Prolegomena to a Future Science of Biolinguistics

W. Tecumseh Fitch

This essay reviews some of the problems that face biolinguistics if it is to someday succeed in understanding human language from a biological and evolutionary viewpoint. Although numerous sociological problems impede progress at present, these are ultimately soluble. The greater challenges include delineating the computational mechanisms that underlie different aspects of language competence, as implemented in the brain, and understanding the epigenetic processes by which they arise. The ultimate challenge will be to develop a theory of meaning incorporating non-linguistic conceptual representations, as they exist in the mind of a dog or chimpanzee, which requires extensions of information theory incorporating context-dependence and relevance. Each of these problems is daunting alone; together they make understanding the biology of language one of the most challenging sets of problems in modern science.

1. Introduction

Prolegomena (*from Greek, plural noun, singular prolegomenon*) — *a preliminary discourse, statement or essay prefixed to a book, etc.*

In the first years of the new millennium, the word 'biolinguistics' has rather suddenly come into use as an umbrella term for various biological approaches to the study of human language. At least three recent books have 'biolinguistics' in the title (Givón 2002, Jenkins 2000, 2004), the journal *Biolinguistics* was founded (www.biolinguistics.eu), and the first Laboratory of Biolinguistics (Riken Brain Science Institute, Japan) is producing its first generation of PhD students. Based simply on the divergent contents of the books mentioned above, this nascent field is broad in its interests and incorporates diverse viewpoints, both about what language is and how it should be studied. Despite numerous disagreements, what the scholars embracing this term all have in common is the core belief that

---

A version of this paper was previously published as: Fitch, W. Tecumseh. 2009. Prolegomena to a science of biolinguistics. In Louise S. Röska–Hardy & Eva M. Neumann–Held (eds.), *Learning from Animals? Examining the Nature of Human Uniqueness*, 15–44. New York: Psychology Press. It is reprinted here with permission of the publisher.

the human capacity to acquire and use language is an aspect of human biology, and that it can thus be profitably studied from a biological perspective. While this core assumption of biolinguistics is not particularly new (Chomsky 1965, Darwin 1871, Lenneberg 1967, Lieberman 1975), it appears to be an idea whose time has come. Biolinguistics is not yet a science — it is more a loosely-defined collection of questions and approaches — but it certainly has the *potential* to become a science. The purpose of the current article is to survey this potential, and to highlight some problems that stand in the way.

2. The Promise of Biolinguistics and Obstacles to Progress

It is certainly an opportune time for scientists interested in human cognition to adopt a biological perspective, since the suite of tools available to support empirical inquiry into all aspects of biology have recently become so powerful. Human brain imaging techniques are now widely available, unthinkable a few decades ago, that allow us to examine neural function noninvasively, in normal subjects. After a decade or so of somewhat self-indulgent neo-phrenology, this field shows signs of maturing into a promising endeavor with important advantages over the patient-based approaches to neurolinguistics that preceded it. These tools will help map the functional circuits underlying language competence, and ultimately help point the way to the underlying neural computations that are of central interest. Behavioral techniques for investigating language and related cognitive functions, including eye tracking and looking time techniques, are unveiling a complex cognitive world in pre-verbal infants and non-verbal animals that stands in sharp contrast to their limited communicative ability. Finally, and perhaps most profoundly, the revolution in molecular genetics has produced genome sequences of humans, chimpanzees, dogs and many other species, and gene sequences of humans turn out to be nearly identical in many cases with homologous genes in chimps, mice, flies and even yeast. Genes involved in diverse aspects of human cognition are being pinpointed, and we can now both observe and control gene expression in animal models.

These and other new techniques are generating a flood of empirical data relevant to age-old questions about the development and evolution of language and the mind. These data often demand fundamental changes in entrenched ways of thinking about these problems. For instance, accumulating results and animal and infant cognition belie the belief that language is a pre-requisite for any form of complex conceptual processing. Similarly, the new results from developmental molecular genetics necessitate profound changes in traditional conceptions of 'innateness'. Thus, more than ever before, the biological approach to language has much to offer the linguist, psychologist, anthropologist, and philosopher. Problems that once seemed insuperable, such as interactions between ontogeny, cultural 'evolution', and phylogeny, are slowly yielding to concerted theoretical and empirical effort (Deacon 1997, Kirby, Dowman & Griffiths 2007, Kirby, Smith & Brighton 2004, Steels 1999, Tomasello 2001).

But there is trouble in this potential interdisciplinary paradise, and despite considerable grounds for optimism, it is by no means certain that the new bio-

linguistic approach will be as successful and productive as it deserves to be. As I see it, the problems facing a future science of biolinguistics come in two flavors. The first, and the less challenging intellectually, are essentially sociological problems concerning terminology, disciplinary turf wars, and struggles for dominance. A reliance on oversimplified models and outmoded distinctions is another important sociological impediment to progress. Although these problems are easily diagnosed, they may be difficult to solve. Fortunately, some of the more deeply-entrenched and recalcitrant disciplinary divides and outmoded debates and dichotomies seem to be breaking down, and I am optimistic that the next generation of young biolinguists, for whom disciplinary boundaries are more fluid, will eventually leave many of these sociological problems behind.

The second class of problems involve far more profound theoretical difficulties, and constitute some of the most serious intellectual challenges of our time, or indeed that science has ever faced. It is on these difficulties that I focus in this essay. I see three broad areas of conceptual challenge, each of them related to the others, and all three demanding fundamental theoretical and empirical progress before we can hope to understand the biological basis for language. The first challenge is *neuroscientific*: Despite huge progress, at a basic level we still do not understand how brains generate minds. This is as true of a dog's brain as for a human's, and it is true of very basic aspects of cognition, such as vision and motor control, along with language. The most fundamental neurolinguistic questions concern the basic computations underlying language use, and their specific neural basis. Current attempts to address this question remain on a shaky theoretical footing. The second challenge concerns genes and *development*: How do genes control the development of a single-celled zygote into the trillions of integrated cells comprising a complex behaving organism? Again, great progress has been made, and the new epigenetic paradigm allows us to reject long-reigning models of the genome as blueprint. However, the complex and circular nature of epigenesis, and the resultant causal indirectness of development, still pose serious conceptual challenges. While we now understand in some detail how physical structures like the vertebrate limb develop, the principles underlying brain development and evolution remain only dimly understood.

Finally, while the neuro-computational and developmental difficulties are basically biological, and apply to *any* aspect of cognition, the last and I fear most profound difficulty concerns language more specifically. This suite of problems concerns questions of *meaning*. Put simply, we have a good theory of information (Shannon information theory), but we lack anything even approaching a good theory of meaning (what I intend with this information/meaning distinction will become clear below). Problems of reference, relevance and context-dependent interpretation remain central unresolved issues in the philosophy of mind. While the first two problems have matured to a stage where they appear to be accepted as problems of the empirical natural sciences, these last problems remain in the philosophical category. (We don't even know how to devise experiments to help sort the issues out.) While these unsolved semiotic challenges pose problems for any aspect of cognition (what is it that happens when an organism interprets some stimulus as 'meaningful'), they become particularly acute when discussing language, which is that aspect of cognition centrally concerned with meaning.

Recent reviews of new approaches and data in biolinguistics are already available (Fitch 2005b, in press, and Johansson 2005). Therefore, my goal here will rather be to outline and clarify the problems facing the field. As one interested in seeing this field flower and grow, I intend my critical comments to be constructive. I have been working in 'biolinguistics' (without knowing it) for the last 15 years, since my decision as a young marine biologist to refocus my efforts on the evolution of language (e.g., Fitch 1994). Although I remain optimistic, I have become acutely aware of the difficulties facing the field, in part because successes in various areas have brought the remaining problems into sharper focus. Through my involvement in a recent interdisciplinary foray in biolinguistics (Hauser *et al*. 2002) and the debate that followed (Pinker & Jackendoff 2005, Fitch *et al*. 2005), I have also developed a healthy (if depressing) awareness of the sociological problems that await attempts at interdisciplinary bridge-building.

In this article I will start by briefly discussing the sociological problems and disciplinary strife that arise from choices in terminology and differing conceptions of 'language'. These pose important but soluble problems for those with a *bona fide* interest in solutions, I think, and will not be my core focus here. In the main part of the article I will outline and clarify some of the deeper intellectual challenges facing biolinguistics, discussing why many currently-popular models and metaphors for understanding genes, brain and language need to be abandoned if we hope to make substantial progress. In some cases I will also try, tentatively, to sketch approaches to the problem that appear to me to offer promise. But I will be satisfied if the reader, accepting my critique of the 'state of the art', rejects my proposals for remediation. Each problem alone is extremely difficult, and combined as they must be in biolinguistics, even guessing at plausible answers is difficult. In this essay, as with any prolegomenon, my focus is making the problems sharp and clear, rather than defending particular solutions.

2.1. *Sociological Challenges: Disciplinary Discord and Terminological Debate*

> *The chance that the key ideas of any professional scholar's work are pure nonsense is small; much greater the chance that a devastating refutation is based on a superficial reading or even a distorted one, subconsciously twisted by a desire to refute.*
>
> (Langer 1962: ix)

This wise insight accurately diagnoses much contemporary 'debate' in biolinguistics, particularly concerning the evolution of language. I know of no other field where scholars seem so ready to champion their own pet hypothesis uncritically, while rejecting those of others as ludicrous. While I confess to finding some proposals in the literature uncompelling on first reading (e.g., the 'Throwing Madonna Hypothesis' (Calvin 1983), or the 'Aquatic Ape Hypothesis' (Morgan 1997)), further reading and thought have convinced me that some valuable insights, and probably germs of truth, are to be found in such ideas — for one willing to put in the work of understanding them. Unfortunately, such willingness is too often in short supply, and debate in the biology and evolution of language frequently reduces to either misrepresentation (dismissals based on straw-man caricatures) or arid terminological debates ("I dislike the term *X* for some trait and propose term *Y* for the same thing"). Often the two are combined.

This syndrome is particularly true of criticisms of Noam Chomsky, whose ideas so many scholars apparently love to hate. In my opinion, once placed in context and properly understood, most of Chomsky's scattered statements about both language evolution and its biological bases either are rather uncontroversial statements that any modern biologist studying, say, limb development would accept as a matter of course (e.g., that there must be various biological constraints upon the development of the language system), or statements of unpopular alternative hypotheses that deserve more careful consideration (e.g., language as a tool for thought rather than communication). Outside of his technical linguistics work, Chomsky's main contribution to biolinguistics is his long championing a scientific approach to language as a *biological* phenomenon (Chomsky 2005). One will search in vain in Chomsky's own writings for the naive conceptions of Universal Grammar for which he is so often mistakenly pilloried — one reason his critics typically quote his few scattered statements out of context, if they quote them at all.

My purpose in this article is neither to champion nor to attack Chomsky's conception of language (for this, see, e.g., Jenkins 2000, Lieberman 2000, Jackendoff 2002, Boeckx 2010) — but rather to argue that such discussions too often miss or leave unmentioned deeper commonalities of viewpoint and approach shared by most contemporary theorists interested in the biology of language. In the next sections, I will try to look past the terminology at some uncontroversial facts about the biology of language, briefly discussing the terminological controversies they have driven. My purpose is to shed the rhetoric and move into the conceptual heart of biolinguistics. This will set the stage for the main part of the article, where I discuss the core outstanding conceptual difficulties in detail.

2.2.   *'The Human Capacity to Acquire Language': The Core Explanandum*

The central research topic in biolinguistics is a characterization and explanation of the human capacity to acquire and use language. That this is an aspect of human *biology* is made clear by the everyday fact that any normal child raised in a human household will quickly, and apparently effortlessly, acquire the language(s) of its family and community, while no non-human animal will do the same. The pet dog or cat may learn quite a bit about the social and practical aspects of life in a human household, and often to recognize a few dozen spoken words of the local language, but its abilities to *express* its own thoughts *using* this language are little different from those of a potted plant in the living room. Perhaps more surprisingly, a chimpanzee raised in a human home will not spontaneously do much better: Even with long and intensive training, young apes learn to produce only an indistinct handful of inarticulate words (Yerkes & Yerkes 1929, C. Hayes 1951). Although use of the manual/visual modality via sign or symbols helps young apes considerably (Gardner & Gardner 1969, Premack 1971), the adult ape still cannot progress to anything like the level of a five-year old child, and its 'linguistic' utterances will be mostly confined to requests for tickles or treats. While not belittling the accomplishments or value of such experiments (cf. Savage–Rumbaugh 1986, Savage–Rumbaugh *et al.* 1993), it is important to acknowledge these limitations as well-replicated biological facts.

Clearly, immersion in a linguistic *environment* is not enough for spoken language to develop in most organisms. There must therefore be *something* about human children which differentiates them from other species, and this something provides one of our core *explananda* in biolinguistics. We might gloss this neutrally as 'the human capacity to acquire language'. In generative linguistics this capacity is traditionally called the 'Language Acquisition Device', and a characterization of its properties termed 'Universal Grammar' (Chomsky 1965, reviving a 17th century term). Universal Grammar (before Chomsky) simply designated those aspects of human language competence which, because they are shared by all humans and all languages, went unmentioned in traditional grammars (Chomsky 1966, Allan 2007). For example, the notion that words exist and have specific meanings does not need to be specified in a grammar of French — it can be taken for granted. But this is precisely the sort of fact that *does* need to be explained by a successful biological approach to language. The original usage of the term made no particular claims about the nature of this competence (e.g., that it was specific to language, or conversely a general aspect of human cognition), nor did Chomsky's revival of the term, which is quite neutral on such questions by my reading. However, both 'Language Acquisition Device' and, especially, 'Universal Grammar' arouse suspicion and rejection from scholars who nonetheless accept that such a human-specific biological capacity exists (e.g., Lieberman 1998a, Tomasello 1999, 2005). A huge amount of ink has been shed rejecting the term 'Universal Grammar', even by people who accept without question that a biologically-based capacity to acquire complex language fully is a uniquely-powerful birth-right of any normal human, but no known animal. The substantive debate concerns not the *existence* of such a human capacity for language acquisition, which is abundantly clear regardless of terminology, but rather its nature (e.g., the degree to which it is specific to language).

There remains, today, no widely-accepted term for this central aspect of human biology, despite the consensus about its existence. A recent attempt to break the resulting terminological logjam by introducing two new terms — the faculty of language in broad and narrow senses (FLB and FLN; Hauser *et al.* 2002) — unfortunately elicited similar reactions (e.g., by Pinker & Jackendoff 2005), although FLB was specifically and explicitly intended to capture a much broader and more inclusive conception of the language capacity than the one connoted by Language Acquisition Device or Universal Grammar. FLN was intended to have a considerably narrower scope, perhaps even denoting an empty set, but has been read simply as 'language' by some and 'Universal Grammar' by others. The term 'language instinct', popularized in Pinker (1994), has been rejected equally vehemently (e.g., Tomasello 1995). Frankly, it is unclear to me whether *any* acronym or shorthand version of the 'human capacity to acquire language' will escape a similar rhetorical assassination. Perhaps the field of biolinguistics will have to do without any such term for the time being (although I would personally vote for 'language acquisition capacity' as a relatively neutral designation).

2.3.   *'Innate Knowledge'*

A similar terminological morass surrounds the term 'innate', and particularly the

concept of 'innate knowledge', although the problems here are at least partly substantive rather than terminological. The deep conceptual problem ultimately stems from the complexity of epigenesis (the complex interaction in the developing organism between developmental programs and the internal and external environment), to be discussed below. But the terminological problem hinges on what we are prepared (or inclined) to call 'knowledge'. Knowledge is prototypically a representational state of adult minds, implemented somehow in their brains. We know enough today to say that this implementation will involve the morphology of individual neurons, their complex interconnections with other neurons, and the computational activities these neural circuits engage in. From this mechanistic perspective, it would be odd to ascribe 'knowledge' to genes, or to the just-fertilized egg. But what about the newborn infant's 'knowledge' of language? Here we are on uncertain ground, for the child is certainly born with a brain, equipped with proclivities to attend preferentially to certain things (like human voices) and not others (like dog barks or engine noises). Even at birth the newborn already expresses preferences for its own mother's voice, or her native language, or a lullaby she sang while the child was still *in utero* (Mills & Melhuish 1974, DeCasper & Fifer 1980, Mehler *et al.* 1988, Hepper 1991, Spence & Freeman 1996) — implying that the fetal environment has already shaped this newborn brain. This constitutes, perhaps, a kind of knowledge. In addition to such rapidly-acquired proclivities, the child manifests constraints on the type of regularities it extracts from linguistic input, and these constraints have been argued by many authors to be important or even necessary components of the child's capacity to acquire language. Do such unconscious proclivities and constraints constitute 'knowledge'?

2.4.   *'An Instinct to Learn'*

Light can be shed on this question by examining the analogous but better-understood situation in birdsong learning, where an elegant and insightful model of a biologically-based cultural capacity has been developed by Peter Marler. Most songbirds (nearly half of roughly 9,000 bird species) learn their song: A young bird must hear exemplars of the song of its species in order to produce a normal song (Catchpole 1973, Marler & Slabbekoorn 2004). Birds raised in an aviary with other species, but without access to conspecific song, will sing either a completely abnormal song, or (in some cases) will learn the song of another species. Crucially, most birds do not simply mimic the song of adults exactly: In many species, individuals create new, novel songs that are built upon but not identical to the songs they heard as nestlings. This creative aspect of birdsong ensures that each generation hears slightly different songs from those of the previous generation. This process of song transmission across generations, with slight novelties introduced by creativity and or erroneous copying, leads to 'dialects' of birdsong: Birds in different regions sing quite different learned songs. This cultural evolution process can quickly 'repair' song in a population experimentally seeded with aberrant song, correcting it toward the species-typical norm (Fehér *et al.* 2009). But just as a human child of Chinese descent can learn perfect English, a young bird exposed to a different dialect than that of its

parents will master the new conspecific dialect.

Equally crucially, young birds exposed to the song of many different species will unerringly hone in on the song of their own species: A songbird appears to be born with a proclivity for the song 'style' of its own species, to which it will attend preferentially. So the bird's propensity to learn is constrained in certain ways: It is not simply a 'general purpose' system that will learn anything it hears. These facts have forced students of birdsong to progress beyond simple-minded nature/nurture dichotomies. Marler's model of birdsong acquisition instead integrates both biological and 'cultural' factors, which are inextricably intertwined in an *"instinct to learn"* (Marler 1991). Songbirds, like human children, are born with a *readiness* to master their species-specific communication system, but they are not born *knowing* this system. Part of this biologically-given readiness is a proclivity to attend to certain types of auditory stimuli (conspecific voices and songs) and not others (dog barking, machine noises, etc). Constraints exist on what can and cannot be learned: There are limits on the sorts of artificial birdsongs a youngster can absorb. These facts show clear parallels with the facts of human language acquisition, and a model of the human cultural capacity as an 'instinct to learn' is an important improvement over currently more popular metaphors. This conceptual model has recently been advanced explicitly to model the acquisition of human language (Doupe & Kuhl 1999, Marler 2000, Okanoya 2002, and Fitch, in press).

Returning, thus equipped, to the term 'innate knowledge', it seems to me somewhat misleading to refer to the constraints on the fledgling bird's song acquisition system as 'knowledge'. These constraints (whatever they might be) are not themselves knowledge but instead influence the knowledge the bird will someday possess. I would make the same terminological caveat, *mutatis mutandis*, about human language acquisition. However, many scholars are perfectly willing to term such innate constraints 'knowledge'. I am happy to accommodate them, so long as the distinctions are kept clear between behaviors that are truly innate (e.g., the acoustic structure of human laughter or cry, and the inborn link between these sounds and pleasure or pain) and those, like speech sounds or birdsongs, for which an innate basis for acquisition exists, but where the behaviors themselves depend on structured environmental input to be acquired and expressed. This distinction illustrates why the term 'language instinct' is misleading. The prototypical cases of instinctual behaviors, such as mammalian crying or suckling, a chick's escape from its shell, or a fly's grooming, really are genetically-coded behaviors, fully-functional at birth. 'Instinct' properly characterizes the child's *acquisition system,* but not the knowledge that system will eventually acquire. We are born with a language acquisition 'instinct' but not language *per se*. Again, the terminology is less important than the crucial underlying principle. What, precisely, is the nature of the capabilities, biases, proclivities and constraints that the human child brings to the problem of language acquisition?

2.5. *Beyond Disciplinary Discord*

Whenever people vehemently reject a proposition, they do so not because it simply

> *does not recommend itself, but because it does, and yet its acceptance threatens to*
> *hamper their thinking in some important way.*   (Langer 1942: 238)

To summarize, the current literature on the biology of language reveals a somewhat depressing disciplinary landscape. Despite agreement about the central interest of the questions, and core *explananda*, and the promise of the diverse approaches and perspectives represented, members of competing factions too rarely cite each other or interact constructively. Theoretical discussions are often dominated by rhetorical battles and ideological or terminological debate rather than constructive attempts to make tangible progress. Much of the criticism that currently divides the relevant fields boils down to "My opponent says we should look to *x* for answers, but I believe we should look to *y* instead". Typically, both *x* and *y* are probably important. Given the large number of open questions, biolinguistics will be better off when individual researchers pursue those topics and approaches they believe are important and promising, and refrain from attacking others who have different interests or try different approaches. There is little to be gained from such attacks, and if my experience is any guide, much to be lost.

One can only hope that, whatever else happens, biolinguistics will shed unproductive rhetoric and get serious about making empirical progress. In addition to the stunning progress in contemporary biology, the grounds for optimism within linguistics include increasing convergence among long-separated theoretical approaches to syntax (e.g., minimalism, tree adjoining grammar, construction grammar, and functionalist approaches) towards heavily lexicalized theory of language, with a few basic and powerful operations (e.g., merge, adjoin or unify — see Joshi *et al*. 1991, Stabler 2004). Neighboring fields like neurolinguistics have proven willing to take insights from generative linguistics and test them empirically (Caplan 1987, Friederici *et al.* 2002, Arbib 2005, and Hagoort 2005b), and biolinguistics as a whole will do well to follow this path. Interest in biological approaches to language seems to be growing rapidly in all disciplines, so those established scholars prepared to indulge in self-destructive turf wars should be equally prepared to watch the incoming neuroscientists and biologists take over the field.

3.   Beyond Evolutionarios: Testing Biolinguistic Hypotheses

I consider it self-evident that the appropriate models for biolinguistics come from the natural sciences, such as physics in the early twentieth century, and cellular and molecular biology or neuroscience today. Theorists in these fields consider the issues, define their terms, and propose hypotheses that generate testable predictions. Experimentalists implement empirical research programs to test the predictions, based on widely accepted norms of good experimental design (e.g., explicit consideration of, and controls for, alternative hypotheses) and inferential statistics. The historical success of this 'normal science' approach hardly needs emphasizing: Our modern lifestyle from computers and cell phones to agriculture and medicine relies upon it, and the future holds, if anything, an

acceleration of progress in understanding the physical and biological world.

There is no reason that theorists and experimentalists should be different individuals, and I think at the present state of play most biolinguists need thorough familiarity with both theory and experiment. This is especially true for evolutionary questions, since generating *testable* predictions is far more difficult than coming up with untestable evolutionary scenarios. 'Evolutionarios' are entertaining but typically offer experimentalists little to work with. Despite the dearth of testable hypotheses, and surfeit of evolutionarios, in current discussions, I think the situation is remediable. The onus is on theory-makers to generate clear definitions of terms and hypotheses, and practically testable hypotheses. Furthermore, progress will be aided by comparing and contrasting *multiple hypothese*s, not simply rejecting implausible null hypotheses in favor of single pet hypotheses. Ultimately, as for physics, what biolinguistics needs most are creative empirical tests of hypotheses.

Since Darwin, evolutionary biologists have been testing functional and phylogenetic hypotheses quite successfully, despite our lack of time machines, using the *comparative method* (Harvey & Pagel 1991). Although Language, writ large, is unique to our species, many (probably most) of the mechanisms *involved* in language have analogs or homologues in other animals (Hauser *et al.* 2002, Fitch 2005b), and their comparative study thus offers biolinguistics crucial insights. Furthermore, new genetic techniques make it possible to roughly date the origins of mutations (e.g., Enard *et al.* 2002). A combination of a broad comparative approach, molecular genetic techniques, and creative examination of individual differences among humans offers many ways to test evolutionary hypotheses. For example, consider two venerable hypotheses about the origins of human speech.

Scholars have debated for centuries whether the lack of speech in other animals results from peripheral anatomy of the vocal tract (H1), or the nature and structure of the central nervous system (H2).

Recognizing that dolphins were mammals with large brains, Aristotle suggested that their lack of speech results from their lack of loose tongues and lips (Aristotle 350 BCE). Similarly, the discovery by Europeans of the (speechless) apes led to renewed consideration of the crucial capacities underlying speech, and anatomist Peter Camper concluded that the lack of speech in orangutans was caused by their large air sacs (Camper 1779). Both of these ideas are special cases of H1. Other scholars, like Darwin, considered these arguments, but sided with H2, that central neural factors must be critical: "The relation between the continued use of language and the development of the brain, has no doubt been far more important" (Darwin 1871).

3.1.   The Descended Larynx

As a modern instantiation of H1, consider the descent of the human larynx (Fitch 2000b). The lowered larynx and tongue root of humans was hypothesized by Philip Lieberman and colleagues (1969) to constitute an adaptation to produce a wider range of speech segments (particularly the point vowels, and the 'super-vowel' /i/, used in vocal tract normalization). At that time, and for the next 20

years, both the descended larynx and vocal tract normalization were believed to be uniquely human (Lieberman 1984). In my PhD thesis (Fitch 1994), I developed a related hypothesis, based mainly on principles from physics and physiology, that human formant perception might build upon a capacity for size estimation predating speech, providing a pre-adaptation for the use of formants in speech (H3). This hypothesis required formants to be tied to body size, and clearly predicted that formant perception, and its use in size estimation, would be more widely present in other animals. Thus, it was based on a number of testable assumptions and made numerous testable predictions, and in the last 15 years my colleagues and I have been busy investigating them. We have found that, as predicted, formants provide a reliable cue to body size in many species, because body size, vocal tract length, and formant frequencies are inter-correlated (Fitch 1997, 2000a, Reby & McComb 2003). Further, it predicts that listeners should *use* this potential source of information as an indicator of body size, as several species do (Fitch 1994, Reby *et al.* 2005, Smith *et al.* 2005, Ghazanfar *et al.* 2007). Finally, these finding spurred a closer look at non-human animal vocal production, revealing a descended larynx (once believed uniquely human) in several non-human species (Fitch & Reby 2001, Weissengruber *et al.* 2002, Frey & Riede 2003).

Most of the studies above were directly spurred by specific theoretical questions about the evolution of speech. Besides demonstrating that the descended larynx is not uniquely human, and suggesting that both formant perception and vocal tract normalization build upon primitive mammalian auditory mechanisms, these data revealed that formant signals are an important component of vertebrate communication, are used to judge size, and that these ancient shared uses are still operative in modern humans. They provide abundant evidence consistent with H3, the pre-adaptive hypothesis of Fitch (1994), which thus becomes a serious contender as the original adaptive force driving the descended larynx in our species. Improved speech is no longer the only plausible evolutionary explanation for laryngeal descent, as previously assumed (Lieberman 1984), and it is possible that the descended larynx in adults evolved *before* spoken language.

Equally importantly, these and other recent data on vocal production in mammals demonstrate that the vocal tract is a highly flexible, reconfigurable system: Any mammal can lower its larynx dynamically (Fitch 2000c). Such data offer strong evidence against H1 in its strong forms. While our vocal tract certainly influences the types of sounds we can make, and has presumably been selected in human evolution for its beneficial effects on mechanical control and/or the speed of information transmission (as argued by Lieberman 2006), peripheral anatomy is not a crucial Rubicon that needed to be passed before humans could evolve spoken language. The descended larynx/tongue root is not the core factor keeping chimpanzees from speaking, and by process of elimination, that factor must rest in their brains, not their tongues. The comparative data indicate that neural factors, rather than peripheral anatomy, provide the core mechanistic basis for human speech capacities. What keeps chimpanzees from talking, but allows some seals to talk (Ralls *et al.* 1985) is the configuration of their brains, and not that of their tongues or vocal tracts. But while we have made tangible

progress by rejecting the peripheral vocal apparatus as the core factor underlying human speech, this research as yet offers little insight into which aspects of the central nervous system *are* different. Rejecting H1, we tentative accept the alternative hypothesis H2.

3.2.   Hypothesis 2: Into the Brain

Any hypothesis based on the idea that it is increased neural control over vocal production that allows humans, and not other primates, to speak must posit some difference in neural circuitry that enables this increased control. Given the complexity of speech, and of motor control, there are likely to be several such differences. One well-documented difference between humans and other primates is that our species possesses *direct connections* from lateral motor regions down to the motor neurons that drive the vocal apparatus, particularly the larynx and diaphragm (Iwatsubo *et al.* 1990, Jürgens 1994). By hypothesis, these connections allow increased voluntary control over the vocal organs, and better coordination between the facial and tongue musculature and phonation (which is crucial to human speech). Of course, although plausible, this 'Kuypers/Jürgens' hypothesis (H4) does not by itself prove anything: There are many small differences between a human and chimpanzee brain, and no guarantee that this one is critical to the known behavioral difference. How could H4 be tested?

The existence of other vertebrates with complex vocal learning open the door to an understanding of the mechanisms of vocal control, at both the neural and genetic levels. Although songbirds are by far the best understood group, mammalian vocal learners include cetaceans, seals (Janik & Slater 1997), bats (Knörnschild *et al.* 2009) and probably elephants (Poole *et al.* 2005). Unfortunately, both birds and tractable cetaceans (e.g., dolphins) have a brain and vocal tract very different from that of humans. Nonetheless, the data from birds are consistent with H4: Birds have direct connections between the telencephalon and the primary motor neurons controlling their phonatory organ, the syrinx. While consistent, this is not perhaps as compelling as we would like.

Among mammals, both seals and bats use a normal mammalian brain to control a normal mammalian vocal tract, and thus provide a unique but mostly untapped source of information into the neural and genetic mechanisms underlying complex vocal control (particularly in phocid seals with complex learned 'song' (Janik & Slater 1997, van Parijs 2003)). Currently, though, most questions one might ask about seal neuro-anatomy and vocal production have a simple answer: Nobody knows, because nobody has looked. The discovery of a species of fruit bats with complex vocal learning is so new that very little is known about neural control in this species (Knörnschild *et al.* 2009). So these are clear, open predictions of the hypothesis, waiting to be tested.

3.3.   Convergence, 'Deep Homology', and the Broad Comparative Method

Although the significance of research on birds, bats, deer or seals is sometimes disregarded by those interested in human language because it does not concern primates, and thus does not reveal homologous mechanisms, this attitude misses

two crucial points about the comparative method. First, convergent evolution (e.g., of complex vocal imitation in humans, birds, seals and bats) allows us to *test* adaptive hypotheses. In convergent evolution, each clade that has evolved a trait constitutes an independent evolutionary data point. This is not true of a group of species that all inherit a homologous trait from their common ancestor: No matter how many species share it, such a trait constitutes a single data point. This is a fundamental insight of Darwin's use of the comparative method, as well as modern statistics for evolutionary hypothesis testing (Felsenstein 1985, Harvey & Pagel 1991).

A second, more surprising, fact about convergent evolution is much more recent. It follows from a central realization in modern molecular biology concerning the profound *conservation of genetic mechanisms* across disparate living organisms. Genes involved in development turn out to be highly conserved (Gehring & Ikeo 1999, Carroll 2000). Even traits that have evolved *convergently* may often rely upon *homologous* genetic and developmental mechanisms (termed 'deep homology' by Shubin *et al.* 1997). This discovery vastly broadens the scope of the comparative method, which has traditionally focused mainly on homology (though see Gould 1976). The new data pouring in from diverse distantly-related species (especially birds and rodents, but including pufferfish, flies, worms, yeast, and slime molds) reveals a stunning consistency in underlying genetic and developmental mechanisms in this diverse assemblage (Carroll *et al.* 2005). Such underlying conservatism of genetic details was unimaginable two decades ago. Even phenotypic traits that evolved convergently (and are thus homoplastic) often share common developmental and genotypic mechanisms. Therefore, a broad comparative approach that incorporates homoplasy in addition to superficial homology has deep insights to offer. Biologists can avail themselves of a much broader range of species than previously thought, and confidently expect that much of the resulting data will be relevant to human traits (Carroll 2003, Carroll *et al.* 2005). For example, the discovery of mammals with a descended larynx opens the door to genetic and physiological research on the mechanisms underlying this trait. Widespread conservation of developmental mechanisms gives hope (though not certainty) that similar mechanisms may underlie laryngeal descent in humans and in species, like deer, amenable to experimental study.

This story is of course far from over: Replications remain scarce, and further data are clearly needed. While a plausible case can now be made against H1, and for the pre-adaptive hypothesis H3, one might suppose that H3 never could be demonstrated, as these events occurred pre-historically but do not fossilize. Fortunately, this is not true: Comparative molecular biology offers a new and exciting path out of this apparent dead end. If we can uncover the molecular genetic basis for the descended larynx *and* for the complex vocal control underlying speech, we can use the techniques developed by molecular evolutionists (e.g., Enard *et al.* 2002) to date the selective events that established the corresponding alleles during human evolution. If the selective sweep leading to 'laryngeal descent genes' preceded that leading to 'control genes' (in quotes because it is unlikely that the alleles in question function exclusively in these domains), this would be strong evidence against Lieberman's hypothesis that

speech preceded (and selected for) the descended larynx (and therefore in favor of the alternative).

In summary, the evolution of speech provides numerous evolutionary hypotheses that can be, and have been, tested. This leads us to a certain amount of optimisim about our ability to move beyond the domains of speech production, and resolve debates about core aspects of language: Syntax and semantics. And it brings us to the first of the 'hard problems' facing bio-linguistics.

4.   Mind and Brain: The Need for Bridging Theories of Neural Computation

Trying to understand perception by studying only neurons is like trying to under-stand bird flight by studying only feathers.   (Marr 1982: 27)

In his book *Vision*, a foundational work in cognitive science, David Marr argued that progress in understanding the visual brain requires research at multiple levels — including the implementational level (neurons and synapses), the algorithmic strategy used to tackle the problem, and the computational-level description of the problem space itself (Marr 1982). He used Chomsky's goal of formulating a computational model of language (what Chomsky termed a 'competence' model), as an exemplar of this approach. While Marr's multi-level approach has been embraced in the computational neuroscience of vision, its application to language remains relatively unarticulated (though see Poeppel & Embick 2005). I think this results at least partially from a failure in the cognitive sciences to fully embrace the insight that progress will require multiple, complementary levels of description, at the computational, algorithmic *and* implementational levels. Most crucially, we need *bridging theories* that go between levels of description, particularly the computational and algorithmic levels.

4.1.   Multiple Levels of Description

Despite a long-running debate between connectionists and symbolists in cognitive science (e.g., the many responses to Fodor & Pylyshyn 1988), a connectionist model at the implementational or algorithmic level is not necessarily in conflict with a symbolic computational model, but rather a potential complement to it, as clear thinkers in this debate have remained well aware. But accepting the need for multiple levels of description unfortunately doesn't provide a road map for how to formulate models at each level, or how to link the levels. For that, lacking a general theory of neural computation, we must currently take a catch-as-catch-can approach, using whatever clues we can find. The problem is particularly sharp given that our most powerful empirical tools at the neural level (e.g., single-unit recording or experimental gene regulation) are unavailable for the study of language because the species employed lack language, and the techniques cannot generally be applied to humans. At the highest computational level of language, our best guides must still come from behavioral studies, both psycholinguistics and traditional theoretical linguistics, with some help from

brain imaging.

Comparative linguistics and typology are important additional elements, since the study of diverse languages can sharpen our focus on the problem by cataloging the diversity of solutions to it. In a few cases (e.g., metrical phonology and stress systems), linguists have already developed quite sophisticated models that seem capable of encompassing most of the diversity of the world's languages (e.g., B. Hayes 1995): Both the required theoretical primitives (such as syllables, stress, feet, and prosodic words) and generalizations (e.g., the 'iambic/trochaic law'), are relatively clear and uncontroversial. Such aspects of language seem ripe candidates for constructing algorithmic models incorporating psychological data (e.g., Cutler 1996) which can ultimately be translated to models of implementation. Unfortunately, however, such oases of clarity and agreement are the exception in linguistics. Consideration of the diversity of languages allows one to exclude certain possible theories (e.g., a theory that syntactic structure assignment relies necessarily on word order is falsified by 'non-configurational' languages like Warlpiri that have free word order; cf. Austin & Bresnan 1996). However, besides a general agreement on such theoretical primitives as words and sentences, and on the need for structure-dependent rules, there seem precious few specific theoretical claims that are beyond dispute in contemporary syntax or semantics.

4.1.1. The Mechanistic Level

Given that research and discussion at the purely computational level have so far failed to converge, perhaps there are lessons to be learned from considering the lower levels of description. A crucial lesson from computational neuroscience has been that progress typically results not from investigations at a single level of description, but by attempts to bridge between levels: It is the intersection of constraints from the different levels that gives us purchase on the problem (Rolls & Deco 2001). Our theory of color vision is informed by the understanding that there are three types of cones, and our theory of motion detection by the discovery of separate populations of cells interested in motion and not color. Similarly, consideration of neural data may help theoretical linguists 'cleave nature at the joints' in their attempts to discover robust and useful computational primitives in language. Current brain imaging techniques (fMRI, PET) provide little insight into the computation ⇔ algorithm linking problem. Knowing *where* brain activity increases in some language-related task (e.g., generating an inflected verb, or imitating a spoken word) provides pointers about where to look. Similarly, systems that provide high temporal resolution (EEG, MEG) can provide indications of *when* certain neural regions are activated, and thus provide better data for testing causal models of language processing. But both still leave open *what* the actual corresponding computation is: What aspect(s) of the circuit diagram are crucial.

4.1.2. The Search for Computational Primitives

Despite the value of brain imaging techniques, we cannot expect them to solve

the central problem. For that, we need to distill what we know from linguistic theory into a set of computational primitives, and try to link them with models and specific principles of neural computation. Unfortunately, appeal to general computational principles may be of limited value. To the extent that vision is best conceptualized as a 'bag of tricks', where each aspect of vision (color, motion, depth perception, etc) has its own unique solutions, there may be no general conclusions available about computations underlying 'vision' in general. The same may be true of 'language'. However, vision is a far more ancient evolved system than language, so this lesson may not generalize, and certain classes of models seem to pop up consistently. Individual neurons are slow and sloppy, and sometimes die, and these basic facts have often led to the evolution of parallel redundant circuits, rather than circuits that seem optimal to electrical engineers who have fast, precise and reliable computing elements available.

This difference between silicon and cell-based computers has led to abstract notions of 'natural computation' (Richards 1988, Ballard 1999) that may hold useful clues for biolinguists building bridges between the algorithmic and computational levels. While a focus on just the computational level ('competence') remains a necessity in everyday work, consideration of 'performance models' (including both algorithmic psychological models and, implementational neural models) should ultimately inform our debates about 'natural' theoretical primitives (Fitch 2005a, Hagoort 2005a, Friederici *et al.* 2006). Thus we need linguistic models that are explicit about the computational primitives (structures and operations) they require, and that attempt to define linguistic problems at a fine enough grain that one can discuss algorithmic and implementational approaches to their solution. We need a list of computations that linguistic theorists deem indispensable to solve their particular problem (e.g., in phonology, syntax, or semantics).

4.2. A Tentative List of Computational Primitives

A non-exhaustive smorgasbord of linguistic computational primitives, based on my reading of the linguistic literature, may help make my point, illustrating the sort of computational structures and operations that any model of language will need to incorporate. While different theorists might give rather different names to them (e.g., Jackendoff 2002), or object to my overly schematic descriptions, experts can hopefully read between the lines to see what I'm getting at. Alternatively, my list may spur the theoretically-inclined reader to generate their own, quite different, list of primitives. This list simply illustrates by example the sort breakdown needed to begin building bridges between computational theories, and the algorithmic and implementational levels.

(1)   Phonology and Syntax: Data structures including trees and related multi-level structures are needed, as are structure-building algorithms that concatenate constituents into tree structures, perhaps by forming temporary links among smaller structures stored in long-term memory (the 'Lexicon'); evolutionary links with motor control seem likely.

(2)   Phonology: 'Natural classes' of phonemes, such as stops or high vowels, are required because many phonological phenomena apply to specific classes (rather than specific isolated phonemes, or broader class such as vowels); evolutionarily, such natural classes may have built upon more general auditory categorization circuits.

(3)   Syntax (structure-dependent rules): Computations that apply to classes of structures (noun phrases or sentences) rather than specific words or broad types such as nouns.

(4)   Syntax and Semantics: Dependencies between words require the equivalent of variables or subscripts that can bind constituents into temporary linkages, such as article agreement, anaphora (binding pronouns to whole noun phrases, in the simplest case), or topic/comment markers in connected discourse.

(5)   Semantics (thematic roles): Distinctions like agent vs. patient are necessary to distinguish the roles of multiple actors in such propositions as 'John likes Mary' vs. 'Mary likes John'; although English does this mainly with word order, many languages have more flexible ways of marking and expressing this key semantic difference (e.g., case-marking).

(6)   Semantics: Complex conceptual structures, built up with embedders, conjunctions, and disjunctions with scope, are needed and combining primitive predicates into larger complexes, with possible attribution of an external referent, or truth or falsity, to the whole complex, is a crucial computation in linguistic thought; despite considerable disagreement about whether this computational capacity is part of syntax, semantics or more general conceptual abilities, there is little disagreement about its basic necessity for both language and other aspects of complex thought.

4.3.   *Examples of Bridging Constructs*

I will focus on the algorithmic ⇔ computational bridge in biolinguistics because we clearly have substantial work to do in attempting to build this specific set of bridges. The good news concerning the other, algorithmic ⇔ implementational bridge is that there is little evidence suggesting that language involves any major discontinuities from other aspects of cognition at low implementational levels. The neocortical circuits involved in language have the same layered arrangement as other non-language circuits, are connected with subcortical systems like thalamus, basal ganglia and cerebellum in the same ways, and use the same types of cells releasing the same neurotransmitters with the same kinds of action potentials. The developmental processes by which these circuits arise follow the same basic principles as the circuits involved in vision or motor control in diverse mammals (Finlay & Darlington 1995). Whatever implementational details differentiate language from other cognitive functions, they appear to be only rather subtly different from those underlying other aspects of cognition. Thus we can confidently expect that most aspects of language implementation will be based

on more general principles of brain development and function, and that good first-order approximations can be built upon shared principles of neural computation (Ballard 1999, Rolls & Deco 2001). We can also expect that such first-try models will uncover some important differences in the details (otherwise, *all* brains, including those of other species, would be able to compute language readily), but these will not rely on wholly new neurophysiology or connectivity. For this reason, I see the algorithmic specification of the various components of language, based upon explicitly stated computational primitives and algorithm models, as a crucial missing link in our attempts to build the larger bridge between mind and brain. For a similar argument see (Poeppel & Embick 2005).

4.3.1. Tree Networks and Algorithms over Trees

To illustrate how a computational primitive might be fleshed out at the algorithmic and neural levels, consider the first computational primitive: Linking trees into larger complexes. First, because tree abstractions appear to be ubiquitous in theoretical models of cognitive phenomena (Simon 1962), not just language, research in other cognitive domains (e.g., chess playing, music perception, object recognition, or motor control) may offer insights into the nature of linguistic trees. Second, since words have a hierarchical internal structure (Kenstowicz 1994), and can be thought of as memorized chunks of structure, the processes by which words are learned, stored and recalled should have much in common with other aspects of long-term memory. Once recalled, such 'treelets' must be temporarily combined into larger structures via some process of binding (either adding a treelet's root to the twig of the larger tree, and thus preserving tree structure topologically, or binding two twigs to create 'tree networks'). This process may inherit aspects of the process whereby automatized motor subroutines are combined into temporary motor plans as we execute complex novel actions (e.g., Lieberman 1998b, Arbib 2005). In the same way that our ongoing plans are sometime interrupted and demand a reconfigured plan, the linguistic tree we have built during an ongoing parse may need to be abandoned and reconfigured (e.g., in garden path sentences). Thus, 'performance' theories about how linguistic trees are stored, recalled and recombined may profit from our pre-existing understanding of the neural basis of memory, motor control and other cognitive domains (as envisioned in Miller & Chomsky 1963).

