Some remarks on the history of transfer in language studies

Daniil M. Ozernyi[*]

Abstract. Transfer is a notion that is pervasive in the modern field of second and third language acquisition. However, the definition of transfer is not easily found. Most of this paper is devoted to a historical overview of the notion of transfer and the changes the definition underwent over the years. At the end of this paper, I sketch a proposal which involves discarding the notion of "transfer" in favor of more computationally efficient options.

Keywords. transfer; cross-linguistic influence; interference; formal sufficiency; second language acquisition; third language acquisition

1. On the history of transfer[1].

1.1. INTRODUCTORY REMARKS. The notion of transfer is crucial to many modern subfields of linguistics.[2] This is primarily due to the fact that one of the main objectives of language acquisition studies beyond L_1 is to establish the role of the previous languages in acquisition of the subsequent ones (cf., e.g., Epstein et al. 1996; Rothman et al. 2019). Trivially, the notion of "transfer" is a cornerstone in such an inquiry, or at least so it has been since studies on second language appeared as a branch of psychology and education studies (since Lado 1957a). Some researchers even refer to these studies as "transfer studies" (e.g., Gass 1988; Puig-Mayenco et al. 2020). Since it hardly lends itself to debate that transfer and the adjacent notions (e.g. cross-linguistic influence (CLI), interference, etc.) are at the very core of the field, one might expect them to be well-defined and their definitions to have been long agreed upon. This, however, does not appear to be the case.

Even a brief survey of the recent papers in the field of third language acquisition shows that there is still much debate, about the term or the concept it represents, if not confusion, surrounding the concept of transfer. One such example is Rothman *et al.*'s distinction between transfer and CLI (Rothman et al. 2019), and Westergaard's rejection of such a distinction (Westergaard 2021b; p. 104); another brief critique of transfer vs. CLI can be found in a review (Ozernyi 2021) of Rothman *et al.* (op. cit.). The principle argument of Ozernyi in his review is that transfer is not well-defined, hence it cannot be distinguished in any precise or meaningful way from cross-linguistic influence – or from anything for that matter – *a priori*, since the distinctions that are being made need to be made metaphysically precise.[3] Clearly, there is debate as to what transfer is and what it entails. It appears as though transfer came to be an umbrella term for any influence of any trace of $L_{1,2,3...n}$ on some L_{n+1}.

---

[1] Note that this section which takes up most of the proceedings paper is an excerpt from a longer and more comprehensive manuscript, namely Ozernyi (2022b). The continuously updated version(s) will be available at github.com/DOzernyi/transfer-hist.

[2] For both parsimony and convenience reasons, I will use "transfer" to mean "the notion of transfer" and "the process of transfer", the two being roughly synonymous. Where I will diverge from this convention, I will specify the intended meaning.

[3] What the requirements for metaphysical precision are is another question, but at the very least it's well-definedness and absence of ambiguity.

1.2. TRANSFER IN EARLY PSYCHOLOGY. One of the first references to transfer cited in later relevant literature (i.e., psychology of the 1890s-1920s; James 1890a) appears to be in Priestley when he discusses the nature of judgement as feeling: "[Judgement is] transferring the idea of truth by association from one proposition to another that resembles it" (Priestley 1790; p. 30). Priestley does not himself give a definition, apparently taking it to be self-evident. This is the problem of many future papers which, implicitly amending or adjusting the term, used it without defining it. There are two crucial elements of transfer which we can infer from the Priestley's use of it: (a) transfer presumes assigning a property which is relevant to one item A to another item B, and (b) it is imperative that A and B are associated, i.e. A "resembles" B.[4]

   One century later, Priestley was quoted by William James in his pioneering *Principles of Psychology* (1890a).[5] This was perhaps one of the entrance points for transfer to appear in psychology (in contradistinction to Priestley's philosophical work). James does not give us a definition either. He starts out with Priestley's words almost *sic erat scriptum*, writing about "transfer of feeling from one object to another, associated by contiguity or similarity with the first" (James 1890a; p. 330). However, it is obvious that later on James expands transfer as he talks about "transfer of relations [...] within a homogenous series" (James 1890b; p. 660).[6]

1.3. TRANSFER BETWEEN PSYCHOLOGY AND LINGUISTICS. Some notable studies in the 1900s and early 1910s looked into transfer in the domains of memory (Winch 1908), sensual, perceptual, or motor function (for a pioneering study, see Woodworth & Thorndike (1901)). However, no definition was given in those studies: transfer maintained its subdoxastic nature. Thorndike and Woodworth refer to a "*mysterious* transfer of practice [...], an unanalyzable property of mental functions" (Woodworth & Thorndike 1901; p. 256, emphasis added). Their phrasing signals a lack of contemporary understanding of the mechanisms of transfer, and *ergo*, its origin.

   The first elaborate and clear definition of transfer we get is in the 1940s from McGeoch:[7]

(1)   Definition of transfer and adjacent notions in McGeoch (1942)

   a.   The influence of prior learning (retained until the present) upon the learning of, or response to, new material has traditionally been called **transfer of training**.[8]

   b.   It appears in experimental measurements as a **transfer effect**, which means the influence of a specified amount of practice or degree of learning in one activity upon the rate of learning of another activity or upon response to another situation.[9]

---

[4] Priestley, of course, did not talk about transfer of language – but transfer in philosophy, perhaps assuming it applies across various domains.

[5] I absolutely do not purport to claim that transfer was not used in-between. Priestley was chosen, however, because only four citations separate him from Lado (1951), and only five citations separate Priestley from Flynn (2021).

[6] James also did not talk of transfer in the context of language, but in the context of contemporary psychology generally.

[7] Those interested in a more comprehensive understanding of McGeoch's elaboration would be well-advised to visit Chapter 10 of *The Psychology of Human Learning, An Introduction* which is entirely devoted to "transfer of training".

[8] Note how McGeoch appeals to tradition rather than citing any of the studies, albeit in the rest of the book he is admirably punctilious about the terminology employed and his definitions (for an example of this see his fn. 12 and elsewhere).

