INTRODUCTION

As a first approximation, modular approaches to the study of bilingualism attempt to analyze the relevant cognitive components that make up a person's knowledge of two languages and his or her ability to use them. The objective would be to describe in what way, in fact, these components might be characterized as autonomous domains and how they interact with other components. Far from a unified approach to the problem of modeling linguistic competence and language proficiency, modularity is a concept that has given rise to a great diversity of views. Perhaps, as a blessing in disguise, this divergence could favor a greater openness for the coming together of new lines of discussion, especially in applied linguistics where major theoretical differences can often be temporally set aside. Why this is true is an interesting question in its own right, I suspect, having something to do with the modularity concept itself. Be that as it may, the study of bilingualism will serve to provide us with some limited common ground to explore the possibilities.

EARLY DEVELOPMENTS AND MAJOR CONTRIBUTIONS

The best known early formulation of the basic concept of a modular mental architecture is Fodor (1983). Subsequently, the idea was introduced to the study of bilingualism; Bialystok & Cummins (1991) proposed the discussion of a research program, based on Jackendoff's (1987) Representation-based modularity thesis, for the purpose of orienting work toward a more "differentiated conception" of bilingual development. Since then, researchers have found it useful for working on a wide range of theoretical and applied problems, the ensuing dispersion of modular hypotheses and proposals offering an interesting framework for mutual exchanges that at first might seem unpromising (Coltheart, 1999; Garfield, et al., 2001; Carruthers, 2002; Marcus, 2004). Specifically within the field of second language learning and bilingualism, a new interest in modular approaches has clearly emerged (Herschensohn, 2000; Foster-Cohen, 2001; Paradis, 2004; Sharwood-Smith, 2004, to mention only a few). Following Jackendoff's (2002) most recent update of a modular Tripartite Parallel Architecture (TPA), Francis (2004) has proposed a Bilingual TPA, an elaboration, at the same time, of Cummins' (2000) Common Underlying Proficiency model.
For our purposes, the concept of modularity might be most easily understood from the point of view of analyzing a given ability, examining the different aspects of a bilingual ability, for example. If by "aspects" we simply mean the various ways that a single undifferentiated and internally homogenous bundle of behavior patterns can be studied, this approach might be characterized as integrativist or holistic. A similar view would conceive of language ability (as a cognitive entity) as uniformly structured through and through, internally equidimensional in all respects (e.g., in some versions, the distinction between conceptual domains and grammatical knowledge would be questioned). Strong holistic theories, then, almost by definition, would be anti-modularist. For instance, in the field of literacy, including bilingual/second language literacy, a strong holistic approach would reject the idea that reading ability should, or even can, be analyzed into interacting component parts (e.g., subskills related to phonological knowledge, processing of phoneme-grapheme correspondence patterns, other aspects of word identification, sentence processing, text comprehension, etc.). If, on the other hand, the "aspects" of an ability correspond to actual cognitive components and networks of mental structure, autonomous in some way one from the other, and interconnected by specialized interfaces, then we are considering a modular-type approach. The idea of component then would imply a degree of autonomy, the constituent or module in question being domain-specific in some way. What characterizes this kind of componental mental architecture is internal heterogeneity and specialization. An ability, then, could be broken down analytically into dedicated knowledge structures and specialized processing mechanisms. In contrast to total holistic integration, modularity conceives of interactivity within a network of component structures. In a modular system (when it is engaged, in language use), interaction among components is not unconstrained; not all components and structures are open, in the sense of "penetrable," in an unrestricted way to other knowledge domains and processing interfaces. For example, what are commonly known as "top-down" influences on lower level processes (e.g. decoding) are held in check or confined to some degree, depending on the nature of the interaction in question.

A few informal examples might help to make the notion of modularity more concrete. We could compare the abilities of two bilingual high school students in a hypothetical 10th grade dual language immersion class, both of whom arrived from their native country three years ago, but from very different kinds of educational system. Student A demonstrates exceptionally high levels of mastery in tasks related to academic-type discourse ability. For example, in his second language (L2) he can produce, coherently and skillfully, a complex narrative with multiple characters and embedded story lines; but at the "sentence level" it is evident that his knowledge of the L2 grammar is rudimentary. Comparing this same ability in Student B's first language (L1), exactly the converse profile is apparent: flawless grammar and rudimentary mastery of
narrative ability. University students sometimes comment on this kind of "double dissociation" in comparing their favorite professor, a non-native speaker of the language of instruction, and the native-speaking professor whose lectures are hard to follow (not because of faulty grammar or difficult pronunciation).

