguistic theory (see Phillips 1996: 275-277 for four different interpretations), and it is not entirely clear what is meant here. Moreover, it is relatively widely acknowledged in current psycholinguistic research that, despite certain misalignments between grammar and processing, one cannot make a complete separation of the grammar and the parser, since the parser operates on representations in the grammar (see e.g., Lewis and Phillips 2015), a view that Rothman et al. (2019) also endorse.[11] Thus, the distinction between transfer and CLE as two different processes has a weak theoretical

---

[11] In fact, parsing is recognized as the general way to acquire a language, while copying exists as an extra mechanism, according to the following quote from Rothman et al. (2019: 23): "Language is … parsed in real time through the grammar, and the interactions that take place at the level of on-line language processing … eventually reconfigure that underlying knowledge. However, this is not the only way in which a grammar can be modified. … [L]inguistic representations can be changed by, or even copied from, other representations … without the direct mediation of language processing. This means that … there are cognitive mechanisms at play that allow [linguistic representations] to be modified or replaced by others without the constant trial-and-error stream of information from input parsing … – a shortcut of sorts, with regard to how grammar restructuring generally works."

foundation, in my view. The LPM, on the other hand, argues that there is only one mechanism responsible for crosslinguistic influence: parsing (cf. section 4.3). That is, using a structure from a previously acquired language to parse L3 input will result in a temporary, unstable L3 representation that is influenced by that language. Further input conflicting with this representation may quickly wash it out, while further supporting input and use will strengthen it and make it a stable property of the L3 grammar. Thus, parsing will gradually lead to stable linguistic representations. And since the LPM does not assume any copying of (stable) grammatical representations, this view reflects the variability that is typical of early L3 grammars.

It has been argued that property-by-property transfer may lead to incoherent grammars, most recently formulated by Schwartz and Sprouse (2019) in the following way: "... PT [piecemeal transfer] in L3A is fundamentally incompatible with generative grammar, because elements of natural language grammars ... necessarily interact, whereas just incorporating one element at a time does not allow for the functioning of a complete and coherent grammar." However, a number of (generative) studies on microvariation show that grammars may differ in a single property, with no consequences for the rest of the grammar (cf. section 2). In fact, as we saw in section 4.1, not even properties assumed to be part of old-style parameters cluster together in L2 acquisition. Furthermore, within-language inconsistencies, which abound in natural languages, may of course be dealt with in generative approaches, e.g., by positing multiple sub-grammars, as in Amaral and Roeper (2014). And finally, it seems that step-wise development needs to be assumed even in models arguing for wholesale transfer, as restructuring (which constitutes the actual acquisition process) must take place property by property.

Property-by-property transfer has also been criticized for what is claimed to be lack of predictive power, "particularly with regard to when non-facilitation is expected" (Rothman 2019:143). Emphasizing the prevalence of non-facilitative transfer attested in L3 acquisition studies, they also ask the question if property-by-property transfer should not translate into "a much lower incidence of non-facilitation" (Rothman et al. 2019: 244). As stated above, the LPM predicts that crosslinguistic influence occurs when the learner uses a structure from a previously acquired language to parse L3 input. This means that the LPM is concerned with the abstract linguistic structure of the three languages involved, at the level of linguistic rules and constructions, not whole grammars, and that this is the level where predictions are made.

