Generative linguistics within the cognitive neuroscience of language

ALEC MARANTZ

Abstract

Standard practice in linguistics often obscures the connection between theory and data, leading some to the conclusion that generative linguistics could not serve as the basis for a cognitive neuroscience of language. Here the foundations and methodology of generative grammar are clarified with the goal of explaining how generative theory already functions as a reasonable source of hypotheses about the representation and computation of language in the mind and brain. The claims of generative theory, as exemplified, e.g., within Chomsky's (2000) Minimalist Program, are contrasted with those of theories endorsing parallel architectures with independent systems of generative phonology, syntax and semantics. The single generative engine within Minimalist approaches rejects dual routes to linguistic representations, including possible extra-syntactic strategies for semantic structure-building. Clarification of the implications of this property of generative theory undermines the foundations of an autonomous psycholinguistics, as established in the 1970's, and brings linguistic theory back to the center of a unified cognitive neuroscience of language.

1. The place of linguistics[1]

The first decade of the 21st century should be a golden era for the cognitive neuroscience of language. Fifty years of contemporary linguistic analysis of language can be coupled with a wide range of brain imaging and brain moni-

1. I would like to thank Noam Chomsky, Dave Embick, Wayne O'Neil, David Poeppel and two anonymous referees for helpful comments on an earlier draft of this article. I have stubbornly resisted some of their suggestions, to the probable detriment of the final product.

0167–6318/05/022-0429

toring machines to test hypotheses and refine theory and understanding. However, there is still a gulf between mainstream linguistics within the generative linguistic tradition and most of those engaged in experimental cognitive neuroscience research. Some have argued that the fault here lies with the linguists, whose generative theories are based in principle on separating the study of linguistic representations from research on the acquisition and use of language in the minds and brains of speakers. A new linguistics is required, it is claimed, to allow experimentation to bear on the theory of linguistic representations and computation. Perhaps the most prominent of these alternatives at the moment is Jackendoff's (2002) parallel architecture with its autonomous syntactic, semantic, and phonological generative engines embedded in a research program that rejects the interpretive phonology and semantics of the standard generative theory. Trends in "Construction Grammar" and other forms of emergent grammar share properties with Jackendoff's system.

As a generative linguist who runs MEG brain monitoring experiments (Pylkkänen and Marantz 2003; Pylkkänen et al. 2004), I experience no gap between generative theory and psycho- or neurolinguistic experimentation. If standard linguistic theory is nevertheless perceived as divorced from cognitive neuroscience, generative grammarians perhaps suffer from a public relations problem rather than a fundamental methodological confusion. A brief reexamination of the generative enterprise should serve to clarify the promise of mainstream linguistics for cognitive neuroscience and question the apparent urgency to back alternative approaches to linguistics and to the study of language.

I aim, then, to re-situate generative grammar within the general program of experimental cognitive neuroscience. It would be another project to sketch the ways that linguistics might be merged with neuroscience – see Embick and Poeppel (2004) for a promising beginning. Nor will I provide a critique of Jackendoff's conception of grammar and related approaches; for such a critique from a point of view compatible with this article, see Phillips and Lau (2004). And I should not be taken as claiming that neurolinguistic results argue in favor of generative grammar over alternative linguistic architectures. Data of any sort would be decisive only to competing hypotheses within a particular broad theoretical framework. Arguments between theoretical frameworks are generally made at the conceptual level; to the extent that proponents of the different frameworks can agree on the nature of the enterprise and the relevance of specific data to the enterprise, the frameworks will converge on the same solutions and explanations as the data collect.

2. Foundations of generative grammar and psycholinguistics

At the birth of generative grammar in the 1950's and 60's, linguistics placed itself within a nascent cognitive science community integrating behavioral psychology, computational theory, trends in philosophy, and structuralist linguistics. Mainstream generative linguistics still operates at the nexus of computation, philosophy of language, and cognitive neuroscience. However, sometime in the 1970's it became legitimate for the study of language in psychology and computer science departments not to explore the ongoing discoveries of the generative linguistic tradition.

Symbolically, the break between linguistics and the general scientific study of language can be dated to Fodor, Bever and Garrett's *The Psychology of Language* (1974). Here a failure is announced: the interpretation of generative grammar as a psycholinguistic theory of linguistic computation was falsified by the empirical failure of the "derivational theory of complexity." "The discovery that psycholinguistics has a subject matter – a body of phenomena which are systematic but not explicable within the constructs manipulated by formal linguistics – is, perhaps, the most important result of the last decade of psycholinguistic research" (Fodor et al. 1974: 369).