Let us compare student A again, but now to a native speaker of his L2. Since he had studied the L2 during the primary and secondary school years, previous to intensive L2 learning in an academic setting in the new country, he has built up a sizeable "vocabulary," more extensive in fact than what is typical of native speakers of the language. However, while the absolute number of lexical entries is large, at the "word level" he still experiences persistent difficulties with inflectional affixes and other aspects of grammar related to the structure of words.

Student C, a native-speaker if A's L2, commands error-free mastery of L1 inflectional morphology, as we would expect, but for reasons that are also easy to explain has access to a "smaller vocabulary." But independent of the absolute number of entries, and in contrast to the L2 learner, each entry in the native speaker's lexicon is more complete and well-formed; see van de Craats (2003) for a discussion of what "lexical knowledge" consists of, in L1 and L2.

A general model of bilingual competence and bilingual proficiency should be able to account for the many ways in which its components appear both autonomous each one from the others and closely interconnected, from three points of view: how knowledge of language and how ability are represented, developmentally in language acquisition, and in regard to how language is processed:

1. How do aspects of the strictly linguistic components of proficiency reveal themselves to be independent of general conceptual knowledge, and at the same time how do aspects of language and general cognition appear to be interdependent? A modular perspective on bilingualism could conceive of two dimensions of differentiation/autonomy: one between the linguistic (grammatical) representations of each language system, and the other between a shared conceptual system (which is hypothesized to be non-linguistic in nature) and the linguistic systems.

2. In child development we need to account for why in some circumstances and in regard to some aspects one linguistic system does not appear to affect the other, and why in other circumstances there seems to be a prolific interaction, influences of L_a on L_n, L_n on L_a. Why under similar conditions is there evidence for both balanced development of L_a and L_n, and unbalanced development (in reference to both acquisition and erosion)? In child bilingualism especially, it is necessary to distinguish between the relationship between L1 and L2, on the one hand, and the possibility that neither language that the bilingual knows is "first," "primary," or "dominant." Thus, L_a and L_n refer to the language systems of the bilingual in situations where "first" or "second" either cannot be determined (e.g., because
both are "L1"), or where the distinction between "L1" and "L2" is not relevant to a given analysis (sometimes it is).

(3) In processing and normal language use, how are bilinguals able to keep their languages separate, and how are they able to allow for the two systems to combine and interact? How, when combining and mixing, is this accomplished in a grammatically systematic way? And how do general cognitive operations intervene in different ways in language use that sometimes reveals separation and at other times a close interaction?

It is important to keep in mind that a number of proposed properties of modular systems are still the subject of ongoing research:
• the extent to which they are innately specified,
• the degree to which they are informationally encapsulated,
• how they are represented neurologically and subject to selective breakdown, and
• what accounts for the fast and automatic processing in some domains and controlled processing in others.

As was suggested earlier, the fact that there are many empirical questions still to be resolved is a good thing for continued dialogue in the language sciences.

**WORK IN PROGRESS: BILINGUALISM IN EXCEPTIONAL CIRCUMSTANCES**

When complex systems suffer breakdowns or respond to external pressures, what before appeared to function as a completely integrated unit reveals itself to be internally structured in a different way, unnoticed when all subsystems operate in relative equilibrium. Compared to monolingual competence and performance, bilingualism affords more opportunities for examining the components of language because aside from simply being a more complex system, during development and use it seems to be more susceptible to different kinds of imbalance and tension. Two remarkable and celebrated cases of exceptional bilingualism give us a glimpse into how language development proceeds under circumstances in which processing must be shifted to another modality, and in which development is abnormal.