The mirror-image design argued in Puig-Mayenco, González Alonso and Rothman (2018) to be the optimal research methodology for L3 studies is therefore not the preferred methodology for the LPM, as this methodology focuses on typological/structural similarity vs. order of acquisition. Instead, relevant studies to test the LPM should make a comparison between L3 learners and two groups of L2 learners (with L1s that are the same as the previously acquired languages of the L3 group), cf. Westergaard et al. (2017). This methodology makes it possible to identify the exact contribution of the additional language involved in L3 acquisition. In such scenarios, the L3 group is predicted to perform better than the L2 group on properties that are similar in the L3 and the language that is lacking in the L2 population (due to facilitative influence from the additional language), while they should perform lower on properties that are similar in the L3 and the shared language but different from the additional language (due to non-facilitative influence); cf. the next section. With respect to the concern about non-facilitation expressed by Rothman et al. (2019), the assumption seems to be that when the L3 learner has a choice between transfer from the L1 or the L2, it is a question of choosing either the right or the wrong option, as would be the case in an 80s-style macroparameter approach to acquisition. But within a view of language that takes microvariation into account, this should no longer hold: Given that the L3 is hardly ever identical to either of the previously acquired languages, transfer will often be non-facilitative either way, even when the L3 learner chooses the option from the structurally more similar language. That is, there are often numerous fine linguistic distinctions between the three languages that may account for non-facilitative transfer. Nevertheless, perhaps there is something to the idea that there should be a "lower incidence of non-facilitation" in L3 acquisition – as is often claimed in more non-formal work on multilingualism (e.g., Jessner 1999, Modirkhamene 2006); see also Bartolotti and Marian (2012) on advantages of bilinguals on further language learning. What property-by-property transfer captures then is the insight that the more languages the learner knows, the easier it will be to learn further languages, since the existing repertoire will be larger. But the learner will nevertheless often be slightly (or even completely) wrong.

5.4 Empirical support for the LPM

The LPM is relatively new in the field of L3 acquisition, and so far, there are only a few studies that have tested this model. The original study in Westergaard et al. (2017) investigates Norwegian-Russian bilinguals learning English as an L3, comparing them to Norwegian and Russian L2 learners. The linguistic properties examined are two word order

phenomena, one where English is similar to Russian (non-V2 in declaratives, i.e., Adv-V) and one where it is similar to Norwegian (S-aux inversion), as illustrated in (18)-(19). The findings show that, although Norwegian is typologically closer to English than Russian is, there was a clear facilitative effect of Russian in the Adv-V condition. Nevertheless, the bilinguals scored significantly lower than the Russian L2 group, indicating an additional non-facilitative effect of Norwegian for the same property. The authors thus argued for the importance of linguistic proximity and against wholesale transfer.

(18)   Susan **always eats** sweets. (Adv-V)

(19)   What **will the little girl** play? (Aux-S)

A further study (Mitrofanova and Westergaard 2018) used Norwegian-Russian bilinguals and Norwegian monolinguals in a sentence/picture-matching experiment using an artificial language (Aliensk), made up of Norwegian lexical items and a grammatical property from Russian (that does not exist in Norwegian), nominative vs. accusative case (and corresponding word order variation). Examples of the linguistic stimuli (presented aurally) are illustrated in (20)-(21). After exposure to only 10 (grammatical) sentence/picture pairs, participants were tested on 60 further items, where understanding the nominative/accusative distinction was crucial to decide whether the sentences matched the pictures or not. Results show that the bilinguals scored significantly higher than the monolingual Norwegians, who only considered word order and displayed no sensitivity to the case feature. Again, it was argued that structural proximity may override superficial typological similarity, this time at an extremely early stage (less than a minute of exposure to the L3).

(20)   *Sebrail   tegner   soppsu*
         sebra.NOM draw.PRES mushroom.ACC
         The sebra is drawing a mushroom

(21)   *Hattsu holder   revil*
         hat.ACC hold.PRES fox.NOM
         The fox is holding the hat.