At a more abstract level, such implementation-informed theoretical constructs could have important implications for how we formulate our overall theory. For instance, if we conceptualize language as a whole as a system that maps between high-dimensional conceptual structures ('thoughts') onto low-dimensional signal structures (phonetically-realized speech or sign streams) it immediately becomes clear that this is an ill-posed problem in the technical sense that there can be no unique solution to the signal $\Rightarrow$ concept expansion problem (due to the greater dimensionality of the target domain), nor perfect solution to the concept $\Rightarrow$ signal compression problem (there being multiple candidate mappings, each omitting something). There can't be enough data in the signal to allow perfect reconstruction of the original thought structure. Given this ill-posed problem, what is remarkable is that language works so well for communication,

and that (in general) we succeed at expressing our thoughts in words, and in reconstructing others' thoughts from their words. The solution demands a massive quantity of shared world knowledge: Far more information is generated by 'reading between the lines' than is literally present in the signal. Pragmatic inference using shared world knowledge is a computational necessity (Sperber & Wilson 1986).

4.3.2. Evaluating Optimality

Syntacticians have long recognized that one aspect of the signal ⇔ meaning mapping process is an element of cyclicity in the application of syntactic rules (Miller & Chomsky 1963). Beyond a certain point of expansion, we become unable to deal with large structures and loose ends dangling: We must 'close' or complete old structures if we are to cope with new ones. The first question one can ask is why this effect occurs at all. One likely answer might be that memory limitations ('performance constraints') simply prevent us from what would otherwise be an optimal solution (in much the same way that many theorists agree that memory limitations prevent easy parsing of arbitrarily center-embedded sentences). But an alternative answer is that the nature of the concept-signal mapping problem makes cyclicity a computational necessity: Even an ideal model would include cyclic application of mapping rather than an 'all-at-once' compression. Both models are logically plausible, and adjudicating between them would require two idealized models with which to compare actual human performance. Contemporary computational linguistic parsers don't provide such a model, because they assign syntactic structures, *not* conceptual structures, to strings. Indeed, mapping strings to concepts remains the major unsolved problem in computer language processing (see below). Thus, contemporary linguistics still lacks an 'ideal communicator' model comparable to 'ideal observer' models in vision, and only once we have such models can we decide whether actual human performance on this task is sadly sub-par, due to memory or processing limitations, or in fact are nearly optimal.

5.   Genes, Bodies and Brains: Biology Comes to Grips with Epigenesis

Another core issue that faces biolinguistics, and biology in general, is development. How can a single cell (the fertilized egg), with two copies of a few giga-bytes of DNA, contain within itself the basis for a newborn's body with 100 trillion cells and a brain with a trillion synapses? How can 25,000 genes possibly possess enough information to specify this process? Alternatively, how could the environment *in utero* provide this information? How could evolution have encoded it? Where does all this information come from?

5.1.   Three Reductios of Naïve Models

Let us first dispense with the obvious possible answers in a series of simple arguments, each a *reductio ad absurdum* of the corresponding oversimplistic

models, by considering the information available for pure nativist or empiricist models more closely.

5.1.1. The Naïve Nativist Model

The human brain is estimated to contain roughly 100 billion = 10^{11} cells, each of which has between 100 and 10,000 synapses, leading to at least 10^{14} synapses in the brain. To specify 1 of 10^{11} cells exactly, you need 37 bits. Therefore, to specify simply the connecting cell corresponding to each synapse you would need 37 x 10^{14} bits (and to specify the synaptic weight you would need at least eight bits per synapse). There are about 3 billion (3 x 10^9) base pairs in mammalian genome, so even if the genome was fully dedicated to specifying brain structure (which it is not) and had perfect coding in an information-theoretic sense, we would have a shortfall of at least 5 orders of magnitude to specify the connections in a human brain: We have 1/10,000th of the DNA we would need to code the detailed wiring of our brains. This 'gene shortage' has led scholars like Paul Ehrlich to conclude that little of our behavior could possibly be innate (Ehrlich 2002). Let us therefore similarly consider an exclusive role for the environment.

5.1.2. The Naïve Empiricist Model

Let us optimistically suppose that we learn something from our environments every second, waking or asleep, of our lives. There are 31 million seconds in a year (3.15 x 10^7). If we live to 100, that's just 3 x 10^9 seconds (roughly the number of base pairs in the genome). The first five years of life, when most language learning is occurring, contain only 15 x 10^7 seconds. Even the most fortunate and well-stimulated baby has this paltry number of environmental inputs available to specify 10^{14} synapses. Although we can hope that many synapses are influenced by each environmental input, this doesn't help unless each input event, is very highly structured, carrying a large amount of optimally coded information. This seems optimistic, to say the least. Thus the naïve empiricist faces the same vast information shortfall as the naïve nativist.

5.1.3. The Naïve Evolutionist Model

Finally, for completeness, consider the plight of a different type of nativist: An idealized 'evolutionary empiricist' who suggests that natural selection alone has programmed behavior. Vertebrate evolution has occupied about a billion (10^9) years. If we optimistically hypothesize (e.g., Worden 1995) a few bits of information per generation to accumulate, that's only a few billion bits again (and of course any particularities of the human brain have had far less time — roughly, 6 x 10^6 years — to accumulate). Again a vast information shortfall exists, of roughly the same order: This one a shortage of evolutionary time.

Are we to conclude from this little exercise that development is impossible? Or that the evolution of the brain could not have occurred? No, such basic considerations force us to reject overly simplistic models, and to conclude that both the naïve nativist (genome as blueprint) and naïve empiricist/evolutionist

(environment as instructor) viewpoints are woefully inadequate models. Such considerations quickly lead all serious thinkers on these problems to realize that understanding any aspect of development and evolution requires understanding the *interactions* between DNA and the world beyond the cell nucleus. Despite its tiresome persistence, 'nature *versus* nurture' is a sterile conceptual dead-end, and any valid answer must consider 'nature *via* nurture' in some form or other (Ridley 2003).

5.2.   Respect for the Cell

An important new insight in our understanding of how genes build bodies and brains is the central role of *cell biology* in all aspects of development (Kirschner & Gerhart 1998, 2005). Crucially, the trillions of cells in our body break down into only 200-odd cell types, and there are only roughly 25 morphologically distinct cell types in the cerebral cortex. What the genome carries is not instructions for individual cells, but instructions for *cell types*. Furthermore, most of the basic behavior of these cells is shared among all cells in the body (as well as with free-living single celled organisms like an amoeba or yeast), so something like half of our genome deals simply with basic cellular behavior, and only the *differences* from this 'average' cell need to be further specified (e.g., proteins like hemoglobin that are expressed only in blood cells). Each of the many trillion cells in our body is a semi-independent living thing: Under optimal tissue-culture conditions individual human cells can live for years on their own. This is not surprising when you consider that the first two billion years of evolution took place at the single-cell level. Since single-celled organisms have much shorter generation times than multicellular organisms, *most of our ancestors were free-living single-celled organisms*. From this long evolutionary history, each of our cells inherits some rather impressive behavioral capabilities. Each cell contains a complete copy of the DNA of the organism of which it is a part: It carries the entire 'recipe book' for the body along with it. Cells may make epic migrations through the body, following gradients of nutrients and responding to signals left behind by earlier pioneers, and each must eventually find a home and a job in order to survive. Individual cells are highly responsive and adaptable, and can deal successfully with evolutionarily novel circumstances (e.g., finding themselves in a damaged brain or mutant limb).

Once we recognize cells as active, adaptive, information-processing entities, we see that they form a crucial intervening level of explanation between the genetic and whole-organism levels. The apparent paradox of genetic and environmental information dissolves. Sewell Wright already recognized this in 1931: "From the view that structure is never inherited as such, but merely types of adaptive cell behavior which lead to particular structures under particular conditions, the difficulty to a considerable extent disappears" (Wright 1931: 147). The technical details allowing us to flesh out this basic insight have only recently become clear. From a genetic viewpoint, much of the overall complexity of organisms arises through local interactions between cells and their immediate organism-internal environments: The genome doesn't need to specify the shape of a human hand or a bats wing, but simply must constrain the overall pattern of

development of a mammalian limb, in a sense 'sculpting' a pre-existing developmental archetype rather than building an iconic 'blueprint' of the final structure (Goodwin & Trainor 1983). As often correctly emphasized (e.g., Dawkins 1986, Ridley 2003), the genome is nothing like a blueprint. It is more like a recipe or program. Like any recipe, it leaves a lot of detail unspecified, and up to individual cells' 'decisions' based on their particular history and circumstances. From an evolutionary viewpoint, there is no need for natural selection to perform a detailed and complete hill-climbing process through a complex, mostly non-adaptive morphogenetic space: It can let robust developmental processes do much of the work. Natural selection simply 'chooses' among the various relatively worse or better-formed, but still functional, options that result from development. This perspective on cells as prime movers in development and evolution is nicely described, with many examples, in (Kirschner & Gerhart 2005), and many of the molecular developmental mechanisms explicated in (Gilbert 2003, Carroll *et al.* 2005).

5.3. Epigenesis

Thus, in a way we are finally beginning to understand, recipes for building bodies are constrained both by the information in our genomes and the separately inherited cellular machinery acting on this information. Equally, development is constrained and informed by the environment, and has been shaped by evolution to respond robustly to it. Despite the apparent shortfall of information in any one of the relevant domains, the reality of *epigenesis* — the close interaction between information in the developmental 'program' and information stored in the environment — is that such interaction is fully adequate to specify bodies along with brains and behavior as special cases (Gottlieb 1992). Environmental stimulation, and even social interactions, turn genes on and off, and development occurs via successive waves of interactions among cells, and between cells and their local environments within the body (themselves structured by previous such interactions). Crucially, the relevant 'environmental information' in epigenetic interaction is mostly the *local environment surrounding each cell*, and *not* that in organism-external world. This local environment has traditionally been left out of both nativist and empiricist models, but is clearly where the action is in development, filling in the information shortfall described above.

Each of the trillion cells in our body or the billion cells in our brain has its own, rather myopic, local environment which informs its DNA regulation and thus developmental decisions. Each second of development, different local environments are separately effecting each cell in our body in parallel. While this local internal environment is, for the most part, dependent upon past decisions made by neighboring and predecessor cells, it is also often influenced in important ways by the organism-external environment. This influence is perhaps most marked in the brain (which is the organ most specialized to process organism-external information), but other systems like the immune system have a similarly rich external-responsiveness.

Epigenetic, interactive developmental models are nothing new: The

concept has been standard in embryology for many years (Waddington 1957). Experimental embryologists like Spemann recognized that cells respond to messages generated by other cells, and that this determines their fate later in development. Huge advances in our understanding of the genetic basis of development in the last decades have brought such ideas to fruition, and now the molecular basis of Spemann's 'organizer' signal, and many other similar cell-cell signaling systems, is becoming clear (see Gilbert 2003). The mechanisms by which DNA expression is regulated, both in classic epigenesis via transcription factors (proteins that bind to DNA), and longer-term changes (e.g., the new epigenetics of 'genetic imprinting' that can span generations) are now becoming clear (Reik 2007). This progress in turn has led to the construction of new bridges between evolution and development — *evolutionary developmental biology* or 'evo-devo' — which promise to finally close the most crucial remaining gap in our understanding of biology (for authoritative introductions see Carroll 2005, Carroll *et al.* 2005). Today's biolinguists can help themselves to some well-developed models of epigenesis, and how development interacts with evolution, before trying their hand at understanding the epigenesis of language. For further implications of this perspective on the evolution of mind see Fitch (2008).

5.4. *Neurons: A Very Special Cell Class*

This cell-based epigenetic perspective, a central tenet of the evo–devo revolution, is as applicable to the development of brains as to the rest of the body. However, neurons are unusual in a number of ways. The most important is that they are specialized for information processing by networks of neurons, over and above the normal cell-cell interactions that influence all cells. In the case of a neuron in the developing brain, 'finding a job' means taking part in a circuit that behaves coherently, and many of the neurons that are born fail to achieve this goal, and undergo programmed cell death as a result. While the primary constraints on a skeletal cell in a developing bone are physical forces (stresses and strains), for a neuron the relevant forces are the complex ebbs and flows of an 'information economy' established by myriad surrounding cells (both neurons and glia) as well as quite distant neurons influencing it through their axonal projections. Thus the local environment of the brain is unusual both in the type of commodity processed (information) and the topology of interactions (including precise long-distance connections, made possible by the unusually elongated neuronal morphology). While there is every reason to believe that insights from the development of limbs or the lung will carry over to the brain, we can also be certain that new principles are involved in brain development and evolution (Striedter 2004).

5.5. *The Way Forward*

> *One of the principal objects of theoretical research in any department of knowledge is to find the point of view from which the subject appears in its greatest simplicity.*
>
> (J. Willard Gibbs)

The revolution underway in developmental biology has important implications for biolinguistics. Anyone interested in understanding the biological basis for human language acquisition must be prepared to jettison simplistic debates about nature vs. nurture, and unhelpful notions of heritability from old-school genetics ("Dyslexia has a heritability of 45%"). Instead we can expect highly complex interactions between cells of different types, and in different brain regions, to provide the link between genetic changes and individual phenotypes. We can expect few if any cellular behaviours or cell signaling molecules that are qualitatively novel, either to our species or to language (Hill & Walsh 2005), but instead seek combinations of conserved cell processes building neural circuits that perform qualitatively novel classes of computation (Szathmáry 2001). We can expect that such circuits, built of 'normal' neurons using standard neurotransmitters, will exhibit properties and connections that are 'standard' in the mammalian brain (e.g., cortico-thalamic loops), but that these same circuits may show patterns of connectivity that are unusual, and perhaps in some cases unique, to our species or to language itself.

It would be hard to overstate the difficulties we face in discovering such subtle implementational differences. Despite a long history of trying (e.g., Braitenberg 1977), even circuits whose structure is already known in detail (e.g., in the hippocampus or cerebellum) have proved remarkably resistant to abstract computational analysis. Although a variety of simple models of memory or motor control exist, computational neuroscientists have yet to converge on models that are adequately comprehensive yet simple enough to understand. And these systems are broadly shared with well-studied 'model' animal species (mice, rats, monkeys, etc).

Integrating our computational and developmental problems, we can expect that *any* simple developmental model of the key neural computations involved language will be incorrect in its details. Nonetheless, progress will be fastest if we attempt to develop explicit simple models of various language mechanisms, amenable to experimental disproof, and then let the data show us where they are wrong. In the same way that Galileo and Newton achieved huge gains in physics by abstracting away from the existence of friction, we may expect that abstract models of neurolinguistic function and development, based on known aspects of neurophysiology and neural development but tailored to the specific computational needs of language, will offer hope of rapid progress. Progress requires posing simple (perhaps over-simplified) models, knowing they will be wrong, and letting the data tell us *where* they are wrong. As Einstein advised, "everything should be as simple as possible, but not simpler", and falsifiable, simple models will be vastly preferable to complex, unfalsifiable models with too many unconstrained variables.

6.   Information and Meaning: The Final Frontier

I will end with a brief look at the aspect of language that I think promises to be most difficult to solve: The problem of meaning. While we have a powerful and well-understood theory of *information*, we still lack a mathematical theory of

meaning, and developing such a theory poses some knotty conceptual and computational problems. Here, more than anywhere else in this article, I attempt only to point out the problems, without offering even sketches of solutions. I think the magnitude of the problem often (or even typically) goes unnoticed in linguistics, where theorists tend to rely on an already-linguistic conception of semantics (a 'language of thought' of some sort) without focusing on the far deeper difficulties for modeling non-linguistic concepts (cf. Millikan 1987). The last thirty years of animal cognition research leaves little doubt that non-linguistic animals *have* complex concepts and can reason with these, and in general can have rich, active mental lives — despite their inability to express their thoughts to others (Vauclair 1996, Hauser 2000, Griffin 2001, Hurford 2007). These cognitive systems predated language, and form the cognitive foundation for word and sentence meanings today. Thus the problems involved in developing an adequate theory of meaning are very broad, and extend far beyond the confines of language or linguistics. Indeed many of the problems have been recognized most clearly in artificial intelligence and robotics, where attempts to build computers that can execute simple but novel motor acts, recognize objects, or recognize basic referents and thus implement even the roughest approximation to 'meaning' have thus far been relative failures. Some of the key missing ingredients of a rich cognitive theory of meaning include a sub-theory of context, and a theory of relevance.

6.1. Shannon Information as a Foundation

Claude Shannon's formalization of 'information' as a quantifiable mathematical entity was a bold, unifying theoretical move, recognized as revolutionary almost immediately upon its publication (Shannon & Weaver 1949). The success of information theory in the domain of technology would be hard to overstate: This formalization was the basis for all subsequent work on digital representation and communication theory, without which today's digital world would be unthinkable, where virtually all communicated material (text, speech, music, images, video and other data) is rendered as a pattern of bits. Shannon's paper introduced the very term 'bit' and the underlying conceptual framework of the digital revolution. Shannon's 'information' was also recognized as deeply interesting theoretically, because its intimate formal connection with the physical concept of entropy offers a link between the inanimate world of particles and probabilities, and the biologically critical worlds of information and meaning. However, Shannon and co-inventor Norbert Wiener both clearly recognized that the revolution they sparked was only partial, because 'information' in this formalization is far from identical with information as normally understood. In particular, Shannon and his popularizer Weaver were both explicit in 1949 that Shannon information fails to incorporate any notion of the *meaning* of a signal. This limitation leads to some non-intuitive propositions in information theory (e.g., that the 'information' in white noise is greater than that in a symphony or speech). Despite Shannon's own clarity on the limitations of 'information' in his sense, this caveat has been largely ignored on two important fronts. From a practical viewpoint the distinction between meaning and information has

become blurred (e.g., in engineering), and from a theoretical viewpoint Shannon's call for an extension of his concepts into the domain of true, biologically-relevant meaning has gone unanswered (e.g., in cognitive science or neuroscience).

Although its incompleteness has periodically led to a call to abandon Shannon information theory entirely (e.g., King 2004), this would be unwise given the manifest success of this theory in all domains to which it has been earnestly applied (both technology and neuroscience), along with the steady improvements in the theory (MacKay 2003). Thus, I think the goal for an eventual theory of meaning should be to build upon Shannon's formalism, incorporating his theorems and extending them. I suggest that two key desiderata for such an extended theory of information, incorporating meaning, are formalizations of context and of relevance.

6.2.   Context and Relevance — 'One Man's Signal is Another Man's Noise'

The same signal may be meaningful in one context and meaningless, or meaningful but irrelevant, in another. At several levels this context-dependence is captured by the phrase quoted above. Meaning must be defined relative to some *context*: A broad temporal-spatial window of data, both organism-internal and -external, much larger than the signal whose information is to be interpreted. This context, provides the data relative to which the meaning of any signal is interpreted. A signal (e.g., white noise) may have a meaning of 0, despite its information-rich high bit rate. This distinction may help to resolve the non-intuitive nature of Shannon information: A signal could have high information and low meaning, or lower information (e.g., speech or music, which are quite redundant) and high meaning. *Relevance*, a basic quantity in any adequate formal theory of pragmatics (Sperber & Wilson 1986), depends not just on current external context, but also on an individual's current cognitive state: Drives, goals, unanswered questions, hypotheses being processed. Relevance is thus in the eye of the beholder, and demands a formalization of external context *and* goal-directed internal context.

We should in principle be able to define ideal observer ('ideal interpreter') models that can extract all the possible connections between all possible signals for a given world and goal context. Unfortunately, the well-known combinatorial explosion that results poses serious obstacles to using such models to control action, because a set of computations subject to unconstrained combinatorial explosion is of little use in real-time computation of meaning. This is the infamous 'frame problem' in artificial intelligence (Ford & Pylyshyn 1996), and the 'solutions' to the frame problem currently on offer in AI all essentially involve *a priori* limits on the extent of this explosion: All variants of what Simon long ago dubbed 'bounded rationality' (Simon 1957), or of Chomsky's innate biases. However, it is unclear that such bounded models can do justice to the seemingly unfettered connection-finding revealed by individual human linguistic creativity, or of the social ramifications of this creativity, as seen both in culture and science. While discussions of the frame problem in technology have grown less central as various work-arounds have been developed, the central

epistemological issue in understanding the mind is not solved, or even obviously confronted, by such 'solutions' (for discussion see Fodor 2000: Chap. 2). Further, no ideal interpreter model alone can capture the relevance of a signal without an additional specification of goals, problem states, current behavioural sequence, current location, etc. There is a considerable amount of explicit computational theory still missing here.

6.3.   The Future: Comparative Cognition Meets Formal Semantics?

Given the problems context-dependent combinatorial explosion causes for contemporary computers and robots, the remarkable fact is that organisms seem to rarely suffer from the frame problem. Indeed, simple motor tasks that seem trivial to us (or to a monkey or a dog) — locomoting around obstacles, negotiating novel paths successfully, or picking up objects without breaking them — remain daunting for today's robots. At the level of perception, perceptual 'mistakes' like illusions are the exception and not the rule, and we seem quite effortlessly to exclude a huge variety of possible interpretations, converging reliably on a relatively accurate but extremely flexible model of the world — again a trick that evades today's best machines and algorithms. In computational linguistics, even simple sentences generate hundreds of possible parses — but we humans rarely even consider more than one of them. One thing that seems common to many of these feats is our ability to use context of various sorts to prune away all but the most probable branches of the tree of possibilities. Our ability to evaluate the relevance of various possible interpretations builds on this more basic context-dependence to explore models of the future or possible worlds. Almost all of this computational generation and pruning is unconscious (perhaps necessarily so, as I have argued in Fitch 2005a, 2008). Furthermore, most of these processing capabilities must have predated the evolution of language, since effortless incorporation of context in decisions of relevance typifies the behavior of a dog or chimpanzee as much as a human. Thus, in some sense, the conceptual and neural basis of 'meaning' is a more basic problem than, and its solution should be logically prior to, an understanding of semantics in natural language. Thus, unfortunately, a general theory of 'meaning' ultimately demands a complete theory of how brains make minds, clearly one of the hardest problems left for science to solve. Ultimately, I believe that new theoretical tools will be necessary to understand meaning in the more general non-linguistic sense I have been discussing, and that the study and modeling of non-linguistic animal cognition will play a crucial role in such an enterprise. For now, an attack on the problem from multiple (hopefully someday converging) perspectives will be required.

Linguistics, in the guise of formal semantics, potentially has something to offer this enterprise. Contemporary semanticists have developed a rather powerful set of theories and formalisms, with truth-value, possible word, and model-theoretic semantics among the prominent theoretical approaches, and a variety of formalisms based upon propositional and predicate calculus and their extensions (Portner 2005). Such approaches are unlikely to solve some of the deeper problems of an embodied (organism-dependent) and context-dependent theory

of meaning, precisely because they intentionally abstract away from such problems (see e.g., Montague 1974). Nonetheless, the tools provided by formal semantics should play an important role in our final understanding of linguistic semantics by providing rigorous definitions of the sorts of problems that must be solved (e.g., logical entailment or scope of quantifiers). Contemporary semantics appears largely to take for granted the existence of non-linguistic models of the world (though work on spatial language provides a welcome, if narrowly circumscribed exception: Landau & Gleitman 1985, Landau & Jackendoff 1993). But real progress in understanding this extra-linguistic context- and relevance-sensitive domain of basic cognition will require considerably more work in this direction, (cf. Jackendoff 2002). Until a well-developed, mathematically-formalized cognitive theory of meaning, applicable to animal cognition and including basic reference and context-dependent relevance, is available, any biologically-based theory of language will remain incomplete.

7.   Conclusions and Some Outstanding Biolinguistic Questions, Framed as Testable Hypotheses

With these prolegomena, I have tried to clarify some core problems that face the new science of biolinguistics. The sociological problems discussed at the outset should be soluble with good-will, mutual respect, and self-imposed restraint. Sober biolinguists will recognize that the core problems facing this field are far too big for any one individual to solve on their own (if only because mastery of all the relevant disciplines is impossible for even the most gifted polymath), and will team up to solve them together. I am thus guardedly optimistic that the fascination of the questions and exciting promise of new techniques and approaches will sweep away many traditional barriers to success.

In contrast, the three problem areas that form the heart of this article pose serious scientific challenges. Each is daunting in its own right. When these challenges are combined, it becomes clear that developing a biological understanding of human language is one of, if not the, most difficult problems in all of contemporary science. Although I have tried where possible to indicate possible solutions to at least some aspects of the problems discussed, my primary motivation in this article was simply to clarify the problems themselves. I think that all researchers interested in biolinguistics can profit from musing over these difficulties, and trying to clarify their nature. At the very least, a meditation on the gravity and breadth of these problems can induce a humility about one's own attempts at solutions, perhaps contributing somewhat towards remediation of the sociological problems that hinder the field. But in any case, a clear understanding and statement of unsolved problems is the best spur to their solution.

I emphasized above that the model for progress in biolinguistics will be empirical testing of theoretical predictions, along the lines of physics or molecular biology. Thus I end this article by taking a dose of my own proposed medicine, recapping one testable hypothesis and presenting six more, spanning the range of the problem spaces discussed in this article. Hypothesis (A) below is

recapped from section 3, as a reminder of the type of multifaceted research program that we will need to find answers to *any* of these questions. I imagine key contributions by researchers in disciplines as diverse as field and laboratory ethology, theoretical, comparative and historical linguistics, developmental biology and psychology, molecular genetics, experimental psychology, computational linguistics, comparative neuroanatomy, sociology and brain imaging. I will make no attempt to flesh out the theoretical underpinnings of these hypotheses, or to detail the experiments that would be involved in testing them. These are left as an exercise to the reader, as a prolegomenon is best summed up with questions, rather than answers. While these questions don't begin to exhaust the list of testable hypotheses in biolinguistics, I hope they give some sense of the potential interest, breadth and promise of this nascent field, and illustrate the future need for broad and productive interdisciplinary collaboration.

(A)   Speech Followed Laryngeal Descent

If size exaggeration was a pre-adaptation for speech (Fitch 2002), human genes controlling male pubertal laryngeal descent should have fixated before those involved in complex vocal control (Hypothesis H3, section 3 above).

(B)   Speech Entails Babbling

If 'closing the loop' between production and perception is a prerequisite for complex vocal learning, all vocal learning species should normally babble (show an early stage of autostimulatory vocal play, e.g., sub-song in birds; Fitch 2006a, 2006b); untested species include pinnipeds, bats, cetaceans.

(C)   Signal Imitation

If vocal and visuomanual imitation both reflect an abstract domain-general capacity for 'mimesis' (Donald 1991), auditory and visual imitation abilities in individual humans should be closely correlated; if they reflect independent, separately-evolved mechanisms there should be no such correlation.

(D)   Syntactic Power

If human sentence-parsing capacities indeed occupy the mildly-context sensitive level of the formal language hierarchy (Joshi *et al.* 1991, Stabler 2004), the additional form of memory involved in processing grammars beyond the finite-state level should have the characteristics of a queue, rather than a stack.

(E)   Language Acquisition

If human language acquisition is just a special case of a general innate capacity for acquiring culture (Tomasello 1999), then individual children's progress in acquiring language should be closely correlated, both temporally and across individuals, with their progress in other aspects of socialization and mastery of non-linguistic culture (cf. Markson & Bloom 1997).

(F)   Semantics and Neuronal Arborization

If natural language has cognitive access to conceptual mechanisms that are encapsulated in other species (e.g., chimpanzees), populations or subclasses of neurons with broadened dendritic or axonal arbors should quantitatively distinguish our brains from a chimpanzee brain, and these arbors should be widely distributed throughout the brain rather than restricted to traditional 'language' areas (cf. Enard *et al.* 2009).

(G)   Plasticity of 'Critical Periods'

If epigenetic interaction between genes and external environment plays a key role in developing the neural circuits underlying language (Bates 1999), 'sensitive periods' (Lenneberg 1967) during which such interactions are possible should be plastic; in particular, some classes of extreme environmental change (e.g., adoption) should be capable of 'resetting' the language acquisition system in young enough children, with a concomitant change in gene expression patterns in the child's brain — this should not be true of epigenetic processes dependent only on the early-developing organism-internal environment.

One could easily generate many more such hypotheses. The difficulties lie not in hypothesis generation but in developing empirical research programmes to test such ideas. If the current essay helps current and future workers in this new field reject, or confirm, any one of these hypotheses, I would be very pleased.

References

Allan, Keith. 2007. 'General' or 'universal' grammar: From the modistae to Chomsky. In Keith Allan (ed.), *The Western Classical Tradition in Linguistics* (Equinox Textbooks & Surveys in Linguistics), 153–187. London: Equinox.

Arbib, Michael A. 2005. From monkey–like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences* 28(2), 105–167.

Aristotle. 350 BCE. *The History of Animals*. 1970 edn. [transl. by A.L. Peck]. London: Heinemann.

Austin, Peter & Joan Bresnan. 1996. Non–configurationality in Australian aboriginal languages. *Natural Language and Linguistic Theory* 14, 215–268.

Ballard, Dana. 1999. *An Introduction to Natural Computation* (Complex Adaptive Systems). Cambridge, MA: MIT Press.

Bates, Elizabeth. 1999. Plasticity, localization and language development. In Sarah H. Broman & Jack M. Fletcher (eds.), *The Changing Nervous System: Neurobehavioral Consequences of Early Brain Disorders*, 214–253. New York: Oxford University Press.

Boeckx, Cedric. 2010. *Language in Cognition: Uncovering Mental Structures and the Rules behind Them*. Malden, MA: Wiley–Blackwell.

Braitenberg, Valentino. 1977. *On the Texture of Brains* (Ecological Studies). New York: Springer-Verlag.

Calvin, William H. 1983. A stone's throw and its launch window: Timing precision and its implications for language and hominid brains. *Journal of Theoretical Biology* 104, 121–135.

Camper, Peter. 1779. Account of the organs of speech of the orang outang. *Philosophical Transactions of the Royal Society of London* 69, 139–159.

Caplan, David. 1987. *Neurolinguistics and Linguistic Aphasiology: An Introduction* (Cambridge Studies in Speech Science and Communication). Cambridge: Cambridge University Press.

Carroll, Sean B. 2000. Endless forms: The evolution of gene regulation and morphological diversity. *Cell* 101, 577–580.

Carroll, Sean B. 2003. Genetics and the making of homo sapiens. *Nature* 422, 849–857.

Carroll, Sean B. 2005. *Endless Forms Most Beautiful*. New York: W.W. Norton.

Carroll, Sean B., Jennifer K. Grenier & Scott Weatherbee. 2005. *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design*. 2nd edn. Malden, MA: Blackwell Science.

Catchpole, Clive K. 1973. The functions of advertising song in the sedge warbler (Acrocephalus schoenobaenus) and reed warbler (A. scirpaceus). *Behaviour* 46, 300–320.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1966. *Cartesian Linguistics: A Chapter in the History of Rationalist Thought*. New York: Harper & Row.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36, 1–22.

Cutler, Ann. 1996. Prosody and the word boundary problem. In James. L. Morgan & Katherine Demuth (eds.), *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*, 87–100. Hillsdale, NJ: Lawrence Erlbaum Associates.

Darwin, Charles. 1871. *The Descent of Man and Selection in Relation to Sex*. 1st edn. London: John Murray.

Dawkins, Richard. 1986. *The Blind Watchmaker: Why the Evidence of Evolution Reveals a Universe without Design*. New York: W.W. Norton.

Deacon, Terrence W. 1997. *The Symbolic Species: The Co-Evolution of Language and the Brain*. New York: W.W. Norton.

DeCasper, Anthony J. & William P. Fifer. 1980. Of human bonding: Newborn's prefer their mothers' voices. *Science* 208, 1174–1176.

Donald, Merlin. 1991. *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*. Cambridge, MA: Harvard University Press.

Doupe, Allison J. & Patricia K. Kuhl. 1999. Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience* 22, 567–631.

Ehrlich, Paul. 2002. *Human Natures: Genes, Cultures and the Human Prospect*. Washington, DC: Shearwater Books/Island Press.

Enard, Wolfgang, Sabine Gehre, Kurt Hammerschmidt, Sabine M. Hölter, Torsten Blass, Mehmet Somel, Martina K. Brückner, Christiane Schreiweis, Christine Winter, Reinhard Sohr, Lore Becker, Victor Wiebe, Birgit Nickel, Thomas Giger, Uwe Müller, Matthias Groszer, Thure Adler, Antonio Aguilar, Ines Bolle, Julia Calzada–Wack, Claudia Dalke, Nicole Ehrhardt, Jack Favor, Helmut Fuchs, Valérie Gailus–Durner, Wolfgang Hans, Gabriele Hölzlwimmer, Anahita Javaheri, Svetoslav Kalaydjiev, , Magdalena Kallnik, Eva Kling, Sandra Kunder, Ilona Moßbrugger, Beatrix Naton, Ildikó Racz, Birgit Rathkolb, Jan Rozman, Anja Schrewe, Dirk H. Busch, Jochen Graw, Boris Ivandic, Martin Klingenspor, Thomas Klopstock, Markus Ollert, Leticia Quintanilla–Martinez, Holger Schulz, Eckhard Wolf, Wolfgang Wurst, Andreas Zimmer, Simon E. Fisher, Rudolf Morgenstern, Thomas Arendt, Martin Hrabé de Angelis, Julia Fischer, Johannes Schwarz & Svante Pääbo. 2009. A humanized version of Foxp2 affects cortico–basal ganglia circuits in mice. *Cell* 137, 961–971.

Enard, Wolfgang, Molly Przeworski, Simon E. Fisher, Cecilia S. L. Lai, Victor Wiebe, Takashi Kitano, Anthony P. Monaco & Svante Pääbo. 2002. Molecular evolution of FOXP2, a gene involved in speech and language. *Nature* 418, 869–872.

Fehér, Olga, Haibin Wang, Sigal Saar, Partha P. Mitra & Ofer Tchernichovski. 2009. *De novo* establishment of wild–type song culture in the zebra finch *Nature* 459, 564–568.

Felsenstein, Joseph. 1985. Phylogenies and the comparative method. *American Naturalist* 125, 1–15.

Finlay, Barbara L. & Richard B. Darlington. 1995. Linked regularities in the development and evolution of mammalian brains. *Science* 268, 1578–1584.

Fitch, W. Tecumseh. 1994. *Vocal tract length perception and the evolution of language*. Ann Arbor, MI: UMI Dissertation Services.

Fitch, W. Tecumseh. 1997. Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *Journal of the Acoustical Society of America* 102, 1213–1222.

Fitch, W. Tecumseh. 2000a. Skull dimensions in relation to body size in nonhuman mammals: The causal bases for acoustic allometry. *Zoology* 103, 40–58.

Fitch, W. Tecumseh. 2000b. The evolution of speech: A comparative review. *Trends in Cognitive Sciences* 4, 258–267.

Fitch, W. Tecumseh. 2000c. The phonetic potential of nonhuman vocal tracts: Comparative cineradiographic observations of vocalizing animals. *Phonetica* 57, 205–218.

Fitch, W. Tecumseh. 2002. Comparative vocal production and the evolution of speech: Reinterpreting the descent of the larynx. In Alison Wray (ed.), *The Transition to Language* (Oxford Linguistics), 21–45. Oxford: Oxford University Press.

Fitch, W. Tecumseh. 2005a. Computation and cognition: Four distinctions and their implications. In Anne Cutler (ed.), *Twenty-First Century Psycholinguistics: Four Cornerstones*, 381–400. Mahwah, NJ: Lawrence Erlbaum.

Fitch, W. Tecumseh. 2005b. The evolution of language: A comparative review. [Invited review]. *Biology and Philosophy* 20, 193–230.

Fitch, W. Tecumseh. 2006a. On the biology and evolution of music. *Music Perception* 24, 85–88.

Fitch, W. Tecumseh. 2006b. The biology and evolution of music: A comparative perspective. *Cognition* 100, 173–215.

Fitch, W. Tecumseh. 2008. Nano–intentionality: A defense of intrinsic intentionality. *Biology and Philosophy* 23, 157–177.

Fitch, W. Tecumseh. In press. *The Evolution of Language*. Cambridge: Cambridge University Press.

Fitch, W. Tecumseh & David Reby. 2001. The descended larynx is not uniquely human. *Proceedings of the Royal Society of London* B, 268, 1669–1675.

Fitch, W. Tecumseh, Marc D. Hauser & Noam Chomsky. 2005. The evolution of the language faculty: Clarifications and implications. *Cognition* 97, 179–210.

Fodor, Jerry A. 2000. *The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology*. Cambridge, MA: MIT Press.

Fodor, Jerry A. & Zenon W. Pylyshyn. 1988. Connectionism and cognitive architecture: A critical analysis. *Cognition* 28, 3–71.

Ford, Kenneth M., & Zenon W. Pylyshyn (eds.). 1996. *The Robot's Dilemma Revisited: The Frame Problem in Artificial Intelligence*. Norwood, NJ: Ablex.

Frey, Roland & Tobias Riede. 2003. Sexual dimorphism of the larynx of the Mongolian Gazelle (*Procapra gutturosa* Pallas, 1777) (Mammalia, Artiodactyla, Bovidae). *Zoologischer Anzeiger* 242, 33–62.

Friederici, Angela D., Jörg Bahlmann, Stefan Heim, Ricarda I. Schubotz & Alfred Anwander. 2006. The brain differentiates human and non–human grammars: Functional localization and structural connectivity. *Proceedings of the National Academy of Sciences (USA)* 103, 2458–2463.

Friederici, Angela. D., Karsten Steinhauer & Erdmut Pfeifer. 2002. Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences (USA)* 99, 529–534.

Gardner, Allen R. & Beatrice T. Gardner. 1969. Teaching sign language to a chimpanzee. *Science* 165, 664–672.

Gehring, Walter J. & Kazuho Ikeo. 1999. Pax 6: Mastering eye morphogenesis and eye evolution. *Trends Genet* 15, 371–377.

Ghazanfar, Asif. A., Hjalmar K. Turesson, Joost X. Maier, Ralph van Dinther, Roy D. Patterson & Nikos K. Logothetis. 2007. Vocal–Tract Resonances as Indexical Cues in Rhesus Monkeys. *Current Biology* 17, 425–430.

Gilbert, Scott F. 2003. *Developmental Biology*. 7th edn. Sunderland, MA: Sinauer.

Givón, Talmy. 2002. *Bio-Linguistics: The Santa Barbara Lectures*. Amsterdam: John Benjamins.

Goodwin, Brian C. & L. E. H. Trainor. 1983. The ontogeny and phylogeny of the pentadactyl limb. In Brian C. Goodwin, N. Holder & Christopher C. Wylie (eds.), *Development and Evolution: The Sixth Symposium of the British Society for Developmental Biology* (British Society for Developmental Biology Symposia), 75–98. Cambridge: Cambridge University Press.

Gottlieb, Gilbert. 1992. *Individual Development and Evolution: The Genesis of Novel Behavior*. New York: Oxford University Press.

Gould, Steven J. 1976. In defense of the analog: A commentary to N. Hotton. In

Bruce R. Masterton, William Hodos & Harry Jerison (eds.), *Evolution, Brain and Behavior: Persistent Problems*, 175–179. New York: John Wiley & Sons.

Griffin, Donald R. 2001. *Animal Minds: Beyond Cognition to Consciousness*. Chicago, IL: Chicago University Press.

Hagoort, Peter. 2005a. Broca's complex as the unification space for language. In Anne Cutler (ed.), *Twenty-First Century Psycholinguistics: Four Cornerstones*, 157–172. Mahwah, NJ: Lawrence Erlbaum.

Hagoort, Peter. 2005b. On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences* 9, 416–423.

Harvey, Paul H. & Mark D. Pagel. 1991. *The Comparative Method in Evolutionary Biology* (Ecology and Evolution). Oxford: Oxford University Press.

Hauser, Marc D. 2000. *Wild Minds: What Animals Really Think*. New York: Henry Holt.

Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The language faculty: What is it, who has it, and how did it evolve? *Science* 298, 1569–1579.

Hayes, Bruce. 1995. *Metrical Stress Theory: Principles and Case Studies*. Chicago: The University of Chicago Press.

Hayes, Catherine. 1951. *The Ape in Our House*. New York: Harper.

Hepper, Peter G. 1991. An examination of fetal learning before and after birth. *Irish Journal of Psychology* 12, 95–107.

Hill, Robert S. & Christopher A. Walsh. 2005. Molecular insights into human brain evolution. *Nature* 437, 64–67.

Hurford, James R. 2007. *The Origins of Meaning* (Language in the Light of Evolution). Oxford: Oxford University Press.

Iwatsubo, T., S. Kuzuhara, A. Kanemitsu, H. Shimada, & Y. Toyokura. 1990. Corticofugal projections to the motor nuclei of the brainstem and spinal cord in humans. *Neurology* 40, 309–312.

Jackendoff, Ray. 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. New York: Oxford University Press.

Janik, Vincent M. & Peter B. Slater. 1997. Vocal learning in mammals. *Advances in the Study of Behavior* 26, 59–99.

Jenkins, Lyle. 2000. *Biolinguistics: Exploring the Biology of Language*. Cambridge: Cambridge University Press.

Jenkins, Lyle (ed.). 2004. *Variation and Universals in Biolinguistics* (North-Holland Linguistic Series: Linguistic Variations 62). Amsterdam: Elsevier.

Johansson, Sverker. 2005. *Origins of Language: Constraints on Hypotheses*. Amsterdam: John Benjamins.

Joshi, Aravind K., K. Vijay–Shanker & David J. Weir. 1991. The convergence of mildly context–sensitive formalisms. In Peter Sells, Stuart M. Shieber & Thomas Wasow (eds.), *Processing of Linguistic Structure*, 31–81. Cambridge, MA: MIT Press.

Jürgens, Uwe. 1994. The role of the periaqueductal grey in vocal behaviour. *Behavioural Brain Research* 62, 107–117.

Kenstowicz, Michael J. 1994. *Phonology in Generative Grammar* (Blackwell Textbooks in Linguistics). Malden, MA: Blackwell.

King, Barbara J. 2004. *The Dynamic Dance: Nonvocal Communication in African Great*

Apes. Cambridge, MA: Harvard University Press.

Kirby, Simon, Mike Dowman & Thomas L. Griffiths. 2007. Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences (USA),* 104, 5241–5245.

Kirby, Simon, Kenny Smith & Henry Brighton. 2004. From UG to universals: Linguistic adaptation through iterated learning. *Studies in Language* 28, 587–607.

Kirschner, Marc W. & John C. Gerhart. 2005. *The Plausibility of Life: Resolving Darwin's Dilemma*. New Haven, CT: Yale University Press.

Kirschner, Marc W. & John C. Gerhart. 1998. Evolvability. *Proceedings of the National Academy of Sciences* 95, 8240–8427.

Knörnschild, Mirjam, Martina Nagy, Markus Metz, Frieder Mayer & Otto von Helversen. 2009. Complex vocal imitation during ontogeny in a bat. *Biology Letters* 5, 1–4.

Landau, Barbara & Lila Gleitman. 1985. *Language and Experience: Evidence from the Blind Child* (Cognitive Science Series). Cambridge, MA: Harvard University Press.

Landau, Barbara & Ray Jackendoff. 1993. What and where in spatial language and spatial cognition. *Behavioral and Brain Sciences* 16, 217–265.

Langer, Susanne K. 1942. *Philosophy in a New Key: A Study in the Symbolism of Reason, Rite and Art*. 3rd edn. Cambridge, MA: Harvard University Press.

Langer, Susanne. K. 1962. *Philosophical Sketches*. Baltimore, MD: Johns Hopkins University Press.

Lenneberg, Eric H. 1967. *Biological Foundations of Language*. New York: Wiley.

Lieberman, Philip. 1975. *On the Origins of Language: An Introduction to the Evolution of Human Speech* (The Macmillan Series in Physical Anthropology). New York: Macmillan.

Lieberman, Philip. 1984. *The Biology and Evolution of Language*. Cambridge, MA: Harvard University Press.

Lieberman, Philip. 1998a. *Eve Spoke: Human Language and Human Evolution*. New York: W. W. Norton.