Out of a peculiar multilingual environment in Nicaragua, including spoken languages, homesign systems and sign pidgins, a generation of previously isolated deaf children gave birth to a fully-formed creole, Idioma de Signos Nicaragüense (ISN) (Nicaragua Sign Language). Apparently without access to a complete language acquisition model, ISN came forward in a rapid two stage development: (Stage 1) consolidation of a peer-group pidgin, more advanced than the primitive home-sign gestural system that children had attained prior to enrollment in boarding school. (Stage 2) From the system of shared signs and rudimentary grammar, a distinct, fully-formed sign language emerged. Significantly, the qualitative leap toward the creole ISN occurred when the peer-group pidgin
became the input to young children. Investigators concluded that with a community of homesigners (pre-pidgin stage), favorable conditions are created that give rise to a new system that is sufficiently rich to support the emergence of a full-fledged creole if sufficient exposure is available during the critical period of language acquisition. The most significant finding, however, was that among the boarding school students who received Stage 1 input, it was the youngest group (immersion before age 7) that was the most successful in surpassing the impoverished model. Older children (8 - 14 years at exposure) made considerable progress, but consistently attained levels of mastery that were less native-like. Late immersion students (after 15) showed permanent deficits typical of early language deprivation (Kegl et al., 1999).

Consistent with previous findings of strong critical period effects among deaf children, this line of research points to the operation of specialized language acquisition mechanisms. Here we would have an example of domain-specificity: how "structure-seeking" or "structure-sensitive" modules, dedicated to processing input specific to natural language, build linguistic competence under restricted maturational conditions, in a different way than general learning proceeds. Specialized acquisition components are programmed for ensuring acquisition under severe conditions of "stimulus poverty" (within certain limits) and within a preset developmental window; see Newport et al. (2001) and Senghas & Coppola (2001) for more discussion. Mayberry & Lock (2003) make the important connection between the failure to develop grammatical competence within the critical period and subsequent degraded language acquisition capacity, crucially for deaf children who typically must attain mastery of a spoken language system in addition to sign language. Unlike other kinds of knowledge, normal attainment of linguistic knowledge appears to be dependent on domain-specific acquisition processors that both obey special restrictions (maturationally) and are specially designed to impose complex structure on seemingly incomplete input.

The second case of exceptional bilingualism involves a single individual as opposed to an entire speech community. Smith & Tsimpli (1995) detail an extensive assessment of polyglot savant Christopher's language abilities. Our interest here is to try to account for sharp imbalances and atypical dissociations. But ultimately, as in the case of the atypical development of ISN among Nicaraguan deaf children, findings from these kinds of case study should help us better understand normal language development. After all, every child, monolingual or bilingual, must overcome the Poverty of Stimulus problem (on a different scale perhaps), and all second language learners, polyglots included, reveal interesting performance imbalances. The "savant" side of Christopher's language ability profile included 16 second languages (in a number of which he had attained intermediate to advanced proficiency). In stark contrast, serious deficits were apparent from performance on tasks related to the use of language for elementary problem solving and communicative tasks that involve increased
processing demands: resolving contradictions, discourse level comprehension and expression (e.g. simple translation), and Theory of Mind tests typically passed by young children. According to the researchers, how some components of language ability come to be super-endowed, and others either irremediably deficient or spared from impairment is impossible to explain from a holistic perspective. In what sense can we speak of an adolescent who in such a short time has mastered one native language and 16 second languages (at last count) as suffering from defective intellectual capacity? Only from a modular point of view would the question even make sense at all. While we might admit that this case is certainly highly exceptional, the pattern of imbalances in consistent with a number of studies that have described how linguistic knowledge can develop normally as general cognitive development is radically impaired (Rondal, 1995). Studies of Specific Language Impairment (SLI), on the other hand, provide evidence for completing the double dissociation: normal intelligence and defective grammar.

**PROBLEMS OF ANALYSIS: TWO DIMENSIONS OF MODULARITY**

Returning to an earlier observation about how bilingual competence and bilingual ability might be structured, the presence of two language systems in the mind allows us to portray modularity along two dimensions more clearly. This is one of the main features of the Bilingual Tripartite Parallel Architecture.