We may also consider the studies included in the systematic review by Puig-Mayenco et al. (2018), where altogether 71 L3 acquisition studies were investigated, involving a number of different languages. Out of these 71 studies, 43 show evidence of typological/structural transfer (60.1%), while L1 and L2 transfer only account for 14.1% and 28.2% of the data respectively, and the CEM as little as 5.9%. The Scalpel Model and the LPM are not considered, but the evidence for typological/structural transfer should in principle also support the LPM, at least in cases where the typological transfer has been attested early in the acquisition process. Effects of both previously acquired languages, which would be evidence for property-by-property transfer, are attested in 17 studies (23.9%), seven of them investigating learners at an early stage. This is not a particularly low number: Most of the studies included in the systematic review test only a few properties of the L3 (in some cases only one), and this of course means that it is somewhat unlikely that one would find hybrid transfer. Property-by-property transfer simply means that it should be possible to find effects of both languages at every stage of L3 acquisition, not that it always obtains, regardless of the nature and number of properties investigated. Moreover, influence from both previously acquired languages has been attested in other domains of language, phonology (e.g., Wrembel 2015, Jabbari and Pourmajnoun 2016, Lloyd-Smith, Gyllstad and Kupisch 2017) as well as the lexicon (e.g., Hammarberg 2001, Ringbom 2001), and as Bohnacker (2006: 479) puts it, "[t]here is no reason why syntax should be exempt from such transfer."

Finally, 16 of the 43 studies that provide evidence for what Puig-Mayenco et al. (2018) have classified as typological transfer involve languages that are genetically unrelated. This means that in these 16 studies, structural similarity rather than surface typological similarity must be the main factor responsible for crosslinguistic influence. Thus, there is considerable evidence in the literature for linguistic proximity as an important factor in L3 acquisition. However, in conclusion, it is pertinent to ask the question whether structural similarity is the only factor responsible for crosslinguistic influence in L3 acquisition or whether other factors play a role.

5.5 Other factors?

As in many, if not most, domains of linguistic research, there is hardly ever just one single factor that can account for all the data. This is also one of the main claims of Slabakova (2017: 653), who points out that there are "many additional factors affecting transfer ..., such as construction frequency, availability of clear unambiguous input, prevalent use, and structural linguistic complexity, among others." Other factors that could be added include age,

recency of use, instruction, and language dominance. I would also not completely exclude an influence of the order of acquisition, in the sense that languages learned after the L1 tend to influence each other (cf. the L2 Status Factor), as many Ln learners anecdotally report what we might refer to as a 'foreign language mode.' This could be the reason for the results presented in Bohnacker (2006), who finds non-facilitative influence of L2 English non-V2 word order (XSV) on the L3 German of L1 Swedish learners, as illustrated in (22), from Bohnacker (2006: 470). In this case, the L1 and the L3 are structurally similar in that they share the V2 phenomenon (and would require the word order XVS); at the same time the L3 (German) would presumably be considered to be more similar to the L1 (Swedish) than to the L2 (English) at all levels of the hierarchy of the TPM. Nevertheless, the influence here is arguably from the L2, and as far as I can tell, neither the LPM nor the TPM can account for these findings.

(22)   ***in Montags ich seh*** *ein Programm um Essen # teve.*

in Mondays I see a programme about food # telly$_{SWE}$

'On Mondays, I watch a programme on television about food.' (Gun1, 4 months)

This means that in further research on L3 acquisition, the main research task should be to focus on the question that is formulated in the black rectangle in Figure 6: That is, to investigate which factors decide which slice of the existing linguistic resources to use at which point in the L3 acquisition process - one or both of the previously acquired languages or the general linguistic endowment?
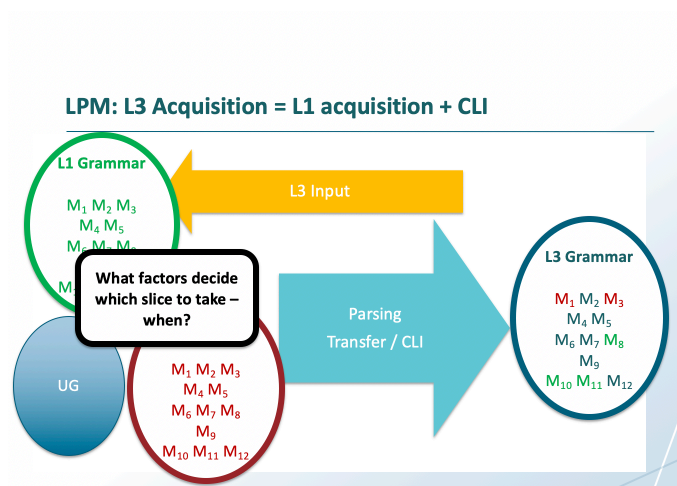

Figure 6: L3 acquisition according to the LPM.