Lieberman, Philip. 1998b. On the evolution of human syntactic ability: Its pre–adaptive bases, motor control and speech. *Journal of Human Evolution* 14, 657–668.

Lieberman, Philip. 2000. *Human Language and Our Reptilian Brain: The Subcortical Bases of Speech, Syntax, and Thought* (Perspectives in Cognitive Neuroscience). Cambridge, MA: Harvard University Press.

Lieberman, Philip. 2006. *Toward an Evolutionary Biology of Language*. Cambridge, MA: Harvard University Press.

Lieberman, Philip, Dennis H. Klatt & William H. Wilson. 1969. Vocal tract limitations on the vowel repertoires of rhesus monkeys and other nonhuman primates. *Science* 164, 1185–1187.

MacKay, David J.C. 2003. *Information Theory, Inference, and Learning Algorithms*. Cambridge: Cambridge University Press.

Markson, Lori & Paul Bloom. 1997. Evidence against a dedicated system for word learning in children. *Nature* 385, 813–815.

Marler, Peter. 1991. The instinct to learn. In Susan Carey & Rochel Gelman (eds.),

The Epigenesis of Mind: Essays on Biology and Cognition (Jean Piaget Symposium Series), 37–66. Hillsdale, NJ: Lawrence Erlbaum Associates.

Marler, Peter. 2000. Origins of music and speech: Insights from animals. In Nils L. Wallin, Björn Merker & Steven Brown (eds.), *The Origins of Music*, 31–48. Cambridge, MA: MIT Press.

Marler, Peter & Hans Slabbekoorn. 2004. *Nature's Music: The Science of Birdsong*. New York: Academic Press.

Marr, David. 1982. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco, CA: W. H. Freeman & Co.

Mehler, Jacques, Peter Jusczyk, Grishlaine Lambertz, Nilofar Halsted, Josiane Bertoncini & Claudine Amiel–Tison. 1988. A precursor of language acquisition in young infants. *Cognition* 29, 143–178.

Miller, George A. & Noam Chomsky. 1963. Finitary models of language users. In R. Duncan Luce, Robert R. Bush & Eugene Galanter (eds.), *Handbook of Mathematical Psychology*, vol. II, 419–492. New York: John Wiley & Sons.

Millikan, Ruth G. 1987. *Language, Thought, and Other Biological Categories: New Foundations for Realism*. Cambridge, MA: MIT Press.

Mills, M., & E. Melhuish. 1974. Recognition of mother's voice in early infancy. *Nature* 252, 123–124.

Montague, Richard. 1974. Universal grammar. *Theoria* 36, 373–398. [Reprinted in Richmond H. Thomason (ed.), *Formal Philosophy: Selected Papers of Richard Montague*, 222–246. New Haven, CT: Yale University Press.]

Morgan, Elaine. 1997. *The Aquatic Ape Hypothesis*. London: Souvenir Press.

Okanoya, Kazuo. 2002. Sexual display as a syntactic vehicle: The evolution of syntax in birdsong and human language through sexual selection. In Alison Wray (ed.), *The Transition to Language*, 46–63. Oxford: Oxford University Press.

Pinker, Steven. 1994. *The Language Instinct*. New York: William Morrow and Company, Inc.

Pinker, Steven & Ray Jackendoff. 2005. The faculty of language: What's special about it? *Cognition* 95, 201–236.

Poeppel, David & David Embick. 2005. Defining the relation between linguistics and neuroscience. In Anne Cutler (ed.), *Twenty-First Century Psycholinguistics: Four Cornerstones*, 103–120. Mahwah, NJ: Lawrence Erlbaum.

Poole, Joyce H., Peter L. Tyack, Angela S. Stoeger–Horwath & Stephanie Watwood. 2005. Elephants are capable of vocal learning. *Nature* 434, 455–456.

Portner, Paul H. 2005. *What is Meaning: Fundamentals of Formal Semantics* (Fundamentals in Linguistics). Oxford: Blackwell.

Premack, David. 1971. Language in chimpanzee? *Science* 172, 808–822.

Ralls, K., P. Fiorelli & S. Gish. 1985. Vocalizations and vocal mimicry in captive harbor seals, *Phoca vitulina*. *Canadian Journal of Zoology* 63, 1050–1056.

Reby, David & Karen McComb. 2003. Anatomical constraints generate honesty: Acoustic cues to age and weight in the roars of red deer stags. *Animal Behaviour* 65, 519–530.

Reby, David, Karen McComb, Bruno Cargnelutti, Chris Darwin, C., W. Tecumseh Fitch, & Tim Clutton–Brock. 2005. Red deer stags use formants as

assessment cues during intrasexual agonistic interactions. *Proceedings of the Royal Society of London*, B, 272, 941–947.

Reik, Wolf. 2007. Stability and flexibility of epigenetic gene regulation in mammalian development *Nature* 447, 425–432.

Richards, Whitman (ed.). 1988. *Natural Computation*. Cambridge, MA: MIT. Press.

Ridley, Matt. 2003. *Nature via Nurture: Genes, Experience, and What Makes Us Human*. New York: Harper Collins.

Rolls, Edmund T. & Gustavo Deco. 2001. *Computational Neuroscience of Vision*. Oxford: Oxford University Press.

Savage–Rumbaugh, E. Sue. 1986. *Ape Language: From Conditioned Response to Symbol* (Animal Intelligence). New York: Columbia University Press.

Savage–Rumbaugh, E. Sue, Jeannine Murphy, Rose A. Sevcik, Karen E. Brakke, Shelly L. Williams & Duane M. Rumbaugh. 1993. Language comprehension in ape and child. *Monographs of the Society for Research in Child Development* 58, 1–221.

Shannon, Claude E. & Warren Weaver. 1949. *The Mathematical Theory of Communication*. Urbana, IL: University of Illinois Press.

Shubin, Neil, Cliff Tabin & Sean Carroll. 1997. Fossils, genes and the evolution of animal limbs. *Nature* 388, 639–648.

Simon, Herbert A. 1957. *Models of Man*. New York: John Wiley.

Simon, Herbert A. 1962. The architecture of complexity. *Proceedings of the American Philosophical Society* 106, 467–482.

Smith, David R. R., Roy D. Patterson, R. Turner, H. Kawahara T. & Irino, T. 2005. The processing and perception of size information in speech sounds. *Journal of the Acoustical Society of America* 117, 305–318.

Spence, Melanie J. & Mark Freeman. 1996. Newborn infants prefer the maternal low–pass filtered voice, but not the maternal whispered voice. *Infant Behavior and Development* 19, 199–212.

Sperber, Dan & Deirdre Wilson. 1986. *Relevance: Communication and Cognition*. Oxford: Blackwell.

Stabler, Edward P. 2004. Varieties of crossing dependencies: Structure dependence and mild context sensitivity. *Cognitive Science* 28, 699–720.

Steels, Luc. 1999. *The Talking Heads Experiment*. Antwerp: Laboratorium.

Striedter, Georg F. 2004. *Principles of Brain Evolution*. Sunderland, MA: Sinauer.

Szathmáry, Eörs. 2001. Origin of the human language faculty: The language amoeba hypothesis. In Jürgen Trabant & Sean Ward (eds.), *New Essays on the Origin of Language*, 55–81. Berlin: Mouton de Gruyter.

Tomasello, Michael. 1995. Language is not an instinct. *Cognitive Development* 10, 131–156.

Tomasello, Michael. 1999. *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.

Tomasello, Michael. 2001. Cultural transmission: A view from chimpanzees and human infants. *Journal of Cross–Cultural Psychology* 32, 135–146.

Tomasello, Michael. 2005. Comment on Everett 2005. *Current Anthropology* 46, 640–641.

van Parijs, Sofie M. 2003. Aquatic mating in pinnipeds: A review. *Aquatic Mammals* 29, 214–226.

Vauclair, Jacques. 1996. *Animal Cognition: An Introduction to Modern Comparative Psychology*. Cambridge, MA: Harvard University Press.

Waddington, Conrad H. 1957. *The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology*. London: Allen and Unwin.

Weissengruber, G. E., Gerhard Forstenpointner, G. Peters, Anna Kübber–Heiss & W. Tecumseh Fitch. 2002. Hyoid apparatus and pharynx in the lion (*Panthera leo*), jaguar (*Panthera onca*), tiger (*Panthera tigris*), cheetah (*Acinonyx jubatus*), and domestic cat (*Felis silvestris* f. *catus*). *Journal of Anatomy (London)* 201, 195–209.

Worden, Robert P. 1995. A speed limit for evolution. *Journal of Theoretical Biology* 176, 137–152.

Wright, Sewall. 1931. Evolution in Mendelian populations. *Genetics* 16, 97–159.

Yerkes, Robert M & Ada W. Yerkes. 1929. *The Great Apes*. New Haven, CT: Yale University Press.

W. Tecumseh Fitch
Universität Wien
Department of Cognitive Biology
Althanstrasse 14
1090 Vienna
Austria
Tecumseh.Fitch@univie.ac.at

Internalism as Methodology

Terje Lohndal & Hiroki Narita

This paper scrutinizes the recent proposal made by Lassiter (2008) that the dichotomy between Chomskyan internalism and Dummett-type externalism is misguided and should be overcome by an approach that incorporates sociolinguistic concepts such as speakers' dispositions to defer. We argue that Lassiter's arguments are flawed and based on a serious misunderstanding of the internalist approach to the study of natural language, failing to appreciate its methodological nature and conclude that Lassiter's sociolinguistic approach is just another instance of externalist attempts with little hope of scientific achievement.

Keywords: externalism; internalism; sociolinguistics

1.   Introduction

In a recent paper, Lassiter (2008) argues that both the Chomskyan internalist approach to human language and the Dummett-type externalism fail to provide an account of some semantic and social facts that he claims are crucial to any linguistic theory. Furthermore, Lassiter claims that we can overcome such difficulties in these two approaches by incorporating some sociolinguistic notions into linguistic theory. We will argue that Lassiter's claim is misguided, and that he misunderstands crucial aspects of the Chomskyan internalist project.

We will first provide an overview of Lassiter's claims, and then examine his critique of the internalist project, and discuss why the alternative he presents is highly problematic. We take the failure of his externalist theory to be rather suggestive of the general feasibility of any scientific investigation of language that rests mainly on an externalist foundation.

2.   Lassiter's Claims

This section is devoted to an overview of the argument presented in Lassiter

(2008). Before we proceed, we have to clarify some of the terminology Lassiter misleadingly adopts. Throughout his paper, Lassiter repeatedly attributes to Noam Chomsky and generative linguistics led by him (which he often calls 'descriptive linguistics') a claim that "a language just is a mental grammar" (p. 619) and calls such a claim 'individualism'. Individualism, as he construes, is a claim that semantic properties, such as reference, of an individual's speech behavior rest exclusively on, and can be explained solely by, the facts and knowledge internal to the individual. This is in itself a serious misrepresentation of the Chomskyan internalism (Chomsky 1995, 2000), as we will argue in the next section, but since his discussion rests heavily on this misrepresentation, we will adopt it and review his arguments with regard to this straw man hypothesis. In order to avoid unnecessary vagueness in terminology, let us tentatively call the 'individualism' under his particular conception sketched here *L-individualism*.

In a nutshell, Lassiter's claim is that (i) there are some sociological facts of human linguistic behavior that can be accounted for only by a linguistic theory that incorporates individuals' intentional contributions to the meaning/reference of linguistic expressions, and (ii) neither L-individualism nor the philosophically dominant tradition of semantic externalism (led by people like Hilary Putnam, Tyler Burge, Michael Dummett, and David Lewis) can satisfy this need. We will articulate these points in what follows.

Basing his argument primarily on observations made by externalists such as those just mentioned, Lassiter claims that there are certain 'crucial' socio-linguistic facts of human speech behavior that highlight the relevance of social contexts in a speech community and also of speakers' intentions. A primary example of this kind is, according to him, individuals' intuitions about the '(in)correctness' of their language use against the standards of the speech community they belong to. Lassiter notes, "sometimes an individual's use of language is just *wrong*, and individuals often acknowledge making mistakes upon reflection or correction" (p. 608). As an illustration, he takes a familiar example from Burge (1979), where we are asked to imagine an English-speaking individual, say Jim, who has rheumatism in his thigh but suspects he has arthritis as a result of having an ailment in his thigh. This individual, not being a doctor, does not know that arthritis is a condition of the joints only, and so when he utters "I have arthritis in my thigh," he is expressing a false belief. In such a circumstance, however, Jim should be able to become aware of his mistake, for example, by being explicitly corrected by a doctor. Given this much, Lassiter argues that "a descriptive theory emphasizing *knowledge* of language" would pre-dict (wrongly, as he argues) that his utterance above is not false but rather is just "true-in-his-idiolect" (p. 609), which he regards as a serious flaw of such a theory. That is, he argues with Burge that the reference of 'arthritis' in such a case is rather fixed by the word's use in the speech community Jim belongs to.

Furthermore, Lassiter claims that L-individualism is seriously flawed in that it does not provide any room for language-external concepts such as speech communities and community standards for normative meaning, in its account of sociologically and contextually varying semantic properties of human speech behaviors, as in Jim's cases above. In this regard, Lassiter more or less endorses Dummett's (1986) claim that an individual's knowledge of his language is merely

"a second-order theory: a partial, and partly incorrect, theory about what the meanings of the expressions are in the common language, that may be represented as a partial theory of what the correct theory of meaning for the language is" (Dummett 1986: 469).

However, Lassiter also claims that we should not take at face value the strong form of externalism that Dummett and other externalists like Putnam and Lewis envisage. His target is what he calls *communitarianism*, a position implicitly or explicitly shared by these externalists' approaches which holds that a language is primarily a social object belonging to a speech community, and that such speech communities exist prior to individual speakers and are capable of determining a unique community language (a *communalect*, in his terms) with or without the individual speakers' cooperation (p. 610). Lassiter argues that we should reject communitarianism, because it makes a clearly unsustainable prediction that 'speech communities' that communalects correspond to should be determinate and isolable real objects that we can find in the world. The problem of this prediction lies in the difficulties in isolating the relevant speech communities in a well-articulated fashion. For example, we cannot define a community corresponding to a communalect language neither by political or institutional boundaries, nor by communicative notions like mutual intelligibility. It is by now an established fact that languages vary both synchronically over geographical and sociopolitical space and diachronically over generations, and moreover that such variation is rather continuous and gradual in most cases, without any sharp divisions in terms of geographic space, sociopolitical boundaries, generation gaps, or mutual intelligibility. Consequently, as Lassiter argues, we should give up any hope to find any objective or absolute criteria for isolating speech communities, which in effect militates against the backbone of communitarianist approaches.

He then goes on to suggest that his sociolinguistic approach to semantics can overcome both the difficulties in L-individualism and those in communitariarism. His alternative theory posits that "the meaning of a word in the mouth of a speaker S is determined by S's dispositions to defer to other speakers with regard to the meaning and use of this word" (p. 623). For example, as for the rheumatism patient Jim in Burge's example, he notes, "if we wish to know what 'arthritis' means in Jim's mouth, we must ask who Jim would defer to with regard to the meaning of this word" (p. 622), say, his doctor, in which case his utterance. "I have arthritis in my thigh" can be still said to be expressing a wrong belief, since "he would be willing to change his use if he *were* to go [to see the doctor—L&N]" (p. 623). This way, we can make room for the effect of the speaker's intuitions of correctness by relocating the effect of normativity to the speaker's dispositions to defer, without making recourse to the dubious communitarianist notion of speech community.

3.   The Externalism vs. Internalism Debate

To begin with, let us first state clearly that we have no problem in accepting Lassiter's counterarguments to L-individualism and communitariarism as such,

since they both seem dubious, although we don't regard his argumentation as particularly strong. What we regard as particularly misguided and problematic is his continuous misrepresentation of Chomsky's internalism (Chomsky 1995, 2000) as L-individualism. In our opinion, L-individualism is a straw man hypothesis that says that the meaning of an individual's speech is to be determined solely by facts internal to the individual. From this, Lassiter subsequently concludes that the dichotomy between Chomsky's position (again, misrepresented as L-individualism) and externalism is misguided and to be remedied by "unify[ing] the two approaches" under the realm of sociolinguistics (p. 607).

Lassiter's "broad outlines of the debate" say, on the one hand, that internalists "believe that the proper object of the scientific study of language is the language of an individual, his idiolect or, in Chomskyan terms, his mental grammar, knowledge of language, or I-language. […] This does not necessarily mean that social aspects of language are unimportant or that they do not admit of a scientific description, though some [internalists] have made this further claim: cf. Chomsky (1975). However, most [internalists] do believe that only individualistic aspects of language can be formalized and used to make predictions (e.g., about entailment and grammaticality)" (p. 608). Elsewhere, he also (correctly) attributes to Chomsky the claim that "only the 'internalist' aspects of language can admit of a truly scientific description" (p. 631). On the other hand, semantic externalists "hold that a language belongs to a community of language users, and that common languages or communalects exist above and beyond individuals. According to this conception, a language has an ontology (e.g., words and grammatical rules, or social practices and/or conventions) and norms (standards of correctness) that are in some sense independent of the linguistic competence of individual speakers" (p. 608). These two "broad outlines" seem to be a sufficiently accurate approximation. Crucially, note that internalism thus understood is primarily a *conjecture about a proper object of the scientific study of language* (which internalists claim to be *I-language*), whereas externalism is rather a philosophical *belief* about the *ontology* of some mind/brain-external aspects of human linguistic behaviors (which would be *E-language* of some sort).

Typically, then, both positions would not, and Lassiter also does not, take issue with the fact that there exist aspects of language purely internal to the human mind/brain (i.e. *I-language*): Thus Lassiter's remark, "I think mental grammars are fully real" (p. 619). Indeed, to the extent that the externalist proposals are even coherent, they presuppose some notion of I-language, as Chomsky repeatedly has stressed (e.g., in Chomsky 2000). An I-language can in many ways be identified as a generative system — an I-language enables a speaker to generate linguistic expressions. It's a (trivial) fact that speakers are able to generate an unlimited amount of hierarchically structured expressions. An I-language makes it possible to make sense of the linguistic creativity we all possess as human beings, a creativity that is unbounded, innovative and uncaused (Chomsky 2009 [1966]; see also McGilvray 2009 on this). This linguistic creativity is in obvious ways individual: It is something that each person uses and how it is used is independent of how other people are using their linguistic creativity. The utterances that are generated are part of what is commonly called E-language.

However, E-language cannot exist without the utterances being generated

in some way or other. Certainly language cannot just appear out of nowhere. Hence individuals have to create it using their linguistic creativity. This creativity, of which I-language is a crucial part, is by definition internal. One can debate what the exact content of I-language is, as people following Chomsky have done since the 1950s, but it is very hard to see that it is possible to debate the reality of I-language. If one wants to study E-language, one has to suppose the existence of some sort of I-language since E-language would not exist without I-language. One can choose not to focus on its existence, but it should be clear that it *has* to exist. More advanced examples can be adduced in support of this, involving central well-known empirical facts about human language that doesn't create themselves (hierarchy, as mentioned above, but also structural relations like c-command and others that are crucial properties of I-languages). Thus, we take it for granted that it should be clear why an E-language approach is presupposing the existence of an I-language.

Surrounding the dichotomy between internalism and externalism are, then, (at least) two separate issues:

(i)   Whether we have reasons to believe in some ontology of 'language' outside of individuals' mind/brain, and

(ii)   Whether we can ever construct a serious scientific theory of such 'language'.[1]

Internalists like Chomsky would typically answer a skeptical "No" to (i) and (ii), as Lassiter correctly acknowledges. Externalists strongly answer "Yes" to (i), rooted in their philosophical belief in the existence of such an 'object'. But they seem less concerned about arguing for an articulated "Yes" to (ii), as far as we can see. This is an important difference. For internalists, (i) is not really an important research question, whereas (ii) is really the question they/we are concerned with. Again, this is an important methodological difference that bears emphasis, as Collins (2009) also underlines.

4.   Evaluating Lassiter's Contributions

Despite the sufficient accuracy of the "broad outlines of the debate" above, which is presented on the second page of his paper, Lassiter continuously misinterprets

---

[1]   McGilvray (2002: 73) provides a nice exposition of what we have in mind for the term *serious science*:

> A serious science is a theory for which there is not only empirical support in the form of the descriptive and explanatory adequacy of a set of formal, explicit principles (adequate to their domain by standards that are universal, although adjusted to a specific domain) and evidence of progress (a history of revision of theories with good reason to think that there have been improvements in adequacy, simplicity, and explicit statement), but some reason to think that the theory's principles can be accommodated to the principles of other, relevant sciences. In the science of mind, the relevant science would, presumably, be some branch of biology — perhaps a much-revised form of neurophysiology.

what internalism is all about elsewhere in his paper, and repeatedly misrepresents it as L-individualism. For example, he makes the following odd claim: "In contrast to the assumptions of thinkers from Chomsky to Putnam, I do not think that externalism and mentalism are incompatible: I think mental grammars are fully real, though I do deny the claim that language just is a mental grammar" (p. 619). Here, Lassiter is attributing to Chomskyan internalism the "claim that language just is a mental grammar." In the same vein, elsewhere he also says, "Individualists hold that an individual's language just is her idiolect" (p. 610).[2] This is a serious misrepresentation of the internalist claim. As noted above, the core claim of internalists who, like Chomsky, seek a naturalistic theory of language is that the proper object of a serious linguistic science should be organism-internal aspects of human language (namely I-language). Internalists never deny that there are phenomena broadly related to language (in particular to language use) that are beyond the narrow confines of the architecture of the human mental grammar (I-language). Such phenomena would surely include prescriptive pressures from the linguistic community, speakers' intentionality for communicative success, and all sorts of other E-language phenomena that Lassiter and others argue for. What internalists doubt is rather the feasibility and/or legitimacy of providing a serious science of any mind-external phenomena such as these. Thus, Chomsky (1995) writes: "[G]eneral issues of intentionality, including those of language use, cannot reasonably be assumed to fall within naturalistic inquiry" (p. 27); see also Chomsky (2000) for much relevant discussion.

Cast in this real internalism vs. externalism debate, we cannot find any compelling reason to believe that Lassiter constructed even a relevant argument for his conclusion that "the choice between individualism [referring to the Chomskyan internalism] and externalism is a false one" (p. 630). He is mostly attacking the 'claim that language just is a mental grammar', i.e. the incorrect L-individualism, a claim never defended by Chomsky, but Lassiter never addresses all the serious issues raised by the Chomskyan internalism that an externalist theory would have to face. Rather, his alleged 'theory' is just another instantiation of externalism, expressing but not quite arguing for his intuitive "Yes" to (i) and (ii).

Let us now turn to the more specific aspects of Lassiter's proposals, where we in particular will focus on important problems surrounding an E-language approach and why Chomsky and others have focused on studying I-language. As we have seen, Lassiter's claims are as follows:

(A)   There are facts that cannot be addressed purely internalistically (such as individuals' intuitions on (in)correctness of language use, e.g., the imprecision of a rheumatism patient's usage of the word 'arthritis', as seen above).

---

[2]   An additional note on Lassiter's terminology may be in order here. Throughout his paper, he uses the notion 'idiolect' as more or less synonymous with grammar or I-language. This is a misunderstanding. Never has Chomsky squared idiolect and I-language; on the contrary, he has been very explicit in numerous writings that idiolect and I-language are very different (e.g., Chomsky 1986, 2000). A notion like 'idiolect' is very much like 'dialect' and 'language'; vague and ambiguous notions that are notoriously hard to define. This fact, together with the numerous other often remarkably vague notions used by Lassiter, makes it quite hard to assess his theory.

(B)   Any linguistic theory must account for these facts.

(C)   We can actually construct (or at least imagine) an explanatory theory of such facts which incorporates externalist (I-language-external) concepts like speakers' dispositions to defer to normativity or authority in their speech community.

As for the claim in (A), we looked at an example borrowed from Burge where the meaning of a rheumatism patient's word 'arthritis' is determined by who the speaker would defer to, and in what fashion. That is, if we want to know what the speaker meant to refer to by the word 'arthritis', we are told by Lassiter to ask who the speaker would defer to with regard to the meaning of this specific word. Internalists would have no problem accepting Lassiter's mundane (and trivially true) claim that such notions as individuals' dispositions to defer, and the E-linguistic system (communalect) that the totality of a person's dispositions to defer in a particular communicative situation map out, obviously go beyond a purely internalistic account. Thus (A) should not be at issue here.

However, internalists may very well be inclined to deny (B) and (C). As for (B), we fail to see any serious justification of this claim in Lassiter's paper, apart from his personal belief that these facts regarding (A), like most phenomena investigated by sociolinguists, are of general interest. It is not clear that an internalist theory needs to take into account the facts of Lassiter's interest, given that the past fifty years of generative investigation have provided more than ample evidence that the I-linguistic mental system can be fruitfully studied purely internalistically, under the abstraction from the external fluctuations from sociological circumstances or intentions of speakers or the like. Thus, nobody would claim that dispositions to defer or other sorts of an individual's social intentions have any influence on the computational properties of the I-linguistic mechanism that generates the mental compositions of hierarchical structures of words and sentences. As long as they can construct and investigate the science of I-language, internalists are fine to admit that they have to leave whatever remains beyond the reach of their I-linguistic science for the time being, such as the facts that Lassiter and other externalists' interest think are very important (see Chomsky 1995, 2000, McGilvray 2009, and Hinzen 2006a, 2006b, among others). In this regard, it is not clear what Lassiter thinks would go wrong if an internalist approach to human language set the facts of his interest aside, and left it to other disciplines such as sociolinguistics to investigate I-language-external facts. Lassiter never articulates his claim on this point, so we do not see any reason to abandon the internalist theory of I-language. In this regard, we completely agree with the following remark by Chomsky (1995: 50): "As for sociolinguistics, it is a perfectly legitimate inquiry, externalist by definition. It borrows from internalist inquiry into humans, but suggests no alternative to it." By contrast, Lassiter somehow believes that not just *some* but *any* linguistic discipline must account for the relevant facts, but he never articulates why that should be.

More to the point, most internalists suspect that we need to understand the I-language much better than we currently do before we can even start to attempt

at pursuing some serious understanding of how I-language is embedded in sociolinguistic contexts, and specifically how the utterances generated with recourse to I-language are used in a given context to refer to things outside the head. Admittedly, internalists have scarcely started to understand how an I-language that an individual possesses contributes to the semantico-pragmatic performance of that individual in a sufficiently comprehensive way, and thus it would be an inextricable leap at this point to broaden the object of study to individuals' varying deference and any other E-linguistic notions; hopelessly complicating the task. We need a more complete understanding of the internal properties of I-language before we can even attempt to try to understand how individuals utilize them to deal with all sorts of E-language phenomena. This is a very different methodology than that of externalists, and a difference that Lassiter seems to have failed to notice. In total, we see that there are both theoretical and metholodogical reasons to be skeptical regarding (B) (cf. Collins 2009).

Furthermore, it is hard to see how one could even imagine a successfully explanatory theory within the framework that Lassiter pursues given that he does not acknowledge the importance of I-language. In this regard, we side with Chomsky (1995, 2000) and McGilvray (1998, 2002, 2009), and many others, in being very skeptical about the feasibility of (C). The claim that we can construct or should be able to construct an explanatory theory of (E-)language by incorporating various sociolinguistic notions is central to Lassiter (2008). Unfortunately, Lassiter never defines crucial notions that are part of his theory, for example, 'deferential dispositions', 'communicative success', 'social identification' or, elsewhere in the paper, important notions such as 'norms', 'correctness' and 'reference', in a sufficiently meticulous way that enables us to derive predictions from his theory. Rather, to address these notions, he seems to borrow heavily from common sense understandings of these terms. However, we have no reason to expect that any commonsense understanding of words like these can merit scientific investigations. Compare Chomsky's (1999: 113) remark: "[T]here is no reason to suppose that common usage of such terms as 'language' or 'learning' (or 'belief' or numerous others like them), or others belonging to similar semantic fields in other linguistic systems, will find any place in attempts to understand the aspects of the world to which they pertain, just as no one expects the common sense terms 'energy' or 'liquid' or 'life' to play a role in the sciences, beyond a rudimentary level" (see also Chomsky 1980). The point is much the same for Lassiter's 'deference', 'social identification', 'norms', etc. Thus, it is hard to assess to what extent we are actually dealing with a theory here.

Related to this point is the fact that crucial aspects of Lassiter's theory fail to provide obvious criteria for falsifiability. Whenever he encounters problematic examples, he stipulates some superficial elaboration of his terms in order to dismiss them. Consider the following illustrative case: "Recall that, in the case of Jim and his community's deviant use of 'arthritis', we came to the conclusion that the deviant usage could be incorrect in certain circumstances (e.g., talking to the doctor), but it could just as well involve dialect-switching in which both usages are correct in different social contexts. In the latter case, some sort of translation manual would be in order" (p. 625). And as above, he never specifies what role

the newly invented terms like 'dialect-switching' and 'translation manual' are supposed to play in his theory. What seems to be going on here is that Lassiter allows there to be multiple 'explanations' for the same phenomenon. Moreover, there seem to be no principles behind these possible explanations. That is, no guidelines can be found that tell us where/when we should use explanation x and where/when we should use explanation y. Without such guidelines or principles, the theory easily becomes vacuous. Though again, it might be eventually possible to develop such principles, but at least they are not stated in his paper.

Furthermore, Lassiter himself admits (correctly) that the crucial external factors he is utilizing would be subject to much fluctuation, in what appears to be unpredictable ways. Thus he even suggests a possibility that "what a term means must be resolved on a case-by-case basis" in reference to speakers' dispositions to defer (p. 622). Thus, even if it were possible to formalize the externalist factors Lassiter is relying on, no systematic account of these changes seems to be on the horizon as his notions are not precisely formulated, let alone explained, which again undermines the scientific significance of the notions employed by Lassiter.

These are all insurmountable problems facing the 'theory' Lassiter proposes. He fails to provide convincing arguments for the feasibility or legitimacy of constructing an externalist linguistic theory of the sort he envisages. For these reasons, we find it particularly puzzling to see Lassiter's remark in his conclusion: "Chomsky […] insists that only the 'internalist' aspects of language can admit of a truly scientific description. I have attempted to provide several counter-examples to this claim in the form of explanations of problems that cannot be addressed or even formulated without externalist concepts" (p. 631). His "counter-examples" (the facts relevant to (A)) are orthogonal to Chomskyan internalism, which just amounts to "the *methodological decision* […] to study *less*, prior to studying *more*: To study the *organism*, prior to the infinitely more complex task of studying how it embeds in a social, physical, and cultural surrounding", to borrow Hinzen's (2006a: 161) words (cf. Collins 2009). Lassiter also fails to demonstrate why internalists have to worry about his "counter-examples", nor does he convince us of how the sociolinguistic theory that he envisages can be explanatory, going beyond case-by-case descriptions.

We take Lassiter's contribution to be somewhat important, since, *contra* Lassiter's own intention, its failure is actually quite suggestive of a much more general conclusion: namely the absence of explanations or even descriptions that go beyond common sense in externalist approaches such as Lassiter's. We suspect that any account that ever tries to address such I-language-external complex phenomena as community standards or speakers' intentions would be relevantly like Lassiter's, and would fail in the same ways as Lassiter's does. This point is plainly another corroboration of the conclusion by Wittgenstein in *Philosophical Investigations*: There is no theory of the domain of language use, apart from just more or less helpful description (McGilvray 1998: 228; cf. Chomsky 1995: 27).

However, as noted, Lassiter's failure has no bearing on the internalist research enterprise. His attempt to articulate a sociolinguistic theory of the sort he envisages is orthogonal to the goal of internalist investigations of the Chomskyan sort, which is to provide a naturalistic scientific theory of I-language.

We should, though, make it clear that this assessment is an assessment of Lassiter's particular sociolinguistic theory, and in particular his unwarranted and misguided intention of replacing the internalist project. It is perfectly possible that a different kind of sociolinguistics might emerge that makes explicit its dependence on, or its supplementary nature to, internalist inquiry. Moreover, it might also turn out to be the case that some of the work that is done by people working on language use will turn out to be grounded in phenomena that can be investigated within an internalist approach to language. In any case, socio-linguistics suggests no alternative to the internalist science of language.

5.   Conclusion

The purpose of this paper was primarily to emphasize the methodological aspect of internalism. To repeat, internalists never deny that there are complicated social aspects in the domain of language use; they just decide not to let these unexplainable aspects of language use enter into their naturalistic theory at the present stage of inquiry: "Naturalistic inquiry is a particular human enterprise that seeks a special kind of understanding, attainable for humans in some few domains when problems can be simplified enough" (Chomsky 1995: 10). Thus, focusing on I-language (i.e. taking an internalist approach) is primarily a metho-dological decision, as we have argued above. And within this domain of study, any I-language-external phenomena such as speakers' intentionality and pre-scriptive pressures by the linguistic community are of rather little interest. Thus internalists decide to abstract away from these complicating factors when they study their object of inquiry, just as physicists abstract away from various factors such as colors and smells when they study motion and movement of physical objects; an abstraction that is not a scientific necessity. We have argued that Lassiter's criticism of internalism is off the point, based on the serious misrepre-sentation of Chomsky's position as L-individualism. Rather, the significance of Lassiter's 'contribution', if any, lies in his demonstration that the I-language-external conceptions of linguistic meaning might well be beyond the reach of naturalistic inquiry.

References

Chomsky, Noam. 1975. *Reflections on Language.* New York: Pantheon Books.
Chomsky, Noam. 1980. *Rules and Representations*. New York: Columbia University Press.
Chomsky, Noam. 1995. Language and nature. *Mind* 104, 1–61.
Chomsky, Noam. 1999. Linguistics and brain science. *University of Maryland Working Papers in Linguistics* 8, 104–117.
Chomsky, Noam. 2000. *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.

Chomsky, Noam. 2007. Biolinguistic explorations: Design, development, evolution. *International Journal of Philosophical Studies* 15, 1–21.

Chomsky, Noam. 2009 [1966]. *Cartesian Linguistics*. 3rd edn. Cambridge: Cambridge University Press.

Collins, John. 2009. Methodology, not metaphysics: Against semantic internalism. *Proceedings of the Aristotelian Society Supplementary Volume* LXXXIII, 53–69.

Dummett, Michael. 1986. A nice derangement of epitaphs: Some comments on Davidson and Hacking. In Ernest Lepore (ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, 459–476. Oxford: Blackwell.

Hinzen, Wolfram. 2006a. Internalism about truth. *Mind and Society* 5, 139–166.

Hinzen, Wolfram. 2006b. *Mind Design and Minimal Syntax*. Oxford: Oxford University Press.

Lassiter, Dan. 2008. Semantic externalism, language variation, and sociolinguistic accommodation. *Mind and Language* 23, 607–633.

McGilvray, James. 1998. Meanings are syntactically individuated and found in the head. *Mind and Language* 13, 225–280.

McGilvray, James. 2002. MOPs: the science of concepts. In Wolfram Hinzen & Hans Rott (eds.), *Belief and Meaning — Essays at the Interface*, 73–103. Frankfurt a.M.: Ontos.

McGilvray, James. 2009. Introduction to the third edition. In Chomsky, (2009), 1–52.

Narita, Hiroki. 2009. How syntax naturalizes semantics; A review of *Syntactic Anchors: On Semantic Structuring* (Juan Uriagereka, 2008, Cambridge: Cambridge University Press). *Lingua* 119, 1767–1775.

Wittgenstein, Ludwig. 1953. *Philosophical Investigations*. New York: Macmillan.

Terje Lohndal   *Hiroki Narita*
University of Maryland   *Harvard University*
Department of Linguistics   *Department of Linguistics*
1401 Marie Mount Hall   *Boylston Hall, 3rd floor*
College Park, MD 20742   *Cambridge, MA 02318*
USA   *USA*
terje@umd.edu   *narita@fas.harvard.edu*

On Parametric (and Non-Parametric) Variation

Neil Smith & Ann Law

This article raises the issue of the correct characterization of 'Parametric Variation' in syntax and phonology. After specifying their theoretical commitments, the authors outline the relevant parts of the Principles–and–Parameters framework, and draw a three-way distinction among Universal Principles, Parameters, and Accidents. The core of the contribution then consists of an attempt to provide identity criteria for parametric, as opposed to non-parametric, variation. Parametric choices must be antecedently known, and it is suggested that they must also satisfy seven individually necessary and jointly sufficient criteria. These are that they be cognitively represented, systematic, dependent on the input, deterministic, discrete, mutually exclusive, and irreversible.

A persistent preoccupation of generative linguistics has been the tension, bordering on paradox, between two questions: "Why are there so many languages?" and "Why are they all so similar?". The tension is sufficiently great that many writers, dazzled by the obviousness of the first, are tempted to deny the truth of the second: Evans & Levinson's (2009) 'The myth of language universals' is a recent example. A resolution of the tension can be found in the framework of 'Principles–and–Parameters' (Chomsky 1981a, 1981b; for overviews and history, see Roberts 1997, Baker 2001, and especially Biberauer 2008a), but making this claim plausible to the skeptics necessitates elaboration and refinement of the theory, in particular of the nature and scope of 'parametric' variation. It is this issue we try to address in the current contribution, suggesting identity criteria for parametric as opposed to non-parametric differences among languages. The situation is reminiscent of the debate about human types: The apparent obvious diversity of different 'races' disguises profound underlying unity, and specifying the nature of the variation is fraught with difficulty. In what follows we spell out our theoretical presuppositions, we present the elements of the Principles–and–Parameters framework and their motivation, and we suggest and defend our identity criteria.

We take seriously the central claim of the Minimalist Program (Chomsky 1995) that in elucidating the nature of the human faculty of language, linguistic theory should restrict itself to what is conceptually necessary or descriptively inevitable. Accordingly, we adopt Hauser *et al.*'s (2002) contrast between the

Faculty of Language in the broad sense (FLB) and the Faculty of Language in the narrow sense (FLN), seeking to identify defining properties of FLN, even if this latter may perhaps consist simply of possible re-combinations of elements of FLB. We have argued elsewhere against Hauser *et al.*'s claim that recursion is the unique property of FLN (see Smith & Law 2007: 2) on the grounds that recursion must be characteristic of the Language of Thought in Fodor's (1975, 2008) sense. However, we are happy to go along with Chomsky's (2009a: 29) suggestion that Natural Language is the same as the Language of Thought except for 'externalization' (cf. Smith 1983). That is, language links the Conceptual–Intentional and Sensori–Motor interfaces, where the former equates to the language of thought and the latter, as used for communication (perception and production), characterizes natural languages, which emerged evolutionarily as the result of being externalized. This external form is anyway the domain of parametric variation, the existence of which may moreover be a defining property of the human language faculty.

The Principles–and–Parameters framework provides simultaneously a solution to Plato's problem and the problem of characterizing typological variety. UG (short for 'Universal Grammar', the innate endowment that the child brings to the task of language acquisition) specifies that human languages consist of a Lexicon and a 'computational system' (referred to as C_{HL}, the computation for human language). The lexicon consists of a set of lexical entries, each of which is a triple of phonological, morpho-syntactic, and semantic features, and with a link to associated encyclopedic information. UG also provides a set of exceptionless principles, such as structure dependence (Chomsky 1971), (strict) cyclicity (Freidin 1999, Chomsky 2002), the Extended Projection Principle (Chomsky 1995), etc., which constrain the operations of the computations and act as a constraint on language acquisition: Children learning their first language have their 'hypothesis space' tightly constrained with the result that they never make mistakes of a particular kind. However, "[…] principles do not determine the answers to all questions about language, but leave some questions as open parameters" (Berwick & Chomsky, forthcoming: 8 [in the 2008 manuscript]).

That is, in addition to a set of universal principles, UG provides a set of parameters which jointly define the limits of language variation. This is typically conceptualized as the setting of a number of 'switches' — on or off — for particular linguistic properties. Examples of such parameters in syntax are the head-direction parameter (whether heads, such as Verb, Noun, and Preposition, precede or follow their complement), the null-subject (or *pro*-drop') parameter (whether finite clauses can have empty pronominal subjects), and the null-determiner parameter (whether noun phrases can have empty determiners). Typical examples in phonology are provided by the stress differences characteristic of English and French, and the possibility of complex consonant clusters found in English but not in Japanese. English stress is 'quantity-sensitive', whereas French stress is 'quantity-insensitive', with the result that words with the same number of syllables may have different stress in English but must have uniform stress in French; in English, words may begin with clusters of consonants in a way which is impossible in Japanese, with the result that English loans into Japanese appear with the clusters separated by epenthetic vowels.

The theory thus unifies two different domains: typology and acquisition. Variation among the world's languages (more accurately the set of internalized I-languages; Chomsky 1986) is defined in terms of parametric differences and, in first language acquisition, the child's task is reduced to setting the values of such parameters on the basis of the stimuli it is exposed to — utterances in the ambient language. Given the strikingly uniform success of first language acquisition, it follows that "the set of possibilities [must] be narrow in range and easily attained by the first language learner" (Smith 2004: 83). By hypothesis, the principles do not vary from child to child or from language to language so, as Chomsky (2006: 183) puts it, "acquisition is a matter of parameter setting, and is therefore divorced entirely from […] the principles of UG".

The theory is at once 'internalist' (i.e. it is a theory of states of the mind/brain), pertaining to knowledge which is largely unconscious, and universalist. An immediate implication of this position is that the range of parametric choices is known in advance and, as a corollary, it claims that acquisition is largely a process of 'selection' rather than instruction (see Piattelli–Palmarini 1989) and that such acquisition is likely to take place in a critical period or periods.

This brief characterization raises a number of problems. The first of these is the issue of deciding which phenomena are to be accounted for by reference to principles and which by reference to parameters, as exemplified in the history of subjacency which began as a universal principle but was later parameterized. More importantly, does this binary choice exhaust the ontology? We argue that parameters account for some of the surface variability — but only some: Much variation is accidental. Accordingly, we need a three-way distinction: Universal Principles, Parameters, and Accidents. Note that even universal principles may have their status obscured by recalcitrant data. For instance, the universality of Merge is not in question even though some items — interjections — do not participate. Similarly, a clear and classic instance of a parameter is 'head direction', even though some examples are problematic like English *notwithstanding*, which can occur before or after its complement, or the occurrence in German of synonymous (and etymologically related) pairs of preposition and postposition (e.g., *längs des Flusses*/ *den Fluss entlang* 'along the river'). Finally, there are 'accidents', exemplified by gaps in morphological paradigms, such as the lack of a past tense form for *beware*; by the (claimed) absence of recursion in Pirahã (Everett 2005), or by the absence of initial consonants in Arrernte syllable structure (Breen & Pensalfini 1999).

Assignment to each of these categories may of course be problematic, with the uncertainty having potentially significant implications for broader considerations such as innateness. Thus Chomsky (2009b: 385), in discussing the optimization of the language faculty in terms of third-factor considerations, writes: "If you take a parameter and you genetically fix the value, it becomes a principle […]. So adding parameters is reducing genetic information". This stance is similar to Janet Fodor's (2009) characterization of principles and parameters as a Minimax solution: 'minimize genetic information' and 'maximize/optimize the amount of learning'.

Reverting to the remarks above about Natural Language and the Language of Thought and the assumption that the syntax of both is the same (but see Smith

2004: 43f. for problems with this position), it is clear that "parameterization and diversity too would be mostly — maybe entirely — restricted to externalization" (Chomsky, in press: 14 [2008 manuscript]; to "language shorn of the properties of the sound system" as Smith 2004: 43 puts it), hence mainly morphology and phonology. One reason for the multiplicity of languages is then that "the problem of externalization can be solved in many different and independent ways" (Chomsky, in press: 15 [2008 manuscript]), where, moreover, these may all be 'optimal' in different ways. The interesting implication is that there is no parametric variation at the Conceptual–Intentional interface (but see below) and perhaps not even any parametric variation in the syntax narrowly construed (C_{HL}).

Despite these observations, we propose for illustrative purposes to pursue with the majority of linguists the possibility that parametric variation (hereafter 'PV') characterizes both syntax and phonology. Further, if there is to be any content to the 'parametric' part of PV, there is need to work out necessary and sufficient conditions for something to count as parametric. That is, we are in explicit opposition to those such as Kayne (2005: 6 and elsewhere), Manzini & Savoia (2007), and Rita Manzini (p.c.), for whom *all* (syntactic) variation is parametric. We reject this stance because of the need to constrain possible parameters. In the absence of such constraints "the term 'parameter' would end up being nothing but jargon for 'language-particular rule'" (Newmeyer 2005: 53) or, as Moro (2008: 107) puts it: "If there were no restrictive generalization on the format of parameters, the theory would be too weak".