[9] This would go on to be at the foundation of contrastive analysis of Lado and Fries. I will talk about their usage of

c. Transfer effects may be (a) **positive**, when training in one activity facilitates the acquisition of a second activity, (b) **negative**, when the training in one inhibits or retards the learning of another, and (c) **zero or indeterminate**, when training in one has no observed influence on the acquisition of a second. (from McGeoch 1942; p. 394, emphasis added)

1.4. TRANSFER IN ROBERT LADO'S WORK. Two figures which pioneered language learning and, specifically, what later came to be known as contrastive analysis (CA), were Charles C. Fries and Robert Lado. Early work of Fries did not concern foreign language learning: his papers were mostly on structure of English as a language and learning English a first language (cf. Fries 1925, 1927; etc.). However, even in later works, he does not appear to use transfer. In his seminal textbook on teaching English as a foreign language, transfer is nowhere to be found (Fries 1945). Instead, what seems to be the very first instance of usage of transfer in a paper on foreign language learning is Lado (1949). In a footnote on this fragment:

> From a psychological point of view we note that the learner will acquire more rapidly those elements of the foreign language that operate on habits already established for the native language, less rapidly those elements that require the acquisition of new habits, and least rapidly those in which the new habits conflict with the linguistic habits already established by the native language (Lado 1949; p. 109)

he refers to McGeoch's transfer (1942:55-59). Approximately at the same time, Fries and Pike in a paper on phonology, mention "transfer from Spanish to English nasals" (Fries & Pike 1949; p. 37) and reference Marckwardt (1946), despite the fact that Marckwardt did not use "transfer" and used "influence" instead (111). Later on, Lado spearheaded the campaign on using transfer, it appears, because a number of his works which I will take a more careful look at below, make exceptionally wide use of the term, including but not limited to Lado (1951, 1956, 1957a, 1957b).

Lado seems to have introduced the term of wholesale transfer which thrives today (albeit hopefully, yet arguably) in a different meaning (cf. Schwartz & Sprouse 2021; Westergaard 2021a):[10] "a **wholesale** transfer of a reading technique into aural comprehension..." (Lado 1951; p. 53), emphasis added). It is clear, however, that at this point, transfer is still not being used in modern, "linguistic" meaning, an example of which would be transfer of parameters or features (properties which can take different shape depending on the theoretical framework one chooses) from the previous language to the target language. Instead, what we see now in Lado's work is the introduction of "psychological" transfer-of-training to linguistics and language learning. The difference between the two will emerge later and will become increasingly pronounced by Zobl (1980). The talk of "structures" appears already in 1950: "those **structures** in the foreign language that are not transferable from the native language are the ones we seek to discover by comparing the two languages in order to have the most effective testing materials" (Lado 1951; p. 19, emphasis added). Here's a useful nascent notion: that of transferability; it is yet another newly-introduced term to account for those structures which are "not transferable". Why, however, any given structure was thought to be "not transferable" is not clarified.

---

transfer in the next section.

[10] Westergaard does not support the wholesale transfer models of L_n acquisition, merely makes wide use of the term (Westergaard 2021a; pp. 2, 12f, 15, etc.).

Later on in his book, we also get a more elaborate description of what "wholesale" transfer is:

> a speaker of one language tends to transfer the entire system of his language to the foreign language[...]. He tends to transfer his sound system, including the phonemes, the positional variants of the phonemes, and the restrictions on distribution. He tends to transfer his syllable patterns, his word patterns, and his intonation patterns, as well (Lado 1956; p. 26).[11]

This view echoes more recent work (e.g., initial work on the TPM model, see references above), and surpasses the initial work of Lado on phonology, augmenting it with "word patterns". Under this definition, transfer is not selective ("the entire system"). Oddly, this view contradicts the earlier (Lado 1951) mention of non-transferrable structure, i.e. those parts of "system" which do not transfer, which would make the "wholesale transfer" or the transfer of "the entire system" simply impossible. Such incoherences are prominent, not only in Lado's work, but overall in transfer literature. In Lado, however, they are particularly pronounced. In addition to the the issue with transferability, for Lado, learners are *conscious* of transfer and unwilling to accept it: "in spite of [them]self [the learner] will transfer those habits to the new dialect and styles [they are] trying to learn" (Lado 1957b; p. 14).[12] Lado also mentions "intent" in earlier work (see Lado op. cit.: fn. 13).

The last paper of Lado we'll consider here is his seminal work on CA – *Linguistics across cultures* (Lado 1957a). Notably, he references transfer at the very beginning, alluding to Fries (1945), despite Fries not using transfer in his book. Apart from re-stating his earlier theses on nature of phonetic and phonological transfer (Lado 1957a; p. 11), Lado expands it to "physically similar phonemes" (12), transfer of morphology (58), even reading habits (94) and writing system (97);[13] he mentions positive/negative transfer as well (109).

The definition of transfer in Lado's work is nowhere to be found. What we find, however, is neither a purely "psychological" transfer ("transfer of training", of habits),[14] nor a purely *ante litteram* linguistic one (*sc.* transfer of structure, perhaps of mental representation, but definitely not of habits or conscious activity or (conscious) metalinguistic competence).[15] Instead, Lado seems to present a kluge of the above two:[16] transferring reading and writing habits obviously are instances of "psychological" transfer; while gender, case, and other morphological features are

---

[11] Flynn (p.c.) notes that "the fact the notion of "transfer" discussed here by Lado was consistent with behaviorism – i.e., the transfer for habits [in Skinnerian sense]. It also occurs within the context of structuralism as a theory of language".

[12] At least, the students of the first-year college courses which he is talking about in the 1951 paper quotes immediately above.

[13] Lado, acknowledges, however, that "we are less clear on how this transfer will affect our learning of a foreign language writing system" (Lado 1957a; p. 97).

[14] This kind of transfer was found in Rugg (1916) and McGeoch (1942).

[15] It is important to note that metalinguistic competence is a notion that is not well-defined. It is taken to mean roughly "conscious insight about language" (take, e.g., Falk et al. 2015). I take it to mean a thinking process along the lines "move out the auxialiary to form a question", a rule which the learner was taught and which is not done unconsciously, but more like math. However, with practice and time, does the learner internalize the grammar and needs not those "conscious rules" or does the learner just get proficient at the rules so that timing shrinks significantly? The answers, and the precise line between acquisition and learning, and the role learning plays in acquisition, hypothesis space for the two processes – all are yet to be clarified.

[16] I take it to mean a clumsy mix of two different notions, a phenomenological chimera of a sort.

obviously a much more subtle, unconscious, "linguistic" transfer.[17] In losing this vital distinction, willingly or unwillingly, Lado leads the reader and the subsequent researchers to confusingly collate, confuse two different definitions of transfer: that of McGeoch (1942) and roughly that of James (1890a). While the latter could deal with subtle, structural properties and employ copying, the former could not. McGeoch's definition is especially unsuitable for linguistics because "linguistic" transfer (in acquiring gender, for example) is not a vague **influence on performance** – recollect "influence" in McGeoch's definition –, it is a much more subtle structural **function**[18] **within competence**.[19] In other words, Lado "merges" under the same definition of transfer:[20]

(2)   Lado's "merge"

    a.   the cases where learners transfer lexicon settings (e.g., gender or some θ-marking) or syntactic parameters (e.g., headedness setting or constraints on any given kind of movement) from a previous language,[21] with

    b.   cases of transfer-of-(conscious)-training along the lines of Rugg (1916).