We will return below to the "derivational theory of complexity," which is just a name for a standard methodology (perhaps the dominant methodology) in cognitive neuroscience. It would be of some historical interest to examine the "body of phenomena" that Fodor et al. thought was beyond the "constructs manipulated by formal linguistics" and ask whether they motivate the establishment of an independent field of psycholinguistics. But history aside, the important question today is whether any cognitive neuroscience of language must have generative linguistic theory at its core. Turning the question around, should generative linguistic theory be responsible for the "body of phenomena" revealed by psychological and neurological approaches to language?

A standard answer to these questions is that the distinction between competence and performance shields linguistic theory from some sorts of evidence and allows for a science of language independent of generative grammar. Linguists, studying competence, would deal with native speakers' intuitions about (primarily) grammaticality, while the cognitive neuroscience of language would build theories that made predictions about reaction times and brain activations. As far as I can tell, this view only makes sense if linguistics were Platonist, as claimed, e.g., by Katz (1981). As a Platonist, for grammatical analysis one would rely exclusively on a certain set of intuitions. But most contemporary linguists would deny Platonist assumptions and all current theory is based on standard scientific methodology, not the methodology of Platonism. That is, although only Platonism justifies the commonplace (and mistaken) interpretation of the competence/performance dichotomy, generative linguistics is not and has never been Plantonist in theory or practice. The conclusion, then, is that the competence/performance distinction, whatever one wants to make of it, does not shield linguists from any sort of data nor does it shield cognitive neuroscientists from the discoveries of linguistic theory. The original conception of a generative grammar rooted in cognitive science both underlies the major successes of linguistics and serves as a sound foundation for the integration of neuroscience with the cognitive science of language. The main confusion over these issues seems to stem from a misunderstanding of the evidence used in standard generative linguistics.

3. Data

On a Platonist linguistics, "language," the target of linguistic theory, is an abstract object about which speakers have intuitions. The evidence for the theory of a grammar for a language, then, are the intuitions of speakers about sentences potentially belonging to the language. Such intuitions can be judgments of grammaticality or judgments of synonymy or logical entailment. Behavioral data from performance in language production and comprehension are irrelevant for linguistic theory proper and fall under the domain of the independent field of psycholinguistics. As observed above, generative linguists are not Platonists; they do not work under the restrictions that Platonism would place on linguistic theory. Rather, for the quotidian linguist, judgments of, e.g., grammaticality **are** behavioral data, and the connection between such data and linguistic theory should follow the standard scientific methodology of cognitive science. However, the linguist generally does not present his/her data as the result of a behavioral experiment. To appreciate the generative linguistic enterprise, one needs to understand the connection between the data relevant to linguistic theory and the example sentences used to represent these data in the literature (for related discussion, see Schütze 1996).

3.1. Judgments of grammaticality

An article on syntax might contain a number of sets of examples annotated with patterns of asterisks ("stars"), question marks and percentage signs, indicating various judgments within or across native speakers about the examples. For the most part, these examples themselves do not constitute "data" in the usual sense from cognitive psychology; rather, they stand in for potential data – they are summaries of the results of experiments the reader could perform at home (if s/he were a native speaker of the language under investigation). Although it is easy to be careless with such examples, there is nothing problematic in general about this kind of "meta-data." However, the use of examples as *representing* rather than *reporting* data is not generally made explicit to those outside the discipline.

First, one must keep in mind that "grammaticality" is technically defined within a linguistic theory: a sound/meaning pair is grammatical according to a grammar if the grammar generates or assigns a representation to the pair that meets all well-formedness constraints and/or conditions within the grammar. Speakers do not have intuitions about "grammaticality" in this sense. One can investigate what sorts of phonological and semantic representations a speaker may assign to a given orthographic or acoustic sequence, asking questions about whether a phrase or sentence sounds OK to the speaker with a particular meaning. One can't ask a speaker whether his/her grammar generates a particular sound/meaning correspondence.

So an example sentence in a linguistics article – a sequence of letters (or phonetic symbols), spaces and punctuation on a page – may stand in for a sound-meaning connection, or a set of such connections, which can be judged acceptable or not by experimental subjects. A reading of the syntax literature over the last thirty years or so would uncover different classes of example sentences used to different ends, but a short review of some of the main types may be instructive here. The first sort represents various examples of "word salad" – sequences of words and/or morphemes that can't actually be assigned a phonological representation (can't be pronounced as a phrase or sentence). These examples are often used to illustrate the consequences of grammatical options that are clearly not motivated for the language in question. For example, (1a) might be cited to show what sentences would look like if English were a head-final language like Japanese; (1b) illustrates a pseudo-English with multiple wh-movement and no locality constraints on movement.