*The Cross Linguistic Dimension (La <-> Lb or L1 <-> L2)*

Strong evidence from childhood bilingual studies favors a model in which L_a and L_b undergo a separation early in development, in response to only partial positive evidence, i.e. with exposure to a minimal threshold level from both languages (Genesee, 2002). There is still dispute on some of the details (how early, from which stage of development), but at least one consensus is that autonomy of each linguistic system is achieved spontaneously without the benefit of awareness on the part of the child of structural differences between the languages, and again based only on examples from each language, without the differences being clearly marked in any way in the input. Evidence of interaction between L_a and L_b in imbalanced bilingual development - from cases of systematic interference of one system upon the other (Sánchez, 2003), and from studies of early L1 attrition (Francis, 2005) - does not call into question cross-linguistic modularity (i.e. autonomy). Research findings from the field of neurolinguistics (specifically related to the different patterns of deficit and recovery from bilingual aphasia) are consistent with this version, calling attention also to the internal modular differentiation within each linguistic system (Paradis, 2004).

An important application of this aspect of cross-linguistic separation and interaction is to the study of language mixing. Even the mixed utterances of
young bilinguals at an early stage of acquisition, and of language attriters in which one language system comes to dominate grammatical patterns, tend to show a strong tendency toward switching and inserting that is systematic and rule-governed. One explanation would emphasize the autonomy of each language representation such that mutual influences operate between systems, not within an undifferentiated network (Meisel, 2001). And a modular approach would perhaps help explain why some aspects of mixing are open to the effects of extra-linguistic factors (metacognition, motivation, general world knowledge), and other aspects are more encapsulated and "sealed off from the outside," and not subject to awareness and monitoring.

The Linguistic System-Conceptual Structures Dimension ($L_{a+b} \leftrightarrow CS$)

The proposal that a Common Underlying Proficiency (CUP) (Cummins 1991) is not "language-bound" introduced the idea of this dimension to the field of bilingual education. Conceptual structures (CS) are non-linguistic; however it is that semantic relations are computed and whatever the conceptual formation rules that determine meaning turn out to be, they are likely to be independent of the formation rules of phonology and morphosyntax (Jackendoff, 2002). How meaning is mapped onto Syntactic Structures (SS), how the parallel structures are correlated and linked up, requires an extensive array of interface mechanisms. The many specific instances of congruence between meaning and sentence structure (e.g. all words for concrete objects are nouns), and the prominence and centrality of the CS-SS interface, suggest sometimes that the linguistic domains of phonology, syntax and morphology are really integratively subsumed into a single undifferentiated holist network along with all aspects of meaning.2

But the nature of bilingual ability in particular prompts us to consider seriously the proposal that a central core of conceptual structure is shared between $L_a$ and $L_b$ (or L1 and L2). The CUP model, for example, assumes that concepts and non-linguistic skills are stored in such a way that they are accessible to the bilingual when they need to be deployed through the medium of one language or the other. The idea would be that concepts and non-linguistic skills need only be stored once, not "belonging" to either linguistic system. For example, the entire set of academic proficiencies, including non-language specific literacy skills, mathematical ability, and scientific principles need not be represented redundantly within the domain of each language that a bilingual or multilingual knows (in the manner of "L1 academic proficiencies," "L2 academic proficiencies," "L3 academic proficiencies" and so forth). In a similar way, other models of bilingual ability that hypothesize this kind of "three component, two level" architecture (Kroll & Tokowicz, 2001), portray lexical and grammatical links communicating L1 and L2 systems within a larger inter-linguistic system and linguistic-conceptual interfaces connecting each separate linguistic representation to a
shared conceptual level. This idea, formulated differently perhaps, has been one of the basic learning-principle justifications for first language instruction in bilingual education for many years now.

Another important practical application of this kind of componential approach is to the area of language assessment of bilingual children. Specifically, what aspect of "language" is an evaluation meant to measure? This is a question that is generally asked as a part of considering the "validity" of interpretations that can be made from test results. If we think about "abilities" as composed, in each case, of different constellations of knowledge structures and processing mechanisms, it should make it easier not to confound one type of performance outcome with another. For example, a common error in school language assessment of bilingual children is to attribute deficient language development to "both languages" (a SLI-type impairment) based on the evaluation of tasks strongly dependent on academic-related competencies and skills that are non-linguistic. In this case low performance in "both languages" would not necessarily imply (typically, it would not) a language development deficit.