While, trivially, these activities share some similarity or perhaps directionality at some level of abstraction, it is impossible to imagine them nesting under the same definition (unless the definition is indefensibly vague). No experimental design could possibly aim to investigate both such "transfers".

We shall not attempt to formalize Lado's definition, since, as we have shown, it is a kluge of two. I will only note that perhaps Lado's lack of care for terminology will lead transfer to become a notion which, in Wenk (1974) and Kellerman (1977)'s words, can mean anything to anyone.

1.5. TRANSFER IN THE 1960-1990S. In the 1960s, we see how Lado's work and collation of the definitions spread to other works. For example, Stephens (1960) writes: "Transfer is more likely to take place when the thing to be transferred is **a generalization**, a conscious insight, a constant error to be dealt with, or a rule that can be understood" (Stephens 1960; p. 1542, emphasis added). Is transfer a generalization now? A conscious insight that came perhaps to be known as metalinguistic competence – awareness of language structure which is akin to awareness that $proj_L(x) = x^{\|}$?[22] This is the first time that generalization gets "transferred". This view (transfer operating over a generalization) contrasts heavily with later work. For example, Libuše Dušková (1969) dichotomized transfer and (over)generalization as a means of acquisition, pointing out that there isn't only one way to acquire language, i.e. it isn't all about transfer. She uses her experiment on acquiring English by Czech student who did not mark plurality in English, albeit Czech

---

[17] My use of unconscious "linguistic" vs. conscious "psychological" distinction is a tad terminologically misleading; however, for the purposes of the paper, take "psychological" to mean studies in psychology of the 1900s-1910s and "linguistic" to be, very roughly, involving I-language.

[18] I use function here in a pre-theoretic sense, roughly as mapping from one set to another.

[19] I am not aiming to define competence vs. performance distinction here, but a relevant introduction is given at the beginning of Chomsky (1965), some discussion relevant to SLA is given in Epstein et al. (1996) and reply to commentaries thereafter.

[20] The use of "merge" here is entirely unrelated to merge in syntax or elsewhere.

[21] It should be noted that perhaps different acquisition processes apply for (i) idiosyncratic properties of language (e.g., parts of lexicon) and (ii) the syntactic, categorical components. More research is needed to speculate on this.

[22] Formula for orthogonal projection of a vector (from Bretscher 2018; p. 61).

marks plurality (Dušková 1969); see also the discussion of this paper by Dušková, including findings contra the CA paradigm (Flynn 1987; pp. 14-17).

Somewhat more principled accounts of transfer developed in the mid-1960s. For example, the notion of hierarchy of learning difficulty was introduced. Stockwell, Bowen, and Martin (1965) write that "assignment of an item [in a hierarchy of learning difficulty] is based on the premise that [positive] transfer from one language to another [...] becomes more difficult as the correspondences weaken" (Stockwell et al. 1965; p. 292). This aligned well with the CA paradigm, but also somewhat refined the boundaries of transfer, being a yet another ancestor of later models of acquisition based on typological relationships between languages.

Around this time, transfer took the central position in second language studies, and critical views on transfer abounded. For example, Politzer, reflecting on transfer, writes that while that "on a beaucoup étudié la question du transfert des connaissances d'une langue étrangère à une deuxième, sans parvenir à des conclusions définitives" (Politzer 1965; p. 1).[23] By the end of 1960s, Jakobovitz claims that while "the literature on transfer (when the term is considered in its broadest sense) is possibly more extensive than that on any other topic in psychology and education [...] careful reviews of the vast literature pertaining to transfer are invariably pessimistic" (Jakobovits 1969; p. 57). Jakobowitz paints a very grim picture, but continues on to work on transfer (see below). While no definition was present in the 1969 paper, Jakobowits proposes that "similarities between two languages in terms of their surface features are more relevant to the operation of transfer effects than deep structure relations" (Jakobovits 1969; p. 55). It follows from this claim that "surface features" are independent of "deep structure relations", for it is otherwise impossible for surface structures to be considered on a separate basis and be thus "more relevant". It is not clear how this aligns with modern views on language architecture, but the overall thesis that superficial differences (e.g. over/covert) are more important for acquisition than underlying structure is trivially inadequate.

Jacobowitz (1969) was also the first one to offer the formalization of transfer. He writes that "a general formulation of the transfer problem must deal with five basic elements: task A, task B, training or practice on task A, training or practice on task B, and the relation between task A and task B" (59). Hence, for him, the transfer effect can primordially be expressed as

(3) $P_{L_2} = f(P_{L_1}, t_{L_2}, R_{L_1 - L_2})$, where
P_{L_1} is proficiency on Task A,
P_{L_2} is proficiency on Task B,
t_{L_2} is training in L2,
$R_{L_1 - L_2}$ is some "relation between L_1 and L_2".

He goes through several modifications of this formula, separating transfer and deducing the formula which he sees as the definition of transfer. The detailed argumentation can be found in Jakobovits (1969. pp. 59ff), but I will only consider some of the problems with his approach. The problem with this formalization is that, while attempting to deal with language (the paper was on "second language learning"), it still operates within this kluge of definitions: we see Jakobowitz using the terms like "proficiency on Task A" which reminds us of Rugg (1916) and McGeoch (1942). Once again, proficiency is a measure of performance (including that on task A) which

---

[23] The question of transfer of knowledge from one language to another has been extensively studied, yet no definitive conclusions have been reached (translation is mine – DMO).

has little to do with competence. While we cannot directly evaluate competence, we do not describe or see acquisition in terms of performance. The (relevant) transfer (if any) generative studies are concerned with cannot be concerned with performance (roughly, E-language) – only with competence (broadly construed).[24] While discussing Lado, we decided against formalizing his definition because it was a kluge of two very different "transfers" – but Jakobowitz went down this road, and the obtained formalization that is questionable. What is "training"? How can it be quantified? What is the function denoted as f? It appears as though "formalization" is no more than a fancy bit notation, void of clarity and precision.

The very first intentional distinction between two definitions Lado fused comes from Selinker:

> I consider the following to be processes central to second-language learning: first, language transfer; second, transfer-of-training; third, strategies of second-language learning; fourth, strategies of second-language communication; and fifth, overgeneralization of TL [target language – DMO] linguistic material (Selinker 1972; p. 216).