(1) a. **Man the book a women those to given has.*
b. **What whom did the man that saw ask which book Mary gave to?*

For "word salad" examples, psycholinguistic experimentation is in general neither required or motivated; the points illustrated by examples such as those in (1) are not advanced by asking a set of subjects to attempt to read them in an experimental setting.

The second sort of example is meant to illustrate non-controversial generalizations about a language. Some consideration of experimental methodology might be useful for syntacticians here, since although the examples themselves should be judged as indicated by any native speaker of the relevant dialect, the question of the sampling space is important – what space of types does this particular token example stand in for? All sentences like *this* example should be judged as indicated (fine or bad), but what exactly does it mean for a sentence to be like *this*? Typical for these sorts of examples are illustrations of local generalizations about word order and uncontroversial statements about agreement and case marking. For example, the contrast in (2a,b) illustrates that adjectives without complements are ordered before the noun they modify in English while adjectives with complements are ordered after. The sentences in (2c,d) exemplify a requirement that verbs agree in number with their subjects in English.

(2) a. *The scared man jumped from his seat.*
**The man scared jumped from his seat.*
b. **The scared of porcupines man jumped from his seat.*
The man scared of porcupines jumped from his seat.
c. *The men are leaving.*
**The men is leaving.*
d. *The man is leaving.*
**The man are leaving.*

While the particular examples in (2) and the modest generalizations they illustrate are uncontroversial, the scope of the generalizations should not be and is not taken for granted. The notions of "adjective" and "complement" relevant to the generalization in (2a,b) and the notions of "subject" and "number" relevant to the generalization in (2c, d) are all up for discussion and research.

Much of the controversy surrounding the use of judgments as data for linguistics concerns a third type of judgment, a contrastive judgment involving neither word salad nor categorical generalizations about a language. In this category one might put some judgments about locality domains for long-distance dependencies (i.e., constraints on wh- movement), judgments about possible co-reference among discourse entities (as for "binding theory") and judgments about semantic scope, e.g., relative scope of quantifiers. Here, for the most part, we have morpheme strings to which speakers can assign both a phonological representation – they know how they are pronounced – and a full or partial semantic representation. At issue are precisely the semantic interpretation associated with a particular phonological interpretation and/or judgments of well-formedness about pairings of sound and meaning. Here, experimentation is possible and useful.

The linguist presenting examples of this sort has already performed an experiment on him/herself or one or more informants. The task is something like a truth-value judgment task of the sort used with children (Crain and Lillo-Martin 1999): given a particular scenario, is this – a particular pronunciation of a morpheme string – an OK way of describing what happened (or, in the case, for example, of judgments of ill-formedness as with violations of locality conditions on long-distance dependencies, the question might be whether any semantic representation can be assigned to the phonological representation recovered). The linguist has made an implicit promise that (i) there is a relevant population of speakers for which the reported judgments hold, (ii) the example sentences provided are representative of a class of sentences as described by the linguist, and (iii) with speakers randomly sampled from the relevant populations and sentences randomly sampled from the relevant class, an experimenter would find more or less the same judgments that the linguist reports.

When looking at this third type of data, we can ask two questions. First, does the amateurish implementation of behavioral experimentation typical of linguistics articles undermine the conclusions of the linguist? Or, to put it differently, would the linguist reach different conclusions about the hypotheses being tested if s/he replaced the informal judgment methodology with carefully controlled experimentation? My own opinion is that most of the generalizations concerning grammatical sound/meaning correspondences that are central to debates in the linguistic literature are solid and that little of the critical literature questioning generative linguistics challenges these empirical generalizations. Nevertheless, all purported data in the literature are open to re-examination, as in any field.

The second question concerning real judgments of sound/meaning correspondences is, would the field of linguistics be better off at this point if linguists explicitly adopted the conventions of experimental psychology in presenting and discussing the third type of data described above? Here, I believe the answer is yes, and for a number of reasons (see Machery et al. 2004 for similar considerations; I should add that I believe I am in the minority among my generative colleagues in this conclusion). First, standard methodology really is effective only for the native speaker linguist, who may perform the relevant and necessary experiments at home. Linguists must always follow the methodology of experimental psychology when the languages discussed lack linguist native speakers. Without a native speaker linguist to perform the experiment on him/herself, the investigator must be careful to sample from a representative set of sentences, explain the nature of the judgments required, check across speakers of the same language/dialect, etc.