Before suggesting such restrictions, it is important to note that the nature of the identity criteria, even the possibility of coming up with any, is dependent on the version of Principles–and–Parameters theory that one adopts. There are several possibilities available in the literature. First, as seen in Rizzi's (2009: 95–96) discussion, there is a conceptual contrast between theories which indulge in overspecification (where UG contains specific statements for certain choices, which must be fixed by experience) and those which indulge in underspecification (where UG has nothing to say — there are gaps, to be filled by experience; cf.: "UG limits the space of possible hypothes[e]s, but does nothing more" (Nevins 2004: 121)).

Second, this distinction cross-cuts that between macro-parametric and micro-parametric variation (for discussion see Baker 2008). 'Macro'-PV is typically exemplified by the head-direction (head-first/head-last) parameter (Chomsky 1981a) or Baker's (1996) polysynthesis parameter which determines the overall morphological structure of the language. Each of these parameters has a wide variety of effects, whereas 'micro'-PV of the sort exemplified by the choice of auxiliary to accompany unaccusative verbs (Perlmutter 1978, Burzio 1986) or case realignment in Albanian causatives (Manzini & Savoia 2007) is characteristically more restricted and has correspondingly fewer repercussions. An emerging consensus seems to be that the 'macro/micro' contrast is not important: "The extent-of-variation question is not well defined or theoretically very interesting" (Baker 2008: 371). We agree, though we wish to argue that the parametric/non-parametric distinction is important both in syntax and in the phonological domain where there is no comparable macro-micro contrast. There is, third, the related issue of whether parameters pertain to principles, as in Chomsky's

original proposal (Chomsky 1981b) or the later, widely accepted, 'Borer–Chomsky Conjecture' (cf. Biberauer 2008a) that all (syntactic) parameters refer to features of functional heads in the lexicon, so that the number of parameters corresponds to the size of the functional lexicon. While we are sympathetic to the restriction implied by the conjecture its apparent irrelevance to phonology makes it less central to our concerns.

At a lower level of abstraction we come, fourth, to the domain or locus of parametric variation. Biberauer (2008a: 32) suggests that the locus of parameters is "the Lexicon and one or more of the Interfaces". We are anxious that our identity criteria should pertain to phonology as well as syntax and if, despite the remarks about externalization above, it proves that there are relevant examples, to semantic choices at the C–I interface (cf. Chierchia 1998), so we are happy to follow this suggestion. At a finer level of detail, Rizzi (2009: 213ff.) observes that syntactic parameters, located within the lexicon, may pertain to any of the three basic computational processes of the syntax: Merge (e.g., head direction), Move (e.g., V to T), and Spell-Out (e.g., Null-subject). Again, there is no obvious phonological counterpart to this taxonomy.

There are many other considerations which are not directly relevant to our concerns or about which we have nothing to contribute. For instance, Nevins (2004: 123) argues on the basis of 'parametric ambiguity'[1] that "variation is the result of maintaining multiple parameter settings simultaneously" (cf. Yang 2002). We are suspicious of this position as it looks like a conceptually undesirable version of 'multiple grammars' (for discussion, see Smith, in press).

We turn now to the main concern of the article: Suggesting, illustrating, and defending a number of criteria which variation has to meet to count as parametric rather than accidental.

The theory of PV hypothesizes that the range of choices is 'antecedently known', and this basic property correlates with a number of others which distinguish PV from non-parametric variation, and allow us to provide identity criteria for it. Being antecedently known may not be as straightforward as we have previously (Smith & Law 2007, in press) assumed. There is both a terminological and a substantive issue. Chomsky (2009b: 395) observes that in many languages the expression used for 'knowing a language' does not involve the word 'know', but rather the equivalent of 'come', 'hear', or 'have'. This has probably underlain some of the philosophical dispute about whether knowledge of language, in the sense of competence, constitutes real knowledge or not, but this terminological concern is of minor importance in the present context. The substantive issue is whether 'antecedently known' entails cognitively 'represented' or could refer simply to 'architectural' (third factor) constraints on the hypothesis space. The strongest position is that all options are laid out — so 'represented' — prior to experience and whatever abilities the child brings to the task of first language acquisition are deployed to select among them. The weaker, architectural, position may be preferable if it allows properties of the language faculty to be derived from more general considerations.

---

[1]   This refers to the situation where several analyses or structures could underlie the data of interest.

Whichever position is correct, we take our **first** criterion to be that variants licensed by parametric choice must be cognitively represented. To make clear what motivates this condition, consider by contrast acclimatization, specifically sweating. We have a critical period for setting our sweating switch: Experiencing hot and humid weather in the first three years of life leads to a different setting from exposure to different conditions, and these settings cannot be significantly altered thereafter (Gluckman & Hanson 2005: 7). Despite a certain superficial similarity, this is not PV because the different states are not (mentally) represented and have no cognitive effects. Further, it is relevant to note that where there is evidence that some linguistic fact is *not* represented there is also evidence that this is not a domain of PV. For instance, Smith (2003, in press) claims that the learning child does not represent its own mispronounced output (e.g., saying [bɔkəl] for *bottle*), but equally such mispronunciation does not constitute the locus of PV.

This leads to our **second** criterion: systematicity. This is implicit in Moro's (2008: 106) remark that the relevant domain is one where variation is "minimal and systematic"; or equivalently, to what Biberauer (2008a: 2) describes as 'non-random' variation. A simple example is provided by irregular morphology of the type exemplified by the impossibility of **amn't* in (most varieties of) English, or the kind of defective paradigm seen in Latin *vis–vim–vi*. We do not consider this to be PV because it is by definition not systematic and hence we could not plausibly acquire knowledge of it by any process of triggering in the way which is plausible for systematic contrasts such as the possibility of null determiners or the absence of codas. Although systematicity and 'potentially triggered' may be extensionally the same the two notions are conceptually distinct so need to be kept separate, but we link them under a single criterion.

Our **third** criterion is dependence on the input; that is, the variant chosen must correspond to a possible state of the adult language, and hence can be illustrated most clearly from first language acquisition. The head-direction parameter clearly reflects properties of the ambient language in a way that is not characteristic of all variation. An example of systematic but input-independent and non-parametric variation is provided by the individual differences in consonant harmony in phonological development (cf. Smith 1973: 163), or the variation in the choice of initial or final negation in syntactic development (cf. Smith 2005: 29). For instance, two children in essentially the same environment may produce the adult *duck* as [gʌk] and [dʌt] respectively. These may both be manifestations of consonant harmony, but they do not count as PV because the particular variants chosen appear to be independent of the input (and consonant harmony is anyway essentially alien to adult phonology). A comparable syntactic example is provided by the development of negation. All children typically go through a stage in which the negator is peripheral, either initial or final. Individual children then differ such that one child learning English may say '*no like cabbage*' and another '*like cabbage no*'. We take such variation to be non-parametric as no language allows only such peripheral negation. This universal exclusion enables us to differentiate this non-parametric variation from UG-licensed errors of the sort described by Crain and his colleagues (cf. Crain & Pietroski 2002). A child may produce a form which never occurs in the input (e.g., '*What do you*

think what pigs eat?') because the structure is licensed by UG and so occurs as a parametric choice in other languages.[2] Despite this potential complication, the case of consonant harmony in phonology and negation in syntax should make the conceptual contrast between parametric and non-parametric variation clear.

Our **fourth** criterion is that PV must be deterministic:[3] That is, the input to the child must be rich enough and explicit enough to guarantee that a parameter such as *pro*-drop or the presence of complex onsets in phonology can be set. If the input does not meet this requirement we are dealing with non-parametric variation. A syntactic example is provided by sequence of tense phenomena where individual variation verges on the random (see Smith & Cormack 2002). A phonological example is provided by Yip (2003: 804) who argues that some speakers treat a post-consonantal glide as a secondary articulation of the consonant, others as a segment in its own right: "[T]he rightful home of /y/ [is] under-determined by the usual data, leaving room for variation". Her conclusion is that "speakers opt for different structures in the absence of conclusive evidence for either". Again that indicates for us that the variation is non-parametric. Determi-nisticness suggests that the process of parameter-setting must be 'reflexive' (cf. Chomsky 2009b: 384) but, as with systematicity and triggering, the notions are conceptually distinct so we keep them apart, though again not as separate criteria.

Our **fifth** criterion is suggested by an observation of Dupoux & Jacob (2007) to the effect that PV in language is 'discrete' (usually binary), whereas in other domains — moral judgment, for instance — one typically finds continuous scales. A linguistic example of the contrast is provided by vowel height. Whether a language displays 2, 3, or 4 degrees of vowel height in its phonological system is a matter of parametric choice(s). The degree to which the particular articulation of some vowel is high — either randomly or as a matter of individual difference (maybe my articulations of [i] are systematically higher than yours) is continuous and could not be parametric.

Our **sixth** criterion is 'exclusivity'. PV gives rise to mutually exclusive possibilities: Languages are either [+*pro*-drop] or [–*pro*-drop] — the choice leaves no room for compromise, no language is both. By contrast, the choice in a [+*pro*-drop] language of using or not using a subject pronoun is non-parametric. The contrast is again most obvious with morality where moral diversity involves "different preference orderings among competing members of a finite set of universal moral values" (Dupoux & Jacob 2007: 377). An extension of mutual exclusivity would be that the choices are exhaustive in that they exhaust the relevant hypothesis space interdependently. That is, the parameters are not independent (as claimed explicitly in e.g., Manzini & Wexler 1987) but are hierarchically nested: The choice of a parameter [±X], gives rise to a range of further choices within each of [+X] and [–X], and apparent exceptions to exclusivity are due to choices being either subordinate or parallel to a given parameter. We do not make this (non-)independence criterial as we know of no

---

[2]   We take it that such over-generalization is a sign that the child has, temporarily, mis-set the relevant parameter.

[3]   Compare Nevins' (2004: 120) approving remark that his (multiple-precedence) view of redu-plication "yields a deterministic output".

cogent evidence either for or against. Similarly, although the phrasing used here in terms of [±X] suggests binarity, which is often presupposed in terms of [strong/weak] in the literature (see e.g., Radford *et al.* 2009: 314), we see no reason to make this essential.

A possible further **seventh** criterion is 'irreversibility': That is, the putative impossibility of the re-setting of parameters in second language acquisition (see e.g., Tsimpli & Smith 1991). The implicit contrast is with the reversible variations found in lexical learning. For instance, despite half a century's exposure to examples like "I didn't see him yet", one of us (NS) still judges them ungrammatical (the only licit possibility is "I haven't seen him yet"). This is in contrast to examples like the second sentence of this article, written without malice aforethought, which begins: "The tension is sufficiently great that many writers […]". This construction was originally ungrammatical for NS (the only licit possibility being "The tension is sufficiently great for many writers to have […]") but has now changed its status. The former contrast is arguably a matter of PV, the latter not.

We summarize and illustrate the foregoing criteria in the following table:

1. *Parametric choices must be cognitively represented*	
<u>Parametric</u>	<u>Non-parametric</u>
stress	sweating
word order	consonant harmony
2. *Choices must be systematic — variations are not accidents*	
<u>Parametric</u>	<u>Non-parametric</u>
±null subject	irregular morphology
3. *Choices must be dependent on the input and hence correspond to a possible state of the adult language*	
<u>Parametric</u>	<u>Non-parametric</u>
quantity-sensitivity	consonant harmony
word order — head direction	word order in early negation
4. *Choices must be deterministic*	
<u>Parametric</u>	<u>Non-parametric</u>
pro-drop	sequence of tense
complex onsets in phonology	Post-consonantal glides
5. *Choices must be discrete*	
<u>Parametric</u>	<u>Non-parametric</u>
number of vowel heights	realization of vowel height
6. *Choices must be mutually exclusive*	
<u>Parametric</u>	<u>Non-parametric</u>
±*pro*-drop	choice of a pronoun (or not) in a *pro*-drop language
7. *Choices must be irreversible*	
<u>Parametric</u>	<u>Non-parametric</u>
temporal adverbial modification	sub-categorization possibilities

A number — a huge number — of issues remain open. We list a few below:

(1) We have in general not committed ourselves to where the parametric choices reside. It is not clear whether there is a single answer, but we assume in the absence of definitive evidence that all such choices are lexical.

(2) It would be helpful to determine which of these criteria might derive from other properties, bearing in mind that nothing so derivable would be part of FLN. In particular, it is desirable to establish which criteria (e.g., deterministicness and mutual exclusivity) might fall out from general properties of complex cognitive systems ('third-factor' considerations, where these include general learning strategies and principles of computational efficiency', as in Chomsky, in press: 15 [2008 manuscript]).[4] An example in principle is provided by the head-direction parameter. It seems clear that the choice between head-first and head-last is a function of the need for linearization imposed by the temporal structure of speech. Given that 'merge' combines A and B it is physically necessary either that A precede B or that B precede A. In such a situation, as Boeckx (2009: 198) observes, appeal to a parameter may be supererogatory. Two points are, however, relevant. First, the physical necessity for linearization may be the ultimate cause of the parameter but the skew distribution of the world's languages and the consistency of head direction within a language suggest that the parameter does exist: The physical constraint has led to grammaticalization of the parameter. Second, although this parameter has a 'third factor' motivation it is only one example and not a criterion for parameterhood whose status is affected. For plausible instances of a criterion being rendered unnecessary we probably need to look elsewhere. We leave the issue for future research.

(3) In earlier work (Smith 2007; Smith & Law 2007, in press) we have investigated whether the criteria for parametric status allow a generalization to other domains, either human or animal, suggesting that our knowledge of music and our moral judgment might be such examples in the former domain, and birdsong in the latter. We are currently less sanguine about the possibility.

The preceding discussion implies that many of the parameters postulated in the literature are, by our criteria, accidents rather than reflecting genuine, but not exceptionless, generalizations. We have already alluded to some of the work of Kayne and Manzini, and Evans & Levinson (2009: 432) explicitly assume that parameters account for all differences: The 'full set of possible combinations'. Our attempt to delineate criteria for PV should not in any way be taken to impugn the value of the work of these authors, but we think it is time for the theory to be put on a more explicit footing. We await corroboration or refutation of our putative criteria with anticipation and apprehension in equal measure.

---

[4] Though we are skeptical of the claim that "[t]o externalize the internally generated expression 'what John is eating what', it would be necessary to pronounce 'what' twice, and that turns out to place a very considerable burden on computation" (Berwick & Chomsky, forthcoming: 11 [in the 2008 manuscript]). The burden seems slight, especially given that in first language acquisition children regularly repeat material 'unnecessarily' (see the examples from Crain & Pietroski 2002 above).

References

Baker, Mark. 1996. *The Polysynthesis Parameter* (Oxford Studies in Comparative Syntax). New York: Oxford University Press.

Baker, Mark. 2001. *The Atoms of Language: The Mind's Hidden Rules of Grammar*. New York: Basic Books.

Baker, Mark. 2008. The macroparameter in a microparametric world. In Biberauer (2008b), 351–373.

Berwick, Robert & Noam Chomsky. Forthcoming. The biolinguistic program: The current state of its evolution and development. In Anna Maria Di Sciullo & Calixto Aguero (eds.), *Biolinguistic Investigations*. Cambridge, MA: MIT Press.

Biberauer, Theresa. 2008a. Introduction. In Biberauer (2008b), 1–72.

Biberauer, Theresa (ed.). 2008b. *The Limits of Syntactic Variation* (Linguistik Aktuell/Linguistics Today 132). Amsterdam: John Benjamins.

Boeckx, Cedric. 2009. Round table — Language universals: Yesterday, today and tomorrow. In Piattelli–Palmarini *et al.* (eds.), 195–199.

Breen, Gavan & Rob Pensalfini. 1999. Arrernte: A language with no syllable onsets. *Linguistic Inquiry* 30, 1–25.

Burzio, Luigi. 1986. *Italian Syntax: A Government–Binding Approach* (Studies in Natural Language and Linguistic Theory 1). Dordrecht: Kluwer.

Chierchia, Gennaro. 1998. Plurality of mass nouns and the notion of 'semantic parameter'. In Susan Rothstein (ed.), *Events and Grammar* (Studies in Linguistics and Philosophy 70), 53–103. Dordrecht: Kluwer.

Chomsky, Noam. 1971. On interpreting the world. *Cambridge Review* 92, 77–93.

Chomsky, Noam. 1981a. *Lectures on Government and Binding: The Pisa Lectures* (Studies in Generative Grammar 9). Dordrecht: Foris.

Chomsky, Noam. 1981b. Principles and parameters in syntactic theory. In Norbert Hornstein & David Lightfoot (eds.), *Explanation in Linguistics*, 123–146. London: Longman.

Chomsky, Noam. 1986. *Knowledge of Language: Its Nature, Origin, and Use* (Convergence). New York: Praeger.

Chomsky, Noam. 1995. *The Minimalist Program* (Current Studies in Linguistics 28). Cambridge, MA: MIT Press.

Chomsky, Noam. 2002. *On Nature and Language*. Cambridge: Cambridge University Press.

Chomsky, Noam. 2006. *Language and Mind*. 3rd edn. Cambridge: Cambridge University Press.

Chomsky, Noam. 2009a. Opening remarks. In Piattelli–Palmarini *et al.* (eds.), 13–43.

Chomsky, Noam. 2009b. Conclusion. In Piattelli–Palmarini *et al.* (eds.), 379–409.

Chomsky, Noam. In press. Some simple evo devo theses: How true might they be for language? In Richard K. Larson, Viviane Déprez & Hiroko Yamakido (eds.), *The Evolution of Human Language: Biolinguistic Perspectives* (Approaches to the Evolution of Language), 45–62. Cambridge: Cambridge University Press.

Crain, Stephen & Paul Pietroski. 2002. Why language acquisition is a snap. *The*

Linguistic Review 19, 163–183.

Dupoux, Emmanuel & Pierre Jacob. 2007. Universal moral grammar: A critical appraisal. *Trends in Cognitive Sciences* 11, 373–378.

Evans, Nicholas & Stephen Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32, 429–492.

Everett, Daniel L. 2005. Cultural constraints on grammar and cognition in Pirahã: Another look at the design features of human language. *Current Anthropology* 46, 621–646.

Fodor, Janet D. 2009. Round table discussion. In Piattelli–Palmarini *et al.* (eds.), 199–206.

Fodor, Jerry A. 1975. *The Language of Thought*. New York: Crowell.

Fodor, Jerry A. 2008. *LOT 2: The Language of Thought Revisited*. Oxford: Clarendon Press.

Freidin, Robert. 1999. Cyclicity and minimalism. In Samuel David Epstein & Norbert Hornstein (eds.), *Working Minimalism*, 95–126. Cambridge, MA: MIT Press.

Gluckman, Peter & Mark Hanson. 2005. *The Fetal Matrix: Evolution, Development and Disease*. Cambridge: Cambridge University Press.

Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298, 1569–1579.

Kayne, Richard S. 2005. Some notes on comparative syntax, with special reference to English and French. In Guglielmo Cinque & Richard S. Kayne (eds.), *The Oxford Handbook of Comparative Syntax* (Oxford Handbooks), 3–69. Oxford: Oxford University Press.

Manzini, M. Rita & Kenneth Wexler. 1987. Parameters, binding theory and learnability. *Linguistic Inquiry* 18, 413–444.

Manzini, M. Rita & Leonardo M. Savoia. 2007. *A Unification of Morphology and Syntax: Investigations into Romance and Albanian Dialects* (Routledge Leading Linguists). London: Routledge.

Moro, Andrea. 2008. *The Boundaries of Babel: The Brain and the Enigma of Impossible Languages* (Current Studies in Linguistics 46). Cambridge, MA: MIT Press.

Nevins, Andrew. 2004. What UG can and can't do to help the reduplication learner. *Plato's Problems: Papers on Language Acquisition — MIT Working Papers in Linguistics* 48, 113–126.

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

Perlmutter, David M. 1978. Impersonal passives and the unaccusative hypothesis. *BLS 5*, 157–189.

Piattelli–Palmarini, Massimo, Pello Salaburu & Juan Uriagereka (eds.). 2009. *Of Minds and Language: A Dialogue with Noam Chomsky in the Basque Country*. Oxford: Oxford University Press.

Piattelli-Palmarini, Massimo. 1989. Evolution, selection and cognition: From learning to parameter setting in biology and in the study of language. *Cognition* 31, 1–44.

Radford, Andrew, Martin Atkinson, David Britain, Harald Clahsen & Andrew Spencer. 2009. *Linguistics: An Introduction*. 2nd edn. Cambridge: Cambridge University Press.

Rizzi, Luigi. 2009. Language invariance and variation. In Piattelli–Palmarini *et al.* (eds.), 211–220.

Roberts, Ian. 1997. *Comparative Syntax*. London: Arnold.

Smith, Neil. 1973. *The Acquisition of Phonology: A Case Study*. Cambridge: Cambridge University Press.

Smith, Neil. 1983. *Speculative Linguistics*. An inaugural lecture delivered at University College London, published by the College.

Smith, Neil. 2003. Representations and responsibilities. *Korean Journal of English Language and Linguistics* 3, 527–545.

Smith, Neil. 2004. *Chomsky: Ideas and Ideals*. 2nd edn. Cambridge: Cambridge University Press.

Smith, Neil. 2005. *Language, Frogs and Savants: More Problems, Puzzles and Polemics in Linguistics*. Oxford: Blackwell.

Smith, Neil. 2007. Unification and the unity of Chomsky's thought. *Chomskyan Studies* 2, 1–21.

Smith, Neil. In press. *Acquiring Phonology: A Cross-generational Case-Study* (Cambridge Studies in Linguistics 124). Cambridge: Cambridge University Press.

Smith, Neil & Annabel Cormack. 2002. Parametric poverty. *Glot International* 6, 285–287. [Reprinted in Smith (2005), 103–108.]

Smith, Neil & Ann Law. 2007. Twangling instruments: Is parametric variation definitional of human language? *UCL Working Papers in Linguistics* 19, 1–28.

Smith, Neil & Ann Law. In press. Parametric variation: Language and birdsong. In Johan Bolhuis & Martin Everaert (eds.), *Birdsong, Speech and Language: Converging Mechanisms*. Cambridge, MA: MIT Press.

Tsimpli, Ianthi-Maria & Neil Smith. 1991. Second-language learning: Evidence from a polyglot savant. *UCL Working Papers in Linguistics* 3, 171–183.

Yang, Charles D. 2002. *Knowledge and Learning in Natural Language* (Oxford Linguistics). New York: Oxford University Press.

Yip, Moira. 2003. Casting doubt on the Onset-Rime distinction. *Lingua* 113, 779–816.

Neil Smith   *Ann Law*
University College London   *King's College London*
Research Department of Linguistics   *Institute of Psychiatry*
2 Wakefield Street   *De Crespigny Park*
London WC1N 1PF   *London SE5 8AF*
United Kingdom   *United Kingdom*
SmithNV@googlemail.com   *Ann.Law@uclmail.net*

On Multidominance and Linearization

Mark de Vries

This article centers around two questions: What is the relation between movement and structure sharing, and how can complex syntactic structures be linearized? It is shown that regular movement involves *internal remerge*, and sharing or 'sideward movement' *external remerge*. Without *ad hoc* restrictions on the input, both options follow from Merge. They can be represented in terms of multidominance. Although more structural freedom ensues than standardly thought, the grammar is not completely unconstrained: Arguably, proliferation of roots is prohibited. Furthermore, it is explained why external remerge has somewhat different consequences than internal remerge. For instance, apparent non-local behavior is attested. At the PF interface, the linearization of structures involving remerge is non-trivial. A central problem is identified, apart from the general issue why remerged material is only pronounced once: There are seemingly contradictory linearization demands for internal and external remerge. This can be resolved by taking into account the different structural configurations. It is argued that the linearization is a PF procedure involving a recursive structure scanning algorithm that makes use of the inherent asymmetry between sister nodes imposed by the operation of Merge.

Keywords: linearization; movement; multidominance; PF interface; (re-) merge

1. Introduction and Overview

Displacement is one of the central tenets in generative grammar. The underlying idea is that a word or phrase may be involved in more than one relationship; therefore, it can be associated with a sentence position where it does not surface. A simple example in English is *wh*-movement, such as illustrated in (1):

(1)   a.   This talented girl should purchase *a new violin*.
       b.   *Which violin* should this talented girl purchase ___?

The unmarked direct object position in English is shown in (1a), where it is occupied by *a new violin*. This phrase is categorically and semantically selected by the verb *purchase*. In (1b), the preposed object *which violin* is thought to be related to the regular direct object position next to the main verb as well, here indicated by an underscore. How does the grammar make sure that the object is pronounced in the higher, operator-related position (leftmost), and not in the lower, thematic position (rightmost)? A fairly standard approach in generative grammar has been the assumption that movement is hierarchically directional, and that a moved phrase leaves an unpronounced trace in the original lower position. In current minimalist theories, specialized traces no longer exist (this follows from the Inclusiveness condition proposed by Chomsky 1995: 225). From the perspective of a bottom-up derivation, it seems that we must make sure that the first occurrence of the relevant phrase (here, *which violin*) remains phonologically silent if, after movement, there will be a second, higher occurrence of it. Clearly, then, the linearization of a sentence structure is a non-trivial process taking place at the interface between syntax and phonology. This article is an attempt to explicate that process, and its preconditions.

An interesting complication is that there appear to be constructions that essentially show the opposite pattern, though not exactly in a mirror fashion. A relevant example is the so-called Right Node Raising (RNR) construction. In (2), *this beautiful Stradivarius* is the object of *admired* as well as *bought*, but here only the rightmost occurrence is spelled out, contrary to the situation in (1b).

(2)   The boy only admired ___, but the girl actually bought *this beautiful Stradivarius*.

Though I do not think that there is rightward or lowering movement, there are reasons to believe that phrases can be structurally *shared*, which could be represented by a multidominance configuration (this will be explained below). The questions we then face are the following:

(Q1)   How are sharing configurations derived, and what is the theoretical relationship with movement?

(Q2)   When and how does the linearization procedure operate, and how does it distinguish between the two different construction types illustrated by (1b) and (2), respectively?

In section 2, I argue that a freely applicable operation Merge gives rise to the possibility of both *internal remerge* and *external remerge*. The concept of movement corresponds to the first, and that of sharing to the second. I should mention right away that this article is not about the correct analysis of RNR, *wh*-movement, or any other particular construction still to be mentioned. Rather, I intend to explore the theoretical consequences of remerge. References to particular

analyses are used for concreteness' sake, and serve as illustrations, mainly.

Since the syntactic configurations arising by applying the two types of remerge are different, the linearization procedure can be made sensitive to it. This is the subject of section 3, which presents a solution to the seemingly contradictory linearization demands briefly introduced here. It is claimed that linearization involves the scanning (traversal) of a full sentence structure. Various complicated construction types are examined. Section 4 presents a more detailed graph scanning algorithm, and discusses the computational load of such a procedure, taking into account the difference between representations and the actual theoretical assumptions. Finally, section 5 is the conclusion.

2.   Internal and External Remerge

2.1.   The Operation Merge: Input and Output

The input for Merge, which I assume to be binary (following standard assumptions dating back to Kayne 1984), is restricted to objects recognizable by syntax, that is, words and phrases — or rather the features associated with these. Nevertheless, judging from general minimalist practice since Chomsky (1995), the *selection* of these objects must be free with respect to their location or history. There are three possibilities, two of which are logically necessary if Merge is the only structure-building device. First, input objects for Merge can be selected from the lexicon or 'numeration'. Of course, syntax would be idle without subject matter. Second, the result of a previous instance of Merge can be selected as the input for a subsequent instance of Merge. This corresponds to the general hierarchical aspect of syntax. Without the recursive application of Merge, objects more complex than two words could never be derived. Non-trivial objects are created in the syntactic workspace. It is not only the active structure itself that is complex after first Merge: Auxiliary structures are also necessary. For example, subjects and adverbial phrases are often complex (notice that even a simple noun phrase like *the man* counts as such, as it consists of more than one element). If they are to be attached to the main projection line, they must have been derived already in an auxiliary derivation.

The third possibility is fairly standard as well, though not undisputed. Let us assume that there is such a thing as displacement, as indicated in the introduction. Displacement from a derivational perspective implies that a constituent ('term') of a derived structure is accessible as a possible input object for another instance of Merge. It follows that a syntactic object can be merged more than once. In this way, we account for the fact that syntactic objects can be involved in more than one relationship, associated with different positions in the structure. Here, it is presupposed that grammatical relationships are a direct or indirect function of Merge, and hence of structure. This is a central insight of generative grammar, and I will not question it.

To sum up, three differently situated kinds of objects may serve as input for Merge: (i) lexical items, (ii) complex items that are the result of previous instances of Merge, (iii) terms of complex items. The option in (iii) normally

corresponds to what is often called Move (Chomsky 1995). However, it is impor-
tant to see that there is only one basic operation, Merge. Depending on the input,
the result may be Move. If Move involves the creation of traces or copies with
special properties, it would constitute a separate, complex operation. However,
according to minimalist reasoning, this cannot be *a priori* assumed. In recent
work, Chomsky refers to iii) as *internal merge*, as opposed to *external merge* for (i)
and (ii), stressing that the possibility of movement simply follows from Merge
(Chomsky 2001a). One could also say that the distinction is between *(first-time)
merge* and *remerge* (that is, Merge again). The first, *merge*, is inevitably external.
But is *remerge* always internal? Standardly, this is tacitly assumed. However, it
does not in any way follow from the definition of Merge, or from the boundary
conditions mentioned so far. This will become clear in a moment.

　　The essence of Merge is that it is structure-building. It combines units into a
larger unit, which then constitutes a new root. In accordance with the two usual
boundary conditions, it combines two distinct syntactic objects (say, A and B)
into a new, larger unit (C), which, by definition, is then also a syntactic object. Let
us notate this as Merge (A, B) → C, which is an operation resulting in the possible
representation [$_C$ A B]. If we go on merging C with an external D (lexical or
complex), we create another syntactic object, call it E: Merge (D, C) → E, resulting
in [$_E$ D [$_C$ A B]]. If instead we merge a term of C, say B, with the root C again, we
create a movement configuration by internal remerge: Merge (B, C) → E, giving
[$_E$ B [$_C$ A B]], where B has now two sisters (that is, Merge-mates), namely both A
and C. We are used to calling the lowest B a copy, but this is misleading: Nothing
in the syntactic system distinguishes the two Bs in the representation (unless
further, complicating assumptions are made). In fact, there is no second B to
begin with. There is just one B that is involved in two relationships created by
Merge. The two Bs are an artifact of the representation. A less misleading way of
representing the result of the two mergers under discussion is the multi-
dominance representation in (3), although it has the disadvantage of being
graphically a little awkward. Notice that we can picture B in its first-merge
position, in its Spell-Out position, or in fact anywhere else on the paper:

(3)　　Merge (A, B) → C
　　　　Merge (B, C) → E
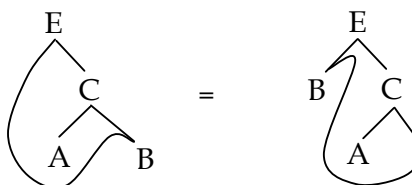
See also Epstein *et al.* (1998), Starke (2001), Gärtner (2002), Zhang (2004), and
Frampton (2004), among others, for further arguments against the copying view
of displacement.[1] For earlier discussion of similar ideas, see Sampson (1975),
Karlgren (1976), McCawley (1982), Peters & Richie (1982), Engdahl (1986), Huck
& Ojeda (1987), Blevins (1990). What should be clear is that the assumption of

---

[1]　　From a completely different perspective, Karttunen & Kay (1985) warn that the amount of
　　computational effort that goes into producing copies is much greater than the cost of
　　'unification' (that is, multidominance) when a graph is being parsed. For this reason, they
　　advocate structure sharing.

copies would require theoretical machinery in addition to the operation Merge *per se*. A different matter is how the phonological interface interprets the result in (3); this will be discussed in detail in section 3.

Movement, as we saw, involves *remerge*, that is, a syntactic object that has been merged before, is merged again. If a previously merged object α is selected as input for Merge, and if the other input object is the root R from which α has been selected, this instance of remerge can be called *internal*. However, as announced before, this does not exhaust the possibilities: α can in principle be remerged with an independent syntactic object, that is, an object that is not R and not embedded in R. This is what I will call *external remerge*; see (4).[2]

(4)   For some constituent α embedded in root R:

   a.   *internal remerge* =_{def} remerge α with R;

   b.   *external remerge* =_{def} remerge α outside R
          (i.e., with some root β not included in R).

Crucially, there is just one operation Merge; labels such as *internal remerge* are just names for the different situations caused by selecting different input objects. This is expressed in (5):

(5)   Merge (α, β) → γ constitutes

   a.   *first-time merge* iff α and β are independent roots before merger;

   b.   *internal remerge* iff β is a root and α is included in β (or the other way around) before merger;

   c.   *external remerge* iff β is included in some root δ, and α is an independent root (or the other way around) before merger.

Notice that heads introduced from lexicon or numeration are (trivial) roots before they are merged. For discussion concerning the strict cycle, I refer to section 2.3..

Although external remerge leads to unconventional structures (see further below), I must be stressed that the possibility of this operation simply follows from the combination of two independently motivated options: The selection of external material as input for Merge (needed for the introduction of lexical material), and the selection of terms (needed for regular movement); see also de Vries (2005c) and van Riemsdijk (2006a). It is of course possible to impose stricter boundary conditions on the input for Merge. For instance, the input could be restricted to roots. The consequence of this would be that remerge is excluded altogether (including regular movement). This point of view is defended in

---

[2]   As far as I know, Barbara Citko, Henk van Riemsdijk, and I myself first published basically equivalent ideas around 2005, independently of each other, and with somewhat differing terminology. In fact, it was predated by a remark in Wilder (1999), and of course inspired by earlier work on interarboreal movement, among others (see further below in the main text). It is perhaps worth mentioning that Chomsky (2007: 8, fn. 10) does not seem to agree: "[External remerge] requires new operations and conditions on what counts as a copy". Further explanation is lacking, and frankly, I fail to see why this would be so. Moreover, the objection is invalid from the present perspective, since a copying mechanism was rejected to begin with.

Koster (2007), among others.[3] The grammar would then be more restricted, but at the cost of an additional rule. If the familiar internal remerge is to be allowed, but the unorthodox external remerge to be excluded, more specific additional conditions must be formulated. However, it may be interesting to put off such stipulations, and allow for remerge in general. Here, I will follow this track, and explore some of the consequences.

Several possible interpretations of what can now be recognized as external remerge have been proposed in the literature. These include 'interarboreal movement' (Bobaljik 1995, Bobaljik & Brown 1997), 'sideward movement' (Nunes 2001), 'multidominance/multidomination/multiple dominance' (McCawley 1982, Ojeda 1987, Blevins 1990, Wilder 1999, Chen–Main 2006, Johnson 2007, Bachrach & Katzir 2009), 'sharing' (Guimarães 2004, Chung 2004, de Vries 2005b, Gracanin–Yuksek 2007), 'grafting' (van Riemsdijk 1998, 2006a), and 'parallel merge' (Citko 2005). Furthermore, external remerge is allowed in some way or another in many theories involving 'parallel structures' (Williams 1978, Goodall 1987, Mu'adz 1991, G. de Vries 1992, Moltmann 1992, Grootveld 1994, te Velde 1997). See also Carnie (2008) for a brief overview. I cannot do justice to all these proposals, but the two central ideas that are relevant here are pictured in (6). Notice that (6a) equals (6a'), and (6b) equals (6b'); apparent differences are only due to the position of the independent two-legged mini-structures on the paper:

(6)   a.

    b.

    a'.

    b'.
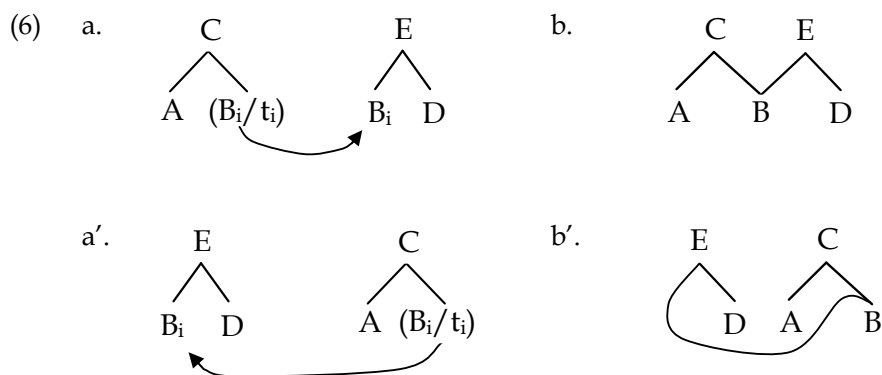
In (6a/a'), B is moved to an independent structure. Let us provisionally call this iMove (short for interstructural movement). This iMove is different from traditional (rightward or leftward) movement, which involves movement to a position within or at the top of the *same* structure. In (6b/b'), B is shared between two structures. Let us call this mDom (short for a hydraic — that is, multi-rooted — multiple dominance configuration), which, like internal remerge as in (3), involves giving up the 'single mother condition' used in previous frameworks

(see Sampson 1975); in a derivational framework, it also involves giving up the 'single root condition' — at least during the derivation (see further section 2.3). However, if structures are derived by Merge, *all* representations in (6) are derived by the following two applications of Merge:

(7)   a.   Merge (A, B) → C
       b.   Merge (B, D) → E

In (7a), B is merged with A, which gives C. In (7b), B is remerged with D, which gives E. Since D is not related to C (the root), the step in (7b) is an instance of external remerge. Thus, the perhaps surprising conclusion must be that it is only the notation that suggests a difference between iMove and mDom, captured as *external remerge:* iMove = mDom.

Without further assumptions (such as special properties of copies/traces and chains, which we must reject *a priori* until strong independent evidence to the contrary comes up, *pace* Nunes 2001 and others), iMove is actually equivalent to mDom. The representations in (6) are just that: More or less successful representations of certain theoretical concepts. What is 'real' is that Merge creates basic relationships between syntactic objects: Grammatical inclusion and grammatical sisterhood (see section 4 for further discussion). A graph that represents such relationships has no independent theoretical status. See also de Vries (2009b) on the issue of notation in syntax, including an unorthodox proposal. Furthermore, I would like to stress that multidominance is independent of *multidimensionality* (e.g., '3D grammar'), despite some suggestive descriptions in the literature. An additional syntactic dimension, in my view, would imply the assumption of an additional basic relationship (next to dominance or sisterhood); see also Groot-veld (1994).[4]

In (7), there is only one B, and this B is engaged in two basic 'triads', if I may borrow an expression from Koster (2007). A triad is the minimum amount of structure, equivalent to what is created by one instance of Merge. Thus, Merge (α, β) → γ relates α, β and γ such that α and β are directly included in γ, and α is the grammatical sister of β. The advantage of using multidominance graphs as in (6) is that they represent the fact that some node (here, B) is involved in a double set of basic relationships, without suggesting that this node itself is magically multiplied. The mDom notation, therefore, can be used to represent remerge in general, and I will stick to it in the remainder of this article.

2.2.   *Potential Examples of External Remerge*

In section 2.3, I will address the status of the strict cycle and some other theoretical issues, but first let me provide some concrete examples of sentence

---

[4]   Such proposals exist for both parenthesis and coordination (see also de Vries 2005a, 2007, 2009b for discussion and further references, some of which are mentioned in the main text). Strict definitions aside, it seems clear that all four combinations of [±3D] and [±mDom] occur: There are 'parallel structures' with and without 'sharing', and there may be remerge in regular hypotactic configurations as well. See the next subsection for some examples of the <+, +> pattern.

structures that may involve external remerge.

A by now almost classic case is RNR (or backward conjunction reduction). A simple example is provided in (8):

(8)   John admires ___, but Jill hates Bush.

The implied object in the first conjunct is *Bush*. McCawley (1982) proposed that this construction can be analyzed by allowing a constituent to be shared between two conjuncts, as is depicted in (9) — my example, with a simplified sentence structure for expository purposes. Here, the object *Bush* is dominated by both verb phrases:

(9)
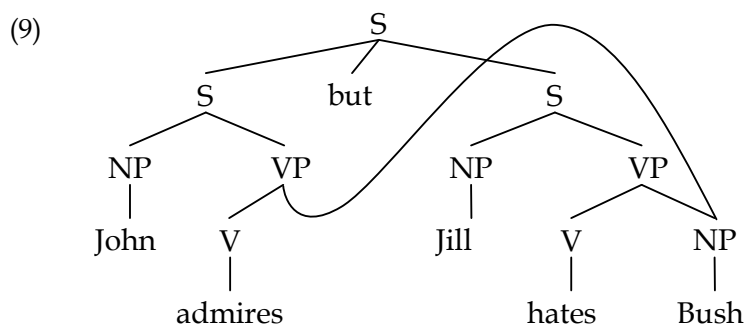
Although it is has not remained uncontested (see Postal 1998, Sabbagh 2007, Ha 2008a, 2008b), the idea of applying multidominance to RNR has been picked up and defended by several authors, for instance, Ojeda (1987), G. de Vries (1992), Wilder (1999, 2008), Chung (2004), de Vries (2005b), Chen–Main (2006), Johnson (2007), Kluck (2007, 2009), Bachrach & Katzir (2009), and Kluck & de Vries (to appear). Even though it is cast in different frameworks and stages of general syntactic theory, the basic idea is still the same. From the present perspective, we would say that the derivation of (9) involves merger of the NP *Bush* with one of the verbs, and then it remerges with the other verb.[5] Temporarily, this leads to a doubly-rooted structure, but since the two conjuncts are united at the top, the problem is resolved. (I will return to this.)

The reason for treating RNR in this special way is that it behaves differently from forward ellipsis/deletion, and also from regular movement and extraposition. For instance, RNR is apparently insensitive to island conditions (see Neijt 1979 and Hartmann 2000, among others; see also below), and it is immune to the Head condition on remnants (Fiengo 1974, Wilder 1997). Both properties fall out naturally from a multidominance approach. Trivially, since there is no ellipsis, there are no remnants, so the head condition does not apply, as

---

[5]   A concern for a theory in which an argument can be shared is that the relevant DP is assigned a theta-role twice (or more). It is conceivable that this is only allowed if these theta-roles are identical. Indeed, it is hard to imagine acceptable instances of RNR involving semantically different types of arguments. Thus, the matching effect induced by structure sharing may in fact serve as an explanation of certain parallelism requirements in reduced coordinated clauses. Notice that the situation is different in amalgams (see below); here, what is shared functions as a predicate in the interrupting clause, so the issue of a double theta-role does not arise (Kluck, in progress).

required. Furthermore, no matter how deeply embedded the shared constituent is (here, the NP *Bush*), it is locally related to each sister (here, the two verbs). That is, the multidominance connection creates a kind of bypass (see also section 2.4). For more discussion concerning RNR *per se*, see Kluck & de Vries (to appear) and the references mentioned.

Other constructions that qualify for external remerge are *wh-amalgams* and *cleft-amalgams*, as discussed in Guimarães (2004) and Kluck (2008), based on earlier work in Lakoff (1974), van Riemsdijk (1998), and Tsubomoto & Whitman (2000) — *pace* Zwart (2006b) and Grosu (2006). These are illustrated in (10a) and (10b), respectively.

(10)   a.   Jack gave [you will never guess which *girl*] a flower.
         b.   Jack gave [I think it was *his girlfriend*] a flower.

Here, the interrupting clause between brackets gives rise to a bracketing paradox, since the *content kernel* in italics is also part of the main clause. A multidominance solution to this problem is informally sketched in (11):

(11)   [interrupting clause   (…)]

   main clause ⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯→ *content kernel*   (…)
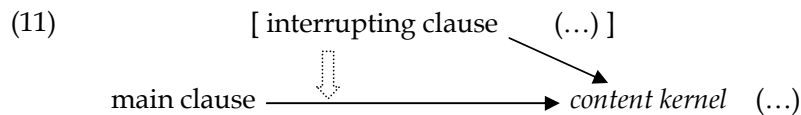
The content kernel is dominated by a projection of the main clause as well as the interrupting clause. The latter is inserted as a parenthetical in the main clause (Kluck, in progress, *contra* Guimarães 2004; see also de Vries 2009b). The details need not concern us here; what is relevant is that the shared constituent needs to be externally remerged.