We are most interested in the seeming difference between the first two: language transfer and transfer-of-training. Selinker explains:

> If it can be experimentally demonstrated that fossilizable items, rules, and subsystems which occur in IL [interlanguage – DMO] performance are a result of the NL [native language – DMO], then we are dealing with the process of **language transfer**; if these fossilizable items, rules, and subsystems are a result of identifiable items in training procedures, then we are dealing with the process known as the **transfer-of-training**... (Selinker 1972; p. 216, emphasis added)

This is an attempt to deal with the problem of collation of definitions mentioned above. Even attempting to separate transfer-of-training and "linguistic" transfer is great. Yet, there are a number of questions in relation to the distinction which Selinker draws. First of all, what is fossilization and what are the "items, rules, and subsystems" which are "fossilizable"?[25] If we turn to the standard definition, only errors are referred to as being fossilized in modern language teaching (not linguistics, though), and fossilized errors are "errors which a learner does not stop making and which last for a long time, even for ever [...]. Fossilization of error often happens when learners [. . .] have no communicative reason to improve their language" (Spratt et al. 2011; p. 63). Fossilization hence appears to be of purely applied nature, i.e. having little to do with primarily unconscious acquisition of language system(s), whatever those systems are.[26] To support that, we see that Selinker operates within performance which we have already taken to be an unreliable narrator of language acquisition.

The distinction which Selinker draws does not appear to be viable. His transfer-of-training is alleged to stem from "identifiable items in training procedures". However, what are these items?

---

[24] This logic holds still even if we stipulate surface structure vs. deep structure distinction of Jacobowitz.

[25] How is fossilizable vs. non-fossilizaable distinction drawn? This is painfully reminiscent of Lado's "transferability".

[26] I remain optimistic about fossilization, in language teaching terms, being just a lag or a bug in externalization, accountable for non-nativelike speech. It is my hope that "fossilizzation" has nothing do to with competence, only with performance: the reason being that if it has to with competence, the learners would have problems parsing the grammatically convention input with the "fossilized" parts of their competence. Receptive bilinguals are a case in point to argue that performance has little to do with competence (cf. Sherkina-Lieber et al. 2011; *inter alia*).

If a learner whose L_1 does not mark plural morphologically overgeneralizes third person singular inflection of English and subsequently uses it for non-conforming forms like oxen and sheep (cf. *oxs or *oxes, *sheeps), is that transfer-of-training? Similar generalization and abstraction processes, no doubt, occur on a much deeper level (cf. Lust 2006; pp. 68, 122, 170 fn30, and section 11.2 generally) and are far from easily identifiable transfer-of-(conscious)-training. Therefore, drawing the distinction by dichotomizing transfer as coming either from L_1 or from training fails to account for the elaborate process of language acquisition.

One positive feature of the definition above is that there are clear elements in it (the four in the original quote). Later works overlooked Selinker's distinction and fused the four elements all over again. As such, Shachter writes that "if the constructions are similar in the learner's mind, [they] will transfer his *native language strategy* to the target language" (Schachter 1974; p. 212, emphasis added). What the strategy is was left undefined. The objects of transfer varied vastly: while Shachter's transfer was that of strategies, Taylor's was again of "structures" (Taylor 1975; p. 75), and so forth.

In subsequent years, attempts to distinguish transfer from interference ensued. One of those was Kellerman's paper on "strategy of transfer":[27]

> the connection between transfer experiments in the laboratory (which is the place where the term "interference" strictly belongs) and transfer in second language learning have been shown to be very tenuous, with many writers being reluctant to link the two in any significant fashion (Kellerman 1977; p. 61)

Similar logic appeared in Kellerman (1979). This could be seen as an attempt to separate language learning or language teaching from language acquisition: while "fossilization", "strategies," and other CA heritage are relevant mostly to the classroom practice (which was perhaps one of the motivations for Lado's CA paradigm), Kellerman attempts to draw a distinction between acquisition inquiry and language classroom. Judging by the work which followed, he did not succeed in his quest: the studies that followed him did not adopt his distinction. Moreover, "transfer of communication strategies" appeared (cf. Zobl 1980; he also called it "reflexation"). In other words, everything was still claimed to have transferred from L_1 to L_2, but nobody really knew what transfer was (cf. lack of definitions in contemporary papers (e.g., Johnson 1989)).[28] Evidently, subdoxasticity persisted and prevailed.

2. Remarks on moving forward with transfer.

2.1. REFLECTING ON HISTORY. Adopting the term "transfer" itself is not too problematic: it is just a terminological convention. It's the behavioristic content and lack of definition that are problematic.[29] The behavioristic scaffolding of transfer seems to have hindered the clarity and transparency of generative acquisition studies too. As such, the situation with "transfer" is rather dire. In what follows I will very briefly sketch out a proposal to abolish transfer in favor of some other terminological notion. What I stipulate is that there is no need for transfer *per se*. Transfer carries with it a certain connotation of "move-from-one-location-to-another", which might

---

[27] I shall not consider these proposals of Kellerman in further detail because the terms "interenfence" and "cross-linguistic influence" deserve a separate investigation, even though a much more narrow one than that of transfer.

[28] The chief problem, in my eyes, is that nobody tried to define L_1 and L_2.

[29] Notably, syntactic literature also uses the term "transfer" (transfer after phases), but there's not as much confusion about it and it is defined well and clearly.

have caused the researchers to talk about "copying" relationships or different "mental represen-tations" for each of n-ary language acquisition. It is this separation and "transfer" from one place to another one that I dispute. Sharwood Smith (2021) aptly noted that "[the] 'move-from-one-location-to-another' notion is misleading and unnecessary" (413).

I see no reason behind assuming such a distinction between different (groups of) mental rep-resentations. What is it based on? It is entirely possible that the mental representations function as sets, with the corresponding nonempty intersections between those sets.[30] In such a contin-uum, the language of an individual is a set of mental representations of properties relating to one or more languages.[31] For example, V2 parameter for an individual who speaks Russian, Ukrainian, and English, and is acquiring German does not transfer from one language to another one, nor is it distinct for each language. Instead, it is one parameter which lies at the intersec-tion of all four subsets of mental representations. Similarly, the SOV parameter for subordinate clauses is only within the German subset, and overt V-T movement is at the intersection of Ger-man and English. The parametric theory or analogous alternatives are, doubtlessly, central to such a framework.