Second, and perhaps more importantly, explaining the relationship between hypothesis and data explicitly in an article requires thinking about the relationship between computation and representations in a way that tightens linguistic theory. As for any cognitive theory, prediction of experimental results requires at least a rudimentary theory of the task in the experiment. When the explicit task for a speaker is, "understand this string of words" (i.e., construct a sound/meaning pairing compatible with the presented string), we need to explain what we suppose is going on in the head of the speaker when s/he is successful at the task but reports a judgment of ill-formedness nonetheless, i.e., where a sentence is interpretable but apparently "ungrammatical." For example, what sort of judgment should we expect from a speaker about a string like that in (3) if constructing the semantic representation that speakers apparently assign to the sentence involves violating a grammatical constraint?

(3) ?**Which man did you ask whether I saw at the park?*

If we suppose that constructing a representation in which "which man" is interpreted as the object of "saw" involves movement violating a locality restriction on long-distance dependencies (whatever accounts for the "wh-island" generalization in English), what do we expect speakers to do as they read or hear the sentence and assign it a phonological representation? What would it mean for the computational system of the language to "do the movement" in violation of the locality condition? One could easily imagine a theory in which sentences like (3) should be predicted to be a word salad, i.e., not assignable a semantic or phonological representation. But this seems contrary to fact – in the literature, sentences like (3) are considered "mild" violations of locality constraints that are fully interpretable. In assessing data such as subjects' interpretation and judgment of (3), the representational and performance issues are only separable under particular accounts of what the subjects are doing in understanding the sentence. For example, one might claim that the subjects generate a fully grammatical structure for (3), but one that should have a phonological interpretation containing a ("resumptive") pronoun "him" after the verb "saw." On this account, speakers understand sentences like (3) as they would speech errors, constructing the "correct" representation and projecting the mistake. This is a testable hypothesis about sentences like (3) and may well be misguided, but we are owed some such account of the relationship between the linguistic theory and the data used to support it, in this case and in general.

3.2. Description and prediction

When properly construed, all judgments of well-formedness and of possible sound/meaning connections are measured behavioral data from experimental subjects. As such, the standard meat and potatoes of the theoretical linguist do not differ from the everyday bread and butter of other cognitive psychologists. Nevertheless, articles in linguistics don't often look like articles in cognitive science, with data presented as the results of experiments performed to test hypotheses. Instead, much work in linguistics aims to be "descriptive" of data rather than predictive of experiments not yet performed. Grammars for various languages are sketched, including accounts of the phonological inventory, the morphological structures, the syntactic constructions, etc. The question arises, what does a grammar describe when it's being descriptive?

The competence/performance distinction is meant to emphasize that a grammar is about the representation and computation of language, not about sentences per se – not directly about the utterances and behaviors of speakers. But one gains the impression from much linguistic writing that grammars in fact are descriptions of data rather than hypotheses about computation and representation. In the American structuralist tradition, a grammar was a reduction of a set of data, where the data were corpora of utterances, perhaps gathered as texts of stories told by speakers of an indigenous tribe. Much work in linguistics is exploratory and, in practice, not easily distinguishable from that in the structuralist tradition. Generalizations about the categories and structures of a language are formed from (distributional regularities within and across) utterances and/or written material gathered from recorded or written sources or via interviews and questionnaires, formal or informal. The structuralist would operate under the methodological constraints that generalizations about collected data were not meant to project beyond the data collected and that the grammar of one language should not necessarily be expected to conform to the grammar of the next language over.

Modern generative grammarians, on the other hand, are conscious of the predictive aspect of grammatical description, expecting their generalizations to cover "similar" examples from a language, and also expecting that many generalizations should follow from linguistic universals true of all languages, whether or not these universals reflect language-specific or even species-specific constraints on language. In practice, however, descriptive linguistics under the structuralist and generative traditions looks very similar, treating distributional generalizations about categories and structures within collected utterances and written sentences.

The standard use of example sentences in linguistic articles blurs the distinction between descriptive and predictive goals. The pattern of judgments used as crucial evidence to decide between two proposed principles, for example, might have been the basis for the formulation of the winning principle rather than a discovery made through hypothesis-testing. Nevertheless, the generative linguistic enterprise involves treating grammatical principles and analyses as predictive and treating utterances and judgments as behavioral data.