6. Summary and conclusion

In this paper I have argued that the genetic endowment for language learning makes humans sensitive to fine linguistic distinctions in monolingual as well as all types of multilingual language development, which means that the acquisition process does not involve setting or re-setting (macro-)parameters or copying whole grammars. Thus, there is no *fundamental* difference between L1 acquisition on the one hand and L2/L3/Ln acquisition on the other. All language acquisition is *learning by parsing*, and learners build the new (L1, L2 or L3) grammar incrementally in a step-wise fashion; i.e., acquisition takes place property by property.

I have briefly outlined the Micro-cue Model of L1 acquisition, focusing on contexts where there is variation in the input. In such cases, young children are found to produce target-consistent utterances from early on, typically making errors of omission rather than errors of commission (under- rather than overgeneralization). That is, children are generally conservative learners, and the generalizations that they do make are small, involving only the addition of a subcategory or subclass, thus minimizing the need for unlearning.

I have also argued against wholesale transfer in L2 and L3 acquisition and instead proposed Full Transfer Potential; that is, *anything may transfer*, not *everything does transfer*. This means that L2/L3 acquisition is learning by parsing, not copying and restructuring. The difference between L1 acquisition and L2/L3/Ln acquisition is simply that in the latter case, the learner has a larger repertoire to choose from when parsing the language to be acquired. In this process, all previously acquired grammars remain active, presumably to varying degrees. Within the Linguistic Proximity Model of L3 acquisition, I argue that the main factor for crosslinguistic influence is structural similarity, especially at later stages of acquisition, while the parser may be influenced by superficial typological similarity at an early stage. Finally, as language acquisition is a complex phenomenon, it is expected that future research will reveal an effect also of other factors.

References

Amaral L and Roeper T (2014) Multiple Grammars and second language representation. *Second Language Research* 30(1), 3-36.

Anderssen M and Westergaard M (2010) Frequency and economy in the acquisition of variable word order. *Lingua* 120(11): 2569-2588. DOI: 10.1016/j.lingua.2010.06.006

Anderssen M, Bentzen K, Rodina Y and Westergaard M (2010) The acquisition of apparent optionality: Word order in Subject and Object Shift constructions in Norwegian. In: Anderssen M, Bentzen K and Westergaard M (eds) *Variation in the Input: Studies in the acquisition of word order*. Dordrecht: Springer, pp. 241-270.

Anderssen M, Bentzen K, Busterud G, Dahl A, Lundquist B and Westergaard M (2018) The acquisition of word order in L2 Norwegian: The case of Subject and Object Shift. *Nordic Journal of Linguistics* 41(3): 247-274. https://doi.org/10.1017/S0332586518000203

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: 459–484.

Bardel C and Falk Y (2012) Behind the L2 Status Factor: A neurolinguistic framework for L3 research. In: Cabrelli Amaro J and Rothman J (eds) *Third language acquisition in adulthood*. Amsterdam: John Benjamins, pp. 61-78.

Bardel C and Sanchez 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: John Benjamins, pp. 85-101.

Bartolotti J and Marian V (2012) Language learning and control in monolinguals and bilinguals. *Cognitive Science* 36(6): 1129-1147.

Berkes É and Flynn S (2012) Further evidence in support of the Cumulative-Enhancement Model: CP structure development. In: Cabrelli Amaro J and Rothman J (eds) *Third language acquisition in adulthood*. Amsterdam: John Benjamins, pp. 143-164.