Another construction that has been argued to involve sharing is Across-the-Board (ATB) movement; see Williams (1978), Goodall (1987), Citko (2005), Mayr & Schmitt (2009), among others. A standard example is (12):

(12)   *Which man* does John admire ___ but Bill hate ___?

The idea is that prior to *wh*-movement the relevant constituent (here, the object *which man*) is shared between positions within two or more conjoined clauses (IPs); see (13):

(13)
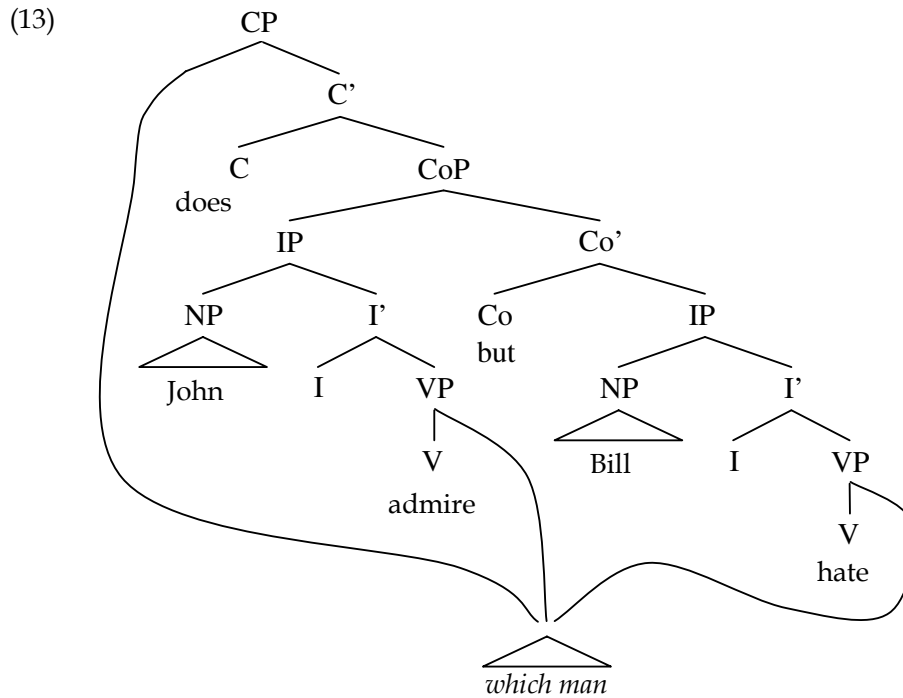
This structure is derived by externally remerging the object from one VP to the other; after that, both conjuncts are completed and joined by means of a coordination phrase;[6] finally, the CP level is added, and regular *wh*-movement takes place. Thus, the ATB-construction combines external and internal remerge. In section 3 it is discussed how it must be linearized.

Let me list some further, interesting proposals that involve externally remerged material (for the record, I am not personally committed to all of these). In chronological order:

— van Riemsdijk (1998, 2006b) on transparent free relatives, where the content kernel (the predicate) of the TFR is shared with the matrix (he also suggests a similar approach to internally headed relative clauses;

— Nunes (2001, 2004) on parasitic gap constructions, where the *wh*-constituent is 'sideward moved' before fronting;

— van Riemsdijk (2001a) on *wh*-prefixes, where the *wh*-word is shared between the matrix and the 'prefix' (for instance, "God knows who…");

— van Riemsdijk (2001b) on bracketing paradoxes as in *a far from simple matter*, where the adjective is part of two different trees;

— van Riemsdijk (2006b) on regular free relatives, where there is sharing of the *wh*-operator between the matrix and the subordinate clause (the purpose of this is to explain Case matching effects).

— Henderson (2007) on relative clauses, where there is sideward movement of the head NP between the relative clause and the matrix;

---

6   For discussion and references concerning coordination *per se*, see de Vries (2005a).

— Gracanin–Yuksek (2007) on coordinated *wh*-constructions, where there is 'bulk sharing' after the *wh*-constituents;

— Meinunger (2008) on bracketing paradoxes in certain complex numerals;

— Heringa (2009, in progress) on appositional constructions, where the appositional core is shared between a parenthetical position and a position in the matrix;

Whether each individual analysis of a particular phenomenon just mentioned will eventually be embraced or discarded does not matter for the purpose of this article. The point is that there is a by now substantial body of literature on structures involving external remerge. This in itself justifies a closer look at the formal properties of sharing, and the problem of linearization in comparison with regular movement.

2.3.   The Strict Cycle

The possibility of *remerge* raises questions about the course derivations can take. In this respect, consider the so-called *extension condition*, also known as *strict cyclicity* (Chomsky 1995: 190, 327). Since Merge is structure-building and not structure-changing, counter-cyclic *merge* or *remerge* is simply impossible. Basically, Merge (X, Y) combines X and Y but leaves the internal structure of X and Y intact. This is worked out in some more detail in (14) and (15), for merge and remerge, respectively. In each case, the projection E is created, but E is not the new root. Instead, E is inserted as the daughter of C, and the original direct inclusion relationship between C and A in (14a) and (15a), and the one between C and B in (14b) and (15b) is destroyed. In each example, the original existence of [$_C$ A B] is the result of a previous instance of Merge.

(14) a. (i) Merge (D, [$_C$ A B]) $-//\rightarrow$ [$_C$ [$_E$ D A] B]

          (ii) [$_C$ A B] and Merge (D, A) $-//\rightarrow$ [$_C$ [$_E$ D A] B]

      b. (i) Merge (D, [$_C$ A B]) $-//\rightarrow$ [$_C$ A [$_E$ D B]]

          (ii) [$_C$ A B] and Merge (D, B) $-//\rightarrow$ [$_C$ A [$_E$ D B]]
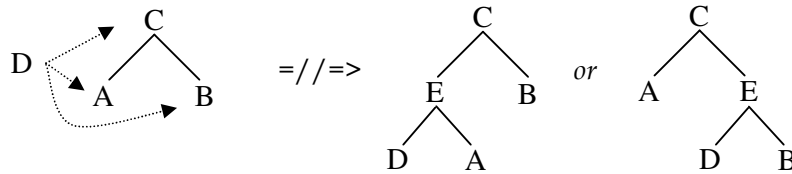
(15) a. (i) Merge (B, [$_C$ A B]) $-//\rightarrow$ [$_C$ [$_E$ B A] B]   (mDom of B)

          (ii) [$_C$ A B] and Merge (B, A) $-//\rightarrow$ [$_C$ [$_E$ B A] B]   (mDom of B)

      b. (i) Merge (V, [$_C$ A [$_B$ U V]]) $-//\rightarrow$ [$_C$ A [$_E$ V [$_B$ U V]] (mDom of V)

          (ii) [$_C$ A [$_B$ U V]] and Merge (V, B) $-//\rightarrow$ [$_C$ A [$_E$ V [$_B$ U V]](mDom of V)

(15a)   (15b)
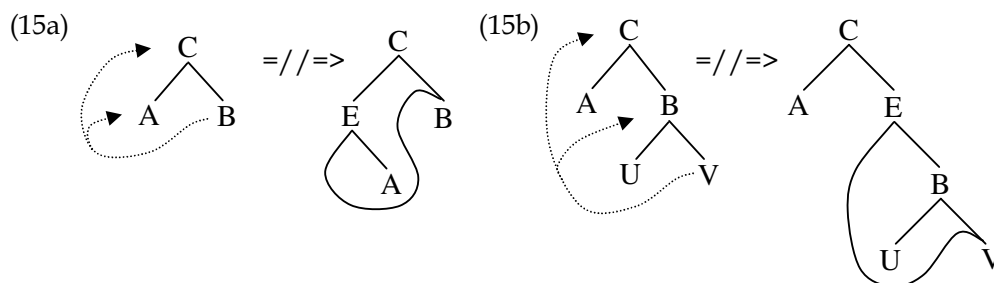
Clearly, if this were possible, it would be an undesirable complication of the theory. A similar reasoning can be found in Chomsky (2005), who introduces the *no-tampering condition*. The no-tampering condition can be considered a derived consequence of the system. It need not be an independent principle of grammar, since tampering is simply not what Merge does, at least not from the most minimalist perspective.[7] That said, the reader may have noticed that instances of external remerge may eventually lead to structures that *seemingly* involve tampering. I will come back to this shortly.

Now we know what the mergers as in (14) in (15) do *not* lead to, let us consider which structures they *do* create. The mergers in (16a–b) are familiar, as they involve *merge* or *remerge* at the root. The option in (16c) constitutes *external remerge*, which leads to a doubly-rooted graph:

(16)   a.   Merge (D, [$_C$ A B])   →   [$_E$ D [$_C$ A B]]

                                                   (regular *first-time merge*)

       b.   Merge (B, [$_C$ A B])   →   [$_E$ B [$_C$ A B]]

                                                   (reg. *internal remerge*: mDom of B)

       c.   [$_C$ A B] and Merge (D, A) →   [$_C$ A B] and [$_E$ D A]

                                                   (reg. *external remerge:* mDom of A)

---

7   However, Chomsky (2000: 137), following Richards (1999), leaves open the possibility of 'tucking in' for 'third Merge', which would be a clear violation of the extension principle. The reason that this might be allowed is that it does not change the relationships of a *head* with respect to its complement and first specifier. Obviously, it does change basic relationships with and between projections of the head. It seems to me that in the absence of overwhelming evidence for tucking in, such a complication of Merge must be rejected. Merge, in its simplest definition, operates on syntactic objects, regardless their internal structure and projection status, creating lasting basic relationships between the input objects and the output object. See also section 4.

(16a)

(16b)

(16c)
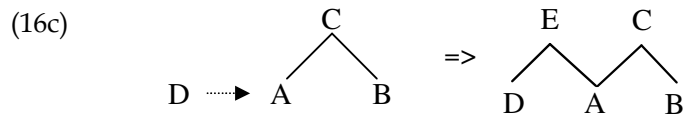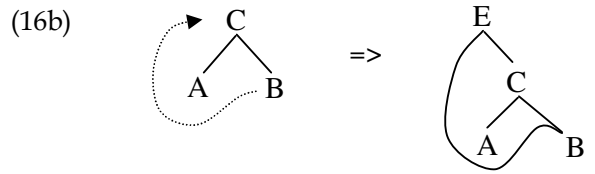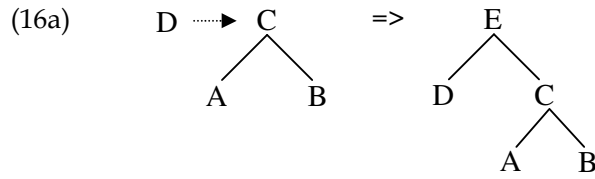
For completeness' sake, note that (16a) replaces (14a–b.i), (16b) replaces (15a.i), and (16c) replaces (14a.ii). The mergers in (15b.i) and (14b.ii) also involve regular internal and external remerge, respectively, and the actual resulting structures can be compared to (16b) and (16c), only then *remerge* concerns the other sister.

The problematic cases are the mergers in (15a.ii) and (15b.ii), which involve remerge with a non-root (namely, in (15a.ii), the term B is remerged with the term A; similarly, both V and B are embedded in (15b.ii)). The result cannot be structure-changing, as in (15), but instead an additional root node will be created, comparable to what happens in (16c). Consider a slightly more sophisticated and illustrative example. In (17), the problematic instance of Merge is accompanied by an exclamation mark. The first merger between brackets is a preparatory sub-derivation. The mergers in grey are a vain attempt to correctly finish the offensive structure.

(17)   (Merge (β, γ) → F)   (to be excluded)

         Merge (A, B) → C
         Merge (D, C) → E
         Merge (F, E) → G
    !   Merge (A, β) → J

         Merge (H, G) → I
         Merge (J, I) → R
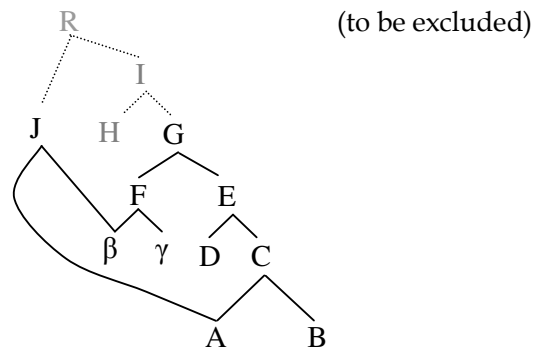
Here, A is remerged with the embedded β; this automatically leads to a tempo-rary second root, J. The reason is that Merge by definition creates a new pro-jection. (The same would apply if A were remerged with F itself, which is also embedded, namely in G.) Eventually, the two temporary roots can be combined into a final single root R, with possible additional material in between (such as H). One could call this 'quirky internal remerge' — internal, since no new

material is selected; quirky, because movement to an embedded position is normally considered ungrammatical. Furthermore, even if it were grammatical, the then intended string of abstract terminals is /H A β γ D B/, but I do not see how this could possibly be read off the structure. I conclude that the theoretical possibility of quirky internal remerge somehow needs to be excluded.

There is a counterpart of the above that we could call 'quirky external remerge', which involves remerge with an embedded position in another structure; see (18):

(18)   (Merge (D, E) → F)   R   (to be excluded)

    Merge (A, B) → C
!   Merge (E, A) → G   H

    Merge (F, G) → H
    Merge (H, C) → R   F   G   C

                                     D   E   A   B
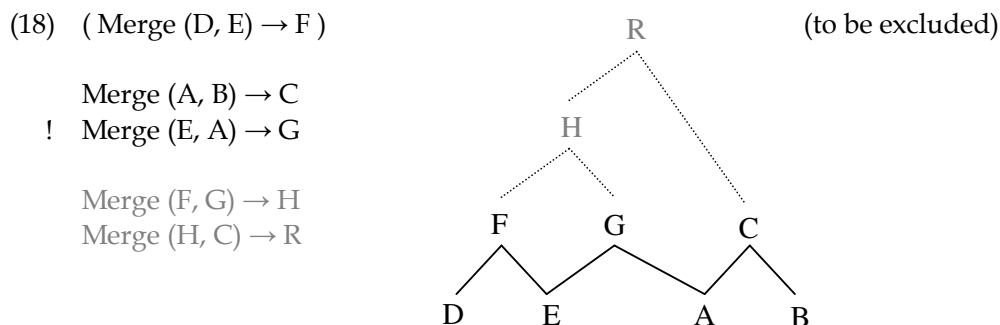
Here, A, which is a term of C, is externally remerged with E, which is embedded in the independently created F. As a result, a *third* temporary root, namely G, is created. Eventually, everything can be combined in one final root R, with possible additional material in between. The problem is that so far, I have not been able to come up with a realistic linguistic interpretation of (18). Moreover, like (17), it is clearly against the spirit of the extension condition, even though it does not involve 'tampering' in the strict sense.

Is there a plausible way to exclude both (17) and (18) at the same time? It seems to me that there is, but of course there is a theoretical cost to this, namely in the form of an explicit condition on the input for Merge. What (17) and (18) have in common is that at some crucial point of the derivation *both* input objects for Merge are terms and not roots (at that stage). A formal condition preventing this can be formulated as follows:

(19)   *Root condition:*
    If α and β are selected as input for Merge, then α or β (or both) must be a root.

There is a clear rationale for this condition. Consider (17) and (18) again. In both cases, the offensive instance of Merge creates an additional root where none was before. But this is not what Merge is for, from a functional perspective. Merge is essentially a combinatory device: it combines lexical items until a final single-rooted structure is created, which can then be pronounced. Bearing in mind that every lexical item itself is a root (of a trivial structure), Merge does the following:

  A. If two lexical items are merged, the result is that the number of roots is reduced by one. Namely, after Merge (A, B) → C, where A and B are lexical items, the new root is C, and A and B have become terms of C.

B. For every other instance of first-time merge (which may involve complex items), the number of roots is reduced by one.

C. For regular internal remerge, the number of roots stays the same. Namely, if some term A (which is not a root) is remerged with the root R_i, R_i becomes a term of the new root R_{i+1}, and A is still embedded.

D. For regular external remerge, the number of roots stays the same. Namely, if some term A of root X is remerged with an independent root R_i, the result is that a new root R_{i+1} is created of which R_i is now a term. X remains a root, and A remains a term. So we start out with two roots (X and R_i), and end up with two roots (X and R_{i+1}).

E. For quirky internal remerge, the number of roots is enlarged by one. Namely, if some term A of root X is remerged with another term B of X, a new root R is created. Before merger, only X is a root; after merger, X and R are roots.

F. For quirky external remerge, the number of roots is also enlarged by one. Namely, if some term A of root X is remerged with some term B of another root Y, a new root R is created, and X and Y remain roots.

Thus, first-time merge is the best way to proceed towards the goal of creating a single-rooted structure.[8] Internal and external remerge are a necessary complication that causes some delay: The number of roots stays the same. But *quirky* internal/external remerge is completely counterproductive from this perspective, and must therefore be excluded. This insight is formalized in (20):

(20) *No proliferation of roots condition*

 If the derivation proceeds from stage i to i+1 through Merge $(\alpha, \beta) \rightarrow \gamma$, then $|\{x \in \{\alpha, \beta, \gamma\}: x \text{ is a root at stage } i+1\}| \leq |\{x \in \{\alpha, \beta\}: x \text{ is a root at stage } i\}|$.

Informally stated: Upon Merge, the number of roots may not become larger.

 The effect of (20) is completely equivalent to that of (19), so there are no two conditions, but just one that can be formalized in different ways, depending on the perspective. It is worth noting that there is a third way of looking at the root condition: one could conjecture that a derivation is always 'active at the top'.[9] Selecting a term is harmless as long as a root is involved as well. In the case of external remerge, the attention shifts from one structure to another. Quirky remerge does not involve any root in the input, and is therefore excluded.

 The attentive reader will have noticed that quirky remerge was not included under the definition of internal/external remerge from the beginning — recall (4) and (5) —, and I will no longer consider it.

2.4. *Remerge: A Discussion of Look-Ahead, Hydras, and Locality*

Let us take RNR as a relevant example of a construction whose derivation

---

[8] This could also be taken as a rationale for Merge-over-Move effects, to the extent that these are real (see, e.g., Castillo, Drury & Grohmann, in press for an overview.).

[9] Compare also Collin's (2002) 'Locus Principle' (bearing on Chomsky's ideas about feature activity), which has largely the same effect, although it is not equivalent.

involves external remerge. A simple sentence such as (21), which is similar to (9) above, can be derived in the following way:

(21)   Mary likes ___, and Jack hates *cars.*

1a   Merge (likes, *cars*) → [likes *cars*]

1b   Merge (hates, *cars*) → [hates *cars*]

2a   Merge (Mary, [likes *cars*]) → [Mary [likes *cars*]]

2b   Merge (Jack, [hates *cars*]) → [Jack [hates *cars*]]

3   Merge (and, [Jack [hates *cars*]]) → [and [Jack [likes *cars*]]]

4   Merge ([Mary [likes *cars*]], [and [Jack [likes *cars*]]]) →
                                        [[Mary [likes *cars*]] [and [Jack [hates *cars*]]]]
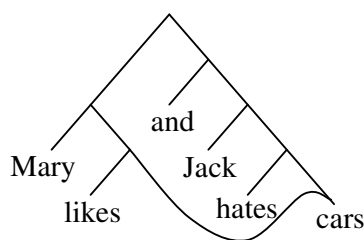

Here, step 1b involves external remerge of the direct object *cars*. During step 2a, 2b, and 3, the structure is doubly-rooted. Step 4 accomplishes a union into a single-rooted structure. Notice that there is no proliferation of roots at any step. In step 1b, we merge one root, [hates], with one term, [cars], and create one new root, [hates cars] — as merge always does. From the perspective of a syntactic workspace that initially contains all activated lexical items required for a particular derivation, we obtain the same result. In 1b, we start out with five roots, namely [Mary], [likes cars], [and], [Jack], and [hates], and we also end up with five roots: [Mary], [likes cars], [and], [Jack], and [hates cars].

A number of other things are worth discussing. First, consider the order of mergers. The derivation in (21) seems to suggest that we start out merging *cars* in the first conjunct, and remerge it in the second. However, before the two clauses are conjoined, there *is* no first and second conjunct: The order between them (or their respective terms) is only established later in the derivation. Are we dealing with an instance of look-ahead here? By no means. It is of no importance whatsoever with which verb *cars* is merged first. The sequence of mergers in 1a/b and 2a/b can be switched around at will. Either permutation (1a–1b–2a–2b, 1b–1a–2a–2b, 1a–1b–2b–2a, or 1b–1a–2b–2a) leads to the same result. Therefore, it is impossible to tell which occurrence of *cars* — the one in the first conjunct, or the one in the second conjunct — is the original and which is the copy. As we said before, there are no copies, just relations. And there should be no need for pre-destination in syntax. It is for PF to decide where *cars* is to be pronounced, independently of how syntax arrived at the structure under consideration. Thus, if a particular structure has more than one possible derivational history, it should be pronounced the same in either case.

The absence of look-ahead implies that every instance of Merge must be motivated in some way or another. It does *not* imply that every structure that *can* be derived by Merge is interpretable at the PF/LF-interface. It is easy to think of licit derivations that are still uninterpretable in the end, that is, incomplete in some sense. For instance, a relevant feature could still be unvalued. Therefore, some possible derivations will survive at the interface, and some will not. This is not the consequence of look-ahead, even though it might seem so from the perspective of a surviving derivation. Within narrow syntax, Merge is an autonomous operation.

Turning back to the case of external remerge, we have noticed that it creates a multi-rooted structure (a 'hydra' or 'forest'). The particular step of Merge itself may very well be motivated: In example (21), the verb *hates* selects a direct object. However, the existence of more than one root is problematic for the linearization procedure at PF (see below for details). So it is convenient that the two clauses are conjoined at a later stage, which resolves the problem. We can derive a heuristic from this: *Every instance of external remerge must be compensated by a joining operation later in the derivation* (surely, this need not be coordination; it can also be parenthetical insertion or subordination). Such a heuristic may suggest look-ahead, but that is misleading. Obviously, the system itself has no meta-modular analytical intelligence. The preferred derivation will survive as long as both the instance of external remerge and the compensating joining instance of Merge are independently motivated within narrow syntax.

Consider the hydraic configuration in (22), where α is the sister of both β_1 and β_2 due to external remerge. The structure projects up to R_1 and R_2.

(22)

R_1 R_2

γ_1 γ_2 β_2

β_1

α
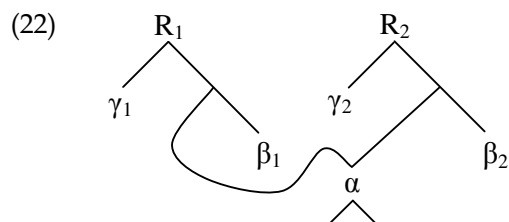
If this structure is sent to PF, how could it be linearized? The answer is that it could not at all. The reason is that the linearization procedure does not know where to start. And even if it randomly chooses one of the roots to be analyzed first, it is intuitively clear that no order between γ_1 and γ_2 can be established. With special additional assumptions, this may be resolved, but not in such a way that the order between the terminals remains invariant with respect to the choice of 'first root'. The two options here are the following: If R_1 is the root taking priority, the string of terminals will have to be /γ_1 γ_2 α β_2 β_1/; if R_2 is the root taking priority, the string of terminals will have to be /γ_2 γ_1 α β_1 β_2/. Clearly then, an asymmetry between the two (or more) temporary roots must be established: one is to be recognized as the matrix, the other as the secondary structure (a 'graft', using van Riemsdijk's terminology). The way to do this is to combine them in syntax. As a consequence, a graft cannot only involve sharing with a constituent of the matrix, the top of the graft must also be syntactically connected to the matrix (*pace* van Riemsdijk 1998, 2006). As I see it, the top connection is not only

required because of PF demands, it also makes sense from a semantic and syntactic perspective. Namely, the way the graft is connected to the matrix determines the relationship between them. For instance, a graft can be a second conjunct, as in (21), or a parenthetical-like insertion, as in cleft-amalgams (10b), or perhaps even a subordinated phrase, as in parasitic gap constructions.

    Next, let us turn to the issue of locality. In section 2.2., it was mentioned that RNR-constructions are insensitive to locality conditions, contrary to *wh*-movement constructions, for instance. This is illustrated in (23), where the dependency crosses (or seems to cross) the boundary of a complex noun phrase (with (23a) RNR and (23b) *wh*-movement our of a relative clause):

(23) a.   Mary likes [men who SELL ___], but she hates [men who BUY *cars*].
      b.   *What does Mary like [men who sell ___]?

External remerge, we said, creates a structural bypass. Let us see in a little more detail why this is so. Below, the derivation passes through the stages (24a–c).

(24) a.

      b.

      c.
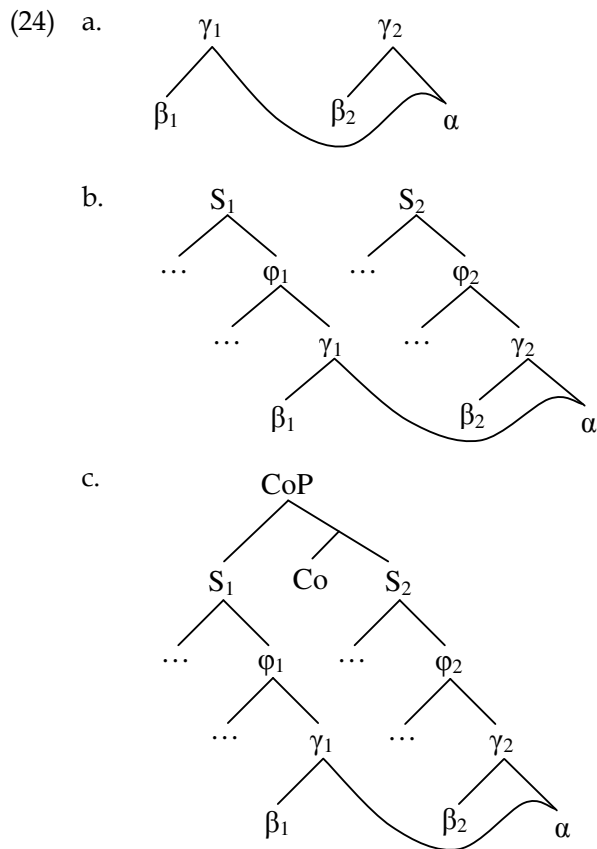
In (24), the constituent α is externally remerged. As no locality boundary is involved, yet, we will assume that this is unproblematic. In (24b), both spines of the structure are expanded by regular Merge up to S_1 and S_2. During this process, the locality boundaries φ_1 and φ_2 are created. In (24c), S_1 and S_2 are united in a coordination phrase. The end result gives the impression that α's relationship to

both β_1 and β_2 crosses the locality boundary φ. However, this is in fact not the case: α is locally related to β_1 and β_2 in step (24a). Whatever happens subsequently to this step cannot undo this local relationship. Put differently, the fact that β_1 and β_2 are not in the same local domain does not imply that some α cannot be locally related to both. This is the surprising consequence of external remerge.

If external remerge can create apparent locality violations in RNR, we may predict non-local behavior to show up in other types of sharing constructions as well. As far as I am aware, this has never been tested. I would like to claim that examples of this kind can indeed be construed. A relevant illustration in Dutch is a complex cleft-amalgam as in (25), where the parenthetical is to be interpreted as *de re*. The content kernel is italicized. Notice that a correct intonation is important: right before the dash, the pitch lingers relatively high in order to create a sense of expectation, the amalgam is pronounced relatively fast, and the content kernel is stressed.

(25) *Dutch*

Joop kuste toen — Piet beweerde dat   hij iemand kende die   zei
Joop kissed then   Piet claimed   that he someone knew   who   said
dat   het *Mieke*   was.
that it   Mieke   was
'Then, Joop kissed — Piet claimed that he knew someone who said that it was Mieke.'

Within the amalgam, *Mieke* is embedded in a complex noun phrase. If it is true that *Mieke* is at the same time part of the matrix (namely, as a direct object), there seems to be a locality problem. But this is only apparently so, and the solution is similar to the one sketched above for RNR. For more examples (in English), see de Vries (2009b).

Does the possibility of bypassing locality boundaries not endanger our theory of locality for regular movement constructions? I do not think this is the case. Consider the following configuration:
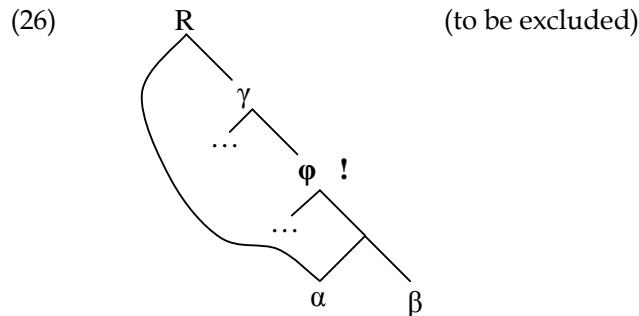
(26)   R   (to be excluded)

In (26), the phrase α is first-merged with β, and internally remerged with γ. If φ constitutes a locality boundary, the derivation is to be excluded. One way to do this is to make the selection of syntactic objects sensitive to structural distance. Thus, selecting a term as input for Merge is allowed as long as it is not too far embedded (where *too far* may be category-sensitive). From the perspective of

phase theory, if φ is a phase boundary, then it seals off its components for further computation. Thus, if the derivation in (26) has reached γ, α cannot be selected anymore since it is embedded in φ. Furthermore, a derivational bypass cannot be established, either: α cannot be externally remerged with γ before φ is reached for the simple reason that γ does not yet exist at that stage of the derivation. In section 3.5.3. I will come back to the issue of phases from the perspective of linearization.

2.5. *Intermediate Conclusion*

If terms of complex syntactic objects are allowed as input for Merge, *remerge* of heads and phrases is possible, as opposed to *(first-time) merge*. This is a way of dealing with the general phenomenon that an item can be involved in multiple (local) relationships, but shows up in only one position. Without ad hoc restrictions, it follows that there are two types of remerge: *Internal* and *external* remerge. The first corresponds to regular movement; the second is much more controversial, and has been characterized as sharing, grafting or sideward movement. Several construction types have been analyzed as involving what we now recognize as external remerge. Naturally, each of these will have to be subject to close scrutiny, and alternatives for some may be more viable in the end. However, what is of interest here is not so much the analysis of individual constructions, but the general mechanism of external remerge, in comparison to internal remerge. Both internal and external remerge can be represented in terms of multidominance. Though it has some graphical disadvantages, this prevents us from inadvertently attributing ad hoc properties to copies or traces. In this respect, it is worth commemorating that there is no inherent directionality in external remerge. If α is to be related to both β and γ, the order of mergers {Merge (α, β), Merge (α, γ)} is irrelevant, and look-ahead should not be necessary.

An automatic consequence of the structure-building characteristic of Merge is that it operates strictly cyclically: Merge creates a new root, and it cannot undo earlier relationships. But even then, some unwanted possibilities remain. These can be excluded by the *No proliferation of roots condition*, which seems a virtual conceptual necessity. As for locality conditions, they can be shown to be bypassed by means of external remerge in certain configurations, but this is never the case for internal remerge. Finally, we have seen that external remerge creates a temporary multi-rooted structure, which needs to be resolved before the structure gets linearized. As there are always asymmetric relationships between the different parts of the sentence, this will be taken care of for independent reasons as well.

3.   The Linearization of Complex Syntactic Structures

Syntactic structures have to be linearized at PF. Structures exclusively composed of relations established by first-time merge are easy to process; the possibility of remerge, however, brings about some complications. This section advances a proposal for the linearization of structures involving internal and external

remerge. The discussion below will be in terms of multidominance graphs. We have to keep in mind, though, that graphs as such are only representations of the underlying relations between syntactic objects created by Merge. Section 4 will examine in some more detail the necessary linearization algorithm and the computational cost it involves.

3.1.   The Problem of Remerge

Comparing *wh*-movement (27a) to RNR (27b), we notice a difference in the position where the displaced constituent (in italics) is pronounced:

(27)   a.   *Which violin* should this talented girl purchase ___?
       b.   The boy only admired ___, but the girl actually bought *this beautiful Stradivarius*.

In (27a), the remerged phrase is realized in the first position in the string; in (27b) the remerged phrase is realized in the last position in the string. This must be due to the different effect caused by internal and external remerge, respectively. In fact, we are facing two complications:

(C1)   Remerged items are only pronounced once.[10]

(C2)   Internally remerged items are pronounced in a different position than externally remerged items.

The abstract configurations corresponding to possible derivations involving internal and external remerge are sketched in (28a) and (28b), represented in terms of multidominance:

(28)   a.   b.

---
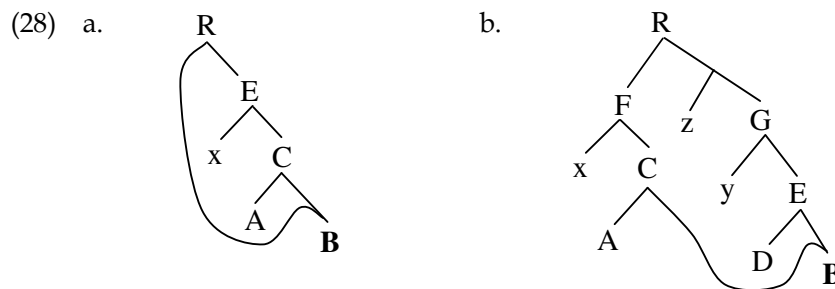
[10]   I am aware of proposals involving spelled-out copies or traces for particular phenomena such as resumptive pronouns and *wh*-copying; for discussion, see Aoun & Li 2003, Grohmann 2003, Nunes 2004, Barbiers *et al.* 2008, and Schippers 2008, among others. This possibility, if correct, is of course exceptional. Moreover, I would like to stress that it does not present additional problems for a remerge approach to movement as compared to a copy/trace approach. Let me quote Starke (2001: 145) on this:

> Questions about multiple traces map onto questions about multiple mergers. Reduplication paradigms are another instance of this logic: To the extent that they are adequately analyzed in terms of spell-out of a trace (i.e. spell-out of multiple 'copies'), they are now reanalysed as spell-out of multiple merger operations […]. No novelty introduced there.

The constituent called B is displaced. It is remerged with E in (28a), and with D in (28b). In (28b), it could also be first-merged with D and remerged with A. Suppose we traverse the graphs in (28) in the usual way, starting at the root (I will come back to this in more detail), we encounter the remerged B twice. Putting the terminals in a string in the order we come across them, we obtain the following picture, where the intended Spell-Out is printed below the other strings. What are terminals in (28), by the way, need not be linguistic heads; only, their possible internal complexity is of no direct interest to us, here.

	(28a)	(28b)
Terminals encountered		
by graph traversal:	/**B** x A **B**/	/x A **B** z y D **B**/
Desired string of terminals:	/**B** x A/	/x A z y D **B**/

From this perspective, it is the first occurrence of B that needs to be overtly realized in (28a), but the second occurrence in (28b). The generalization can be stated as follows:

Internal remerge of X   → overtly realize only the first occurrence of X while linearizing the structure.

External remerge of X   → overtly realize only the last occurrence of X while linearizing the structure.

Below, I will discuss what happens if the two interact.

There is a body of literature on the linearization of syntactic structures involving movement, and there are also some publications on the linearization of sharing constructions (especially RNR). But as far as I know, the issue sketched here has not received any explicit attention in the literature, apart from Chen–Main (2006), Gracanin–Yuksek (to appear), and some brief remarks in Wilder (2008).

I have no intention of negatively reviewing other linearization proposals in detail — and no doubt each has its own merits —, but let me indicate in general terms why the course taken here is somewhat different. The most straight-forward objection to all previous proposals I am aware of is the lack of general applicability: theories about movement (for instance, Fox & Pesetsky 2005) are unfit for sharing and *vice versa* (for instance, Wilder 1999); so the least we can say is that some adaptations are necessary.

Wilder (2008) and Gracanin–Yuksek (to appear) try to develop Kayne's (1994) Linear Correspondence Axiom (LCA) in order to manage sharing [external remerge] (see also Johnson 2007 for discussion), and briefly investigate if some extended version of the LCA is also fit for movement in terms of multidominance [internal remerge]. As they note themselves, it runs into trouble with moved constituents that are complex, that is, simply phrasal (Wilder), and with shared material that is not the most deeply embedded (Gracanin–Yuksek). Apart from that, Wilder explicitly derives the right periphery condition on RNR from the linearization procedure. To the extent that this is successful (see however Kluck

& de Vries, to appear, for critique, which also applies to Bachrach & Katzir 2009), it may be considered unfortunate, since it prevents the generalization of the idea of structure sharing (external remerge) to other constructions in which the presumed shared part is not necessarily right-peripheral.[11] Possible examples involve what van Riemsdijk (1998) called 'saddle grafts', where the shared material is not peripheral inside the graft. A cleft-amalgam in Dutch is (29):

(29)   *Dutch*

Joop heeft [ik vermoed dat het *een   Bugatti* is] gekocht.
Joop has I presume that it a   Bugatti is bought
'Joop bought — I presume it's a Bugatti.'

A further construction that is interesting from this perspective is ATB-movement, where external remerge seems to feed internal remerge (see Nunes 2004 and Citko 2005, among others). In the Dutch example in (30), the two gaps represent indirect object positions; these are certainly not clause-final:

(30)   *Dutch*

Wie   heeft hij ___ een boek gegeven en zij ___ een cd   ontnomen?
who   has he   a   book given   and she   a   CD   taken.away
'Who did he give a book and she take away a CD from?'

Citko explicitly states that movement following sharing is necessary for successful linearization at PF. Her theory, therefore, is quite limited, since it excludes an analysis of RNR and amalgams in terms of sharing.

Chen–Main's (2006) analysis is the most extensive one; it combines an inherent asymmetry between sisters (precedence) with LCA-like demands on the linearization. The proposal has the following basic characteristic: Multidominance of left branches leads to the pronunciation of the highest, leftmost occurrence, whereas multidominance of right branches leads to the pronunciation of the lowest, rightmost occurrence. The first corresponds to regular movement, the second to sharing in RNR-constructions. A serious problem discussed by Chen–Main herself is *wh*-movement of direct objects, which is excluded by the system. She cleverly turns this into a partial advantage by using it to explain Holmberg's generalization: Namely, the configurational problem is resolved if the verb is moved as well. But of course Holmberg's generalization is far from universal; moreover, I think the problem concerns movement of right branches in general, not only of direct objects. Chen–Main discusses various complicated interactions between multidominance links in syntactic graphs. It turns out that some structures can be linearized, some cannot, and some are linearized

---

[11]   The particular right-periphery condition for RNR must then be explained in another way. In fact, this conclusion is corroborated by Kluck (2007, 2009), who shows that purely syntactic approaches to derive the periphery effect fail; instead she proposes to combine Hartmann's (2000) theory on the semantics and prosody of the RNR-construction with a multidominance approach. Put differently, the periphery condition is an interface effect clearly related to contrastive focus; a successful explanation must somehow take that into account.

in a way we would not expect. This raises a more fundamental question: Do we want phonology to restrict syntax in such an intricate and hard-to-predict way? Without lessening any of the value of Chen–Main's discussion, my answer right now would be negative.

This answer confirms my general skepticism towards LCA-based analyses. Let me briefly formulate some general considerations indicating why I am not convinced that this kind of approach is attractive, apart from the issues mentioned above. I should mention beforehand that this does not imply that I am against a universal Spec–Head–Comp order, and hence right-branching graphs. The reason is that this universal is based on conceptual reasoning as well as empirical generalizations. In the absence of a phonological principle such as the LCA, it could simply be hard-coded in syntax.

First, the LCA presupposes that syntax does not directly encode an asymmetry between sisters that can be interpreted as precedence by phonology. Therefore, the necessary asymmetry at PF must be calculable from hierarchical information. However, the mentioned presupposition has been questioned in the literature. In section 3.2.1., I will come back to the idea that Merge produces ordered pairs.

Second, the LCA is often used to linearize (spell out) structures, but strictly speaking, the LCA states a necessary *property* of syntactic structure; it is not a *procedure* to arrive at the demanded (transitive, antisymmetric, and total) linear order from the syntactic structure. Kayne (1994) explicitly formulates the LCA as an *axiom*, not an *algorithm*. Nevertheless, suppose that we formulate such a procedure on the basis of the LCA (see section 4 for an impression how to construct a linearization algorithm). Even then we are not there, yet: What PF wants is not a mathematical linear order, but a string of words. Of course, we can formulate another mapping procedure that translates the set of ordered pairs of terminals that constitutes a linear order into a string, but it should be clear that this is an additional step:

Linearization: syntactic structure $\xrightarrow{\text{LCA-procedure}}$ linear order $\xrightarrow{\text{mapping of complex set onto string}}$ word string
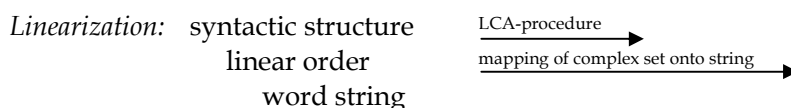
From a practical perspective, if we can go directly from syntactic structure to a word string, this seems preferable over the state of affairs sketched above.

Third, the LCA involves a very intricate definition of c-command. In the extended approaches it is even further enriched by the notion of full dominance/ unique paths (see Wilder 2008, for instance). But c-command is a very general syntactic tool, used for many more things than linearization; essentially, it identifies possible dependencies. It is doubtful that such a fundamental notion could be so complex. Rather, it seems likely that c-command is a direct function of Merge, as initially proposed by Epstein (1999) — put briefly, if A and B are merged, then A c-commands B and every term of B.

Fourth, c-command is restricted to full categories (heads or maximal projections); segments (or X-bar nodes, in Chomsky 1995) do not count. As a consequence, a specifier asymmetrically c-commands the components of its sister,

but not *vice versa*. It seems to me that if the asymmetrical behavior between specifiers and their sisters needs to be (indirectly) stipulated in this way, the goal of deriving linear order from hierarchy is not convincingly attained.

Fifth, the LCA is hard to combine with Bare Phrase Structure; see Chomsky (1995), Uriagereka (1999), and others. In particular, there is a lack of asymmetric c-command between the very first two (lexical) items that are Merged in a derivation.

Sixth, given that movement is now viewed as internal remerge, configurational complications arise if more than one phrase is remerged. If I am not mistaken, LCA-based proposals inevitably run into trouble with remnant movement and roll-up movement (see section 3.5.1. for abstract illustrations of such constructions).

In the next sections, I will formulate an alternative approach that does not suffer from the limitations discussed above. It is fair to say, though, that it is not free of stipulations that are in want of a deeper explanation. My main objective here is to make the required linearization procedure fully explicit.

3.2.   Linearization as a Process: Theoretical Preliminaries

Linearization is the process of turning a hierarchical structure into a string of terminals. In line with Kural (2005), Kremers (2009), and others, I assume that the most straightforward way to do this is by means of graph traversal. The details of this process will be discussed from section 3.3. onwards. But first, a number of theoretical preliminaries need to be addressed.

3.2.1. Asymmetrical Syntax

Graph traversal implies that the linearization procedure can make use of the precedence relation between sisters (known as *direct precedence*, *immediate precedence*, *strict precedence*, or *sister precedence*). This is in line with Frampton (2004), Chen–Main (2006), and many others, now, and especially in the 1980s, when the idea that directionality between heads and complements was considered a language parameter by most people. However, it goes against Kayne (1994) and Chomsky (1995), who state that syntax is about hierarchy and not about order. From the perspective of a derivational grammar, this means that Merge produces a complex whose components are unordered (an unordered set, according to Chomsky).

However, even when we grant the idea that syntax should not be preoccupied by linear order between sisters, the conclusion that Merge produces an unordered pair does not logically follow. After all, it is very well possible that Merge produces an ordered pair that encodes a syntactic or semantic asymmetry. If this is so, it may be the case that the relevant syntactic asymmetry can directly be mapped onto direct precedence at the PF interface. But that does not mean that linear order is part of syntax. It is simply a misconception to equate asymmetry between sisters in syntax with linear/temporal precedence, even if it will eventually lead to this in the phonological component.