2.2. MOVING FORWARD FORMALLY.

2.2.1. SUMMARY. In what follows, I essentially propose that acquisition process is not redun-dant. The proposals involves two key ideas. First, the properties (e.g., parameter-settings) instan-tiated in both target and previous language are *never separate*. This means that there are *not two copies* of a pro-drop setting within mental representation of an English-French learner or bilin-gual. Instead, there is a function `Identify` to that end. For those settings not shared by L_1/L_2 (e.g., pro-drop settings for French-German bilingual), I still propose that there are no copies. I in-sist that two settings are integrated in one system with the `Consolidate` function. A *very brief* sketch of this proposal is below.

2.2.2. FORMALISM. The merits of formalization have been discussed elsewhere, we'll omit this discussion (though see Chomsky 1990). Ozernyi (2022a) offers a rough sketch of a formaliza-tion for a non-redundant variant of "transfer" – the function he calls `Identify`, style it $\mathscr{I}$. The exact formalization is as follows (with minor edits):

> Define $E = \{P_1^e...P_n^e \mid P^e$ is an property of internalized grammar$\}$ and define $T = \{P_1^t...P_n^t \mid P^t$ is a target language property to be internalized$\}$. Trivially, at least $E \cap T \neq \emptyset$ and perhaps $E = T$. Let further stipulate that for some P_a, $P_a^e = P_a^t$. Then, the acquisition will proceed by an operation which identifies P_a^e with P_a^t but the resultant sets are applied to respective domains (E vs T), call it `Identify` ($\mathscr{I}$). Then, $\mathscr{I}(P_a^e, P_a^t) = \langle P_a^e, P_a^t \rangle = P_a^{\langle e,t \rangle}$. Trivially, while $P_a^e \in E \backslash T$ and $P_a^t \in T \backslash E$, $P_a^{\langle e,t \rangle} \in E \cap T$.[32]

---

[30] A particular case of this set-theoretic conception is when there is no intersection between sets of L_1 and L_2 (and L_3...) – but I find that this case is repealed through mere existence of language universals. A stronger case comes from recent Chomsky's work which assumes that cross-linguistic variation is limited to externalization. Then, all we have is one set. This a very strong claim with very serious implications for acquisition prompting us to rethink what it is we are trying to investigate.

[31] Depending on the parametric framework one chooses, this could be e.g., parameters, etc.

[32] See the details of the proposal in Ozernyi (2022a).

Ozernyi doesn't clarify, but E is roughly for "existing language" and T for "target language" respectively.[33] The notion of language is purposefully not used in formalization above because it is rejected: it is not clear what the boundaries of "a language" are, so remaining on the level of individual properties – implied parameters – is more desirable. It is also worth noting that $\mathscr{I}$ does not deal with acquisition of new properties of language *per se*: i.e., it says nothing about the acquisition of SOV word order in subordinate clauses of German by English speakers. This was stated via "for some P_a, $P_a^e = P_a^t$". Consider an example. Take G to be a set of "properties" of German, one of them being the aforementioned SOV, and A being the analogous set for English, VSO does not operate within $G \cap A$, hence is out of the domain of $\mathscr{I}$.

Such formalization is easily adjustable for, for example, the Micro-cue model of Westergaard. Property in that case is defined by "abstract piece of linguistic structure". I think this is somewhat underspecified, but it is a viable hypothesis and given the lack of successful theories of cross-linguistic variation, perhaps one of the few available options. Yet, identification is trivially more economical than any "copying" or "transfer" (wholesale, property-by-property, piecemeal, partial, or any other one might come up with): there's just no transfer or copying, only identification which needs to happen for "transfer" anyway in one sense or another. The costs of $\mathscr{I}$, as well as precise mechanisms guiding it are a good question, but many paths are available, one being reward/penalty system, as used in Yang (2000).

Expanding on Ozernyi (2022a) and following Ozernyi (in prep; 2022b), let me also suggest a version of formally sufficient process for acquisition of target properties not instantiated in the set of internalized ("existing" already) ones, i.e. $T \backslash E = \{p : p \in T \land p \notin E\}$. In order to state this formalization, one does need a comprehensive theory of cross-linguistic variation which is readily formalizable. One such theory is Biberauer's emergent parameters. Take, for example, pro-drop parameter (PDP) hierarchy as it is given in Biberauer (2018). I will not repeat it here, but only suggest that a learner of, for example, Icelandic might have the pro-drop parameter settings like those in (4). Alternatively, a learner of German would conform to settings in (5). Both were obtained from Biberauer's tree by changing "yes" to "1" and "no" to "0".
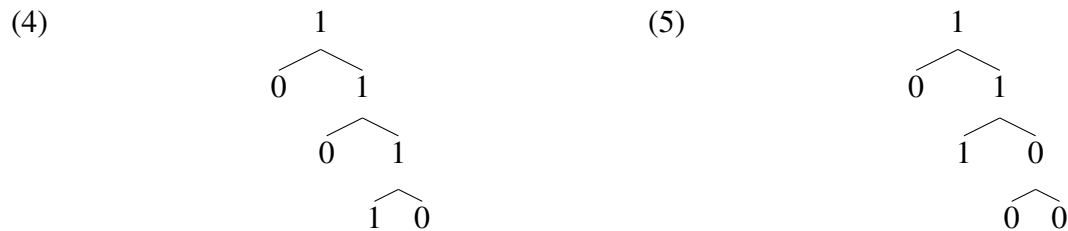
Transfer, whatever it is, can operate in multiple ways, which could be formalized using these trees. For example, wholesale transfer would imply making a copy of (4) and revising the values to (5). Property-by-property transfer would imply copying parts of (4) which coincide with (5), and rebuilding the tree as the acquisition goes along. More elaborate proposals would include both version of transfer at different stages like development vs. initial state, if such stages are taken to exist. All of these could and ought to be formalized elsewhere, and their computational efficiency is to be compared to find the most economical one.

---

[33] That is, E is the internalized grammar at any given point, and T is the input, which means T is only constrained by UG.

I suggest a very different approach. Again, let me state it in a sketch of somewhat informal formalization, complete formalization to follow elsewhere. Take S be a set (or a set of sets) of parameter settings such that it consists of valuations of parameter-settings (zero and one) and covers the entirety of cross-linguistic variation (is constrained by UG). As such, there could be s_{PDP} for PDP and it could be (4) or (5). Trivially, s_{PDP} is a set of values, and it is a structured (and ordered) set, but I will not pursue investigation of either such $s_a \in s_{PDP}$ or ordering of this set(s).[34] For simplicity, let's say $S = \{s_1...s_n : v(s) = 1 \oplus v(s) = 0\}$.[35] Each "property" p from $E \cup T$ above has a set of valuations, like pro-drop above. Trivially, properties of both "existing" and "target" sets of properties have (consist of) parameter-settings, so we obtain respective sets $s(T) = T_S, s(E) = E_S$, etc. Note that there's a bijection between a set of properties and a set of valuations, i.e., no "property" can have two different valuations.[36] This seemingly contradicts the facts that there is cross-linguistic variation, i.e. there *are* different valuations, e.g., (4) and (5). I offer a solution which, rather than stipulating that there are copies,[37] as wholesale or partial or any other transfer seems to do, suggests a more economic way to go.