Bley-Vroman R (2009) The evolving context of the fundamental difference hypothesis. *Studies in Second Language Acquisition* 31(2): 175-198.

Biberauer T and Roberts I (2012) Towards a parameter hierarchy for auxiliaries: diachronic considerations. *Cambridge Occasional Papers in Linguistics* 6: 267–294.

Blumenfeld H and Marian V (2013) Parallel language activation and cognitive control during spoken word recognition in bilinguals. *Journal of Cognitive Psychology* 25(5): 547-567. http://dx.doi.org/10.1080/20445911.2013.812093

Bohnacker U (2006) When Swedes begin to learn German: from V2 to V2. *Second Language Research* 22(4): 443–486.

Bowerman M (1988) The 'no negative evidence' problem: How do children avoid constructing an overly general grammar? In: Hawkins JA (ed) *Explaining language universals*. Basil Blackwell, pp. 73-101.

Brown R (1973) *A First Language: The Early Stages.* Cambridge, MA: Harvard University Press.

Chomsky N (1965) Aspects of the theory of syntax. Cambridge, MA: MIT Press.

Chomsky N (1981) *Lectures on government and binding*. Dordrecht: Foris.

Chomsky N (1986) *Knowledge of language: Its nature, origin and use.* New York: Praeger Publishers.

Clahsen H, Meisel JM and Pienemann M (1983) *Deutsch als Zweisprache: Der Spracherwerb ausländischer Arbeiter*. Tübingen: Narr.

Clahsen H and Muysken P (1986) The availability of universal grammar to adult and child learners — a study of the acquisition of German word order. *Interlanguage studies bulletin (Utrecht)* 2(2): 93-119.

Cunnings I (2017) Parsing and working memory in bilingual sentence processing. *Bilingualism: Language and Cognition* 20(4): 659-678.

De Houwer A (2005) Early bilingual acquisition: Focus on morphosyntax and the separate development hypothesis. In: Kroll J, DeGroot A (eds) *Handbook of Bilingualism: Pycholinguistic Approaches.* New York, NY: Oxford University Press, pp. 30–48.

Dekydtspotter L and Renaud C (2014) On second language processing and grammatical development: The parser in second language acquisition. *Linguistic Approaches to Bilingualism* 4(2), 131-165.

Del Maschio N and Abutalebi J (2019) Language Organization in the Bilingual and Multilingual Brain. *The Handbook of the Neuroscience of Multilingualism*. Wiley, pp. 199-213.

Donaldson B (2011) Left dislocation in near-native French. Studies in Second Language Acquisition 33: 399-432.

Eubank L (1993/94) On the transfer of parametric values in L2 development. *Language Acquisition* 3: 183–208.

Falk Y, Lindquist 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.

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 of relative clauses. *The International Journal of Multilingualism* 1: 3–16.

Fodor JD (1998) Unambiguous triggers. *Linguistic Inquiry* 29(1): 1-36.

González Alonso J and Rothman J (2017) Coming of age in L3 initial stages transfer models: Deriving developmental predictions and calling for theory conservatism. *International Journal of Bilingualism* 21(6): 683-697. doi:10.1177/1367006916649265.

Green DW and Abutalebi J (2013) Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology* 25.5: 515-530.

Grüter T (2006) Another Take on the L2 Initial State: Evidence from Comprehension in L2 German. *Language Acquisition* 13:4: 287-317. DOI: 10.1207/s15327817la1304_2

Hammarberg B (2001) Roles of L1 and L2 in L3 production and acquisition. In: Cenoz J, Hufeisen B and Jessner U (eds) *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives*. Multilingual Matters, pp. 21–41.

Hartsuiker RJ, Pickering MJ and Veltkamp E (2004) Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological Science* 15(6): 409-414.

Hermas A (2014) Multilingual transfer: L1 morphosyntax in L3 English. *International Journal of Language Studies* 8(2): 10–24.