So now we face two questions: First, is there indeed a consistent syntactic

or semantic asymmetry between sisters? In other words, does Merge automatically produce ordered pairs? Second, can this syntactic asymmetry consistently be mapped onto direct precedence at PF? The first question has been answered positively by several authors, albeit on somewhat different (but not necessarily contradictory) grounds: Jaspers (1998), Koster (1999, 2003, 2007), Di Sciullo (2000), Langendoen (2003), Zwart (2004, 2006a), Di Sciullo & Isac (2008). If the second question is to be answered positively, a much-preferred condition, I presume, is a universal base order. The most likely is a universal Spec–Head–Comp order, as advocated by Kayne (1994), Zwart (1994), and many others since (*pace* alternative configurations in Fukui & Takano 1998 and Haider 2000). In accordance with several of the cited works, let me sketch an account in terms of *dependency*.

The operation Merge creates sisters. From a compositional-semantic point of view, the objective of Merging, say, X and Y is to relate X to Y, thereby giving rise to a combined meaning. Crucially, sisters are never in a symmetrical relationship. For one thing, it is generally assumed that complements are c- and s-selected by heads. If it is indeed the case that complements universally follow heads, then the PF-mapping of this syntactic asymmetry on a phonological direct precedence relationship is unproblematic.

How about the combination of phrasal constituents? According to Koster's *configurational matrix*, there is always a left–right asymmetry (presupposing universal SPEC-left). Semantically, the righthand sister is 'relatively about' the lefthand sister, which is normally more salient: For instance, the predicate is interpreted with respect to the subject; a comment is about the topic, and so on. Syntactically, anaphoric dependencies (in the broadest possible sense of the word) are from the lefthand sister to (a term of) the righthand sister, *modulo* reconstruction effects due to A-bar movement, as is well-known. Thus, in a configuration [$_\gamma$ α β] it is always β (and indirectly, its terms) that is dependent on α. Zwart's (2006a) hypothesis is that dependency is a function of Merge. Thus, Merge (α, β) produces an ordered pair <α, β> such that β is the dependent. If Zwart and Koster are right, then the asymmetrical merger of phrases can be mapped onto direct precedence at the PF interface, as well. The generalization is that if α and β are merged, whether they are heads or phrases, then the direction of dependency can directly be mapped onto direct precedence, such that the dependent always follows the non-dependent. Let us call this the Uniformity of Mapping hypothesis:

(31) *Uniformity of Mapping Hypothesis*

At the PF interface, generalized syntactic dependency is directly mapped onto phonological precedence, such that in a basic syntactic triad <$_\gamma$ α, β>, α will directly precede β.

In the next sections, I will use (31) as a background assumption. In principle, the necessary mapping could as well be performed by a more intricate rule system that makes reference to language-specific directionality parameters, but let us stand by the simplest solution.

Before we return to the main subject, a note on the widespread idea of Spec–Head agreement may be in order. Crucially, a specifier is a sister of a

projection of the head: In [$_{XP}$ Spec [$_{X'}$ X Comp]], Spec is the sister of X'. Since a projection of X contains all the features of X per definition (this is the idea of percolation), the Spec–Head relation can be reduced to a sisterhood relationship. One could say that the head is the nearest (most local) term of the sister of the specifier, which is probably the reason why the Spec–Head relation is important. Therefore, it comes as no surprise that it is often morphologically encoded. Nevertheless, there are indications that Spec–Head is only a special case of the more fundamental sisterhood relationship. An interesting example in this respect is subject-predicate agreement in Swahili. Consider the example in (32), taken from Carstens (2003: 395):

(32)   *Swahili*

   Juma   a-li-kuwa   a-ngali   a-ki-fanya   kazi.
   Juma   SA-PST-*be*   SA-*still*   SA-PROG-*do*   *work*
   'Juma was still working.'

Here, SA is subject agreement, PST past tense, and PROG progressive. As is evident from the gloss, the subject agreement morpheme is spelled out on several elements within the predicate, including an adverb. Zwart (2006a) suggests that we can analyze this example as follows: Not just the finite verb, but the predicate as a whole, being the sister of the subject, is marked as the dependent of the subject. This dependency can then be spelled out on several terms of the predicate; which one(s), that concerns a language-particular morphological choice.

   A line of research with very similar characteristics is Matushansky's (2008) approach to Case marking. In her view, Case marking involves a sisterhood dependency, which may eventually be spelled out on a term of the dependent sister (normally a DP). In this way, different Case features may accumulate on a single DP, which leads to language-particular morphological choices.

   Finally, let me briefly comment on Chomsky's conception of set-Merge. Chomsky (1995, and subsequent work) advocates a dominance-only grammar. In his notation, the result of Merge (a, b) is an unordered set {a, b}, which is then type-lifted to {a, {a, b}} in case *a* projects. At first sight, the issue of the label is interesting in the light of the by now famous Wiener–Kuratowski convention (Wiener 1914, Kuratowski 1921), which states that an ordered pair <x, y> is equivalent to {{x}, {x, y}}. This complex set is often thought to be equivalent to {x, {x, y}}, for instance in Cormen *et al.* (1990: 80); see also Quine (1945) and Schneider (1977) for discussion. Such set-theoretic reductionism has been criticized by several philosophers (see Armstrong 1986, Forrest 1986, Goodman 1986, Sider 1996), mainly because it is arbitrary. Nevertheless, the convention is widely used. When applied to Chomsky's notation, it would follow that {a, {a, b}} is an ordered pair <a, b>. Thus, one might suppose that Merge (a, b) produces the ordered set <a, b> in case *a* projects, and <b, a> in case *b* projects. We will see that this is of no advantage for the linearization at PF.

   Suppose we map the asymmetry of projection onto direct precedence at PF. That would produce an unfortunate result: Either we obtain a universal order /Head Comp Spec/ (if projecting elements precede non-projecting elements) or we obtain /Spec Comp Head/ (if the reverse is the case). The first is assumed by

no one; the second is actually proposed by Fukui & Takano (1998), but it is in clear contrast with assumptions by Chomsky himself, and of course Kayne (1994); see also Yasui (2004) for discussion on related issues, and Zwart (2005), who provides additional evidence for a universal /Head Comp/ structure based on a cross-linguistic typology of noun phrase conjunction. I conclude that the asymmetry of projection is not the asymmetry we are looking for. In this respect, recall that it is not certain that {x, {x, y}} can be equated with <x, y>. Also, I am not convinced that Chomsky is right that Merge (x, y) produces {x, {x, y}} to begin with, apart from the issue of asymmetrical dependency discussed above, and the question if projection labels are necessary at all (Collins 2002). The idea that the projection label equals the head is strange from the perspective of compositionality. Logically, the whole cannot equal one of the parts. One might, not unreasonably, object that the label x is not the same as the head x: These are different categories of things. But then of course set reduction of $\{x_{label} \{x_{head}, y\}\}$ to <x, y> is impossible.

In short, I contend that Merge produces asymmetrical pairs that encode syntactic dependency. This asymmetry can be mapped straightforwardly onto direct precedence at the PF interface. If this is so, one might wonder why we still need a linearization procedure involving graph traversal. The reasons are clear-cut. Even though a simple structure <a, <b, c>> transparently yields the string /a b c/, it should be noted that this does not directly follow from the transitivity of precedence, as it is mediated by the inclusion relationship. The next subsection shows that the usually implicit 'Inheritance of Precedence' assumption is untenable as soon as we take *remerge* into account (whether for regular movement or for sharing). Syntactic structures are not simple trees, they are graphs. Furthermore, recall that establishing a (total) order, for example, {{b<c}, {a<b}, {a<c}}, where < means *precedes* — is only halfway the goal of producing an actual word string (see also the second objection against the LCA near the end of section 3.1.).

3.2.2. Dominance/Inclusion and Precedence

Merge (α, β) → γ produces three relationships: γ directly includes α, γ directly includes β, and β is directly dependent on α. At PF, the last relationship can be reinterpreted as α directly precedes β.

Both inclusion and precedence are transitive relations: If αRβ and βRγ then αRγ. Both are irreflexive: αRα is excluded. I assume the notion *dominance* to be synonymous with inclusion. Reflexive dominance/inclusion plays no role in the discussion here, and I will abbreviate *proper dominance* to dominance.

Direct precedence and direct dominance resulting from the same instance of Merge are mutually exclusive: if α directly precedes β, then it cannot be the case that α directly dominates β or *vice versa*. It is often assumed that precedence and dominance *in general* are mutually exclusive. Whether this is really the case, however, depends on further assumptions. Such an assumption is the inheritance of precedence, which can be formulated as in (33):

(33) *Inheritance of Precedence* (to be rejected)

If x directly precedes y, then x and all nodes dominated by x precede y and all nodes dominated by y.

In a regular tree, this makes sense. However, in multidominance graphs, it gives rise to inconsistencies. Indeed, versions of (33) are known as the Non-tangling condition, which is used to prevent crossing branches and so on (see also Gärtner 2002, Carnie 2008, and Fortuny 2008 for further discussion and references). Consider (34), where β has been internally remerged:

(34)
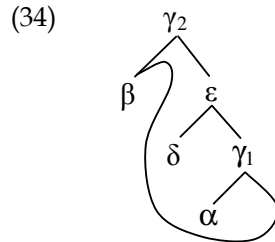
Here, β directly precedes ε; by (33), β would also precede the descendants of ε, namely δ, $γ_1$, α and β. But then β precedes itself, which is odd. Even worse, δ precedes $γ_1$ and, by (33), α and β. Thus, β precedes δ (the descendant of ε), and δ precedes β (the descendant of $γ_1$), which is a contradiction. Moreover, notice that ε and $γ_1$ dominate β *and* are preceded by β at the same time. Given that (34) is a regular movement structure (involving internal remerge), it seems to me that we must simply reject the stipulation in (33). I conclude that the precedence relation is independent of the dominance relation. Therefore, there is also no basis for the mutual exclusion of dominance and precedence in general.

As we have seen in section 2, the absence of stipulative conditions on the input of Merge leads to the possibility of remerge, which in turn leads to multidominance representations. Traditional restrictions on possible tree representations, such as the Single mother condition and the Non-tangling condition, are no longer wanted. It should be clear that it is not the case that anything goes. For instance, Merge makes sure that the resulting graphs are fully linked with respect to (transitive) dominance in the sense that there is a path from every node to every other node (that is, if we allow for a 'change of direction'). Furthermore, we saw in section 2.3. that the PF interface cannot possibly interpret multi-rooted structures (forests); thus, the traditional Single Root condition will be maintained on principled grounds (but notice that this only concerns the end result of the derivation). Also, I argued that the proliferation of roots during the derivation is unwelcome (section 2.3.).

In the previous section, an LCA-based approach to linearization was abandoned on principled and practical grounds. Instead, let us try to develop an alternative in terms of graph traversal, which is inherently a procedure. In the next sections, let our guiding principle be the following: *Every single-rooted structure that can be produced by Merge can be linearized*. In other words, grammatical syntactic structures do not crash when they are linearized at PF.

3.3.   *Tree Traversal*

Before we turn to multidominance graphs, let us have a look at tree traversal, which is less complicated. A standard order-sensitive top–down depth-first

recursive tree traversal procedure yields a string of node contents. First, consider the basic triad in (35), were P is the root, LCh the leftmost child, and RCh the rightmost child.

(35)

P
LCh   RCh
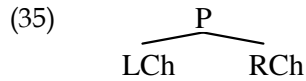
As for the terminology, note that what linguists often call *mother* is also known as *parent* or *(immediate) ancestor*; *daughter* is also known as *child* or *(immediate) descendant*.

In principle, there are three ways of traversing (35): (i) the so-called *preorder* traversal, which lists the parent first, and then the children from left to right; (ii) the *inorder* traversal, which lists the leftmost child first, then the parent, and then the rightmost child; and (iii) the *postorder* traversal, which lists the children before the parent.

Tree traversal is a recursive algorithm, consisting of three basic steps (in a binary tree): Select the leftmost child, select the rightmost child, and perform some action, such as listing the present node content. If a complex child is encountered, interrupt the activity in the present layer and start scanning the child first, returning to this higher layer later (this is called recursive depth-first scanning). The core of this procedure is stated in (36), the three possible positions of undertaking the action are indicated.

(36)   *Scan Triad*

↓   (← list present node: Preorder)
↓ select left child; if complex => start *scan triad* of child
↓   (← list present node: Inorder)
↓ select right child; if complex => start *scan triad* of child
↓   (← list present node: Postorder)

The results of scanning the more complicated abstract tree in (37) are given in (38).

(37)

a
b   c
d   e   f   g
h   i
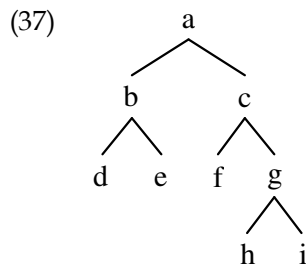
(38)   preorder: /a b d e c f g h i/   ('spell out before going down')
       inorder: /d b e a f c h g i/   ('spell out before going down the second time')
       postorder: /d e b f h i g c a/   ('spell out before going up')

A linguistic linearization requires a list of end nodes (terminals) only. In (39), the strings from (38) are repeated, with the terminals printed in boldface.

(39)   preorder: /a b **d e** c **f** g **h i**/
        inorder: /**d** b **e** a **f** c **h** g **i**/
        postorder: /**d e** b **f h i** g c a/

Interestingly, the required ordering of terminals /d e f h i/ is obtained in *each* of the three cases. Thus, ordering terminal nodes is much less arbitrary than ordering projection nodes: the difference between preorder, inorder, and postorder is irrelevant in practice.[12] What we do is recursively call upon a procedure like *scan triad*; the action of spelling out is restricted to terminals. It seems to me that this is a welcome conclusion.

Just to be concrete, let me describe precisely what happens if we linearize the tree in (37). We start at the root, which is *a*. This node has children; we turn to the preceding one, *b*, first; *b* is complex as well; we turn to child *d*, which does not include members. The content of node *d* is therefore added to some string that is initially empty. We return to *b* and scan the rightmost child *e*, which does not include members; hence it is added to the string. We return to *a* (via *b*) and scan the rightmost child *c*, which is complex; we turn to child *f* and add it to the string. We return to *c* and start scanning the rightmost child *g*, which is complex. We scan child *h* and add it to the string, return to *g*, scan the rightmost child *i* and add it to the string. We return (four times) and end the procedure. The obtained string is /d e f h i/, as required.

In conclusion, traversing a tree in order to produce a string of terminals is straightforward. No puzzling stipulations are necessary.

3.4. Traversing Multidominance Graphs: The Issue of Internal and External Remerge

Let us now turn to the effect of remerge on the linearization. It will become clear that we have to combine traversal with structural conditions. In section 3.1., we saw that a remerged node is encountered twice when traversing the graph at PF, whereas it is only pronounced once. Dealing with regular movement, Frampton (2004) proposed the following structural condition:

(40)   The linearization of x (x = α or β) in $_\alpha\diagup^\gamma\diagdown_\beta$ is omitted if x has a parent outside γ. (to be revised)

Here, *outside* means 'not dominated by'. The effect is that α is pronounced in its highest position. From this perspective, consider the graph in (41):

---

[12]   An interesting alternative is proposed in Yasui (2002, 2004), who defines syntactic structures without projection nodes (e.g., [_will_ it [_be_ raining]] 'it will be raining'). Different ways of scanning such structures produce different word orders. See also Kural (2005) for a proposal in which the difference between preorder, inorder, and postorder traversal for non-terminals is exploited to account for word order variation.

(41)
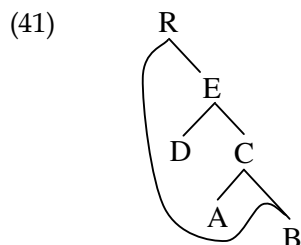
The remerged node is B. Its two parents are C and R. Traversing the graph, we arrive at B directly from R. As B has no other parent outside of R, the linearization is *not* omitted; hence, let us assume that the first occurrence of B is linearized (spelled out). We go on, and spell out the nodes D and A. Then we arrive at B for the second time, now from C. In this case, there *is* another parent outside C, namely R; therefore, the linearization is omitted the second time, as required. Setting aside the possible internal complexity of A, B, and D for a moment, the produced string is /B D A/.

So far, so good. But now consider the case of external remerge (which is not discussed by Frampton). An abstract example that corresponds to an RNR-configuration is given in (42), where the intended string of terminal syntactic objects is /F A Co H D B/:

(42)
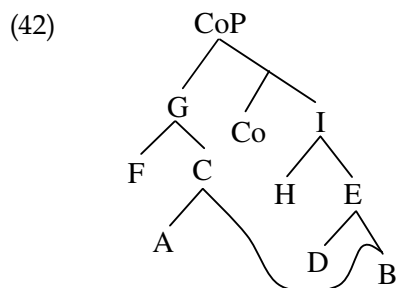
Let us see what happens if we apply the condition in (40). The relevant remerged node is B again, which has two parents, C and E. When we arrive at B from C, we have to check if there is another parent outside C. This is the case: There is another parent, E, which does not dominate C. Therefore, the linearization of B is omitted at this point. Later, when we arrive at B from E, we determine that there is another parent that is not dominated by E, namely C, and again B is not linearized, although it should be. Thus, B will not be spelled out at all.

How can we improve on Frampton's condition? We have to take into account several things; so let us proceed step by step. As a first preliminary, let us change perspective from omitting the linearization of some node to spelling it out. After all, in a linearization procedure (and more generally), we would rather want to know under which conditions a certain action is to be performed than when we have to do nothing. The absence of events is a universal default; actions need to be specified. In language, successive cyclic movement and multiple RNR show that silence rather than pronunciation is the default:

(43) a.   *What* did John say ___ that Mary thought ___ that Bill ___ bought?
      b.   John hates ___, Bill likes ___, Mary admires ___, and Jack detests *the president*.

In principle, only one occurrence of a linguistic object is overtly realized, whereas the number of silent occurrences is unbounded.

Furthermore, during the traversal, we have to keep track of where we came from. Otherwise, we do not know what a potential *other* parent is. Consider (44), where α has been internally remerged twice (its three parents are γ_1, γ_2, and γ_3; its sisters are β_1, β_2, and β_3):

(44)
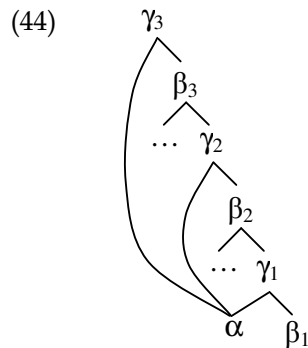
At some point during the traversal the node under consideration is α. How is it determined if α is to be spelled out at this point? That depends on the path from which α is most recently arrived at. Node α must be spelled out in the triad $[_{\gamma_3} \alpha \beta_3]$, but not in $[_{\gamma_2} \alpha \beta_2]$ and $[_{\gamma_1} \alpha \beta_1]$ because there is a parent, γ_3, that is outside (in fact, dominates) γ_2 and γ_1. But how do we know in which triad we are at the moment? Each parent γ_1, γ_2, and γ_3 is equally local to α. Therefore, we need to keep track of the traversal history in some way. To this end, let us define the notion of *current parent*:

(45) *Current Parent*

      The current parent of α is the most recently traversed parent during the linearization procedure.

We can now reformulate (40) as follows:

(46) *Spell-Out of (Internally) Remerged Nodes* (preliminary version, to be revised)

      Linearize an α with more than one parent if the current parent dominates every other parent.

Here, I take *dominance* to be a transitive, non-reflexive relation; an *other parent* is a parent that is not the current parent. Note that the condition in (40) has much in common with the 'Connected ancestor condition' discussed in Barker & Pullum (1990: 22).

We have to be aware that *Spell out/Linearize* α is not always equivalent to

adding α to the string to be pronounced, although it can be. This has to do with the difference between heads and phrases. A phonological string of words or morphemes is a string of heads. So if α is complex, it must be analyzed by further traversal. Of course movement often concerns phrases; therefore, consider (47):

(47)

γ_2

β_2

δ γ_1

α β_1

x y
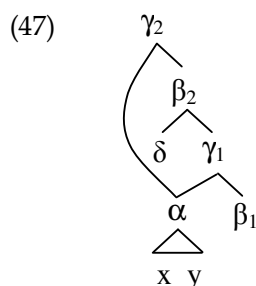
Here, the phrase α containing two heads x and y has been internally remerged; the intended string of terminals is /x y δ β_1/. Without a condition such as (46), x and y are encountered twice during the traversal. This would be problematic, since x and y themselves are relatively *in situ* heads (with only one parent), so their Spell-Out should be uncompromised, which leads to a double realization. According to (46), the linearization of α in (47) is executed only if γ_2 is the current parent. If γ_1 is the current parent (later during the traversal), α will not be spelled out, which implies that its components will not be traversed; hence, x and y are not encountered a second time.

Before we turn to external remerge, let me summarize the assumptions so far:

(A1) The linearization of a syntactic object involves recursive graph traversal.

(A2) Traversal history needs to be monitored.

(A3) For each encountered node, further analysis of its components is conditional: If the number of parents is zero (for the root) or one, further analysis is called upon in any case; if the number of parents is more than one (the consequence of remerge), the configuration between these parents, relative to the current parent, is decisive.

(A4) Further analysis means further traversal if the relevant object is complex. If it is not, the node content is added to the string of words/morphemes.

(A5) Traversal provides a continuously shifting perspective, where the current node and indirectly its current parent are the center of attention. But notice that the interaction with structural conditions implies that the rest of the structure can be inspected at any time.

We are now in a position to analyze (48), which involves external remerge of α. Its parents are γ_1 and γ_2; the intended string of terminals is /δ_1 β_1 δ_2 β_2 α/:

(48)
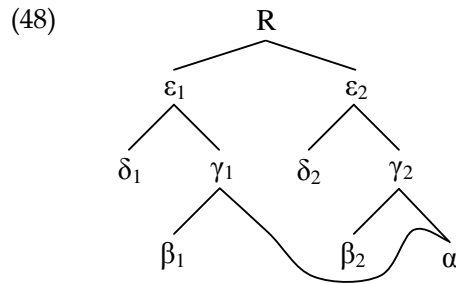
Each of the two parents is 'outside' the other; none dominates the other. The configuration seems symmetrical, but not if the course of the traversal is taken into account. Since we already established before that the traversal history needs to be monitored, I take this to be a possibility. The crucial decision depends on the direct precedence relationship between two of the ancestors of α, namely ε_1 and ε_2, which have been the input for the 'root-uniting' instance of Merge. Parent γ_1 is part of the left branch of the graph, which is traversed before γ_2 in the right branch of the graph. When α is encountered the first time during the traversal, via γ_1, the other parent γ_2 has not yet been traversed; however, when α is encountered the second time, via γ_2, the other parent γ_1 has already been traversed. Therefore, the necessary Spell-Out condition can be formulated as follows:

(49)   *Spell-Out of (Externally) Remerged Nodes* (preliminary version)

Linearize an α with more than one parent if

(i)   every parent has been traversed, *and*

(ii)   the current parent is not dominated by any other parent.

As required, α will be spelled out if γ_2 is the current parent. The second proviso in (49) is necessary to make sure that Spell-Out of the last occurrence is not applied to configurations resulting from *internal* remerge, where one parent dominates the others. Let us now combine the two conditions in (46) and (49):

(50)   *Spell-Out of Remerged Nodes* (to be revised)

An α with more than one parent is linearized if and only if

(i)   the current parent dominates every other parent; *or*

(ii)   every parent has been traversed, *and*
        the current parent is not dominated by any other parent.

No contradictory linearization demands can be imposed on a multidominated node. First, consider some basic possible configurations. In each case, α is the shared node, γ_i is a parent, a subscript c indicates the current parent at a particular stage of the traversal, and the plusses and minuses indicate which parents have been traversed at this stage. For ease of exposition, all other sentence material is omitted.

        The repeated structure in (51) represents movement via an intermediate landing site, which results from applying internal remerge to the same syntactic

object twice:

(51)

spell out α   omit α   omit α
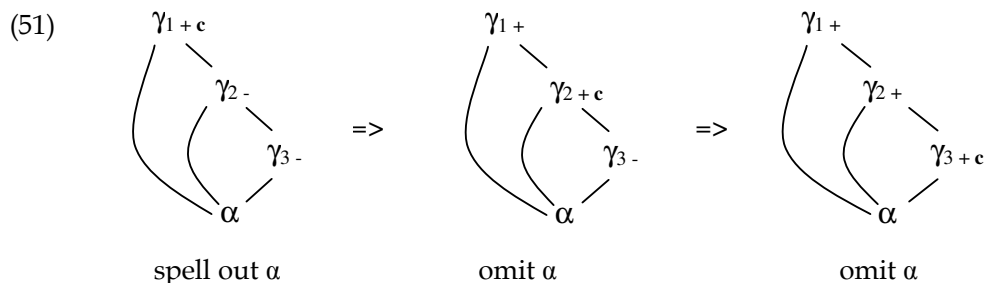
The three structures correspond to different stages where α is reached during the traversal. The decision whether it spelled out at this point is printed below the structure. In the first structure, we arrive at α from the highest parent γ_1. According to (50i), α is spelled out; (50ii) does not apply because the other parents have not been traversed yet. In the second and third structure, α is arrived at from γ_2 and γ_3, respectively, and the linearization of α is to be omitted. Indeed, (50i) no longer applies, since γ_2 and γ_3 are not the highest parent; and (50ii) does not apply because they are dominated by γ_1.

The case of external remerge is sketched in (52). Again, α is remerged twice (a concrete example could be (43b) above). Recall that all parents $\gamma_{1/2/3}$ are locally related to α, which is the center of attention three times during the traversal.

(52)

omit α   omit α   spell out α
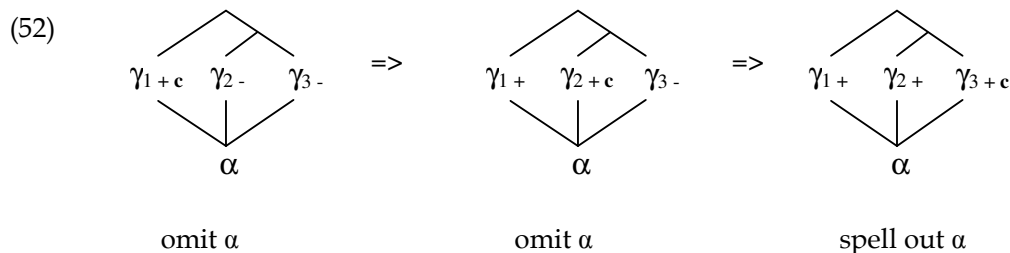
In the first two situations, γ_1 and γ_2 are the respective current parents. They do not dominate all other parents, so (50i) does not apply. Furthermore, not every parent has been scanned, yet, so (50ii) does not apply either, and the linearization of α is omitted. In the third situation, where α is encountered the third time, now via γ_3, every parent has finally been traversed, and α is spelled out.

We also have to check what happens when internal and external remerge are combined. There are two basic possibilities. The first is pictured in (53), where a sharing configuration is embedded in a movement configuration; this results from internal remerge, after root-union, of a constituent that has already been externally remerged. Concretely, this may correspond to ATB-movement (see section 2.2.).

(53)

γ_{1+c} => γ_{1+} => γ_{1+}

γ_{2-} γ_{3-} γ_{2+c} γ_{3-} γ_{2+} γ_{3+c}

α α α

spell out α omit α omit α
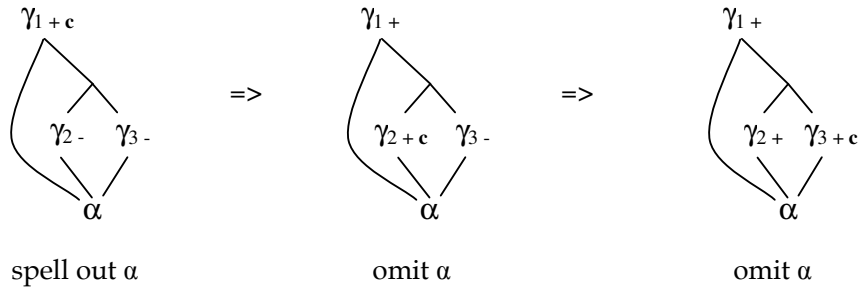
Since there is one parent, γ_1, that includes all other parents of α, the structure will be handled on a par with standard movement configurations as in (51), and α is spelled out in the first position accessed, as required.

The second configuration is pictured in (54). Here, internal remerge is followed by external remerge, or at least root union has to apply after internal remerge. Intuitively, it seems clear that the third occurrence of α needs to be spelled out. This is the position where α is moved to inside the relevant dependent substructure. For symmetry reasons I have added internal remerge in the matrix (or first conjunct) as well (a concrete example may be a 'saddle graft' such as (29), where the object is moved to the middle field.)

(54)

=> => =>

γ_{1+c} γ_{3-} γ_{1+} γ_{3-} γ_{1+} γ_{3+c} γ_{1+} γ_{3+}

γ_{2-} γ_{4-} γ_{2+c} γ_{4-} γ_{2+} γ_{4-} γ_{2+} γ_{4+c}

α α α α

omit α omit α spell out α ?!   omit α
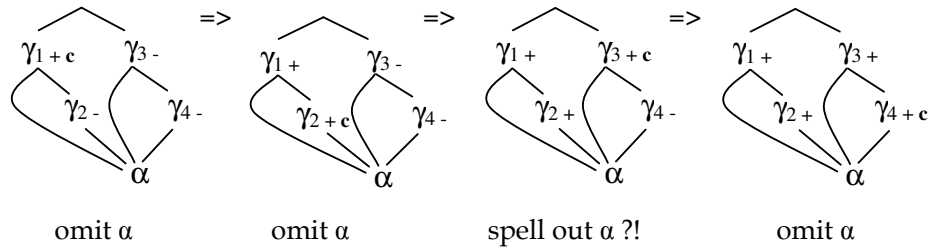
However, according to (50), α will never be spelled out: (50i) never applies since there is no parent that dominates every other parent (note that γ_1 dominates γ_2 but not the other parents; similarly, γ_3 dominates γ_4 but not the other parents), and (50ii) never applies because only when γ_4 is the current parent, all parents have been traversed, but γ_4 is a parent that is dominated by another parent (γ_3). How can this omission be repaired? The answer is that (50i) must be relativized according to the traversal status of the dominated nodes, as is shown in (55). The second proviso in (55i) is necessary to prevent the Spell-Out of α in intermediate landing sites.

(55)   *Spell-Out of Remerged Nodes* (correct, pre-final version)

An α with more than one parent is linearized if and only if

(i)   the current parent dominates every other parent <u>that has not been traversed</u>, *and*
the current parent is not dominated by any other parent; *or*

(ii)   every parent has been traversed, *and*
the current parent is not dominated by any other parent.

In (54), α will now be spelled out if the current parent is γ_3, since this parent dominates all parents not yet traversed, namely γ_4 alone. Crucially, the first occurrence of α, via γ_1, does not lead to Spell-Out, since at this point of the traversal, γ_3 and γ_4 are not yet traversed, and they are not dominated by γ_1.

The second proviso in (55i) equals the second proviso in (55ii); therefore, a more compact formulation is possible:

(56)   *Spell-Out of Remerged Nodes* (final version)

An α with more than one parent is linearized if and only if

(i)   the current parent is not dominated by any other parent, *and*

(ii)   – every parent has been traversed, *or*
        – the current parent dominates every other parent that has not been
          traversed

This concludes the basis of the proposal. The next section discusses some complex cases and potential problems.

3.5.   Complex Structures and Potential Problems

In a number of separate subsections, I will briefly discuss crossing and nesting dependencies, roll-up movement, roll-out movement, remnant movement, RNR without coordination, head movement, and the issue of phases.

3.5.1. Multiple Instances of Internal Remerge

Internal remerge of more than one constituent is possible. Merge itself facilitates both crossing and nesting movement configurations. Let us investigate if these are generally spelled out correctly according to the present linearization proposal. Two relevant structures are depicted in (57a–b). In both cases, α and β have been internally remerged. Graphically, I positioned them where they ought to be spelled out.

(57)   a.
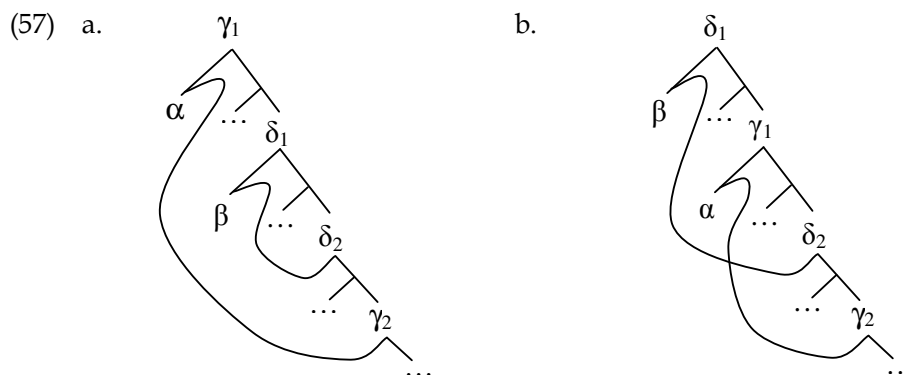
A brief look at (56) will reveal that these structures present no particular problem. Both constituents are linearized the first time they are encountered

because then the current parent dominates the other parent (which is not yet traversed). Notice that α and β are not related by dominance (neither includes the other). In this respect, (57) is no different from (58), where the two instances of remerge do not interfere in any sense:

(58)
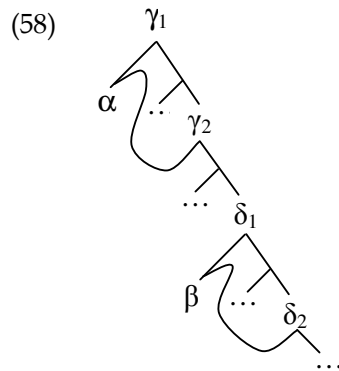
More interesting cases arise if α is included in β. First, consider movement within a moved constituent, as is depicted in two synonymous ways in (59). In a traditional notation, this would correspond to $[_\delta [_\beta \ldots \alpha \ldots t_\alpha \ldots] \ldots t_\beta \ldots]$:

(59)
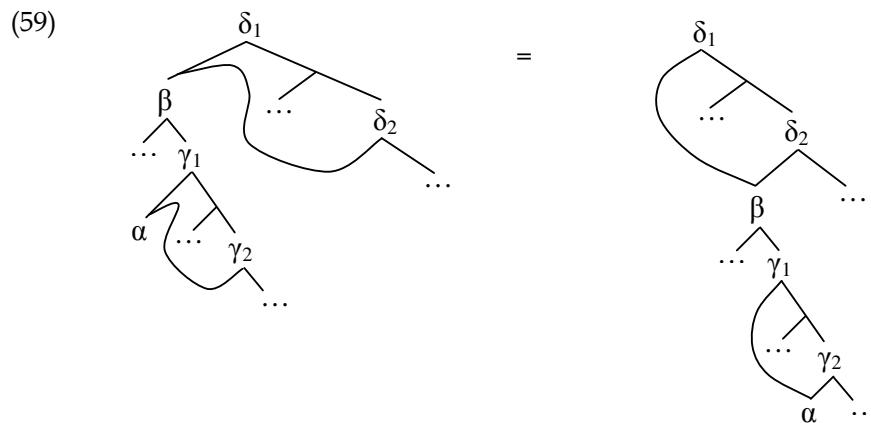
The first occurrence of β is to be linearized. This is performed as in simple movement constructions. Within β, the terminals will be added to the string of words/ morphemes. Of these, only the first occurrence of α is to be spelled out; since the parent γ_1 dominates the other parent γ_2, this is indeed the case. The second time β is encountered, when δ_2 is its current parent, the current parent is dominated by the other parent δ_1, so according to (56i), there will be no linearization of β this time. This implies that none of the contents of β will be traversed again, as required.

A special case of iterative internal remerge of the embedding type is so-called roll-up movement; see Barbiers 1995 and Brody 1997, for instance, who use it to mirror the order of PPs across the verbal right sentence bracket, and the order of adjectives across a head noun. An abstract example in traditional notation is $[_{ZP} [_{YP} [_{XP} X] [_{Y'} Y t_{XP}]] [_{Z'} Z t_{YP}]]$. As a result of these movements, the order of terminals has become the mirror order of how they were first-merged. In

a multidominance graph, the example can be pictured nicely as follows:

(60)
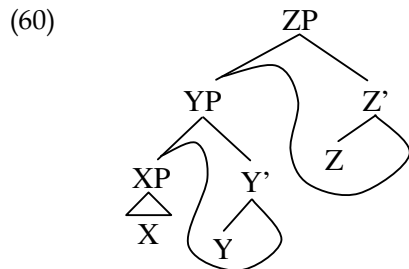
When traversing this graph, we encounter YP, which has two parents, ZP and Z′. As the first parent dominates the second, only the first occurrence of YP will be linearized. Inside YP, the situation is similar for XP.

In a sense, the structural reverse of roll-up movement is roll-out movement: after internal remerge of a complex constituent, a term of this constituent is remerged even higher, and so on. Abstractly, it looks like (61), where the required order of terminals is /X Y Z/, again a reversal of the order in which the heads were first-merged:

(61)
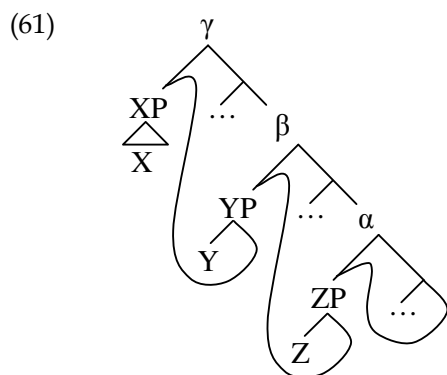
Here, we recognize a violation of the Freezing principle (Wexler & Culicover 1980). However, one should ask whether it is the task of the phonological interface to exclude such constructions. The answer in this article is negative (recall the discussion in sections 3.1. and 3.2.). If there need to be syntactic constraints concerning sub-extraction, so be it, but that is of no direct concern to the linearization procedure in principle. Apart from that, there are many documented exceptions to the Freezing principle, such as Dutch/German *wat voor/was für*-splits, *wh*-movement from a scrambled constituent, or 'smuggling' in English (for discussion, see Corver 1990, Müller 1998, and Collins 2005, among others). An interesting example of what could in fact be a double violation in German is taken from Ott (2009), who discusses a kind of 'multiple NP split' in detail (proposing an alternative solution in terms of scattered deletion, which goes back to Fanselow & Ćavar 2002). The final trace in (62) is not Ott's but mine, in accordance with a general Head–Comp approach.

(62) *German*

Bücher$_i$ wurden [so   richtig gute t$_i$]$_k$ in diesem Jahr nur wenige
books   were   PRT really   good   in this   year only few
t$_k$]$_n$ rezensiert t$_n$.
      reviewed
'As for books, only few really good ones have been reviewed this year.'

Returning to the abstract structure in (61), we note that XP has two parents: γ and
YP. Since YP is dominated by γ, XP (hence X) is spelled out when it is first en-
countered, that is, if γ is the current parent. In turn, YP has two parents, β and ZP,
where β dominates ZP; therefore YP is linearized in the highest position as well.
When traversing YP, we spell out Y, and encounter XP for the second time, but
here XP cannot be linearized because the current parent, YP is included by an-
other parent, namely γ. And so on. I conclude that the linearization of both roll-
up and roll-out movement is correctly performed by the conditioned traversal
proposed in (56).

Finally, let us turn to remnant movement, which is perhaps the most
complicated of all. An abstract illustration in traditional notation is (63), where α
is originally a term of β. A well-known concrete example involves topicalization
of a remnant VP after movement of an object to the middle field (for discussion,
see den Besten & Webelhuth 1990, Koopman & Szabolcsi 2000, and Müller 2001,
for instance).

(63)   (…) [$_β$ … t$_α$ …] … α$_i$ … t$_β$ (…)

A corresponding multidominance structure is (64), where α has parents γ$_1$ and γ$_2$,
and β parents δ$_1$ and δ$_2$. Concretely, α could be an object, and β a verb phrase.

(64)
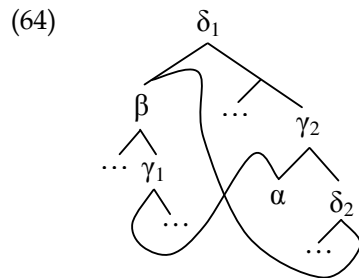
When scanning this structure, we first encounter β via δ$_1$. As in regular move-
ment constructions, β is linearized in the highest position, since δ$_1$ dominates the
other parent δ$_2$. Within β, we encounter α via γ$_1$. Since γ$_1$ does not dominate α's
other parent γ$_2$, and not all parents have been traversed yet, the linearization of α
is omitted here. Later, when α is encountered the second time and γ$_2$ is the cur-
rent parent, the remnant α will be linearized because γ$_2$ is not dominated by the
other parent and all parents have been traversed by that time. In retrospect, what
makes remnant movement special is that an instance of internal remerge (here, of
α) gets the structural appearance of external remerge because of an additional
(internal) remerger of another category (β) that dominates the lower position of α.

As a consequence, only the second occurrence of α will be spelled out.

Needless to say, δ_1 in (64) could be input for further roll-up movements. This poses no particular problem. As a final (theoretical) worst-case scenario, consider iterative remnant movement, which would result from moving γ_2 across δ_1 in (64); this is shown in (65), where γ_2's parents are called ε_1 and ε_2:

(65) $[_{\varepsilon 1} [_{\gamma 2} \alpha_i \ldots t_\beta] \ldots [_{\delta 1} [_\beta \ldots t_\alpha \ldots] [_{\varepsilon 2} \ldots t_{\gamma 2}]]]$
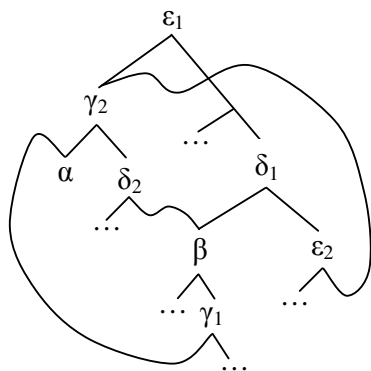
Here, γ_2 will be linearized upon its first encounter (because ε_1 dominates ε_2), and α will be spelled out in the highest position since γ_2 dominates γ_1 (via β). The contents of β are omitted when first encountered via δ_2, and spelled out when δ_1 is the current mother, as required. It is worth noting that there are no actual loops in the multidominance structure in (65), despite its circular appearance at first sight. Therefore, the linearization procedure will have no trouble in ending.

In conclusion, the structures resulting from all possible instances of multiple internal remerge are spelled out correctly, and rather straightforwardly.

3.5.2. Some Issues Concerning External Remerge

The application of external remerge in a derivation may eventually lead to a structural representation that *apparently* involves a violation of the strict cycle. Consider (66), and recall from section 2.3. that such a structure cannot involve *internal* remerge, for the simple reason that Merge functions inherently cyclically (it creates a new root, and does not tamper with the existing structure):

(66)
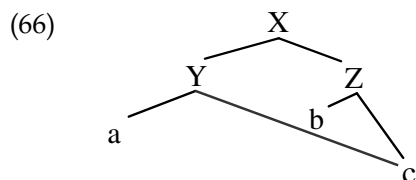
This abstract representation *can* be derived cyclically by first merging a with c, then externally remerging c by applying Merge (b, c), thereby creating a second root node Z, and finally merging the two temporary roots, Y and Z. We could also start with merging b and c, and then remerge c by Merge (a, c).

The derivation of (66) necessarily involves external remerge, and this has

consequences for the linearization: Applying the conditioned linearization in (56) gives the string of terminals /a b c/ (and not /a c b/). Crucially, therefore, it is impossible to analyze c in the representation of (66) as being internally remerged with (moved to) the embedded position within Y, as was discussed in section 2.3. Instead, (66) is a type of sharing construction.

A well-known example of sharing is RNR, but section 2.2. suggested that there are many more possibilities. It does not automatically follow from Merge that structures like (66) are only possible in coordinative constructions (that is, where the node joining the double-rooted substructure is a coordination phrase). Of course, we could stipulate such a limitation as a syntactic constraint, but this is not the most interesting way to go. We already briefly touched upon paren-thetical-like insertions called amalgams. Here, let me present two examples that fit the pattern in (66). The first involves syntactic subordination of a phrase that contains a shared constituent; see (67), in Dutch, where the shared constituent is printed in italics:

(67) *Dutch*

Het kan moeilijk zijn om syntactische ___ van semantische
it   can   difficult   be   to   syntactic   from semantic
factoren te onderscheiden.
factors   to distinguish
'It can be hard to distinguish syntactic ___ from semantic *factors*.'