Why, then, do the data in linguistic articles look so different from that in much of the literature in cognitive science? For example, reaction time data of the sort most commonly gathered by psycholinguists – in a variety of experimental situations often involving priming or interference paradigms – do not play a major role in the work of most generative grammarians. One reason for this discrepancy is that many linguists are more concerned with static aspects of linguistic representations – the inventory of categories and their structural relations – than with computational mechanisms. Distributional data of the sort extractable from recorded utterances provide rich material for theories of representation. But another reason to disfavor, e.g., reaction time data from controlled behavioral experimentation is pragmatic: there has been an abundance of cheap data to keep linguistics busy without relying on experiments of the sort that involve significant expenditure of time and money relative to the hypothesis-testing value they provide.

Progress in linguistics has made it less rewarding recently to rely on standard distributional evidence. First, as explained in the next section, our understanding of linguistic representations currently emphasizes the dynamic nature of such representations such that the computations involved in their generation are crucial to their well-formedness. Evidence about dynamic processing, then, becomes more central to the questions of categories and their distribution than previously thought. Second, in many areas we are running out of new cheap distributional data. Years of research have yielded massively exemplified generalizations, at both a language particular (for well-studied languages) and universal level, about the inventories and distribution of sounds and phonological features, about morphemes and words, and about phrases and sentences. Linguistic issues now frequently arise that are difficult to think about and to settle with distributional data and with judgments of well-formedness and meaning; the competing theories involved all account for the numerous known generalizations about these data and do not obviously differ in their predictions about similar data.

4. Representations and dynamic models

The strongest tradition in psycholinguistics, one stemming in part from Fodor, Bever and Garrett (1974), supposed that although the linguistic representations supported in linguistic theory were "psychologically real" (part of a speaker's knowledge of his/her language), the generative mechanisms proposed to create these representations, although perhaps one route speakers could take to the representations, did not constitute the sole computational means available to speakers for creating the representations of their language. Rather there might be psycholinguistic strategies for structure-building that to some degree bypass the syntactic rules and computations of linguistic theory. Progress in all areas of linguistics, particularly on the syntax/semantics interface, has clarified the claims that linguists make about computations and representations. Today, the Minimalist Program (Chomsky 2000, 2001) illustrates perhaps the most straightforward interpretation of generative linguistic theory. In this approach to grammar, there is only one generative engine of language – the syntax – and only one route to grammatical representations – through the computational mechanisms of syntax. Therefore, were there in fact "psycholinguistic support" (i.e., evidence of any sort) for "strategies" for building linguistic representations without using the computations of syntax, these would constitute an alternative hypothesis about linguistic knowledge, not a supplement to the generative theory. That is, if the strategies are right, the theory is wrong (in need of modification).

In Chomsky's Minimalist Program and related approaches, the minimal combinatory units of language subject to both semantic and phonological interpretation, traditionally known as morphemes, combine via a recursive "merge" operation, creating hierarchical constituent structures. Each element and each merger involves a promise to be interpreted both phonologically and semantically. Within a local domain (called now by Chomsky a "phase"), the derived structure is submitted for interpretation in sound and meaning. Whenever a speaker or listener is operating with a linguistic representation, the representation must be built via the computation machinery described above. Thus in every psycholinguistic experiment, the mechanisms of the computational system should be evident.

Recall that a motivation for an independent psycholinguistics was the apparent failure of the "derivational theory of complexity" (DTC) – the hypothesis that the number of operations that the grammar uses to generate a sentence should correlate with, e.g., the reaction time of speakers processing the sentence in some psycholinguistic task. One could reexamine both the generative theories under evaluation by DTC experiments and the experiments themselves to see whether the experiments in fact disconfirmed the theories and whether the theories were well-motivated by other data. But, history aside, linguists really have no choice but to embrace the derivational theory of complexity, since it's essentially just a name for standard methodology in cognitive science and cognitive neuroscience. All other things being equal, the more complex a representation – the longer and more complex the linguistic computations necessary to generate the representation – the longer it should take for a subject to perform any task involving the representation and the more activity should be observed in the subject's brain in areas associated with creating or accessing the representation and with performing the task (see, e.g., Phillips et al. 2005 for a recent application of this reasoning). The Minimalist Program, in denying multiple routes to linguistic representations, ties linguistic theory closely and comfortably to this standard methodology, but hypotheses within all generative theories are testable via standard experimental techniques that correlate representational and computational complexity with behavioral and neurological dependent variables, with more or less complication spelling out the "all other things being equal" assumptions.