Hulk A and Müller N (2000) Bilingual first language acquisition at the interface between syntax and pragmatics. *Bilingualism: language and cognition* 3(3): 227-244.

Ivanova I, Pickering MJ, Branigan HP, McLean JF and Costa A (2012). The comprehension of anomalous sentences: Evidence from structural priming. *Cognition* 122:193–209.

Jabbari AA and Pourmajnoun F (2016) The acquisition of onset Consonant Clusters by Persian Speakers (L1), Learners of English (L2) and French (L3): An optimality Account. *IOSR Journal of Humanities and Social Science* 21(9): 13-21.

Jensen M (2017) *Investigating the Bottleneck Hypothesis in Second Language Acquisition: The acquisition of narrow syntax and functional morphology among Norwegian L2 learners of English.* Master thesis, UiT The Arctic University of Norway.

Jessner U (1999) Metalinguistic awareness in multilinguals: Cognitive aspects of third language learning. *Language awareness* 8(3-4): 201-209.

Jin F (2009) Third language acquisition of Norwegian objects: Interlanguage transfer or L1 influence? In: Leung YI (ed) *Third language acquisition and universal grammar*. Bristol, UK: Multilingual Matters, pp. 144-161.

Julien M (2005) *Nominal Phrases from a Scandinavian Perspective.* Amsterdam: John Benjamins,

Köpke B, Schmid MS, Keijzer M and Dostert S (eds) (2007) *Language attrition: Theoretical perspectives*. John Benjamins Publishing.

Kroll JF, Dussias PE, Bogulski CA and Valdes-Kroff J (2012) Juggling two languages in one mind: What bilinguals tell us about language processing and its consequences for cognition. In: Ross B (ed) *The psychology of learning and motivation* 56. San Diego, CA: Academic Press, pp. 229-262.

Lardiere D (2009) Some thoughts on the contrastive analysis of features in second language acquisition. *Second language research* 25(2), 173-227.

Leung YI (1998) Transfer between interlanguages. *Proceedings of the 22nd Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press, pp. 477-487.

Leung YI (2003) Failed features versus full transfer full access in the acquisition of a third language: Evidence from tense and agreement. In: Liceras JM (ed) *Proceedings of the 6th Generative Approaches to Second Language Acquisition Conference*. Somerville, MA: Cascadilla Press, pp. 199–207.

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. doi:10.1017/S1366728904002044

Lewis S and Phillips C (2015) Aligning grammatical theories and language processing models. *Journal of Psycholinguistic Research*, 44(1): 27-46.

Lightfoot D (1999) *The Development of Language: Acquisition, Change and Evolution.* Malden, MA and Oxford: Blackwell.

Lightfoot D (2006) *How New Languages Emerge.* Cambridge: Cambridge University Press.

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): 131-162.

Lohndal T, Westergaard M and Vangsnes ØA (Forthcoming) Verb Second in Norwegian: Variation and Acquisition. In: Biberauer T, Woods R and Wolfe S (eds), *Rethinking V2.* Oxford: Oxford University Press.

Lødrup H (2011) Norwegian Possessive Pronouns: Phrases, Words or Suffixes? In: Butt M and King TH (eds) *Proceedings of the LFG11 Conference*. Stanford: CSLI Publications, pp. 383-403.

Meisel JM (1989) Early differentiation of languages in bilingual children. In: Hyltenstam K and Obler LK (eds) *Bilingualism across the lifespan: Aspects of acquisition, maturity and loss*. Cambridge: Cambridge University Press, pp. 13-40.

Mitrofanova N and Westergaard M (2018) Wholesale vs. property-by-property transfer: Acquisition of morphological case in an artificial L3. Poster presented at 2nd International Symposium on Bilingual and L2 Processing in Adults and Children (ISBPAC), Braunschweig, May 24-25.

Modirkhamene S (2006) The reading achievement of third language versus second language learners of English in relation to the interdependence hypothesis. *International Journal of Multilingualism* 3(4): 280-295.

