Language transfer: a useful or pernicious concept?

Michael Sharwood Smith
Heriot-Watt University, UK

Abstract
Westergaard’s microcue account raises the question of the exact nature of language transfer in the acquisition of languages as well of how L1/Ln input interacts with the principles of universal grammar (UG) during processing. In order to consider in more detail the actual representation building, processing mechanisms that would be involved, her approach will be spelled out in terms of the Modular Cognition Framework (see Sharwood Smith, 2017; Sharwood Smith and Truscott, 2014; Truscott and Sharwood Smith, 2004, 2019). As well as identifying the way in which microcues would be constructed, this analysis in particular will have the effect of undermining the concept underlying the transfer metaphor as a way of characterizing crosslinguistic influence.

Keywords
crosslinguistic influence, L3, language processing, language representation, language transfer, modular cognition

I Introduction
There are two salient themes that characterize Westergaard’s discussion (Westergaard, 2021). The first is the nature of language transfer, a problematic concept in the literature that never gets resolved. The second theme is the interaction between input of first language and additional languages (L1/Ln) and Universal Grammar (UG). Given the similarity that Westergaard acknowledges between her approach and the approach to cognitive representation and processing developed by Sharwood Smith and Truscott, called the Modular Cognition Framework (MCF), I will interpret her approach using MCF architecture (see Sharwood Smith, 2017; Sharwood Smith and Truscott, 2014; Truscott and Sharwood Smith, 2004, 2019). This will permit her line of argument to be spelled out in terms of the representation-building, processing mechanisms involved. The spotlight
will be on crosslinguistic influence, but I also consider how Westergaard’s UG-plus-microcues account would operate within this framework.

II Transfer dismissed

Westergaard has her own reservations about transfer as a concept but is still content to keep the term as a handy metaphor. In line with Sharwood Smith and Truscott (2006), I take a more extreme position. Whether meant literally or metaphorically, the concept is every bit as misleading as it is handy. Unease with the term is not new. Back in the 1970s, the proponents of creative construction downplayed its relevance, at least as regards grammar, associated as it was with behaviourism (Dulay et al., 1982: 96ff). Nonetheless, albeit in different ways, it was maintained by others as a cognitivist means of describing interactions between new and established language systems in the mind (Selinker, 1972; White, 1985). Unease with the concept persisted, however. The term ‘crosslinguistic influence’ was introduced to cover all kinds of such interaction including avoidance of interaction without actually rejecting the transfer term itself (Sharwood Smith, 1983). More recently, as Westergaard mentions, the underlying concept was subjected to criticism by Sharwood Smith and Truscott (Sharwood Smith and Truscott (2006) and I will reiterate its basic argument.

What, then, is so very wrong about transfer, even as a metaphor? First, transferring in any sense can hardly mean switching a grammatical property from one location to another since it deprives the host grammar of precisely that property. At the very least, it must mean ‘copying’ or ‘cloning’ a property that can then be used independently to build the new grammar without ‘damaging’ the old one. Otherwise, if the identical property serves both grammars then it cannot have been ‘transferred’. This property must exist in a system that makes it simultaneously available to the grammars of any language system being acquired: no transfer of properties is required.

The reason why the transfer concept has lasted so long, apart from its graphical convenience, might be due to the clear separation that has persisted between processing, on the one hand, and representation on the other. What might be called the ‘P/R distinction’ has influenced both thinking and research specialization. Despite recent trends taking processing seriously in studies still mainly concerned with representation, this barrier between apparently independent dimensions of linguistic development persists. Its persistence is traceable to Chomsky’s original competence/performance distinction (Chomsky, 1965). For many years this made processing-based investigations and development through time appear irrelevant for the pursuit of generative research. Developmental studies, however, cannot avoid dealing with real time and, arguably, the long-standing reluctance to include processing in generative explanations of development has proved to be unreasonably restrictive. Contributing factors may have included lack of interest and expertise in processing and/or the fact that compatible psycholinguistic ‘P’ research has lagged behind progress in linguistic ‘R’ research.

The question now is: Is it logically possible to maintain the P/R distinction in accounts of acquisition and so treat processing and representation as two interdependent aspects of the same, ‘P+R’ developmental account? The answer from an MCF point of view is resoundingly positive (see Jackendoff, 2002: 200–30).
The moment that processing is taken seriously in accounts of language acquisition, it becomes incumbent on researchers in this area to specify in detail the relationship between representations and what happens to them online when exposure to utterances in L1/L2/Ln causes them to change, however minimally, as a consequence. This means a necessary step down in level of abstraction. Looking exclusively at linguistic structural outcomes permits linguistic properties and constructions to be described completely *in abstracto* hence (1) ignoring considerations of time and space and, thereby, (2) licensing the free use of temporal and spatial metaphors like ‘transfer’. However, given the necessary close relationship between development and online performance, it becomes counterproductive not to consider the growth of grammatical representation as an outcome of processing in real time. This still means developing explanations concerning representation at a certain level of abstraction but down one notch from the level maintained when focusing purely on linguistic-theoretic aspects of developmental grammars (see related discussion in Marr, 1982: 25; Sharwood Smith, 2015, 2019). This means that abstract representational properties themselves may need some redescription to permit detailed accounts of their development over time and, suddenly, the notion of ‘transferring’ properties seems to require a new explanation as to how movement of ‘R’ elements between locations would be manifested in a time-sensitive ‘P’ system.