Embracing the DTC in its most general sense should help linguists demystify the nature of linguistic representations and computations. In addition to making predictions about complexity, linguistic theories make claims about similarity and identity between representations and between their constituent pieces. These claims are straightforwardly tested in standard priming and interference experiments (see the research discussed in Pylkkänen and Marantz 2003 and Pylkkänen et al. 2004 for some work along these lines).

5. Building on discovery

The short explication of linguistic methodology provided above should clarify how generative linguistic theory serves as a theory of language within cognitive neuroscience. The categories and operations of generative grammar are hypotheses about the representations and computations in the minds and brains of speakers. The Minimalist Program makes the claims of generative theory more explicit and thus allows for more straightforward testing and falsification of linguistic hypotheses. In essence, this "program" is built on the claim that there are no "dual routes" to linguistic representations; the syntactic computations described in the theory are necessary to the representations that they derive and thus speakers and listeners must carry out these computations whenever they have access to and manipulate the representations.

In contrast to the Minimalist Program, some currently popular theories of grammar have endorsed versions of a parallel structure account of language in which autonomously generated syntactic, semantic and phonological structures are linked via mapping rules. The very existence of these rules, one supposes, is meant to explain the appearance of isomorphism among the levels of structure, while the notion of a "lexicon" of stipulated mappings among levels accounts for the apparent mismatches among structures at different levels as apparently observed, for example, in idioms (but see, e.g., McGinnis 2004).

One might be tempted to weigh various proposals about the architecture of grammar on the grounds of restrictiveness. So, for example, Jackendoff's parallel architecture would seem less restrictive than the Minimalist Program in allowing more degrees of freedom in the analysis of phenomena. However, at the level at which details of different grammatical "programs" are filled in, a comparison on restrictiveness would be slippery at best. Rather, here I will explain the relationship between the structure of generative linguistic theory and three fundamental insights into the structure of language on which all contemporary linguistic theories are built. The first insight is that of (weak) compositionality: the meanings of linguistic expressions are built up recursively from the meanings of the constituents of the expressions, where a recursive definition of "constituent" decomposes phrases eventually to atomic sound/meaning connections (atomic "morphemes"). The second insight is that of interpretive phonology: phonological well-formedness at all levels of analysis requires reference to syntactic structure, i.e., to morphemes and their combination. The final insight is that of locality: semantic and phonological interpretation are negotiated within local structural domains. Parallel architectures generally require independent constraints to account for these insights while in the Minimalist Program and related generative approaches, these insights follow from the basic structure of grammar.

The well-formedness of a linguistic structure is understood to be recursively defined. That is, one asks about a structure C whether it contains pieces or is an atomic unit. If it is an atomic unit, one searches one's list of atomic units, and if C occurs on this list, the structure is well-formed. If C contains pieces, it is well-formed if each of the pieces is well-formed and the method of composing the pieces into C is licensed/well-formed. Each of the pieces constituting C might itself be atomic or consist of other pieces. This recursive definition of well-formedness assumes a bedrock of listed atoms for composition. It also implies a hierarchical constituent structure, with levels of embedding of complex (non-atomic) constituents.

Within the Minimalist Program, hierarchical structure is generated via a general merger operation in the syntax. Two elements, A and B, are "merged" to create a constituent C. These elements may be atomic or already the result of a previous merger. If A is also an internal constituent of B, then A has "re-merged" (merged again, in what Chomsky calls "internal merge"). The possibility of re-merger allows for the generation of what is sometimes called syntactic "movement" or displacement, where a single constituent acts within the grammar as if it is occupying two structural positions. Every description of constituent structure, be it generative rule or passive constraint, implies at least the basic structure-building (or structure-licensing) notion of merger. The claim embodied in generative grammar is that semantic and phonological constituent structures are "interpretations" of a single syntactic structure, not structures built or licensed by independent rules or constraints that might be described via the general merge operation. That is, phonological and semantic well-formedness are defined recursively with respect to syntactic structure, not with respect to phonological and semantic representations.

Although every atomic element in the syntax and every merger of elements promises to be interpreted both phonologically and semantically, standard generative grammar, unlike versions of, e.g., Montague grammar, does not claim that the interpretation happens as soon as the elements are merged. Rather, interpretation is cyclic, within domains now called phases. Phase-based interpretation allows for some apparent mismatches between semantic and phonological structure, including those associated with the displacement (movement) generated via re-merger. It also allows for the contextual interpretation of atoms within a local (phase) domain – contextual allomorphy in the case of phonological interpretation (plural is pronounced "-s" after "cat" but "-en" after "ox") and contextual polysemy in the case of semantic interpretation (consider the different meanings of "paper" in "term paper" vs. "torn paper").