Montrul S and Slabakova R (2003) Competence similarities between native and near-native speakers: An investigation of the preterite-imperfect contrast in Spanish. *Studies in second language acquisition* 25(3): 351-398.

Mykhaylyk R, Mitrofanova M, Rodina Y and Westergaard M (2015) The Linguistic Proximity Model: The Case of Verb-Second Revisited. In: Grillo E and Jepsen K (eds) *Proceedings of the 39th Boston University Conference on Language Development* 2. Somerville, MA: Cascadilla Press, 337-349.

Na Ranong S and Leung YI (2009) Null objects in L1 Thai-L2 English-L3 Chinese: An empirical take on a theoretical problem. In: Leung YI (ed) *Third language acquisition and universal grammar*. Bristol, UK: Multilingual Matters, pp. 162-191.

Newmeyer FJ (2005) *Possible and Probable Languages: A Generative Perspective on Linguistic Typology*. Oxford: Oxford University Press.

Paradis M (2009) *Declarative and procedural determinants of second languages*. Amsterdam: John Benjamins.

Phillips C (1996) *Order and Structure.* PhD dissertation, Massachusetts Institute of Technology.

Platzack C (1996) The Initial Hypothesis of Syntax. A minimalist perspective on language acquisition and attrition. In: Clahsen H (ed) *Generative perspectives on language acquisition*. Amsterdam: John Benjamins, pp. 369–414.

Puig-Mayenco E, González Alonso J and Rothman J (2018) A Systematic Review of Transfer Studies in Third Language Acquisition. *Second Language Research*. Epub ahead of print, 22 November 2018. DOI: 10.1177/0267658318809147.

Radford A (1992) The acquisition of the morphosyntax of finite verbs in English. In: Meisel JM (ed) *The Acquisition of Verb Placement: Functional Categories and V2 Phenomena in Language Acquisition*. Dordrecht: Kluwer Academic Publishers, 23-62.

Ringbom H (2011) Lexical transfer in L3 production. In: Cenoz J, Hufeisen B and Jessner U (eds) *Cross-linguistic influence in third language acquisition: psycholinguistic perspectives*. Multilingual Matters, pp. 59–68.

Roeper T (1999) Universal bilingualism. *Bilingualism: Language and Cognition* 2(3): 169–86.

Rothman J (2011) L3 syntactic transfer selectivity and typological determinacy: The typological primacy model. *Second Language Research* 27: 107–127.

Rothman J (2013) Cognitive economy, non-redundancy and typological primacy in L3 acquisition: Evidence from initial stages of L3 Romance. In: Baauw S, Dirjkonigen F, Meroni L and Pinto M (eds) *Romance languages and linguistic theory*. Amsterdam: 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: Language and Cognition* 18(2): 179–190.

Rothman J and Cabrelli Amaro J (2010) What variables condition syntactic transfer? A look at the L3 initial stage. *Second Language Research* 23: 189–218.

Rothman J, González Alonso J and Puig-Mayenco E (2019) *Third Language Acquisition and Linguistic Transfer*. Cambridge: Cambridge University Press.

Rowland CF and Pine JM (2003) The development of inversion in wh-questions : A reply to Van Valin. *Journal of Child Language* 30: 197–212.

ScanDiaSyn, project on Nordic Dialect Syntax. Information available here: http://websim.arkivert.uit.no/scandiasyn/

Schwartz B and Sprouse R (1996) L2 cognitive states and the full transfer/full access model. *Second Language Research* 12(1): 40-72.

Schwartz B and Sprouse R (2019) *Contra* Piecemeal Transfer: A reply to Slabakova (2017) and Westergaard et al. (2017). Talk presented at GASLA XV, Reno, March 22-24.