A much more feasible alternative in the light of current thinking about mind and brain is to express crosslinguistic influence as a matter of alternative connections within a network. Grammatical properties would then be ‘untransferable’: that is, they would neither admit of movement nor of copying elsewhere. Furthermore, properties located in what we only as outside observers think of as separate grammatical systems will exist within a single network. Describing the relationships between, say, L1, L2 and L3 grammars in this way allows a more realistic representation of how a modular mind accommodates different grammars within one language-neutral grammar but still interconnected within that system in different ways.

**III MCF in a nutshell**

MCF processing and representational architecture has been extensively described elsewhere. The advantages of conducting research within such a framework that is based on a network of functionally specialized systems, each with their own processor and (memory) store, were summed up in Sharwood Smith (2019) roughly as follows. It provides (1) an integrated account of representation and processing while keeping the two concepts distinct; (2) a developmental theory for any type of cognition; (3) a set of clearly defined mechanisms for both within-system and cross-system processing operations; and (4) an open architecture that allows the embedding of theoretical and empirical research from a range of different disciplines within a shared framework. The compatibility question here would be as follows: How plausible would Westergaard’s proposals for L3 acquisition look when formulated in MCF terms? To answer this, I will follow the Sharwood Smith (2019) account of grammatical gender by treating development simplistically as a unidirectional set of stages: input, development (acquisition) and consolidation but still keeping in mind the essentially to-and-fro character of both parallel incremental processing and consequently development itself as well.
IV Developing representations

In the input stage, environmental stimuli begin to regularly provide perceptual (visual and auditory) systems with evidence of syntactic features. This environmental input may count as evidence for outside observers, but it is input to perceptual systems alone: these will respond by matching patterns of stimuli with representations written in their own particular (e.g. visual or auditory) code. Each cognitive system has its own construction principles instantiated in its processor; any structural property, however combined, will therefore conform to these principles. The visual system, for example, produces only representations that reflect properties and principles of human vision. In the case of phonology and syntax, the processors are governed by UG, meaning that the idea behind UG actually reflects the biological norm for all cognitive processing.

Now take spoken input: speech events will provide input to the auditory system, which will try to represent any (perceptible) type of sound in the form of activated auditory ‘structures’ (i.e. representations/properties). This auditory system has an interface with the phonological system: any currently activated auditory representations (AS) will provide input via that interface. This is the stage where linguistic (phonological or syntactic) development may begin: any phonological representations (PS) that are activated will trigger a parallel response from the syntactic system potentially initiating syntactic development.

The other system adjoining syntax, the conceptual system, will be independently generating conceptual (meaning) structures: responding to inputs from various systems including the syntactic system. The forward/backward flow of inputs building associations over the various interfaces is depicted here, with the specifically linguistic systems marked in boldface. Note that in each case the final ‘S’ in AS, PS, SS and CS stand for ‘system’ but in appropriate contexts may also stand for ‘structure’:

\[ \text{acoustic input} \rightarrow \text{AS}\leftrightarrow\text{PS}\leftrightarrow\text{SS}\leftrightarrow\text{CS} \]

The syntactic system will respond to its inputs by activating a number of candidate structures within its syntactic store, which will enter into competition with one another producing, temporarily activated to different degrees, a number of competing candidate chains of structure across the four systems.

Any syntactic structure will include syntactic properties that we, again only as outside observers, would associate with one or more of the grammars in question (L1, L2 or L3), but here they just happen to have been, or should be associated with, particular phonological representations (PS). These PSs, we (again only as outside observers) can identify as belonging one or other or all these three languages, e.g. /ike/ (Norwegian), /not/ (English) and /niçt/ (German).

Each SS property that has been activated will also coactivate associated representations in any systems with which it shares an interface: here, this means that not just phonological but also conceptual representations (CS). Note crucially, that in this interaction between different systems, no information is carried across the interfaces. Interfaces operate only to associate and coactivate particular structures in adjoining stores. Nevertheless, identifying crosslinguistic influence remains perfectly possible. While no representation (feature, property) like V2 or masculine or Negative can be ‘transferred’ or ‘copied’, it can be shared by what we linguists recognize as different
linguistic systems (L1/L2/Ln). The possibility still remains of that shared property being
assembled with other syntactic properties in different combinations.