The prepositional phrase *van semantische factoren* 'from semantic factors' is part of the (extended) predicate. This possibility has been noticed before in Huybregts & van Riemsdijk (1985), among others. Examples like (67) give the impression of RNR at the constituent level (note that the adjectives *syntactic* and *semantic* are contrasted). And in fact, verbs like *distinguish*, *compare*, as well as comparative constructions and comitative constructions, are semantically related to coordi-nation. In each case, two or more items with the same selectional properties are used. Syntactically, however, the prepositional connection is subordinative. This leads to the interesting idea that syntactic subordination/ coordination and semantic subordination/coordination are independent of each other. For further discussion, see Postal (1993), Culicover & Jackendoff (1997), van der Heijden (1999), and Lechner (2001).

Another example that is reminiscent of RNR involves a parenthetical-like insertion hierarchically *above* the phrase that contains the shared part; see (68):

(68) *Dutch*

Joop is, ofschoon een schuldbewuste gebruiker van ___,
Joop is although   a   contrite   user   of
niettemin principieel gekant tegen *de intracontinentale luchtvaart.*
nevertheless principally opposed against the intracontinental aviation
'Joop is, although a contrite user of ___, nevertheless principally opposed to intracontinental aviation.'

Again, there is an implied contrast, and the two prepositions *van* 'of' and *tegen* 'against' must be stressed. The details of the construction need not concern us, here. What is immediately clear is that the structure involves the pattern in (66).

In my analysis, external remerge leads to the pronunciation of the second occurrence of the remerged item — unless it is followed by internal remerge, as in ATB-movement (see sections 2.2. and 3.4.). From this, it can be predicted that instances of *forward* deletion are not sharing constructions involving external remerge, but actual cases of deletion (or ellipsis). Whether this prediction is correct or not must be substantiated by empirical research. What I can say at this point is that indeed the large majority of sharing analyses concern backward deletion constructions, or constructions where there is no overt clue and the analysis could go either way. An important exception, which inspired some others, is Goodall (1987), who also discusses forward gapping from a sharing perspective. However, many differences have been reported between forward and backward deletion in coordination constructions (see Wilder 1997 and de Vries 2005b, for instance), which raises doubts concerning this particular proposal.

Finally, I should add a note on head movement. As the topic is beyond the scope of this article, I will limit myself to a few remarks. A much-discussed problem is that head movement appears to be counter-cyclic (Watanabe 1995). That is, in a structure [$_{YP}$ Y [$_{XP}$… X …]] the head X should not be able to move to Y because Y is embedded in YP. Instead of simply relegating head movement to phonology, many people feel that we should try to find an answer to this problem. One potential solution, first proposed by Bobaljik & Brown (1997), I believe, involves 'sideward movement' — and hence external remerge.[13] The idea is illustrated in (69):

Merge (X, …) → [X …]
Merge (…, [X …]) → XP

   Merge (X, Y) → Y

      Merge (Y, XP) → YP
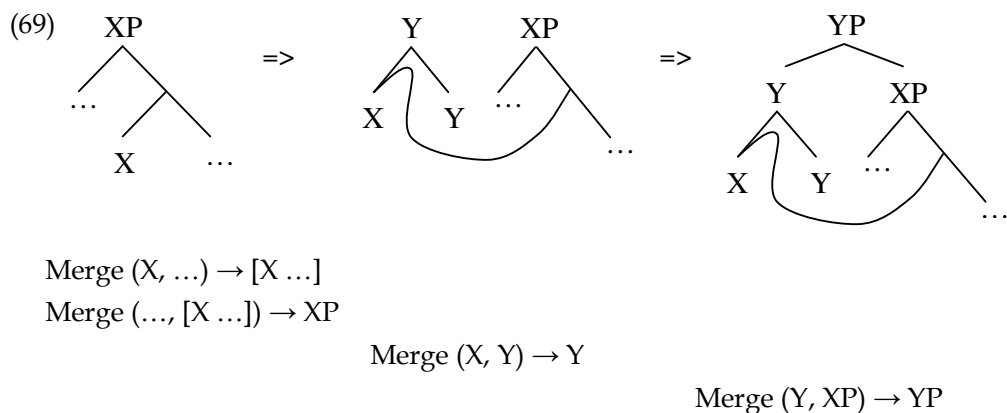
In the crucial step, the head X is externally remerged with an independent item (head adjunction). The combination of both is then merged with XP, such that Y projects.

According to the present linearization proposal, (69) will be treated as a sharing construction. That is, without further assumptions, the second occurrence

---

[13]   Others have argued that head movement does not involve a derived adjunction structure; see, for instance, Koeneman (2000) and Matushansky (2006) for discussion and alternative solutions.

of X (the first-Merge position inside XP) will be pronounced, contrary to fact (at least for 'overt' head movement). Does this mean that we have to abandon the sideward movement approach to head movement? Not necessarily so. Suppose, tentatively, that the final representation in (69) can somehow be transformed into (70) before the actual linearization commences. This would mean that the morphological component is activated at PF. If indeed it recognizes the head incorporation structure created by syntax as a word, it can simplify it into a syntactic unit X+Y (that is, morphological fusion by combining more than one feature bundle under a single category heading). As a consequence, the structure is reinterpreted as a regular movement configuration, with the required linear result.

(70)
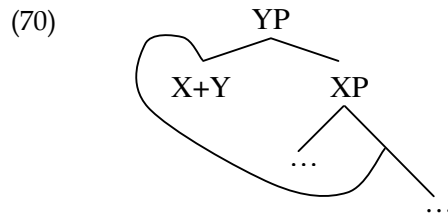
Needless to say, this is a non-trivial operation, whose validity needs to be evaluated carefully. I leave the issue open for further research.

3.5.3. What about Phases?

So far, we have discussed the linearization of completed syntactic derivations. However, one may wonder if this approach is compatible with the idea that syntax operates in cycles or phases (see Chomsky 2005 in particular). To the extent that phases can be equated with domains of syntactic locality, the idea is not very controversial. Chomsky also suggests that phases reduce the computational load. Though this is intuitively not implausible, it remains to be shown how it would work out exactly. Importantly, material inside a previous phase of the derivation is no longer accessible for further syntactic computation. Whenever a phase is completed, the contents are transferred to the interfaces. Crucially, it does not follow from this that the phonological component and the semantic component work in phases themselves. It might very well be that the materials transferred by syntax are accumulated until they are complete, and then further processed. Evidence that the phonological and semantic component work in phases (what is more, the *same* phases as syntax does, and in the same direction) can only come from phonology and semantics, respectively, and not from syntax.[14] Furthermore, linearization is probably not a part of phonology proper, but an interface process.

Bearing in mind, then, that the linearization procedure at PF may have access to the complete structure without contradicting the idea of syntactic

---

14   Interestingly, there are indications that certain phonological processes work in cycles. To which extent these correspond to syntactic domains is currently debated. For some discussion and further references, see Kratzer & Selkirk (2007).

phases, let us nevertheless examine the possibility that the linearization works in similar phases.

Linearization is essentially a top–down procedure, starting at the root *within each cycle*.[15] By contrast, the syntactic derivation is standardly thought of as bottom-up (here, I cannot discuss alternative proposals, but see, among others, Phillips 2003, Chesi 2007, and Zwart, in press). At the end of each syntactic cycle, the structure is handed over to the PF interface. Suppose that this substructure is linearized right away. When the next cycle is entered, the previous cycle becomes opaque (ideally, in every respect). As long as there is no remerge, this seems possible, indeed. In fact, for each step of Merge $(\alpha, \beta) \rightarrow \gamma$ it is the case that *linearization* (γ) = *linearization* (α) + *linearization* (β). Suppose β is a phase, then the outcome of linearization (a string) is already stored, and the structure of β need not be scanned again. If each projection counts as a phase, that is, if phases are as small as possible, then of course the linearization can be said to behave bottom-up for all practical purposes; see Uriagereka (1999) for a fundamental discussion of 'multiple Spell-Out'.[16, 17]

If, however, remerge is at stake, non-trivial problems arise. The basic trouble, as discussed in section 3.2.2., is that we can no longer rely on Inheritance of precedence.

Per definition, internal remerge across a cycle boundary is impossible, unless the designated escape hatch (the edge) of the phase is used. For external remerge we can simply state that it must take place before the cycle is closed (giving rise to a hydraic structure, as discussed). Let us restrict the discussion to internal remerge for the moment. Suppose some object δ has been remerged (that is, 'moved' to the edge). It is normally assumed that the edge of a cycle is not passed on to PF at the interface (for the obvious reason that what is in the edge will be used in the higher cycle). It is the complement of the phase head that is transferred ('spelled out'). But then the question arises how the actual spelling out of δ within the lower cycle can be prevented. Its linearization should be

---

[15] For LCA-based approaches this may not be entirely true, depending on the implementation. However, recall from sections 3.1 and 3.2.1. that establishing a total order is not the final step in a linearization procedure. In order to map a set of asymmetrical (precedence) relationships between pairs of terminals onto an actual word string, this list of relationships needs to be scanned in a way comparable to what top–down graph scanning amounts to when viewed from the perspective of basic relationships (see section 4).

[16] Interestingly, for Uriagereka, in an attempt to adapt the LCA to Bare Phrase Structure, PF linearization crucially makes use of multiple Spell-Out (with minimum-sized phases) in order to cope with complex left branches. This conclusion is almost opposite to the one in this section, for reasons explained immediately below.

[17] Another approach making crucial use of phases is Fox & Pesetsky (2005). They claim that the order of elements in one cycle is preserved in the next. Contrary to how it is often referred to, this theory is not an actual linearization *procedure*, as far as I understand it. The central *condition* states that the relative positions of elements must be preserved across cycles. But how these relative positions are actually determined, and how and when it is decided which occurrence of a remerged item is overtly realized are secondary issues from this perspective. I guess any procedure that does the job would be fine, including, perhaps, an adapted version of the present proposal. Notice, however, that the apparent incompatibility of remerge and PF-cycles discussed in the main text constitutes a serious problem for this. For critical comments concerning the idea of order preservation *per se* and the relevant data involved, see Nilsen (2005), among others.

omitted, as Frampton (2004) calls it (meaning that the lower occurrence is not to be pronounced), but the information that δ has another, dominating parent, which would lead to this decision, is no longer present after transfer, during the linearization procedure targeting the structure generated in the lower cycle. Unfortunately, Frampton does not discuss this issue. According to Chomsky (2007: 16), who describes movement in terms of copies instead of remerge, there does not seem to be a problem at all: "[…] all copies are formed by IM [internal merge] at the phase level, hence identifiable for Transfer". To me, honestly, this is more of a mystery than an explanation of how it would work exactly. Perhaps the phase head, which has the higher copy in its specifier position, is able to discern and mark the lower copy as 'lower copy', information that PF can then use after transfer. Clearly, such a move would not be possible in a remerge account, where a lower and higher occurrence involve one and the same item.

I conclude that the introduction of cycles in PF is highly problematic under a remerge account of displacement (disposing of copies or traces), even for standard internal remerge. Not very surprisingly, the situation further deteriorates if external remerge is taken into account:[18]

(D1) External remerge leads to a temporarily multi-rooted structure.

(D2) External remerge shows apparent non-local characteristics, which complicates the need to recover the structural relationship between the relevant parents of a remerged element.

(D3) An externally remerged element must then be first-merged in the cycle where its Spell-Out is to be omitted, but in a bottom-up derivation the relationship between the higher structural parts is not yet fixed, so it cannot be decided which part is, say, the first conjunct, and which the second.

In short, PF linearization must be exempt from the opacity restrictions imposed by phases. This implies that the full syntactic structure must be made accessible to the linearization procedure after the completion of the full sentence. This conclusion is not an argument against phases in syntax: it just means that an essential procedure on the way to phonology does not operate in the same way.

4.   A Computational Perspective on Linearization

The linearization of syntactic structure is a procedure performed at PF. Let us examine in some detail what it amounts to. The semi-formal algorithm in (71) combines the basic idea of recursive traversal in (36) with the conditions in (56) for remerged elements. I supplemented it with some comments between braces. Needless to say, if the input is a syntactic object, the output will be a word string.

Lines 1–6 are the initialization and ending. Lines 9–13 are the conditions for

---

18   Bachrach & Katzir (2009) claim that their analysis of RNR (which is also in put terms of remerge) involves Spell-Out in cycles. However, they seem to overlook that their notion of complete dominance implies that the linearization procedure is able to 'see' the complete structure (which is necessary if some node Y has parents in different phases).

linearization. Here, line 9 refers to single-merged items; lines 10–13 apply to re-merged nodes. Lines 15–21 involve recursive graph scanning. Terminals are pronounced; non-terminals are subject to further analysis, which means going over the entire procedure again at a lower hierarchical level. The elsewhere case in line 22 implies omitting the linearization of the relevant substructure.

The traversal history can be monitored in a very simple way. Each node that is currently inspected is assigned an index; with each subsequent step in the graph, the index is raised by 1. The effect will be that the current node always has the highest index, and the current parent (if relevant) is always the parent with the highest index. Furthermore, each node that has been traversed has an index greater than zero. Thus, it is easy to see that lines 10–13 are equivalent to the conditions in (56).

(71)   *Linearization Algorithm for Syntactic Graphs Involving Remerged Nodes*

algorithm *linearization*   {performs the conditioned linearization of a binary branching graph
                        with the possibility of multidominance resulting from internal or external remerge}

1   create a new, empty string
2   create a numeral index i with an initial value of 0
3   assign an indexical value of zero to each node   {corresponding to "not yet scanned"}
4   select the root;
5   *scangraph*   {start the scanning procedure with the root}
6   end

procedure *scangraph*   {conditioned scanning procedure based on recursive traversal}

7   add 1 to the numeral index
8   mark the present object with the numeral index i   {keep track of which node
                                                                is scanned when}

9   **if**   (the present object has ≤ 1 parents)   {single parent or no parent (for the root)}
10   OR ((there is no parent that dominates the parent with the highest index)
11   AND ((every parent has an index i ≥ 1)   {external remerge: last occurrence}
12   OR (the parent with the highest index dominates
13   every parent with a zero index)))   {internal remerge: arrival
                                                                from hierarchically highest parent}
14   **then**   {depth first recursive traversal; add to the string if a terminal is reached}
15   **if** the present object directly includes members   {is it a non-terminal?}
16   **then**   {traverse both daughters in order of precedence}
17   select the preceding member
18   *scangraph*   {recursive step}
19   select the other member
20   *scangraph*   {recursive step}
21   **else** add the present object to the string   {"spell out terminal"}
22   **else** — (do nothing)   {"omit linearization"}
23 return

The algorithm refers to syntactic structure as if it involves a graph. This is a metaphor, since a graph is only a *representation* of an underlying array of basic syntactic relationships brought about by a series of mergers. Merge creates direct inclusion relationships, and dependency relations between sisters (which is directly mapped onto precedence at PF). We are used to the (ordered) set, bracketed structures, and graph notations, but that implies the following convention, where left–right order, top–down lines, and subscripts are interpreted in a specific way:
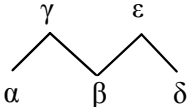
(72) *Basic syntactic triad convention*

$<_\gamma \alpha, \beta>$ $=_{def}$

$[_\gamma \alpha \beta]$ $=_{def}$ (γ directly includes α) $\land$ (γ directly includes β)
                                                        $\land$ (β is a direct dependent of α /
                                                             α directly precedes β)

$$\gamma$$
$$\diagup\diagdown$$
$$\alpha \quad \beta$$ $=_{def}$

A sequence of mergers produces a list of triads, and hence a list of basic relationships. A full graph or complex set can rather straightforwardly be composed on the basis of this list. Nevertheless, if we want to estimate the computational cost of the algorithm in (71), we have to examine what it does in terms of such a list of local relationships.

In formalized grammars it is often the case that a designated node is defined as the root node. In principle, this is completely unnecessary; because it can be derived which node is the root node by going over the complete list of basic relations (just once). For instance, the list in (73b), which results from the mergers in (73a), can only be interpreted such that ε is the root, as it is the only node that is not included by any other node.

(73) a.   Merge $(\alpha, \beta) \rightarrow \gamma$
            Merge $(\delta, \gamma) \rightarrow \varepsilon$
      b. γ directly includes α, γ directly includes β, α directly precedes β,
            ε directly includes δ, ε directly includes γ, δ directly precedes γ

If we find more than one node that satisfies this criterion, the list is not linearizable. For example, the list in (74b), which corresponds to the hydraic structure in (74c), produces two possible roots: γ and ε.

(74) a.   Merge $(\alpha, \beta) \rightarrow \gamma$
            Merge $(\beta, \delta) \rightarrow \varepsilon$
      b. γ directly includes α, γ directly includes β, α directly precedes β,
            ε directly includes β, ε directly includes δ, β directly precedes δ
      c.

$$\gamma \qquad \varepsilon$$
$$\diagup\diagdown\diagdown\diagup\diagdown$$
$$\alpha \qquad \beta \qquad \delta$$

Thus, (74) would crash at the PF interface, contrary to, say, a completed RNR- or ATB-movement construction, in which a uniting instance of Merge has taken place.

When there is a detectable root, the linearization procedure *scangraph* in (71) can be started. In line 9 it is checked if the present object (PO) has one parent or less. This amounts to going over the list and find all items of the type *p directly includes PO*. The parents are temporarily stored in a set $\{p_1, \ldots, p_n\}$, whose cardinality can then be determined.

Line 10 checks if there is no parent that dominates the current parent (which is the parent with the highest index). This is the most costly condition because of the transitive character of dominance. A straightforward way to proceed is as follows. For each parent p_i separately, go over the list and check for items of the form p_i *directly includes* $x_{i1/2}$. Notice that the binary character of Merge makes sure that there will always be two such statements, if applicable. Temporarily store every $x_{i1/2}$ in a set. If this set is not empty, go over the list again and check for items of the form $x_{i1/2}$ *directly includes* y_k. Temporarily store every y_k in a set. If this set is not empty, go over the list again, etc. Finally, check if the current parent is present in the unified set of dominated items $\{x_1, \ldots, x_{n(x)}, y_1, \ldots, y_{n(y)}, \ldots\}$. If this is *not* the case, the condition is fulfilled.

Line 11 checks if all parents have been traversed, which simply involves checking the index of each p_i.

Lines 12–13 check if the current parent dominates every untraversed parent. Thus, we go over the list and check for items of the form $p_{current}$ *directly includes* x_i, and temporarily store every x_i. As before, we go over the list again, now checking for x_i *directly includes* y_k, etc. The set of dominated items can be combined into a unified set. Now, if the subset of parents with a zero index is also a subset of the combined set of items dominated by $p_{current}$, then the condition is fulfilled.

Line 15 checks if the present object is a terminal. This condition is fulfilled if the list of relations contains the items *PO directly includes* x_i. In order to select the preceding or the other child (lines 17 and 19) it is necessary to go over the list again and find the item x_1 *precedes* x_2 or the other way around.

Let us try to estimate the number of steps required to linearize a graph relative to the number of mergers involved in deriving the structure. Note that the number of nodes in a single-rooted graph is directly proportional to the number of mergers (to be precise, two times the number of first-time mergers plus one time the number of remergers plus one), which in turn is one third of the cardinality of the list of basic relations.

After finding the root, the procedure *scangraph* is passed through for each node. In doing so, it is necessary to go over the list of basic relationships a few times (lines 9 and 15–21) for one-parented nodes. Furthermore, because of the possibility of remerged nodes, we must also check the conditions in lines 10–13, which involves going over the list multiple times, thereby looping over the graph depth. Thus, we obtain the formula in (75), which shows polynomial growth:

(75) $CS = c_0 + c_1 \cdot M + M \cdot (c_2 \cdot M + c_3 \cdot M^2) = c_0 + c_1 \cdot M + c_2 \cdot M^2 + c_3 \cdot M^3$

Here, CS is the number of computational steps, M is the number of mergers involved in the derivation, and c_0 through c_3 are constants (depending on units of computation, etc.). The term $c_1 \cdot M$ involves finding the root; $c_2 \cdot M^2$ is essentially for scanning the entire graph and finding the terminals; $c_3 \cdot M^3$ is the toll for allowing remerge/multidominance.

In fact, the estimate above is a worst case scenario, in which all the information that is necessary for the linearization is calculated at the PF interface. It is also conceivable that each node α carries information such as a complete set of nodes included by α. A set like this can be established by Merge during the derivation, simply by adding the 'inclusion sets' of the two relevant input nodes. With this information available, (75) would reduce to linear growth, if I am not mistaken: $CS = c_0 + c_1 \cdot M$.

5.   Conclusion

Complex syntactic objects are derived by the operation called Merge. The input for Merge is free in the sense that input objects can be selected from the lexicon (or numeration), from the syntactic workspace (including syntactic objects that are the output of earlier instances of Merge), and from within complex syntactic objects in the workspace. The last option leads to the effect that an item can be merged more than once, that is, *remerged*. I showed that remerge can be *internal* as well as *external*. Internal remerge, which corresponds to regular movement, is commonly accepted, but external remerge is not. However, both possibilities simply follow from the core system; both involve the basic operation Merge. If one or both are to be excluded, this has to be stipulated explicitly. I argued against such stipulations, not only because minimalist guidelines urge us to, but also since there are sensible interpretations of structures involving external remerge, including several types of amalgams, RNR, and ATB-movement. The sheer variety of relevant construction types can be taken to be a warning to refrain from hastily building in additional constraints on Merge. Furthermore, we must make sure that the linearization procedure at the phonological interface can handle such structures.

Both internal and external remerge can be represented in terms of multidominance. The difference between 'sharing' and 'interarboreal movement' is an artifact of the notation. The unification of internal and external remerge makes it particularly clear that there is no need for a copying mechanism, for traces or for (movement) chains. Thus, the syntactic apparatus can be kept to a minimum. This is not to say that there are no differences between constructions involving internal remerge and those involving external remerge. The structural *effect* of remerging with the dominating root or remerging with an external object is quite different. The parents of an internally remerged syntactic object are in a dominance relation, whereas this is not the case for the parents of an externally remerged syntactic object. In the last case, apparent non-local behavior may show up, which can be explained as the possible consequence of a structural 'bypass' between complex substructures that are united at the top. Furthermore, there is no inherent directionality in external remerge. The structural result of first-

merging α with β and remerging it with γ is the same as first-merging α with γ and remerging it with β. Both derivations yield the same configuration with the same structural relationships, and both should be spelled out in the same way. For internal remerge the situation is different: Remerging α with the root is fine, but remerging α with a hierarchically lower position is inherently impossible, since Merge operates strictly cyclically by its very nature, and does not allow 'tampering' of existing relationships.

External remerge creates a temporary multi-rooted structure; this need not be a problem, provided that it is compensated by a uniting instance of Merge at a later stage of the derivation. There are several reasons for such a union; one is that the structure must eventually be linearized, and PF can only interpret a single-rooted object. Thus, even though remerge may lead to unconventional structures, the possibilities are not unconstrained. I argued that there is an additional limitation, dubbed the No proliferation of roots condition, which says that Merge may not create an additional root, which would be counterproductive from a functional perspective. This does not generally exclude external remerge, since the external structure simply keeps having a root of its own, but is does exclude the possibility of remerging an item with another item that is embedded (internally or externally), which would result in undesirable structures. As a consequence, it must always be the case that one of the input elements for Merge is a root at the relevant stage of the derivation.

Evidently, the possibility of remerge complicates the linearization of completed syntactic objects, which must be performed at the PF interface. It is not only the case that remerged items are normally only pronounced once, there also seem to be contradictory linearization demands for internal and external remerge: Informally put, movement is to the left, sharing is to the right. These issues are insufficiently discussed in the literature. All the linearization proposals I am aware of can adequately deal with only a subset of the relevant data. I also argued against LCA-based proposals on principled and practical grounds. Moreover, I showed that the linearization procedure is not likely to operate in cycles. Therefore, I proposed an alternative solution in the form of a conditioned linearization algorithm, which makes use of the different structural configurations created by the two types of remerge. First, it was presented as a graph traversal procedure, in combination with relative structural conditions. I showed that it can handle all kinds of intricate structural patterns involving multiple instances of remerge, including roll-up movement, combinations of sharing and movement, and iterative remnant movement. Subsequently, I showed what this procedure amounts to in terms of a list of basic relationships brought about by a series of mergers, and I estimated the computational load of such a process.

References

Aoun, Joseph & Audrey Li. 2003. *Essays on the Representational and Derivational Nature of Grammar: The Diversity of Wh-Constructions* (Linguistic Inquiry Monographs 40). Cambridge, MA: MIT Press.

Armstrong, David M. 1986. In defense of structural universals. *Australasian Journal of Philosophy* 64, 85–88.

Bachrach, Asaf & Roni Katzir. 2009. Right-node raising and delayed spellout. In Kleanthes K. Grohmann (ed.), *InterPhases: Phase-Theoretic Investigations of Linguistic Interfaces* (Oxford Studies in Theoretical Linguistics 21). Oxford: Oxford University Press.

Barbiers, Sjef. 1995. The syntax of interpretation. Leiden: Universiteit Leiden dissertation.

Barbiers, Sjef, Olaf Koeneman & Marika Lekakou, 2008. Syntactic doubling and the structure of chains. *West Coast Conference on Formal Linguistics (WCCFL)* 26, 77–86.

Barker, Chris & Geoffrey Pullum. 1990. A theory of command relations. *Linguistics and Philosophy* 13, 1–34.

den Besten, Hans & Gert Webelhuth. 1990. Stranding. In Günter Grewendorf & Wolfgang Sternefeld (eds.), *Scrambling and Barriers* (Linguistik Aktuell/ Linguistics Today 5), 77-92. Amsterdam: John Benjamins.

Blevins, James. 1990. Syntactic complexity: Evidence for discontinuity and multidomination. Amherst, MA: University of Massachusetts dissertation.

Bobaljik, Jonathan David. 1995. In terms of merge: Copy and head movement. *MIT Working Papers in Linguistics* 27, 41–64.

Bobaljik, Jonathan David & Samuel Brown. 1997. Interarboreal operations: Head movement and the extension requirement. *Linguistic Inquiry* 28, 345–356.

Brody, Michael. 1997. Mirror theory. *UCL Working Papers in Linguistics* 9, 179-222.

Castillo, Juan Carlos, John Drury & Kleanthes K. Grohmann. In press. Merge over move and the extended projection principle: MOM and the EPP revisited. *Iberia* 1. [Revised and expanded version of Castillo, Juan Carlos, John Drury & Kleanthes K. Grohmann. 1999. Merge over move and the extended projection principle. *University of Maryland Working Papers in Linguistics* 8, 63–103.]

Carnie, Andrew. 2008. *Constituent Structure* (Oxford Surveys in Syntax & Morphology). Oxford: Oxford University Press.

Carstens, Vicky. 2003. Rethinking complementizer agreement: Agree with a Case-checked goal. *Linguistic Inquiry* 34, 393–412.

Chen–Main, Joan. 2006. On the generation and linearization of multi-dominance structures. Baltimore, MD: John Hopkins University dissertation.

Chesi, Cristiano. 2007. An introduction to phase-based minimalist grammars: Why move is top-down and from left-to-right. *CISCL Working Papers on Language and Cognition* 1, 38–75.

Chomsky, Noam. 1981. *Lectures on Government and Binding: The Pisa Lectures* (Studies in Generative Grammar 9). Dordrecht: Foris.

Chomsky, Noam. 1995. *The Minimalist Program* (Current Studies in Linguistics 28). Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik*, 89-155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001a. Beyond explanatory adequacy. *MIT Occasional Papers in Linguistics* 20.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A Life in Language* (Current Studies in Linguistics 36), 1-52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2004. Beyond explanatory adequacy. In Adriana Belletti (ed.), *Structures and Beyond: The Cartography of Syntactic Structures*, vol. 3, 104–131. Oxford: Oxford University Press.

Chomsky, Noam. 2005. On phases. Ms., Massachusetts Institute of Technology, Cambridge.

Chomsky, Noam. 2007. Approaching UG from below. In Uli Sauerland & Hans-Martin Gärtner (eds.), *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax–Semantics*, 1–29. Berlin: Mouton de Gruyter.

Chung, Dae-Ho. 2004. Multiple dominance analysis of right node raising constructions. *Language Research* 40, 791–812.

Citko, Barbara. 2005. On the nature of merge: External merge, internal merge, and parallel merge. *Linguistic Inquiry* 36, 475–496.

Collins, Chris. 2002. Eliminating labels. In Samuel David Epstein & T. Daniel Seely (eds.), *Derivation and Explanation in the Minimalist Program* (Generative Syntax 6), 42–64. Malden, MA: Blackwell.

Collins, Chris. 2005. A smuggling approach to the passive in English. *Syntax* 8, 81–120.

Cormen, Thomas, Charles Leiserson & Ronald Rivest 1990. *Introduction to Algorithms*. Cambridge, MA: MIT Press.

Corver, Norbert. 1990. The syntax of left branch extractions. Tilburg: Katholieke Universiteit Brabant dissertation.

Culicover, Peter & Ray Jackendoff. 1997. Semantic subordination despite syntactic coordination. *Linguistic Inquiry* 28, 195–217.

Di Sciullo, Anna Maria. 2000. Asymmetries: Consequences for morphological configurations and paradigms. *Acta Linguistica Hungarica* 47, 81–101.

Di Sciullo, Anna Maria & Daniela Isac. 2008. The asymmetry of merge. *Biolinguistics* 2, 260–290.

Engdahl, Elisabet. 1986. *Constituent Questions: The Syntax and Semantics of Questions with Special Reference to Swedish* (Studies in Linguistics and Philosophy 27). Dordrecht: Reidel.

Epstein, Samuel David. 1999. Un-principled syntax and the derivation of syntactic relations. In Samuel David Epstein & Norbert Hornstein (eds.), *Working Minimalism*, 317–345. Cambridge, MA: MIT Press.

Epstein, Samuel David, Erich M. Groat, Ruriko Kawashima & Hisatsugu Kitahara. 1998. *A Derivational Approach to Syntactic Relations*. New York: Oxford University Press.

Fanselow, Gisbert & Damir Ćavar. 2002. Distributed deletion. In Artemis Alexiadou (ed.), *Theoretical Approaches to Universals* (Linguistik Aktuell/Linguistics Today 49), 66–107. Amsterdam: John Benjamins.

Fiengo, Robert. 1974. Semantic conditions on surface structure. Cambridge, MA: MIT dissertation.

Forrest, Peter. 1986. Neither magic nor mereology: A reply to Lewis. *Australasian Journal of Philosophy* 64, 89–91.

Fortuny, Jordi. 2008. *The Emergence of Order in Syntax* (Linguistik Aktuell/Lingu-

istics Today 119). Amsterdam: John Benjamins.

Fox, Danny & David Pesetsky. 2005. Cyclic linearization of syntactic structure. *Theoretical Linguistics* 31, 1–45.

Frampton, John. 2004. Copies, traces, occurrences, and all that: Evidence from Bulgarian multiple wh-phenomena. Ms., Northeastern University, Boston.

Fukui, Naoki & Yuji Takano. 1998. Symmetry in syntax: Merge and demerge. *Journal of East Asian Linguistics* 7, 27–86.

Gärtner, Hans-Martin. 2002. *Generalized Transformations and Beyond: Reflections on Minimalist Syntax*. Berlin: Akademie Verlag.

Goodall, Grant. 1987. *Parallel Structures in Syntax: Coordination, Causatives and Restructuring* (Cambridge Studies in Linguistics 46). Cambridge: Cambridge University Press.

Goodman, Nicolas. 1986. Intensions, Church's thesis, and the formalization of mathematics. *Notre Dame Journal of Formal Logic* 28, 473–489.

Gracanin–Yeksek, Martina. 2007. About sharing. Cambridge, MA: MIT dissertation.

Gracanin–Yeksek, Martina. To appear. In Theresa Biberauer & Ian Roberts (eds.), Linearizing Multidominance Structures. *Challenges to Linearization*. Berlin: Mouton de Gruyter.

Grohmann, Kleanthes K. 2003. *Prolific Domains: On the Anti-Locality of Movement Dependencies* (Linguistik Aktuell/Linguistics Today 66). Amsterdam: John Benjamins.

Grootveld, Marjan. 1994. Parsing coordination generatively. Leiden: Universiteit Leiden dissertation.

Grosu, Alexander. 2006. An amalgam and its puzzles. In Hans-Martin Gärtner, Sigrid Beck, Regine Eckhardt, Renate Musan & Barbara Stiebels (eds.), *Between 40 and 60 Puzzles for Krifka*. [http://www.zas.gwz-berlin.de/fileadmin/material/40-60-puzzles-for-krifka/index.html.]

Guimarães, Maximiliano. 2004. Derivation and representation of syntactic amalgams. College Park, MD: University of Maryland dissertation.

Ha, Seungwan. 2008a. On ellipsis features and right node raising. *Conference of the Student Organisation of Linguistics in Europe (ConSOLE)* XV, 67–90.

Ha, Seungwan. 2008b. Ellipsis, right node raising, and across-the-board constructions. Boston, MA: Boston University dissertation.

Haider, Hubert. 2000. OV is more basic than VO. In Peter Svenonius (ed.), *The derivation of VO and OV* (Linguistik Aktuell/Linguistics Today 31), 45–67. Amsterdam: John Benjamins.

Hartmann, Katharina. 2000. *Right Node Raising and Gapping: Interface Conditions on Prosodic Deletion*. Amsterdam: John Benjamins.

Heijden, Emmeke van der. 1999. Tussen nevenschikking en onderschikking. Nijmegen: Katholieke Universiteit Nijmegen dissertation.

Henderson, Brent. 2007. Matching and raising unified. *Lingua* 117, 202–220.

Heringa, Herman. 2009. A multidominance approach to appositional constructions. Ms., Rijksuniversiteit Groningen.

Heringa, Herman. In progress. Appositional constructions. Groningen: Rijksuniversiteit Groningen dissertation.

Huck, Geoffrey J. & Almerindo E. Ojeda. 1987. Introduction. In Geoffrey J. Huck

& Almerindo E. Ojeda (eds.), *Discontinuous Constituency* (Syntax and Semantics 20), 1–9. New York: Academic Press.

Huybregts, Rini & Henk van Riemsdijk. 1985. Parasitic gaps and ATB. *North East Linguistic Society (NELS)* 15, 168–187.

Jaspers, Dany. 1998. Categories and recursion. *Interface* 12, 81–112.

Johnson, Kyle. 2007. LCA + alignment = RNR. Paper presented at the workshop on *Coordination, Subordination and Ellipsis*, Tübingen. [Eberhard-Karls-Universität Tübingen, 7–8 June 2007.]

Johnson, Kyle. 2009. Why movement? Ms., University of Massachusetts, Amherst.

Karlgren, Hans. 1976. Why trees in syntax? *Studia Linguistica* 30, 1–33.

Karttunen, Lauri & Martin Kay. 1985. Structure sharing with binary trees. *ACL Proceedings* 23, 133–136. [Reprinted in Stuart Shieber, Lauri Karttunen & Fernando C.N. Pereira (eds.), *A Compilation of Papers on Unification-based Grammar Formalisms*, Part II, 5–16. Report no. CSLI-86-48, Center for the Study of Language and Information, Stanford, CA.]

Kayne, Richard S. 1984. *Connectedness and Binary Branching*. Dordrecht: Foris.

Kayne, Richard S. 1994. *The Antisymmetry of Syntax* (Linguistic Inquiry Monographs 25). Cambridge, MA: MIT Press.

Kluck, Marlies. 2007. The perspective of external remerge on right node raising. *Proceedings of CamLing 2007*, 130–137.

Kluck, Marlies. 2008. Intertwined clauses, interacting propositions: A note on the interpretive aspects of sentence amalgamation. *Conference of the Student Organisation of Linguistics in Europe (ConSOLE)* XVI, 77–101.

Kluck, Marlies. 2009. Good neighbors or far friends: Matching and proximity effects in Dutch right node raising. *Groninger Arbeiten zur germanistischen Linguistik* 48, 115–158.

Kluck, Marlies. In progress. The structure and meaning of amalgams. Groningen: Rijksuniversiteit Groningen dissertation.

Kluck, Marlies & Mark de Vries. To appear. Cumulative rightward processes. In Heike Walker, Manfred Sailer & Gert Webelhuth (eds.), *Rightward Movement from a Cross-Linguistic Perspective*. Amsterdam: John Benjamins.

Koeneman, Olaf. 2000. The flexible nature of verb movement. Utrecht: Universiteit Utrecht dissertation.

Koopman, Hilda & Anna Szabolcsi. 2000. *Verbal Complexes* (Current Studies in Linguistics 34). Cambridge, MA: MIT Press.

Koster, Jan. 1999. De primaire structuur. *Tabu* 29, 131–140.

Koster, Jan. 2003. All languages are tense second. In Jan Koster & Henk van Riemsdijk (eds.), *Germania et Alia: A Linguistic Webschrift for Hans den Besten*. [http://www.let.rug.nl/koster/DenBesten/contents.htm.]

Koster, Jan. 2007. Structure preservingness, internal merge, and the strict locality of triads. In Simin Karimi, Vida Samiian & Wendy K. Wilkins (eds.), *Phrasal and Clausal Architecture: Syntactic Derivation and Interpretation* (Linguistik Aktuell/Linguistics Today 101), 188–205. Amsterdam: John Benjamins.

Kratzer, Angelika & Elisabeth Selkirk. 2007. Phase theory and prosodic spellout: The case of verbs. *The Linguistic Review* 24, 93–135.

Kremers, Joost. 2009. Recursive linearisation. *The Linguistic Review* 26, 135–166.

Kural, Murat. 2005. Tree traversal and word order. *Linguistic Inquiry* 36, 367–387.

Kuratowski, Kazimierz. 1921. Sur la notion de l'ordre dans la théorie des ensembles. *Fundamenta Mathematicae* 2, 161–171. [Reprinted in Kuratowski, Kazimierz. 1988. *Selected Papers*. Warszawa: Polish Scientific Publishers.]

Lakoff, George. 1974. Syntactic amalgams. *Proceedings of the Chicago Linguistic Society (CLS)* 10, 321–344.

Langendoen, Terence. 2003. Merge. In Andrew Carnie, Heidi Harley & MaryAnn Willie (eds.), *Formal Approaches to Function in Grammar: In Honor of Eloise Jelinek* (Linguistik Aktuell/Linguistics Today 62), 307–318. Amsterdam: John Benjamins.

Lechner, Winfried. 2001. Reduced and phrasal comparatives. *Natural Language and Linguistic Theory* 19, 683–735.

Matushansky, Ora. 2006. Head movement in linguistic theory. *Linguistic Inquiry* 37, 69–109.

Matushansky, Ora. 2008. A case study of predication. In Franc Marušič & Rok Žaucer (eds.), *Studies in Formal Slavic Linguistics: Contributions from Formal Description of Slavic Languages 6.5*, 213–239. Frankfurt am Main: Peter Lang.

Mayr, Clemens & Viola Schmitt. 2009. Order and the coordinate structure constraint. Ms., Harvard University, Cambridge, MA & Universität Wien.

McCawley James. 1987. Some additional evidence for discontinuity. In Geoffrey J. Huck & Almerindo E. Ojeda (eds.), *Discontinuous Constituency* (Syntax and Semantics 20), 185–200. New York: Academic Press.

McCawley, James. 1968. Concerning the base component of a transformational grammar. *Foundations of Language* 4, 243–269.

McCawley, James. 1982. Parentheticals and discontinuous constituent structure. *Linguistic Inquiry* 13, 91–106.

Meinunger, André. 2008. Complex numerals as grafts. Paper presented at the *Third Brussels Conference on Generative Linguistics (BCGL 3)*, Brussels. [Hogeschool-Universiteit Brussel, 21–23 May 2008.]

Moltmann, Friederike. 1992. Coordination and comparatives. Cambridge, MA, MIT dissertation.

Mu'adz, Husni. 1991. Coordinate structures: A planar representation. Tucson, AZ: University of Arizona dissertation.

Müller, Gereon. 1998. *Incomplete Category Fronting: A Derivational Approach to Remnant Movement in German* (Studies in Natural Language and Linguistic Theory 42). Dordrecht: Kluwer.

Müller, Gereon. 2001. Two types of remnant movement. In Artemis Alexiadou, Elena Anagnostopoulou, Sjef Barbiers & Hans-Martin Gärtner (eds.), *Dimensions of Movement: From Features to Remnants* (Linguistik Aktuell/ Linguistics Today 48), 209–241. Amsterdam: John Benjamins.

Neijt, Anneke. 1979. *Gapping: A Contribution to Sentence Grammar* (Studies in Generative Grammar 7). Dordrecht: Foris.

Nilsen, Øystein. 2005. Some notes on cyclic linearization. *Theoretical Linguistics* 31, 173–183.

Nunes, Jairo. 2001. Sideward movement. *Linguistic Inquiry* 32, 303–344.

Nunes, Jairo. 2004. *Linearization of Chains and Sideward Movement* (Linguistic Inquiry Monograph 43). Cambridge, MA: MIT Press.

Ojeda, Almerindo. 1987. Discontinuity, multidominance, and unbounded depen-
dency in generalized phrase structure grammar: Some preliminaries. In
Geoffrey J. Huck & Almerindo E. Ojeda (eds.), *Discontinuous Constituency*
(Syntax and Semantics 20), 257–282. New York: Academic Press.

Ott, Dennis. 2009. Multiple NP split: A distributed deletion analysis. *Groninger Ar-
beiten zur germanistischen Linguistik* 48, 65–80.

Phillips, Colin. 2003. Linear order and constituency. *Linguistic Inquiry* 34, 37–90.

Postal, Paul M. 1993. Parasitic gaps and the across-the-board phenomenon.
Lingu-istic Inquiry 24, 735–754.

Postal, Paul. 1998. *Three Investigations of Extraction* (Current Studies in Linguistics
29). Cambridge, MA: MIT Press.

Quine, W.V.O. 1945. On ordered pairs. *The Journal of Symbolic Logic* 10, 95–96.

Richards, Norvin. 1999. Featural cyclicity and the ordering of multiple specifiers.
In Samuel David Epstein & Norbert Hornstein (eds.) *Working Minimalism*,
127–158. Cambridge, MA: MIT Press.

van Riemsdijk, Henk. 1998. Trees and scions – science and trees. In *Chomsky 70th
Birthday Celebration Fest-Web-Page*. Retrieved from http://cognet.mit.edu/
Books/celebration/essays/riemsdyk.html.

van Riemsdijk, Henk. 2001a. Wh-prefixes: The case of *wäisch* in Swiss German. In
Chris Schaner–Wolles, John Rennison & Friedrich Neubarth (eds.), *Natural-
ly! Linguistic Studies in Honour of Wolfgang Ulrich Dressler, Presented on the
Occasion of his 60th Birthday*, 423–431. Torina: Rosenberg & Sellier.

van Riemsdijk, Henk. 2001b. A far from simple matter: Syntactic reflexes of
syntax–pragmatics misalignments. In István Kenesei & Robert Harnish,
(eds.), *Perspectives on Semantics, Pragmatics, and Discourse: A Festschrift for
Ferenc Kiefer* (Pragmatics & Beyond New Series 90), 21–41. Amsterdam:
John Benjamins.

van Riemsdijk, Henk. 2006a. Grafts follow from merge. In Mara Frascarelli (ed.),
Phases of Interpretation (Studies in Generative Grammar 91), 17–44. Berlin:
Mouton de Gruyter.

van Riemsdijk, Henk. 2006b. Free relatives. In Martin Everaert & Henk van
Riemsdijk (eds.), *The Blackwell Companion to Syntax*, vol. II, 338-382. Oxford:
Blackwell.

Ross, John R. 1967. Constraints on variables in syntax. Cambridge, MA: MIT
dissertation. [Published revised as *Infinite Syntax!* (Language and Being),
Norwood, NJ: Ablex, 1986.]

Sabbagh, Joseph. 2007. Ordering and linearizing rightward movement. *Natural
Language and Linguistic Theory* 25, 349–401.