Suppose that for some properties $(\exists p_E \in E)(\exists p_T \in T)(p_E = p_T)$.[38] That is, the property is the same (e.g., pro-drop). Now stipulate further that their parameter-settings are different (as in (4) vs. (5) above): $(\exists s_E \in E_S)(\exists s_T \in T_S)(s_E \neq s_T)$. So, we get same parameter but different settings. That be the case, a new valuation is created using an operation, call it `Consolidate` (styled $\mathscr{C}$) such that (6).

(6) $\quad \forall s_E \forall s_T (s_E \neq s_T \to (\forall v_{s(E)} \in s_E)(\forall v_{s(T)} \in s_T)$
$$((v_{s(E)} \neq v_{s(T)} \to (\mathscr{C}\{v_{s(E)}, v_{s(T)}\} = \langle v_{s(E)}, v_{s(T)} \rangle)))$$
$$\longleftrightarrow (v_{s(E)} = v_{s(T)} \to (\mathscr{I}\{v_{s(E)}, v_{s(T)}\} = v_{\langle E,T \rangle}))))$$

That is, if there is a set of parameter-settings s_E and it does not match s_T,[39] then it means that there are specific valuations which do not match; further, those valuations consolidate into an ordered pair. Recall also that $v(s) = 0 \oplus 1$. So, if there exists such $v_{s(E)} = 1$ and learner is faced with input such that $v_{s(T)} = 0$, they ought to $\mathscr{C}(1_E, 0_T) = \langle 1_E, 0_T \rangle$.

This could be demonstrated clearly on (4) and (5), because manifestly there are different valuations. The tree, according to $\mathscr{C}$ and $\mathscr{I}$ will look as in (7),[40] and a simplified version of it is in (8) (comparing this to (4) and (5) makes the proposal intuitively clear).

Identification operation does not change anything, it simply identifies value 1 from E with value 1 from T, obtaining $1_{\langle E,T \rangle}$. This is trivial, and ordered pair is largely irrelevant at this point.

---

[34] I do not introduce necessary hierarchy here, but s_{PDP} itself perhaps consists of a set(s) of parameter settings, which in turn consist of valuations.

[35] Below, I will use $v(s_E)$ as $v_{s(E)}$ for convenience.

[36] Approximately $\forall p_a \forall p_b \forall s (p_a \in T \cap E \wedge p_b \in T \cap E \wedge s \in (T \cap E)_S \wedge R(p_1, s) \wedge R(p_b, s) \to p_a = p_b)$.
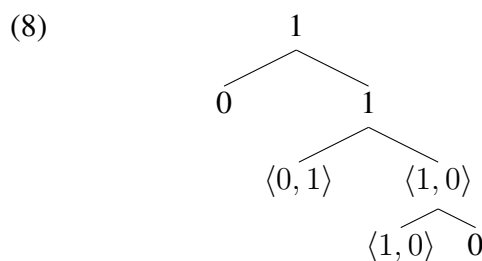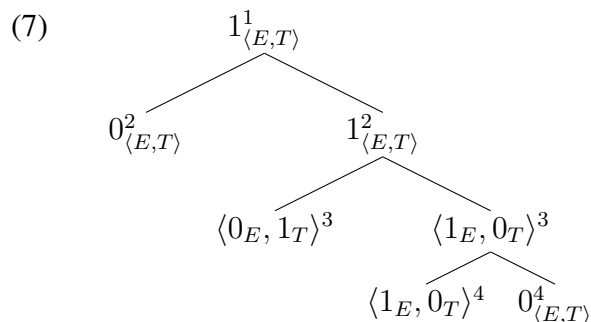
[37] This could be perhaps stated as $\forall p_a \forall p_b \forall s_a \forall s_b (p_a \in T \wedge p_b \in E \wedge p_a = p_b \to \neg R(p_a, p_b) \wedge \neg R(s_a, s_b))$ or alternatively as $\forall E \forall T \forall E_S \forall T_S ((\forall p_a \in E)(\forall p_b \in T)(\forall s_a \in E_S \vee s_b \in T_S)(p_a = p_b \wedge s_a \neq s_b \to \neg R(p_a, p_b) \wedge E_S \cap T_S = \emptyset))$. One property cannot have two parameter-settings in one set, forcing us into two sets.

[38] A keen reader might observe that with Biberauer's hierarchy, $(\nexists p_E \in E)(\nexists p_T \in T)(p_E \neq p_T)$. All of them are instantiated, hence we will need to adjust $\mathscr{I}$ respectively, as in (6). Acquiring one language then means exhausting UG hypothesis space with valuations of 0 or 1. This is a strong and questionable assumption to be explored elsewhere.

[39] While I take it that $(\exists s_E \in T_S)(\exists s_T \in T_S)(\exists p_i \in T \cup E)(R(s_E, p_i) \wedge R(s_T, p_i))$.

[40] The superscripts are added to index level and are only for the reference purposes in fn. 41 below.

Consolidation yields, e.g., $\langle 1_E, 0_T \rangle$ which are ordered and E-setting is used to parse E-input, conversely for T. Of course, none of this happens instantly.[41]

(7)
$$1^1_{\langle E,T \rangle}$$
$$0^2_{\langle E,T \rangle} \qquad 1^2_{\langle E,T \rangle}$$
$$\langle 0_E, 1_T \rangle^3 \qquad \langle 1_E, 0_T \rangle^3$$
$$\langle 1_E, 0_T \rangle^4 \qquad 0^4_{\langle E,T \rangle}$$

(8)
$$1$$
$$0 \qquad 1$$
$$\langle 0,1 \rangle \qquad \langle 1,0 \rangle$$
$$\langle 1,0 \rangle \qquad 0$$

So, the bulk of acquisition could proceed with a match of identifying and consolidating. Further, it is notable that such a view aligns with evidence for cumulativity of acquisition.[42] It also aligns with the ubiquitous evidence that language systems even for different sets of properties ("languages") are integrated: code-switching. Naturally, the formalization above is eclectic and at this point is just a bit of fancy notation, but I hope it conveyed the general idea. Crucially, identification and consolidation are flexible whether the initial system S of parameter-settings is one big tree of which (7) is a fragment, or parameters are not connected (sec. Yang's theory).[43]