In the course of acquiring new language, as utterances in the target language are pro-
cessed, the flow of inputs described above will lead to the parallel activation of AS/PS/
SS/CS chains in working memory. Candidate chains compete for the best fitting combi-
nations. Inevitably those structures in any of these four systems, which happen to be the
most ‘consolidated’, i.e. in MCF terms those ones that have achieved via regular process-
ing the highest resting activation level, will have a relatively better chance of winning
while they remain in working memory. In an L3 scenario, this includes both L1 and L2
syntactic properties and especially ones that shared by both Ls because, by dint of fre-
quent activation, these will be particularly competitive: this would explain apparent
‘typological’ effects, even though the syntax system remains blind to language identity.
Only current activation levels count. It will then be not at all surprising that – as the
learner’s mind assembles different configurations of syntactic properties when process-
ing L3 – crosslinguistic influence from either L1 or L2 will manifest itself (see also
Truscott, 2006). While any property, reflecting Westergaard’s Full Transfer Potential
principle, can in principle be shared, what is actually shared will depend on the outcome
of competition between candidates in this typologically neutral arena. Westergaard’s
‘potential’ relates to the ability of any property to take part in processing and develop-
ment irrespective of its observer-based typological identity.

A great deal of MCF-related detail has been left out of this article for reasons of space,
but it should begin to be clear that MCF can supply the appropriate mechanisms for what
we have misleadingly tended to call ‘transfer’ in a way that permits formulation of
Westergaard’s analysis and proposals. It also underlines why the ‘move-from-one-loca-
tion-to-another’ notion is misleading and unnecessary.

Development is driven following the MCF equivalent of Westergaard’s ‘learning by
parsing’. In the case of syntax, the syntax processor will match its inputs from adjoining
systems by activating various combinations of syntactic structure. What principles con-
trol how this processor operates are defined by the chosen theory of UG, so in this case
it would be Westergaard’s microcue version. Whatever metaphors that informed observ-
ers of the output may use to characterize parsing, the syntactic processor itself makes no
generalizations and no predictions but automatically applies its principles in response to
the continuing flow of input by assembling matching structures (SS) in its working mem-
ory in line with UG and, following Westergaard’s account, constructing microcues.
Inevitably, in the early stages well-established (frequently activated) configurations of
SS – whatever language or dialect they happen to ‘belong’ to – will have a competitive
advantage until repeated input processing builds up the resting activation levels of the
optimal configurations thus ‘consolidating’ them. In other words, the continual elabora-
tion of microcues, as Westergaard characterizes syntactic development, in line with her
‘Full Crosslinguistic Potential’ principle, now rephrased, is eminently expressible using
MCF architecture.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship,
and/or publication of this article.
Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD
Michael Sharwood Smith | https://orcid.org/0000-0002-7477-8283

References
Chomsky N (1965) *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
Dulay H, Burt H, and Krashen K (1982) *Language two*. Oxford: Oxford University Press.
Jackendoff R (2002) *Foundations of language*. Oxford: Oxford University Press.
Marr D (1982) *Vision: A computational approach*. San Francisco, CA: WH Freeman.
Selinker L (1972) Interlanguage. *International Review of Applied Linguistics* 10: 209–31.
Sharwood Smith M (1983) Cross-linguistic aspects of second language acquisition. *Applied Linguistics* 4: 192–99.
Sharwood Smith M (2015) On virtual versus real spatio-temporal explanations of linguistic development. *Linguistic Approaches to Bilingualism* 5: 537–40.
Sharwood Smith M (2017) *Introducing language and cognition: A map of the mind*. Cambridge: Cambridge University Press.
Sharwood Smith M (2019) Representing representation: Three perspectives and the case of grammatical gender. In: Slabakova R, Corbet J, Dominguez L, Dudley A, and Wellington A (eds) *Explorations in second language processing and representation*. Cambridge: Cambridge Scholars Publishing, pp. 2–40.
Sharwood Smith M and Truscott J (2006) Full transfer full access: a processing oriented interpretation. In: Unsworth S, Parodi T, Sorace A, and Young-Scholten M (eds) *Paths of development in L1 and L2 acquisition: In honor of Bonnie D. Schwartz*. Amsterdam: John Benjamins, pp. 201–16.
Sharwood Smith M and Truscott J (2014) *The multilingual mind: A modular processing perspective*. Cambridge: Cambridge University Press.
Truscott J (2006) Optionality in second language acquisition: A generative, processing-oriented account. *International Review of Applied Linguistics in Language Teaching* 44: 311–30.
Truscott J and Sharwood Smith M (2004) Acquisition by processing: A modular perspective on language development. *Bilingualism: Language and Cognition* 7: 1–20
Truscott J and Sharwood Smith M (2019) *The internal context of bilingual processing and acquisition*. Amsterdam: John Benjamins.
Westergaard M (2021) Microvariation in multilingual situations: The importance of property-by-property acquisition. *Second Language Research* 37(3): 379–407.
White L (1985) The acquisition of parameterized grammars: Subjacency in second language acquisition. *Second Language Research* 1: 1–17.