Compositionality, of a "weak" variety as this is usually understood in the semantics literature, is demanded by this grammatical system in a number of ways. First, the semantic interpretation of each phase must be used in the interpretation of any phase in which it is embedded. Contextual interpretations can't reach into a completed phase. Second, any semantic structure that results from the interpretation of syntactic structure must always be built via the syntactic structure. Thus, although it may be possible for semantic interpretation to add meaning not represented by syntactic pieces or structure, this would have to be a type of meaning that is never represented by syntactic pieces or structure. For example, it might be possible that the type of causation implied by resultatives such as, "John hammered the metal flat," is not represented syntactically by any causative head and thus by any syntactic relation between a causative head and the predication between "the metal" and "flat." If semantic interpretation adds this type of meaning to structures like these resultatives, then this type of causative meaning should never be syntactically represented with a causative head and the relevant additional syntactic structure – not in English nor in any other language.

The compositionality of semantics is mirrored in phonology, although certain obvious consequences of the cyclic phonological interpretation of syntactic structures are not often highlighted. Since syntactic structures are equivalently interpreted in both phonology and semantics, the consequences of compositionality may be illustrated in considerations of the relation between semantics and phonology (as mediated, of course, by the syntax). The structure of grammar prevents phrasal suppletion, where a simplex phonological structure expresses a complex semantic structure, at the same time ruling out the reverse situation, where a complex phonological structure serves as the expression of a simplex semantic structure. The case of phrasal suppletion would involve something like the past tense of the phrase, "kick the tires," being pronounced as "blick" and blocking the regular, "kicked the tires" (in the way that suppletive "went" expresses the past tense of "go" and blocks "goed"). Such suppletion would be allowed by most parallel architectures that map between independently generated semantic, syntactic, and phonological structures. Within a phase-based generative grammar, suppletion is contextual allomorphy, necessarily restricted to a local domain.

The case of a complex phonological structure serving as the expression of a simplex semantic structure would be illustrated by the idiom, "kick the bucket," if it really had the same semantic structure as "die." I certainly believe that the best evidence supports the hypothesis that phrasal idioms like "kick the bucket" always involve interpretation of their complex syntactic structure (see McGinnis 2002 for a discussion). But what's crucial to the predictions of generative theory here is clarification of the notion, "lexical entry." For someone like Jackendoff, idioms are lexical entries, like simplex words, that connect semantic, phonological, and syntactic structures. On such a theory, the literal bucket that holds water involves a different lexical entry from the bucket in the idiom – they are homophones, and a speaker interpreting a sentence with the phonological form of "bucket" in it must decide which lexical entry s/he heard. For a generative theory with cyclic interpretation within a local domain, the idiom, "kick the bucket," must contain the same lexical entry for "bucket" as that in "the bucket is full of water," and the relation between the literal and idiomatic buckets must be more like the polysemy relation between the papers of "term paper" and "torn paper" than the homophony relation between the banks of "river bank" and "savings bank." A speaker hearing "kick the bucket" does not need to decide which "bucket" s/he heard; there is only one lexical entry for "bucket." There's a growing psycho- and neurolinguistic literature showing clear implications for the polysemy/homophony distinction (see, e.g., Beretta et al. 2005; Pylkkänen et al. in press); the experimental paradigms already available may be extended to idioms to test the conflicting predictions drawn from the competing theories. I personally am already convinced by the arguments against the lexical theory of idioms presented in Egan (2004), who sketches a version of the polysemy view with interesting consequences for the behavior of phrasal idioms already observed in the literature.

In addition to compositionality, a second fundamental property of language is the "syntactic" foundation of phonological well-formedness. From the first formal studies of phonological structure, it was observed that even such low-level properties of a language as phonotactic generalizations depend on morphological structure: the same sequences of phonemes that are fine across a morpheme boundary (e.g., in English the /-ksθs/ of "sixths") might be ill-formed internal to a morpheme. From the point of view of acquisition, this dependence of phonological structure on syntactic structure is a good thing; phonotactic regularities can be used by the child to segment and organize syntactic atoms independent of a developed knowledge of vocabulary and other language-specific features of the language. In general, phonological well-formedness has always been assessed as the phonological well-formedness of a structure derived from the syntax. Even apparent independent generative mechanisms like the "prosodic hierarchy" are used in practice as constraints on phonological structures derived from the syntax, not as generators of independent parallel structures.