It is trivially true that language learning is not a clean process with no mistakes, and factors which lead to both `Identify` and `Consolidate` are not clear, i.e. what are the exact criteria for applying those, etc. I am agnostic that triggers and ability to parse plays a crucial tole here, in the sense of Gibson & Wexler (1994) and Sakas & Fodor (2012). It is likely there's another operation, call it `Update` (styled $\mathscr{U}$) that amends misapplied identification or consolidation based on the input: it is evident misapplication does happen in the acquisition process. Further investigations and clearer, fuller formalizations and computational explorations ought to make all of this clear.[44]

2.3. FUTURE STEPS. Ozernyi (2022b) argues that the history of transfer carries Skinnerian, behavioristic heritage which ought to be done away with. The next step in this pursuit is the ample and rigorous formalization of transfer. It is only in formal guise that various conceptions of transfer can be theoretically evaluated.[45] Such theoretical and empirical (to the extent that algorithmic efficiency serves as a measure of empirical adequacy) evaluations constitute one area of investigation. Another area which I see as particularly valuable is aligning $\mathscr{I}$ and $\mathscr{C}$ suggested here (and $\mathfrak{I}$ generally) with the broader formalization of minimalist syntax, e.g., in Collins & Stabler

---

[41] I am agnostic about the order, but perhaps something along the liens of $\mathscr{I}(0^4_E, 0^4_T) \prec \mathscr{C}(1^4_E, 0^4_E) \preceq \mathscr{C}(1^3_E, 0^3_T) \prec \mathscr{C}(0^3_E, 1^3_T) \preceq \mathscr{I}(1^2_E, 1^2_T) \preceq \mathscr{I}(0^2_E, 0^2_T) \preceq \mathscr{I}(1^1_E, 1^1_T)$. Note that $\prec$ and $\preceq$ are used very loosely with temporal context.

[42] Notably, the more pairs $\langle v_{s(E)}, v_{s(T)} \rangle$ are consolidated, the less $\mathscr{C}$ will apply and the more $\mathscr{I}$ will: the more languages you know, the easier it gets.

[43] It seems to me that the connected version is more likely, but both hypotheses ought to be entertained and empirically tested (if they give rise to different predictions, as they should).

[44] Since the notion of "a language" is rejected and all $\mathscr{I}, \mathscr{C}, \mathscr{U}$ operate continuously, without initial states or anything of the like, call the state (set) $\mathfrak{I}$ of internalized grammar within which these operations take place *the intermediate*.

[45] By various conceptions of transfer I mean wholesale vs. partial vs. "combined" (*sensu* Rothman et al. 2019) vs. none (i.e., cumulativity suggested in this paper).

(2016).[46] It is (some parts of) syntax (construed *very* broadly) that are being acquired, so the formalization of acquisition can only be an extension of the formalization of syntax. Acquisition is that of syntax since syntax is core, and so it is presumptuous to separate the two by attempting investigations of acquisition without (at least to some degree) solid conception of syntax.[47] We lack such a conception presently which attests to the fragility of acquisition studies: a limitation to be acknowledged and to be *acutely* aware of.

References

Biberauer, Theresa. 2018. Pro-drop and emergent parameter hierarchies. In Federica Cognola & Jan Casalicchio (eds.), *Null subjects in generative grammar*. Oxford: Oxford University Press.

Bretscher, Otto. 2018. *Linear algebra with applications*. London: Pearson.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1990. Topic...comment: On formalization and formal linguistics. *Natural Language & Linguistic Theory* 8(1). 143–147. https://doi.org/10.1007/BF00205534.

Collins, Chris & Edward Stabler. 2016. A Formalization of Minimalist Syntax. *Syntax* 19(1). 43–78. https://doi.org/10.1111/synt.12117.

Dušková, Libuše. 1969. On sources of errors in foreign language learning. *International Review of Applied Linguistics in Language Teaching* 7(1). 11–36. https://doi.org/10.1515/iral.1969.7.1.11.

Epstein, Samuel David, Suzanne Flynn & Gita Martohardjono. 1996. Second language acquisition: Theoretical and experimental issues in contemporary research. *Behavioral and Brain Sciences* 19(4). 677–714. https://doi.org/10.1017/S0140525X00043521.

Falk, Ylva, Christina Lindqvist & Camilla Bardel. 2015. The role of L1 explicit metalinguistic knowledge in L3 oral production at the initial state. *Bilingualism: Language and Cognition* 18(2). 227–235. https://doi.org/10.1017/S1366728913000552.

Flynn, Suzanne. 1987. *A parameter-setting model of L2 acquisition: Experimental studies in anaphora*. Dordrecht: Springer. https://doi.org/10.1007/978-94-009-3747-5.

Flynn, Suzanne. 2021. Microvariation in multilingual situations: The importance of property-by-property acquisition: Pros and cons. *Second Language Research* 37(3). 481–488.

Fries, Charles Carpenter. 1925. The periphrastic future with *shall* and *will* in Modern English. *Publications of the Modern Language Association (PMLA)* 40(4). 963–1024. https://doi.org/10.2307/457534.

Fries, Charles Carpenter. 1927. The meanings of words. *The English Journal* 16(8). 602–606. https://doi.org/10.2307/803491.

Fries, Charles Carpenter. 1945. *Teaching and learning English as a foreign language*. Ann Arbor: University of Michigan Press.

Fries, Charles Carpenter & Kenneth L. Pike. 1949. Coexistent phonemic systems. *Language* 25(1). 29–50. https://doi.org/10.2307/409907.

---

[46] Albeit perhaps with some minor adjustments or emendations of the mentioned formalization, particularly on the use of Kuratowski's definition of ordered pair for labelling and generally.

[47] Yet, as acknowledged in fn. 30, it might be that acquisition is not really that of syntax: this would be the case if cross-linguistic variation was something syntax-external (again, broadly construed) so as to be learnt rather than acquired (whether Lado-style, skill learning, or belief fixation). Then $L_{2...n}$ learning is relevant to E-language as opposed to I-language and is linguistically uninteresting. A valid hypothesis, this is to be investigated further.

Gass, Susan M. 1988. Second language acquisition and linguistic theory: The role of language transfer. In Suzanne Flynn & Wayne O'Neil (eds.), *Linguistic theory in second language acquisition*, 384–403. Dordrecht: Springer. https://doi.org/10.1007/978-94-009-2733-9_21.

Gibson, Edward & Kenneth Wexler. 1994. Triggers. *Linguistic Inquiry* 25(3). 407–454.

Jakobovits, Leon A. 1969. Second language learning and transfer theory: A theoretical assessment. *Language Learning* 19(1-2). 55–56. https://doi.org/10.1111/j.1467-1770.1969.tb00933.x.