Sharwood Smith M (1983) On first language loss in the second language acquirer: Problems of transfer. In: Gass S and Selinker L (eds) *Language Transfer in Language Learning. Issues in Second Language Research*. Rowley, MA: Newbury House Publishers, pp. 222-231.

Sharwood Smith M and Truscott J (2006) Full transfer full access: A processing oriented interpretation. In: Unsworth S, Parodi T, Sorace A, Young-Scholten M (eds) *Paths of development in L1 and L2 acquisition: In honor of Bonnie D. Schwartz*. Amsterdam: John Benjamins, 201-216.

Sharwood Smith M and Truscott J (2008) MOGUL and Crosslinguistic Influence. *Morphosyntactic Issues in Second Language Acquisition* 29: 63-85.

Sharwood Smith M and Truscott J (2014) *The multilingual mind: A modular processing perspective*. Cambridge: Cambridge University Press.

Slabakova R (2017) The scalpel model of third language acquisition. *International Journal of Bilingualism* 21(6): 651-665.

Snyder W (2007) *Child language: The parametric approach*. Oxford University Press.

Thornton R (1990) *Adventures in long-distance moving: The acquisition of complex wh-questions*. PhD dissertation, University of Connecticut.

Tomasello M (2003) *Constructing a language: A usage-based theory of language acquisition.* Cambridge, MA: Harvard University Press.

Truscott J and Sharwood Smith M (2004) Acquisition by processing: A modular perspective on language development. *Bilingualism: Language and Cognition* 7(1): 1-20.

Ullman M (2001) The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition* 4(2): 105–122.

Vainikka A and Young-Scholten M (1996) Gradual development of L2 phrase structure. *Second Language Research* 12(1): 7–39.

Vangsnes ØA (2006) Microparameters for Norwegian *wh*-grammars. *Linguistic Variation Yearbook* 5. Amsterdam: John Benjamins, pp. 187-226.

Westergaard MR (2003) Unlearning V2: Transfer, markedness, and the importance of input cues in the acquisition of word order in English by Norwegian children. *EUROSLA Yearbook* 3: 77-101.

Westergaard M (2009a) *The acquisition of word order: Micro-cues, information structure and economy*. [Linguistik Aktuell/Linguistics Today 145]. Amsterdam: John Benjamins.

Westergaard M (2009b) Microvariation as Diachrony: A View from Acquisition. *Journal of Comparative Germanic Linguistics* 12(1): 49-79.

Westergaard M (2009c) Usage-based vs. rule-based learning: The acquisition of word order in wh-questions in English and Norwegian. *Journal of Child Language* 36(5): 1023-1051.

Westergaard M (2011) Subject positions and information structure: The effect of frequency on acquisition and change. *Studia Linguistica* 3: 299-332.

Westergaard M (2014) Linguistic variation and micro-cues in first language acquisition. *Linguistic Variation* 14(1): 26-45.

Westergaard M and Bentzen K (2007) The (non-) effect of input frequency on the acquisition of word order in Norwegian embedded clauses. In: Gülzow I and Gagarina N (eds) *Frequency Effects in Language Acquisition: Defining the Limits of Frequency as an Explanatory Concept.* Berlin/New York: Mouton de Gruyter, pp. 271-306.

Westergaard M, Mitrofanova N, Mykhaylyk R and Rodina Y (2017) Crosslinguistic influence in the acquisition of a third language: The Linguistic Proximity Model. *International Journal of Bilingualism* 21(6): 666:682. DOI: 10.1177/1367006916648859.

White L (1985) The "pro-drop" parameter in adult second language acquisition. *Language learning* 35(1): 47-61.

White L (1990) The verb-movement parameter in second language acquisition. *Language acquisition* 1(4): 337-360.

White L (2003) *Second language acquisition and Universal Grammar.* Cambridge: Cambridge University Press.

Wrembel M (2015) In search of a new perspective: Cross-linguistic influence in the acquisition of third language phonology. Available from https://repozytorium.amu.edu.pl/handle/10593/14647.