As already explained, the Minimalist Program, following in the tradition of standard generative grammar, builds compositionality and the dependence of phonological structure on syntactic structure into the architecture of the theory. Each atom of syntactic combination will be interpreted (or explicitly not interpreted) both in phonology and semantics, as will be also every licensed combination of elements, atomic or derived, formed via the recursive structure-building operations of the syntax. The syntax is the sole generative engine of the grammar, responsible for the recursive hierarchical structure of words and sentences. Both phonology and semantics are interpretive rather than generative. The locality of dependencies in phonology and semantics is partially a consequence of the cyclic (phase by phase) interpretation of syntactic structure (stronger locality restrictions may be a property of certain computations within a phase). Syntactic structures within a phase are interpreted both phonologically and semantically, where the interpretation of each phase is a necessary component of the interpretation of a phase containing it. The particular dependence of phonological well-formedness on syntactic structure is captured at a general level through the interpretative nature of the phonological component – the syntactic structure is fed to phonological interpretation – and at a more specific level through derivation by phase – the cyclic domains of phonological structure are provided by the syntax.

6. Conclusion

Standard generative linguistics as instantiated, e.g., in the Minimalist Program is already well-integrated into cognitive neuroscience. Through standard experimental methodologies like the "derivational theory of complexity," the well-developed representational and computational hypotheses of linguistics may be used to learn about how the brain stores and generates symbolic representations (this is of course true about any well-developed and empirically well-supported linguistic theory). In return, cognitive neuroscience will help us flesh out our linguistic theories and provide additional rich sources of data to supplement what is cheaply available through standard work with informants. In light of the remarks above, recent claims that generative grammar needs radical surgery to participate in cognitive neuroscience should seem quite unmotivated and misdirected.

Massachusetts Institute of Technology

References

Beretta, Alan, Robert Fiorentino, and David Poeppel (2005). The effects of homonymy and polysemy on lexical access: An MEG study. *Cognitive Brain Research* 24 (1): 57–65.

Chomsky, Noam (2001). Derivation by phase. In *Ken Hale: A Life in Language*, Michael Kenstowicz (ed.), 1–52. Cambridge, Mass.: MIT Press.

— (2000). Minimalist inquiries: The framework. In *Step by Step*, Roger Martin, David Michaels, and Juan Uriagereka (eds.), 89–155. Cambridge, Mass.: MIT Press.

Crain, Stephen, and Diane Lillo-Martin (1999). *An Introduction to Linguistic Theory and Language Acquisition.* Oxford: Blackwell Publishers.

Egan, Andy (2004). Pretense for the complete idiom. Ms., Australian National University, Canberra.

Fodor, Jerry A., Thomas G. Bever, and Merrill F. Garrett (1974). *The Psychology of Language*. New York: McGraw Hill.

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

Katz, Jerrold J. (1981). *Language and Other Abstract Objects*. New Jersey: Rowman and Littlefield.

McGinnis, Martha (2002). On the systematic aspect of idioms. *Linguistic Inquiry* 33 (4):665–672.

Machery, Edouard, Ron Mallon, Shaun Nichols, and Stephen P. Stich (2004). Semantics, cross-cultural style. *Cognition* 92 (3): B1–B12.

Phillips, Colin and Ellen Lau (2004). Foundational issues [Review article: Jackendoff (2002), Foundations of Language.] *Journal of Linguistics* 40: 571–591.

Phillips, Colin, Nina Kazanina, and Shani Abada (2005). ERP effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research* 22 (3): 407–428.

Poeppel, David and David Embick (in press). Defining the relation between linguistics and neuroscience. In *Twenty-First Century Psycholinguistics: Four Cornerstones*, Anne Cutler (ed.), Lawrence Erlbaum Associates.

Pylkkänen, Liina, Sophie Feintuch, Emily Hopkins, Alec Marantz (2004). Neural correlates of the effects of morphological family frequency and family size: an MEG study. *Cognition* 91 (3): B35–B45.

Pylkkänen, Liina and Alec Marantz (2003). Tracking the time course of word recognition with MEG. *Trends in Cognitive Sciences* 7: 187–189.

Pylkkänen, Liina, Rodolfo Llinás, and Gregory Murphy (in press). Representation of polysemy: MEG evidence. *Journal of Cognitive Neuroscience*.

Schütze, Carson (1996). *The Empirical Base of Linguistics: Grammaticality Judgements and Linguistic Methodology.* Chicago: The University of Chicago Press.