James, William. 1890a. *The principles of psychology*, *Vol. 1*. New York: Henry Holt.

James, William. 1890b. *The principles of psychology*, *Vol. 2*. New York: Henry Holt.

Johnson, Janice. 1989. Factors related to cross-language transfer and metaphor interpretation in bilingual children. *Applied Psycholinguistics* 10(2). 157–177. https://doi.org/10.1017/S014271640000850X.

Kellerman, Eric. 1977. Towards a characterisation of the strategy of transfer in second language learning. *Interlanguage Studies Bulletin* 2(1). 58–145.

Kellerman, Eric. 1979. Transfer and non-transfer: Where we are now. *Studies in Second Language Acquisition* 2(1). 37–57. https://doi.org/10.1017/S0272263100000942.

Lado, Robert. 1949. The relation of entrance level to rate of progress in aural comprehension. *Language Learning* 2. 105–112. https://doi.org/10.1111/j.1467-1770.1949.tb01161.x.

Lado, Robert. 1951. Testing control of the structure of a foreign language. *Language Learning* 4(1-2). 17–35. https://doi.org/10.1111/j.1467-1770.1951.tb01180.x.

Lado, Robert. 1956. A comparison of the sound systems of English and Spanish. *Hispania* 39(1). 26 – 29. https://doi.org/10.2307/335188.

Lado, Robert. 1957a. *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor: University of Michigan Press.

Lado, Robert. 1957b. Sentence structure. *College Composition and Communication* 8(1). 12–16. https://doi.org/10.2307/354410.

Lust, Barbara C. 2006. *Child language: Acquisition and growth.* Cambridge: Cambridge University Press.

Marckwardt, Albert H. 1946. Phonemic structure and aural perception. *American Speech* 21(2). 106–111. https://doi.org/10.2307/486482.

McGeoch, John A. 1942. *The psychology of human learning: An introduction*. Oxford, UK: Longman.

Ozernyi, Daniil M. 2021. Jason Rothman, Jorge González Alonso, & Eloi Puig-Mayenco: Third language acquisition and linguistic transfer. Cambridge University Press, 2019. *Applied Linguistics.* https://doi.org/10.1093/applin/amab040.

Ozernyi, Daniil M. 2022a. Linguistic transfer, or there and back again: A chronological study of terminological meandering. Poster presented at the 96th Annual Meeting of the LSA. https://doi.org/10.13140/RG.2.2.27507.86560.

Ozernyi, Daniil M. 2022b. The rise and fall of linguistic transfer. Manuscript. Northwestern University. https://doi.org/10.5281/zenodo.6237257.

Politzer, Robert L. 1965. Some reflections on transfer of training in foreign language learning. *International Review of Applied Linguistics* 3(3). 171–177.

Priestley, Joseph. 1790. *Hartley's theory of the human mind, on the principle of the association of ideas; with essays relating to the subject of it* (2nd ed.). London: J. Johnson.

Puig-Mayenco, Eloi, Jorge González Alonso & Jason Rothman. 2020. A systematic review of transfer studies in third language acquisition. *Second Language Research* 36(1). 31–64.

Rothman, Jason, Jorge González Alonso & Eloi Puig-Mayenco. 2019. *Third language acquisition and linguistic transfer* (Cambridge Studies in Linguistics 163). Cambridge: Cambridge University Press.

Rugg, Harold Ordway. 1916. *The experimental determination of mental discipline in school studies*. Baltimore: Warwick & York.

Sakas, William Gregory & Janet Dean Fodor. 2012. Disambiguating syntactic triggers. *Language Acquisition* 19(2). 83–143. https://doi.org/10.1080/10489223.2012.660553.

Schachter, Jacquelyn. 1974. An error in error analysis. *Language Learning* 24(2). 205–214. https://doi.org/10.1111/j.1467-1770.1974.tb00502.x.

Schwartz, Bonnie D. & Rex A. Sprouse. 2021. The Full Transfer/Full Access model   and L3 cognitive states. *Linguistic Approaches to Bilingualism* 11(1). 1–29. https://doi.org/10.1075/lab.20055.sch.

Selinker, Larry. 1972. Interlanguage. *International Review of Applied Linguistics in Language Teaching* 10(1-4). 209–232. https://doi.org/10.1515/iral.1972.10.1-4.209.

Sharwood Smith, Michael. 2021. Language transfer: A useful or pernicious concept? *Second Language Research* 37(3). 409–414. https://doi.org/10.1177/0267658320941035.

Sherkina-Lieber, Marina, Ana T. Perez-Leroux & Alana Johns. 2011. Grammar without speech production: The case of Labrador Inuttitut heritage receptive bilinguals. *Bilingualism: Language and Cognition* 14(3). 301–317. https://doi.org/10.1017/S1366728910000210.

Spratt, Mary, Alan Pulverness & Melanie Williams. 2011. *The TKT Course* (2nd ed.). Cambridge & NewYork: Cambridge English.

Stephens, J. M. 1960. Transfer of learning. *Encyclopedia of Educational Research* 1535–1543.

Stockwell, Robert P., J. Donald Bowen & John W. Martin. 1965. *The grammatical structures of English and Spanish contrastive structures*. Chicago: University of Chicago Press.

Taylor, Barry P. 1975. The use of overgeneralization and transfer learning strategies by elementary and intermediate students of ESL. *Language Learning* 25(1). 73–107. https://doi.org/10.1111/j.1467-1770.1975.tb00110.x.

Wenk, B. 1974. Interference. Unpublished paper.

Westergaard, Marit. 2021a. Microvariation in multilingual situations: The importance of property-by-property acquisition. *Second Language Research* 37(3). 379–407. https://doi.org/10.1177/0267658319884116.

Westergaard, Marit. 2021b. The plausibility of wholesale vs. property-by-property transfer in L3 acquisition. https://doi.org/10.1075/lab.20091.wes.

Winch, W. H. 1908. The transfer of improvement in memory in school-children. *British Journal of Psychology, 1904-1920* 2(3). 284–293. https://doi.org/10.1111/j.2044-8295.1908.tb00180.x.

Woodworth, R. S. & E. L. Thorndike. 1901. The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review* 8(3). 247–261. https://doi.org/10.1037/h0074898.

Yang, Charles. 2000. *Knowledge and learning in natural language*. Cambridge, MA: MIT Dissertation.

Zobl, Helmut. 1980. Developmental and transfer errors: Their common bases and (possibly) differential effects on subsequent learning. *TESOL Quarterly* 14(4). 469–479. https://doi.org/10.2307/3586235.