Sag, Ivan & Janet Fodor. 1994. Extraction without traces. *West Coast Conference on
Formal Linguistics (WCCFL)* 13, 365–384.

Sampson, Geoffrey. 1975. The single mother condition. *Journal of Linguistics* 11, 1–
11.

Schippers, Ankelien. 2008. Partial wh-movement and wh-copying in Dutch: Evi-
dence for an indirect dependency approach. Ms., University of Groningen.

Schneider, Hubert. 1977. Some remarks concerning a definition of ordered pairs.
The American Mathematical Monthly 84, 636–638.

Sider, Theodore. 1996. Naturalness and arbitrariness. *Philosophical Studies* 81, 283–

301.

Starke, Michal. 2001. Move dissolves into merge: A theory of locality. Geneva: Université de Genève Dissertation.

te Velde, John. 1997. Deriving conjoined XPs: A minimal deletion approach. In Werner Abraham & Elly van Gelderen (eds.), *German: Syntactic Problems — Problematic Syntax* (Linguistische Arbeiten 374), 231–259. Tübingen: Niemeyer.

Tsubomoto, Atsurô & John Whitman. 2000. A type of head-in-situ construction in English. *Linguistic Inquiry* 31, 176–182.

Uriagereka, Juan. 1999. Multiple spell-out. In Samuel David Epstein & Norbert Hornstein (eds.), *Working Minimalism*, 251–282. Cambridge, MA: MIT Press.

de Vries, Gertrud. 1992. On coordination and ellipsis. Tilburg: Katholieke Universiteit Brabant dissertation.

de Vries, Mark. 2005a. Coordination and syntactic hierarchy. *Studia Linguistica* 59, 83–105.

de Vries, Mark. 2005b. Ellipsis in nevenschikking: Voorwaarts deleren, maar achterwaarts delen. *Tabu* 34, 13–46.

de Vries, Mark. 2005c. Merge: Properties and boundary conditions. *Linguistics in the Netherlands* 22, 219–230.

de Vries, Mark. 2007. Invisible constituents? Parentheses as b-merged adverbial phrases. In Nicole Dehé & Yordanka Kavalova (eds.), *Parentheticals* (Linguistik Aktuell/Linguistics Today 106), 203–234. Amsterdam: John Benjamins.

de Vries, Mark. 2009a. Asymmetric merge and parataxis. *Canadian Journal of Linguistics* 53, 355–386.

de Vries, Mark. 2009b. Unconventional mergers. Ms., University of Groningen.

Watanabe, Akira. 1995. Conceptual basis of cyclicity. *MIT Working Papers in Linguistics* 27, 269–291.

Wexler, Kenneth & Peter Culicover. 1980. *Formal Principles of Language Acquisition*. Cambridge, MA: MIT Press.

Wiener, Norbert. 1914. A simplification of the logic of relations. *Proceedings of the Cambridge Philosophical Society* 17, 387–390. [Reprinted in van Heijenoort, Jean (ed.). 1967. *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*. Cambridge, MA: Harvard University Press.]

Wilder, Chris. 1997. Some properties of ellipsis in coordination. In Artemis Alexiadou & Alan Hall (eds.), *Studies on Universal Grammar and Typological Variation* (Linguistik Aktuell/Linguistics Today 13), 59–107. Amsterdam: John Benjamins.

Wilder, Chris. 1999. Right node raising and the LCA. *West Coast Conference on For-mal Linguistics (WCCFL)* 18, 586–598.

Wilder, Chris. 2008. Shared constituents and linearization. In Kyle Johnson (ed.), *Topics in Ellipsis*, 229–258. Cambridge: Cambridge University Press.

Williams, Edwin. 1978. Across-the-board rule application. *Linguistic Inquiry* 9, 31–43.

Yasui, Miyoko. 2002. A graph-theoretic reanalysis of bare phrase structure theory and its implications on parametric variation. Paper presented at *Linguistics and Phonetics 2002*, Urayasu. [Meikai University, 2–6 September 2002.]

Yasui, Miyoko. 2004. Syntactic structure without projection Labels. Ms., Dokkyo University.

Zhang, Niina. 2004. Move is remerge. *Language and Linguistics* 5, 189–209.

Zwart, C. Jan-Wouter. 1994. Dutch is head initial. *The Linguistic Review* 11, 377–406.

Zwart, C. Jan-Wouter. 2004. Een dynamische structuur van de Nederlandse zin. Deel 1 en 2'. *Tabu* 33, 55–71 and 151–172.

Zwart, C. Jan-Wouter. 2005. Some notes on coordination in head-final languages. *Linguistics in the Netherlands* 22, 231–242.

Zwart, C. Jan-Wouter. 2006a. Local agreement. In Cedric Boeckx (ed.), *Agreement Systems* (Linguistik Aktuell/Linguistics Today 92), 317–339. Amsterdam: John Benjamins.

Zwart, C. Jan-Wouter. 2006b. Over het enten van interpolaties. *Tabu* 35, 163–180.

Zwart, C. Jan-Wouter. In press. Prospects for top-down derivation. *Catalan Journal of Linguistics* 8.

Mark de Vries
Rijksuniversiteit Groningen
Department of Linguistics
P.O. Box 716
9700 AS Groningen
The Netherlands
Mark.de.Vries@rug.nl

Metrical Combinatorics and the Real Half of the Fibonacci Sequence

William J. Idsardi & Juan Uriagereka

Languages with stress group syllables into metrical feet (Halle and Idsardi 1995, Hayes 1995)—non-exhaustive groups of contiguous syllables. The size of feet in natural languages ranges from unary (a single syllable) to unbounded (as many syllables as possible); in addition syllables can also remain unfooted. Under these conditions, the number of possible metrical footings for a string of n syllables is known to be Fib($2n$) (Idsardi 2008), where Fib(n) is the Fibonacci sequence, as in (1)[1] where the nth element is the sum of the previous two elements (for example, $13 = 8 + 5$).

(1) n: 0 1 2 3 4 5 6   7 …
        Fib(n): <u>1</u> 1 <u>2</u> 3 <u>5</u> 8 <u>13</u> 21 …

For example, a string of two syllables (here notated with 'x's) can be non-exhaustively footed in five ways (= Fib(4)): (xx), (x)(x), (x)x, x(x), and xx. In contrast, if footing were required to be exhaustive (that is, if every syllable had to belong to some foot) then a string of two syllables could only be footed in two ways: (xx) and (x)(x). It is easy to see from the bracketed grid representations that the number of possible exhaustive footings of a string of n syllables must be 2^{n-1} as every exhaustive footing must begin and end with foot-boundaries and between each pair of x's we have a binary choice between having a foot juncture and not having one. Since there are two choices for each space between x's and there are n-1 spaces between n x's, it follows directly that there are 2^{n-1} distinct exhaustive footings.

As a consequence, only half of the Fibonacci numbers (those underlined in (1): 1, 2, 5, 13, …) are solutions to the task of creating non-exhaustive footings; the other half (3, 8, 21, …) are not. An intriguing question is: Why is it the one half of the sequence and not the other? We venture some speculations about potential answers.

In 1680, Cassini (1733) discovered a relation among successive members of the Fibonacci sequence, expressed in (2):[2]

(2) $\text{Fib}(n)^2 - \text{Fib}(n-1) \cdot \text{Fib}(n+1) = (-1)^n$

---

[1]   The Fibonacci sequence can also be defined to start with 0: 0, 1, 1, 2, 3, …
[2]   This relation was independently discovered by Simson (1753).

That is, the square of any Fibonacci number is equal to the product of the two flanking Fibonacci numbers, give or take one. For example, $\text{Fib}(4)^2 - \text{Fib}(3) \cdot \text{Fib}(5) = 25 - 3 \cdot 8 = 1 = (-1)^4$ and $\text{Fib}(5)^2 - \text{Fib}(4) \cdot \text{Fib}(6) = 82 - 5 \cdot 13 = -1 = (-1)^5$. Rearranging (2) gives (3):

(3) $\text{Fib}(n)^2 - (-1)^n = \text{Fib}(n{-}1) \cdot \text{Fib}(n{+}1)$

The left-hand side of (3) has two possible expansions, depending on whether n is odd or even, as in (4):

(4)   a. n is even: $\text{Fib}(n)^2 - 1$
         b. n is odd: $\text{Fib}(n)^2 + 1$

We can now see that (4a) has the form $(x^2 - 1)$ and (4b) has the form $(x^2 + 1)$. Elementary algebraic polynomial factorization (Herstein 1977) shows that (4a) has real-valued roots, (5a), whereas (4b) only has complex-valued roots, (5b):

(5)   a. $\text{Fib}(n)^2 - 1$ $=$ $[\text{Fib}(n) - 1][\text{Fib}(n) + 1]$
         b. $\text{Fib}(n)^2 + 1$ $=$ $[\text{Fib}(n) - i][\text{Fib}(n) + i]$ (where $i^2 = -1$)

Thus, for example, $\text{Fib}(3) \cdot \text{Fib}(5) = 3 \cdot 8 = 24 = 4 \cdot 6 = [\text{Fib}(4) - 1][\text{Fib}(4) + 1]$. Only the even-numbered Fibonacci numbers (here, Fib(4)) show up in the real-valued roots, and this is the same Fibonacci subset that characterizes the number of valid metrical groupings of strings of n syllables.

In conclusion, the 'metrical' half of the Fibonacci sequence is also the 'real-valued' half of the sequence (in the sense of (5)). Evidently, the Fibonacci character of footing arises just when we allow for non-exhaustive footing, as exhaustive footings can be counted as a simple set of independent binary choices. Generally, the Fibonacci sequence is associated with a number of 'edge of chaos' effects, especially systems which illustrate *dynamical frustration* (Binder 2008); systems in which opposing forces cannot reach an equilibrium solution. We speculate that the 'forces' operative here in defining non-exhaustive footings could be the local coherence of syllables into feet clashing with word-level properties of footing. Another potential view of the emergent complexity observed here would be that sequences of footed syllables can be metrically distinct — for example, $(x)(x) \neq (xx)$ — whereas all sequences of unfooted syllables are the same; thus we have asymmetric growth patterns in the footed and unfooted portions of syllabic strings resulting in Fibonacci complexity.

References

Binder, Philippe M. 2008. Frustration in complexity. *Science* 320, 322–323.
Cassini, Jean-Dominique. 1733. Une nouvelle progression de nombres. *Histoire de l'Acaémie Royale des Sciences*, volume 1. Paris.

Halle, Morris & William J. Idsardi. 1995. General properties of stress and metrical structure. In Eric S. Ristad (ed.), *Language Computations*, 37–69. Providence, RI: American Mathematical Society.

Hayes, Bruce. 1995. *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.

Herstein, Israel N. 1977. *Topics in Algebra.* New York: John Wiley and Sons.

Idsardi, William J. 2008. Combinatorics for Metrical Feet. *Biolinguistics* 2, 233–236.

Simson, Robert. 1753. An explication of an obscure passage in Albert Girard's commentary upon Simon Stevin's works. *Philosophical Transactions of the Royal Society of London* 48, 368–376. [doi:10.1098/rstl.1753.0056]

William J. Idsardi / Juan Uriagereka
University of Maryland
Department of Linguistics
1401 Marie Mount Hall
College Park, MD 20742
USA
idsardi@umd.edu / *juan@umd.edu*

★ REVIEWS ★

Climbing up the Eiffel Tower

Berger, Ruth. 2008. *Warum der Mensch spricht. Eine Naturgeschichte der Sprache.* Berlin: Eichborn. (Why man speaks. A natural history of language.)

by Michael Crombach

On étonne toujours en disant que le problème de l'origine du langage n'est pas un problème d'ordre linguistique. (Vendryes 1921: 6)

Ruth Berger presents an ambitious attempt on the origins and evolution of human language, she even maintains a website intended for regular updates on the topic (http://www.sprache-und-evolution.de). She lists over 400 references, and thereby shows great research efforts, although a sorted bibliography would be of great help. The book reminds of Kenneally (2007), reviewed in *Biolinguistics* recently (Jenkins 2008), although they arrive at completely different solution scenarios. Ruth Berger's is popular science writing that, although not authored by an expert in the field, provides a lot of insight into an ever-growing field. *Warum der Mensch spricht* ('Why man speaks') is somehow unusual for German popular science as it is written in a witty, colloquial style that is fast and easy to read, and it does not contain any obvious errors. Still, some of the pert statements are annoying and should not go uncommented — therefore this contribution. Another reason to review the book is to create awareness that research in the field of biolinguistics is also pursued in languages other than English. Hardly ever a work is translated like Oudeyer (2006); works on language (evolution) not published in English go mostly unnoticed by the academic community.

This has been the case with the anthropological-psychiatric couple Doris and David Jonas, who despite being an English–German melange decided to publish their book on the origins of language in German (Jonas & Jonas 1979). They did not receive much attention although they originally presented their results in English (Jonas & Jonas 1975); maybe the mid to late Seventies have not been the time to present a 'female' origin of language. Anyway, I can not follow Berger (212f.) in her claim that the arguments for a 'motherese' origin of language brought forth by Jonas & Jonas would be more elaborate than Falk (2004, 2009). The trouble with both explanations of the origin of language is that they miss the bigger picture. The loss of body hair, neoteny, and upright walk led to the evolution of language, the lateralization of the brain, etc. — these are the basic arguments of Jonas & Jonas (1975), Jonas & Jonas (1979), Falk (2004, 2009), but these premises of language development also need an explanation, otherwise it is not too prudent to base the explanation of language origins on it.

According to Berger, Falk represents the currently most popular

explanation for the origin of language. Still she devotes only seven lines of text in parentheses to one of the essential premises of this thesis, the loss of body hair (p. 211). I would expect more arguments, not only the hint that genetic evidence points to hairlessness in *homo ergaster*, why hominids have lost their body hair early in their evolutionary development. The *relative chronology* of events is essential to all evolutionary assumptions; if the loss of body hair is a premise of language evolution, because it lead to "putting the baby down" (Falk 2009: 58), then there should be an explanation why it is so sure that hominids lost body hair before they started to talk, and why they lost it at all (Dawkins 2004: 67 considers gain and loss of body hair as trivial; for possible scenarios, see e.g., Kingdon 2003, Morgan 1997, and Niemitz 2004). Otherwise the argument is not complete. Anyway, I think that the motherese hypothesis has some good points that should get more attention, and with the methodological precautions mentioned here this is a promising direction.

I read *Warum der Mensch spricht* in Finland, but while I read it I thought a lot about the Eiffel Tower. The escalator that takes visitors from the ground to the first platform slowly moves upwards passing steel beams, rivets, and paint. If you pass a box of rivets in a hardware store, you'll surely call them 'rivets', but you'll never see them as a part of the Eiffel Tower, although you might explain to your child that these are things are used for the Eiffel Tower. The major error of all gradualist approaches is that they ignore the insight of Aristotle's *Metaphysics* (1041b), that the whole is more than the sum of its parts. There is no way to declare a pile of some 18,038 iron beams, 2.5×10^6 rivets, and 60 tons of paint the Eiffel Tower — being the Eiffel Tower is a quality in its own right and therefore can be studied independently. Human language shares features with (other) animal communication systems, but as Keller (2003: 44) explains:

> Die Fähigkeit zur Kommunikation ist dem Sprachbesitz logisch vorgängig. Eine Sprache erleichtert das Kommunizieren, ist aber nicht Bedingung seiner Möglichkeit.[1]

In the end, we have to accept that human language is unique, and there is no way to bridge the gap between the communication of animals and language of humans (see e.g., Bickerton 2008). Berger dares to claim that there is no principle difference between animal communication and language (p. 250), while a few pages later humans "slowly crossed a Rubicon" (p. 254). That is confusing, but I would say that hominids in curse of their evolution approached the Rubicon, crossed it, and now there is no turning back. It is even hard for us to imagine a communication system that is not 'language' — we use language, and even our thinking is essentially language, so we can hardly imaging being without it. It's a bit like the pink elephant you can impossibly not think of if asked to do so.

Rather irritating within the book is something I would call 'Chomsky bashing' (pp. 18–27). Berger even announces "the failure of the Chomsky-project" (p. 27). I do not have any objections on critique towards Chomsky and his theories; as a matter of fact, I do think there is a lot of room for improvement. But

---

[1]   'The ability to communicate is a precursor to the possession of language. Language makes communication easier but is not a premise to the possibility of communication.'

Berger devotes a lot of energy in showing how wrong Chomsky was, though I can not see where she offers an alternative. Again the Eiffel Tower: Many single explanations do not make a theory; the theory as a whole has a value in its own right, and as long as there are only details changed, the theory can stay in place. Chomsky has been and will remain to be a controversial figure, but I do not think that Chomsky is to blame for everything that went wrong in linguistics in the last 60 years or so. All linguists have had the opportunity to make up their own minds. And as long as there is no convincing alternative to Generative Grammar as a theory, I would not abandon it. When Berger states in the end that Chomsky was right with the assumption that the language faculty is innate (p. 255), this somehow leads back to the start — so, why first condemn Chomsky, his ideas, and his pupils?

The claim by Berger that linguists are "traditionally left" (p. 130) must not go uncommented either, as it leaves out a lot of dark history of linguistics. To cut this short, I dare to claim that it is one of the biggest achievements of Noam Chomsky to have freed linguistics of the racist, fascist heritage and suspicion that was burdened onto it by men from Arthur de Gobineau to Walter Wüst. The mindless mixture of linguistic terminology with racism has discredited generations of linguistic research. It is essential that science keeps sound distance to ideology; ideology is non-scientific. Berger embeds her statement in the discussion, whether or not all "native speakers" share the same competence, as if degree of language skills was of any relevance to the basic principle described by the innateness of language.

Berger creates the impression that the emergence and evolution of language are a biological necessity (pp. 252–254). She does not make a clear distinction between 'target' and target (Jenkins 2008). Language is the 'target' (i.e. the result) of human evolution, but not the target (i.e. the goal). There is no teleological necessity in the rise of language. Berger correctly understands the evolution of language as a process stretching over millions of years, but she arrives at the rather strange conclusion that "language was at the beginning and not at the end of the human evolution" (p. 259). She believes in a developmental continuum of language evolution stretching over 2.6 million years one the one hand, while on the other hand, language capability was the prerequisite for the evolution of *homo sapiens*. This contradiction asks for a solution I cannot offer here.

The book starts with the observation that language origins have mainly been the concern of linguists and not of biologists. This is true, and a pity — as the evolution and the origin of language are not a problem of linguistics as already observed by Vendryes (1921). The evolution of language is a biological problem and has to be solved by biological means. But as long as biologists do care so little about this issue, linguists/linguistics have/has to stand in for them. Berger does not explain why biologists ignore language evolution, although the blurb states that she studied "Turk languages, Hebrew and English, general linguistics and biology". Maybe the true solution is the interdisciplinary co-operation proposed and demonstrated by Hauser *et al.* (2002). But still, I think the evolution of language is focused too narrowly on language. Jonas & Jonas and Falk try to add other aspects, but still are focused on language. In my opinion,

explanations need to be found for all of the small but significant differences between humans and apes — from the loss of body hair to bipedal locomotion, from the white in the eyes to art and music, etc. — to really understand the origins and evolution of language, and what it means to be human.

Berger's book, like Kenneally's, is a valuable source for resources. Although Berger maintains her website for updates, she should have stressed the fact that most of the things dealt with in the study of language evolution are in permanent movement and the main skill for any scientist is to balance probabilities, exclude impossibilities, and reorder events. Or as Bickerton (2007: 524f.) has put it:

> [The study of language evolution is] still a paper-and-pencil field, though with immeasurable amounts of reading and thinking involved. It is, accordingly, an ideal field for any ambitious young scholar itching to make his academic bone. But take care, it's a minefield out there.

References

Bickerton, Derek. 2007. Language evolution: A brief guide for linguists. *Lingua* 117, 510–526.

Bickerton, Derek. 2008. Two neglected factors in language evolution. In Andrew D.M. Smith, Kenny Smith & Ramon Ferrer i Cancho (eds.), *The Evolution of Language: Proceedings of the 7th International Conference (Evolang 7)*, 26–33. New Jersey, NJ: World Scientific Publishing.

Dawkins, Richard. 2004. *The Ancestor's Tale: A Pilgrimage to the Dawn of Evolution.* Boston: Houghton Mifflin Harcourt.

Falk, Dean. 2004. Prelinguistic evolution in early hominins: Whence motherese? *Behavioral and Brain Science* 27, 491–503.

Falk, Dean. 2009. *Finding Our Tongues: Mothers, Infants, and the Origins of Language*. New York: Basic Books.

Hauser, Marc, Noam Chomsky & W. Tecumseh Fitch. 2002. The faculty of ;anguage: What is it, who has it, and how did it evolve? *Science* 298, 1569–1579.

Jenkins, Lyle. 2008. Not the only Word. *Biolinguistics* 2, 342–347.

Jonas, Doris F. & A. David Jonas. 1975. Gender differences in mental function: A clue to the origin of language. *Current Anthropology* 16, 626–630.

Jonas, Doris F. & A. David Jonas. 1979. *Das erste Wort: Wie die Menschen sprechen lernten*. Hamburg: Hoffmann und Campe.

Keller, Rudi. 2003. *Sprachwandel.* 3rd edn. Tübingen: UTB.

Kenneally, Christine. 2007. *The First Word: The Search for the Origins of Language.* New York: Penguin Books.

Kingdon, Jonathan. 2003. *Lowly Origin: Where, When and Why Our Ancestors First Stood Up.* Princeton, NJ: Princeton University Press.

Morgan, Elaine. 1997. *The Aquatic Ape Hypothesis.* London: Souvenir Press.

Niemitz, Carsten. 2004. *Das Geheimnis des aufrechten Gangs. Unsere Evolution verlief anders.* München: C. H. Beck Verlag.

Oudeyer, Pierre-Yves. 2006. *Self-Organization in the Evolution of Speech.* Oxford: Oxford University Press.

Vendryes, Joseph. 1921. *Le langage : Introduction a l'histoire.* Paris: La Renaissance du livre.

Michael Crombach
Kierlinger Strasse 106
A–3400 Klosterneuburg
Austria
michael.crombach@gmx.at

Everything You Wanted to Know About the Genetics of Language (and Beyond)

Benítez–Burraco, Antonio. 2009. *Genes y Lenguaje: Aspectos Ontogenéticos, Filogenéticos y Cognitivos* [*Genes and Language: Ontogenetic, Phylogenetic, and Cognitive Issues*]. Barcelona: Editorial Reverté.

by Víctor M. Longa

As Cedric Boeckx and Kleanthes Grohmann pointed out in 'The *Biolinguistics* Manifesto', which opened this journal, there are two different senses of the term 'biolinguistics', a weak one and a strong one. Their own words illustrate: "The weak sense of the term refers to 'business as usual' for linguists, so to speak, to the extent that they are seriously engaged in discovering the properties of grammar" (Boeckx & Grohmann 2007: 2). With regard to the second (strong) sense, it "[…] refers to attempts to provide explicit answers to questions that necessarily require the combination of linguistic insights and insights of related disciplines (evolutionary biology, genetics, neurology, psychology, etc.)" (p. 2).

The book reviewed is one of the most important references in the second (strong) sense of 'biolinguistics' which has been published to date. This piece of impeccable scholarship pursues two main aims. Firstly, it provides the reader with an impressive and completely up-to-date overview on the genetic and molecular (and, by extension, biological) foundations of language. In this respect, it suffices to say that whereas the discussion about the genetics of language is usually restricted to the role of the 'famous' *FOXP2* gene, the book refers to (and analyzes) more than 150 genes which recent research has somehow linked to language. Fundamental as this enterprise would be by itself, the book is not confined to it. As a second aim, the aforementioned overview is the input for sophisticated and in-depth discussion about key issues having to do with the biology of language. These include (i) how to manage the relationship between genes and behavior; (ii) what the true significance is of genes, their properties, and their products for understanding human language; (iii) what genes can reveal for topics such as language organization in the brain, language phylogeny (evolution) or language ontogeny (development); (iv) how the relationship between language and cognition should be characterized; and (v) what degree of convergence exists between discoveries coming up from the genetics of language and proposals which theoretical linguistics (especially, the Minimalist Program, henceforth, MP) has brought to the fore. For these reasons, I consider the book by Antonio Benítez–Burraco (henceforth, ABB) to be an essential reference (to put it more precisely, a true hand-book), which everybody interested in the biological

seat of language should consult.

I began the review by mentioning the two senses the term 'biolinguistics' is endowed with, according to Boeckx & Grohmann. In order to go deeper in discussion on the strong sense of the term, it should be noted that there are in principle two different strategies for such a strong sense to be fulfilled — a multidisciplinary approach and an interdisciplinary approach. Although both strategies are usually conflated, a great difference opposes them. A multi-disciplinary approach means that the same problem is studied from several disciplines, but this approach does not necessarily connect achievements gained by each of them. However, with regard to an interdisciplinary approach, quite the opposite applies. Knowledge offered by different disciplines is integrated (i.e. merged), the outcome being a shared body of knowledge. Needless to say, an in-terdisciplinary approach is much harder to be obtained than a multidisciplinary one. It is perhaps for that reason that, according to Newmeyer (1997), linguists have been traditionally reluctant to seriously consider issues which transcend linguistics itself (for example, clinical, behavioral, cognitive or biological evidence). In that regard, one of the many merits of ABB's book is that it clearly surpasses a multidisciplinary approach (a perspective which does not ensure the property of consilience, or unity of knowledge, as stated by Wilson 1998), to become truly interdisciplinary. ABB is an especially suitable scholar for achieving such a task, given his (really welcome) academic training both in molecular biology and theoretical linguistics.

I will offer a brief outline of the organization of the book and of the main topics the chapters deal with, although this is not an easy task considering the denseness of the book. After a brief introduction (pp. 1–3) where the *raison-d'être* of the book is outlined and its main objectives are advanced, Chapter 1 (pp. 5–33) is devoted to the anatomical and physiological foundations of language. It critically discusses a number of models aiming at explaining the anatomical and functional organization of language. The neurolinguistic discussion is summar-ized in an appendix which gathers all the brain areas involved in linguistic processing, with an indication of the key references for each of them.

Chapter 2 (pp. 35–53) analyzes the polemic issues of innateness and learning in language ontogeny, and the controversies surrounding them. The author makes the point that, for nativism to be truly justified, the need exists to consider a wider range of evidence than the linguistic one, thus broadening the evidence with which linguists have been mainly concerned. According to ABB, genetic and molecular evidence is suitable for such an objective to be achieved. Nevertheless, what I take to be the main contribution of the chapter is the discussion of what the very notion of 'innate' means, and how it has been reformulated within MP, as opposed to the previous generative tradition. Mini-malism has reduced the role of the genetic endowment for language (i.e. the linguistic genotype or 'first factor', following Chomsky 2005) which was sup-posedly required for language acquisition to take place (cf. Longa & Lorenzo 2008 and Lorenzo & Longa 2009). Accordingly, MP reduces the specifically linguistic (i.e. specifically grammatical) component of the human mind (cf. Lorenzo & Longa 2003), and considers the faculty of language to be the outcome of epigenetic processes rather than the product of purely genetic processes. It is

for those reasons that minimalism redefines the very notion of innateness. As opposed to preceding generative models (paradigmatically GB, that is, Government–and–Binding Theory) and to the Neo-Darwinian framework as well, the notions of genetic trait and innate trait are no longer conflated in MP (cf. Longa 2006). Therefore, minimalism argues for a phenotypic notion of innateness, not a genotypic one (Longa & Lorenzo 2008). The faculty of language would lose its genetic character, but not its congenial or innate nature, in such a way that it would be innate attending to its propensity to arise irrespective of the foundations of its development, those foundations not requiring to be purely genetic (cf. Maclaurin 2002 and Moore 2001: Chap. 13, for a defense of such a view from a strictly biological point of view). ABB's discussion of that issue is well taken and illuminating.

Chapters 3 and 4 develop a wide analysis of the currently known genetic and molecular mechanisms which are responsible for how the neural circuits related to language develop and function. Chapter 3 (pp. 55–81) approaches the molecular bases of development and plasticity of the brain linguistic areas, and how those areas work. The chapter aims at exploring both the structural and functional development, and, furthermore, it seeks to integrate them. As usual in every chapter, an appendix is offered (pp. 80–81) where the different genes referred to so far are summarized: Gene name, chromosome localization, protein function, and main scientific literature.

As of chapter 4, it could well be an independent book by itself on the basis of its length alone (pp. 83–281), and it is undoubtedly one of the most valuable chapters of the book. Had I to highlight one of the chapters, it would be this one. To put it simply, it is impressive. As far as I know, it offers the most extensive overview of the genetic bases of language to date, and it is this overview which makes the chapter so innovative. It begins by presenting the essentials of the different methods and strategies available for cloning genes (comparative, functional, and positional cloning), and then it goes on to characterize the problems which arise when trying to define the linguistic phenotype and its impairments. After those introductory topics, the main goal of the chapter is approached in which the author provides us with both a structural and functional characterization of the currently known genes which are somehow linked to language. For this goal to be achieved, ABB carried out a large and detailed search in many scientific journals, and applied further analysis and synthesis. This has allowed him to collect up to and characterize more than 150 genes which recent research has shown to be related to language. The genes are arranged according to three general categories:

(A)   genes involved in exclusively linguistic impairments (although ABB acknowledges the controversy surrounding the specifically linguistic nature of those deficits);

(B)   genes involved in general cognitive impairments which also affect language, and

(C)   genes involved in cognitive impairments which do not seem to affect language, but are relevant anyway in order to characterize the genetic bases of language.

Although the extensive analysis of the *FOXP2* gene ABB offers should be high-lighted, specific treatment of the remaining genes is worth considering as well. The overall picture offers a comprehensive overview about the regulatory mechanisms which are responsible for how language areas involved in language are organized and function.

Although the appendices are not unusual in the book, I feel obliged to stress a very extensive appendix (pp. 240–281) which ends chapter 4. Its purpose it to collect all the genes discussed in the chapter and their main properties: gene name, chromosome localization, protein function, linguistic impairments associated to the gene mutation, clinical name of the syndrome, and selected scientific literature.

For the reasons specified so far, the chapter is a 'bedside reading' reference about the genes which are somehow related to language ontogeny. As mentioned above, the overall picture spectacularly surpasses the usual conception which conflates the genetics of language with just *FOXP2*.

Chapter 5 (pp. 283–337) is devoted to the other side of the coin, language phylogeny. ABB claims that the range of traditional evidence on language evolution (mainly the vocal tract, symbolic artefacts, and paleoneurological evidence related to brain size and cranial reconstructions) should be broadened with new types of evidence; especially, those that are offered by genetic and molecular analyses. According to ABB, this type of evidence could help us judge more traditional ones, which suffer from an intrinsically ambiguous nature. The goal of the chapter is therefore quite similar to that of chapter 4, but referred to at the phylogenetic level, which is to discuss the evolution of the known genes (related to language) whose expression levels have been modified over the evolutionary course. In a quite similar vein to chapter 4, chapter 5 offers an exhaustive picture of the relevant genes, those genes being arranged according to several categories: (i) genes related to brain size, (ii) brain metabolism, (iii) brain lateralization, and (iv) neural structures (circuits or areas) which have to do with language. The properties of the genes are summarized in a valuable appendix on pp. 332–337.

Finally, chapter 6 is the clearest example of the truly interdisciplinary (not multidisciplinary) nature of the book. Although ABB has chosen to simply name it 'Conclusions', the chapter is really much more than what its title suggests. In fact, the chapter develops a wide discussion (pp. 339–364) about the linguistic significance of the biochemical and genetic evidence analyzed in the preceding chapters. The discussion seeks to unravel the ontogenetic, phylogenetic, and cognitive implications of the genes involved (in several ways) in human language. Many topics of main concern from a theoretical point of view are confronted, and sophisticated attempts are made to offer answers for them. To give some hints of the relevance of the chapter, some of the topics it is concerned with are: (i) how genes really work (far from simplifying assumptions about the direct relationship between genes and phenotypic traits); (ii) an assessment of how the relationship between nature and nurture should be addressed; (iii) a discussion on language evolution; and (iv) how the notion of modularity should be understood in the light of how genes work and are organized. In addition, (v) a proposal is made that the language organ derives from a double developmental program (one being more general and the other one being more specific), and (vi) claims are

made about the non-specific nature of the 'genes of language' (this expression is systematically endowed with quotation marks through the whole book). Let us take into account that this point was already fully advanced by Lenneberg (1967: Chap. 6), when he wrote that it was not necessary to make doubtful claims about 'genes of language'. Finally, (vii) an assessment is made of how the genetic issues considered in the book, and the non-specific nature of the genes themselves, fit in with proposals suggested by MP.

The book ends with an impressive reference section, of more than 80 pages (pp. 365–449) listing more than 2,000 references, showing the immense work put in by the author. A very detailed thematic index is offered as well (pp. 451–478).

Although the book is highly technical (it becomes obvious that such a book could not be jargon-free), ABB's effort to make its reading and use easier should not go unnoticed. Beyond the aforementioned appendices, the book is endowed with 20 tables and no less than 115 figures (the vast majority in color).

I hope that the brief presentation of the main contents of the book will allow to shed light on at least some of its many merits. In addition, it seems necessary to highlight that, although the book focuses on genes, ABB's view is, much to my delight, very far from the primacy of the 'genetic program' metaphor (and, consequently, very far from the primacy of the genes themselves) which has been at the heart of Neo-Darwinian thinking (and which can still be perceived in works such as Carroll 2005 and other practitioners of Evo–Devo). To put it in other words, the author is well aware of the dangers a strictly reductionist perspective has meant for the biological study of organisms (cf. Kaufmann 2000 and Lewontin 2000, among many other references). ABB's own words clearly illustrate: "[…] the genetic approach to the study of language should not be understood from a strictly reductionist perspective, which considers the gene to be the final point of any analysis of language" (p. 364; own translation — VML). As Oyama (2000: 40) puts it, "[…] a gene initiates a sequence of events only if one chooses to begin analysis at that point". For that reason, the author contends that knowledge gained from genetic and molecular analyses should be integrated in an overall picture. As ABB himself acknow-ledges, his view is not far from 'probabilistic epigenesis' as developed by Gilbert Gottlieb (cf. Gottlieb 2001), an influential scholar close to Oyama's Develop-mental Systems Theory (cf. Oyama *et al.* 2001a, 2001b). Consequently, according to ABB, genes are not the main biological entities, but they strongly interact with the whole range of developmental resources and levels (cellular, neural, behavioral, environmental, and so on) of which the system is composed. Such a position concerning the role attributed to the genes is in full agreement with theoretical stances which strongly depart from Neo-Darwinian assumptions; in fact, in several passages of the book, ABB suggests that his proposal fits in well with Developmental Systems Theory (cf. p. 363, among others), and with the view sustained by MP as well. Thus, for ABB the genome cannot be conceived of as an encapsulated entity.

Other hints also make it clear that the author departs from the biological establishment (i.e. Neo-Darwinism); for example, this is demonstrated in his con-ception of heredity. Such a conception goes beyond the usual (Neo-Darwinian) stance, according to which genes are the only biological elements which can be

inherited. ABB recognizes the role of other types of heredity (maternal, epi-genetic, social or even behavioral; cf. p. 84), in full agreement with positions which defend that "there is more to heredity than genes" (Blumberg 2005: 148, Jablonka & Lamb 2005: 1), as can be seen in the four dimensions of heredity developed by Jablonka & Lamb (2005) (for a synthetic presentation, see Jablonka 2001), or the even wider notion of 'extended heredity' argued for by Develop-mental Systems Theory (cf. Griffiths & Gray 2001 and Oyama *et al.* 2001a, 2001b).

Another aspect of the book I fully agree with is the status ABB confers to MP, an ontological one rather than merely methodological. This means that the author does not share the 'consensus view' (cf. Boeckx 2006, Freidin & Vergnaud 2001, and Hornstein *et al.* 2005, among many others) by which minimalism would be no more than an extension, or a mere refinement, of the Principles–and–Parameters model which creates an opening for simplicity, naturalness, and so on (cf. Longa & Lorenzo 2008 for a discussion of the differences between GB and MP). The following words by Hornstein (2009: 178) illustrate: "[…] MP is a continuation of the GB research program […]. MP starts from the assumption that GB is roughly correct." It seems to me that this position is based on a metho-dological (i.e. weak) minimalism, and it does not jibe with an ontological (i.e. strong) consideration of minimalism. On the contrary, ABB considers MP to be an important (or even radical) break with regard to GB (cf. chapters 3 and 6), and I think that his view is accurate. It should be noted that the biological position adopted by GB was based on the 'consensus view' on organisms and organismal development which Neo-Darwinism brought to the fore (cf. Lorenzo & Longa 2009 for discussion). Robert (2004: 39) characterizes that 'consensus view' according to three main features: (1) genetic informationism (the information required for the development of an organism is contained within its genes), (2) genetic animism (such information consists on a genetic program), and (3) genetic primacy (genes are the vehicles by which the information is inherited, the main promoters of development). The solution GB provided to Plato's Problem was to fully assume that 'consensus view', to assume a genetically encoded state of linguistic knowledge (Universal Grammar) or 'linguistic genotype' (Chomsky 1980, Lightfoot 1982, 2006), which was taken to be a direct expression of the genes. Therefore, the strong geneticist view of GB can be summarized in the notion of a genetic program (Chomsky 1980, Wexler 1999) (cf. Longa 2008 for critical discussion of that notion). However, the strong geneticism (which has been the focal point in every generative model except the minimalist one) has been removed from the agenda, since minimalism advocates the need to reduce the role of the genetic endowment, and argues for the non-specific nature of the principles the language faculty is composed of. The book reviewed clearly favors an ontological minimalism, and, interestingly, ABB shows that the conclusions reached from the analysis developed in the book are consistent (both in phylogeny and ontogeny) with the framework of (ontological) minimalism, as the author himself acknowledges.

To sum up, the book provides us with a delicious cocktail: Biology and theoretical linguistics side by side (i.e. merged in a truly interdisciplinary way). There is no room for doubt: The field of biolinguistics has many reasons to cele-brate the publication of ABB's book. I am sure it will become an indispensable

reference for anyone seriously interested in the biolinguistic approach. For this reason, given that the book has been published in Spanish, an English translation would be highly desirable as soon as possible. The lack of such a translation would be an important disservice to the field.

References

Blumberg, Mark. 2005. *Basic Instinct: The Genesis of Behavior*. New York: Thunder's Mouth Press.

Boeckx, Cedric & Kleanthes K. Grohmann. 2007. The *Biolinguistics* manifesto. *Biolinguistics* 1, 1–8.

Boeckx, Cedric. 2006. *Linguistic Minimalism: Origins, Concepts, Methods, and Aims*. New York: Oxford University Press.

Carroll, Sean B. 2005. *Endless Forms Most Beautiful: The New Science of Evo Devo and the Making of the Animal Kingdom*. New York: W.W. Norton.

Chomsky, Noam. 1980. *Rules and Representations*. New York: Columbia University Press.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36, 1–22.

Freidin, Robert & Jean-Roger Vergnaud. 2001. Exquisite connections. Some remarks on the evolution of linguistic theory. *Lingua* 111, 639–666.

Gottlieb, Gilbert. 2001. A developmental psychobiological systems view: Early formulation and current status. In Oyama *et al.* (2001b), 41–54.

Griffiths, Paul E. & Russell D. Gray. 2001. Darwinism and developmental systems. In Oyama *et al.* (2001b), 195–218.

Hornstein, Norbert, Jairo Nunes & Kleanthes K. Grohmann. 2005. *Understanding Minimalism*. Cambridge: Cambridge University Press.

Hornstein, Norbert. 2009. *A Theory of Syntax: Minimal Operations and Universal Grammar*. Cambridge: Cambridge University Press.

Jablonka, Eva & Marion J. Lamb. 2005. *Evolution in Four Dimensions. Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life*. Cambridge, MA: MIT Press.

Jablonka, Eva. 2001. The systems of inheritance. In Oyama *et al.* (2001b), 99–116.

Kaufmann, Stuart A. 2000. *Investigations*. New York: Oxford University Press.

Lenneberg, Eric H. 1967. *Biological Foundations of Language*. New York: John Wiley.

Lewontin, Richard C. 2000. Foreword. In Oyama (2000), vii–xv.

Lightfoot, David. 1982. *The Language Lottery: Toward a Biology of Grammars*. Cambridge, MA: MIT Press.

Lightfoot, David. 2006. *How New Languages Emerge*. Cambridge: Cambridge University Press.

Longa, Víctor M. 2006. No solo genes: El Programa Minimalista y la reformulación de la noción de innatismo. *Ludus Vitalis* XIV, 141–170.

Longa, Víctor M. 2008. Una visión crítica sobre la noción de 'programa genético' desde la biología y la lingüística: Consecuencias para la conceptualización de la ontogenia del lenguaje. *Verba* 35, 347–385.

Longa, Víctor M. & Guillermo Lorenzo. 2008. What about a (really) minimalist theory of language acquisition? *Linguistics* 46(3), 541-570.

Lorenzo, Guillermo & Víctor M. Longa. 2003. Minimizing the genes for grammar: The Minimalist Program as a biological framework for the study of language. *Lingua* 113, 643–657.

Lorenzo, Guillermo & Víctor M. Longa. 2009. Beyond generative geneticism: Rethinking language acquisition from a developmentalist point of view. *Lingua* 119, 1300–1315.

Maclaurin, James. 2002. The resurrection of innateness. *The Monist* 85, 105–130.

Moore, David S. 2001. *The Dependent Gene: The Fallacy of 'Nature vs. Nurture'*. New York: Henry Holt.

Newmeyer, Frederick J. 1997. Genetic dysphasia and linguistic theory. *Journal of Neurolinguistics* 10, 47–73.

Oyama, Susan. 2000. *The Ontogeny of Information: Developmental Systems and Evolution*. 2nd edn. Durham, NC: Duke University Press.

Oyama, Susan, Paul E. Griffiths & Russell D. Gray. 2001a. Introduction: What is developmental systems theory? In Oyama *et al.* (2001b), 1–11.

Oyama, Susan, Paul E. Griffiths & Russell D. Gray (eds.). 2001b. *Cycles of Contingencies: Developmental Systems and Evolution*. Cambridge, MA: MIT Press.

Robert, Jason Scott. 2004. *Embriology, Epigenesis, and Evolution: Taking Development Seriously*. Cambridge: Cambridge University Press.

Wexler, Kenneth. 1999. Maturation and growth of grammar. In William C. Ritchie & Tej K. Bhatia (eds.), *Handbook of Child Language Acquisition*, 55–109. San Diego, CA: Academic Press.

Wilson, Edward O. 1998. *Consilience: The Unity of Knowledge*. New York: Knopf.

Víctor M. Longa
Universidade de Santiago de Compostela
Department of Spanish Literature, Literary
        Theory and General Linguistics
Plaza Isabel la Católica, 2, 2º E
ES–36204 Vigo
Spain
VictorManuel.Longa@usc.es

★ FORUM ★

Notice

We would like to use this opportunity to thank all those involved in creating the third volume of *Biolinguistics*. Our special gratitude goes to the reviewers that have served us throughout 2009, who are listed below (colleagues who reviewed more than one submission are suffixed by an asterisk). For everything else, we thank our supporters as well as all the members of the *Biolinguistics* Advisory Board, the *Biolinguistics* Editorial Board, and the *Biolinguistics* Task Team that are not specifically mentioned by name for active participation and feedback. We would also like to thank those board members who are now retiring after more than three years of service.

Reviewers

David Adger
Eric Baković
Sjef Barbiers
Josef Bayer
Robert C. Berwick
Damir Ćavar
Carlo Cecchetto
Gema Chocano
Barbara Citko
John Collins*
John E. Drury
Naama Friedmann
Martin Haiden
James R. Hurford
Gunnar Hrafn Hrafnbjargarson
William J. Idsardi
Tommy Leung

Terje Lohndal*
Víctor M. Longa
James McGilvray*
David P. Medeiros*
Andrea Moro
Salikoko S. Mufwene
Hiroki Narita
Massimo Piattelli–Palmarini*
Anne Reboul
Marc D. Richards
T. Daniel Seely
Tobias Scheer*
Robert Truswell*
George Tsoulas
Charles D. Yang
Hedde Zeijlstra