In the domain of informal classroom assessment, a better understanding of the components of language knowledge and language ability should help teachers respond more consistently to second language learners. In fact, the reliability of teacher corrective feedback has been pointed out as a persistent problem for language learners. Students may produce a text or utterance that is syntactically well-formed but may violate a semantic constraint (e.g., the L1 optionally allows for one or two arguments for a verb depending on the intended meaning, while the L2 only allows for one). For sure, syntactic and semantic patterns, together, form what are considered grammatical sentences, but being able to specify the learner’s difficulty helps teachers be more selective and systematic. And depending on the circumstances, no corrective feedback may be the most appropriate response, for example, in the case of syntactic well-formedness, and local semantic acceptability that happens to violate a pragmatic injunction or subtle collocation restriction. In this case, what clearly sounds non-native to the teacher (and in fact is) may require a different kind of observation, or none at all, than an error related to malformed morphology or syntax. All or most of these examples may fall under the broad category of “ungrammatical,” but distinguishing among the components of grammar allows for providing more helpful corrective feedback and being selective in calling attention to errors in this way is also another way to give second language learners credit for what they have mastered.

At this stage, it would be fair to say that the proposal of a \( L_{a+b} \leftrightarrow CS \) modularity dimension is still more of an empirical question, in comparison to the stronger evidence for the \( L_a \leftrightarrow L_b \) dimension. Future research will have to sort out a number of difficult methodological problems and clarify how the different analytical categories should be understood.
FUTURE DIRECTIONS

In an broad attempt to make contact with the cognitive sciences and other theoretical models in linguistics, recent proposals from a current within Universal Grammar (Newmeyer, 2003; Culicover & Jackendoff, 2005) have laid out an attractive program of research. From the point of view of the present discussion, this important opening should also be able to make contact with the growing diversity of perspectives on modularity, that also have evolved in some interesting ways. This would include lines of investigation that have, until now, eschewed consideration of standard versions of modularity because of their historical association with mainstream generative grammar and strong innatist positions. The old manner in which debates have lined up (now largely stagnant) might to be in for a major realignment. One direction around which new lines of cross-discipline sharing of findings might emerge is related to a shift among a number of UG-oriented linguists toward the idea that more of language acquisition can be explained by the concurrence of general learning strategies. According to this view, generative approaches, by and large, have tried to account for too much by positing tightly encapsulated, very specific hardwired constraints that are highly determinate in the way they automatically trigger different components of the child's grammar. The new approach should allow for a broader range of acquisition mechanisms, thus reducing the burden on a single genetically pre-programmed Language Acquisition Device.

This is all related to the L<sub>abh</sub> <--> CS modularity dimension that we proposed for bilingualism. Modularity should also imply that not all cognitive domains are equally encapsulated and autonomous, that carry out computations in the same mandatory way, and that interface with other components with the same highly constrained degree of interactivity (Pinker & Ullman, 2002). Another way of considering this possibility is that we should allow for "degrees of modularity." Thus, some aspects of language development might unfold in a highly modular way (closed-ended, bottom-up, and "vertical"), and others to a lesser degree or in a way that is highly interactive (more open-ended), and more dependent on cognitive domains associated with conceptual structure. Inductive learning and hypothesis testing, for example, would fall under the latter category. These domains would be more "central" in the sense (among other senses) that they are in fact more global, holistic, "horizontal" and integrated - properties that strong integrativist models apply indistinctly to all aspects of language ability. The proposal for this kind of internal cognitive diversity, different kinds of subsystem for different kinds of computation, should be especially useful for studying the greater degree of diversity that reveals itself in bilingualism, of the many different kinds. It should also lend itself to opening up the discussion, to some degree, hopefully, among researchers working from different theoretical perspectives,
because different models of language and cognition might be right about different kinds of knowledge and ability.

**REFERENCES**


College of Education
Northern Arizona University
Flagstaff, Arizona 86011, USA
norbert.francis@nau.edu
http://jan.ucc.nau.edu/nf4/