

## The Internal and the External in Linguistic Explanation

BRIAN EPSTEIN

*Croatian Journal of Philosophy,*  
Vol. VIII, No. 22, 2008

---

*Chomsky and others have denied the relevance of external linguistic entities, such as E-languages, to linguistic explanation, and have questioned their coherence altogether. I discuss a new approach to understanding the nature of linguistic entities, focusing in particular on making sense of the varieties of kinds of “words” that are employed in linguistic theorizing. This treatment of linguistic entities in general is applied to constructing an understanding of external linguistic entities.*

In a recent interview Peter Ludlow conducted with Chomsky, Chomsky mentions B.F. Skinner’s work with pigeons to illustrate his views on internalist explanation. The science of pigeon psychology, Chomsky asserts, is a matter of understanding strictly the internal states and pathways of the pigeon’s mind, including how those pathways react when they receive signals or impressions of various kinds. A scientist who wants to carry out the program of understanding a pigeon, Chomsky says, will in the end forget about the external environment in which the pigeon is embedded, but will investigate “what’s going on in the pigeon’s head on the occasion of sense.”<sup>1</sup> Similarly, it is unnecessary, unscientific, and probably incoherent to look outside the mind of the individual for explanations of the workings of language for that individual, on Chomsky’s view.

---

\* My thanks to audiences at the 2007 Dubrovnik Workshop on the Philosophy of Linguistics and at Virginia Tech, and for an ITSG grant from Virginia Tech CLAHS.

<sup>1</sup> Chomsky and Ludlow (2003)

Interestingly, even in his brief discussion of the pigeon experiment, Chomsky attacks Skinner in two different ways. His main criticism does not actually address internalism, but rather, it is criticism of Skinner on what we might call the “ontology of pigeon psychology.” He asserts, presumably correctly, that Skinner mistakenly categorizes the pigeon behaviors he observes as “bar-pressings,” when Skinner gets the pigeons to peck at bars in response to various stimuli. As Chomsky notes, what the pigeons are doing actually ought to be divided into two kinds of pecking categories, corresponding to two different instinctual behaviors: pecking for water, and pecking for a seed. In the pigeon brain, those involve different pathways, and a shortcoming of Skinner’s experiment is that it fails to distinguish these.

What is interesting about his criticism of Skinner, of course, is that it invokes features of the outside world, namely, seeds and water. Now, Chomsky immediately goes on to say that pigeon psychology ought to focus on the internals and forget about the seeds and the water; but at least for this moment in the analysis of the psychological pathways of the pigeon, it seems that the seeds and the water play an explanatory role.

Most attacks on Chomsky’s internalism, and his claims of the irrelevance of public language to linguistic explanation, have largely focused on the individuation of linguistic entities and properties. Semantic externalists have argued that certain if not all mental states are externally individuated, and teleofunctional accounts of certain linguistic entities have recently provided a concrete story of how such individuation might go. Inasmuch as linguistics traffics in externally individuated entities, it cannot be exhausted by an investigation of internal psychology. Externalist attacks are the ones that Ludlow addresses in detail in chapter 4 of his forthcoming book, and equally the metaphysics of linguistic entities is the topic that occupies much discussion of Michael Devitt’s *Ignorance of Language* and in recent work by Georges Rey on the existence of standard linguistic entities.<sup>2</sup>

In this paper I wish to pursue a different avenue to vindicating the role of

---

<sup>2</sup> Devitt (2006); Rey (2006b); Rey (2006a); Ludlow (2007).

external entities in linguistic theory. My intention here is to put off the questions of externalism, in the interest first of arguing for the role of external entities in linguistic explanation. It is not accidental that the seeds and water are involved in the explanation involving the pigeon, nor, I shall argue, will it be the case that in an ideal or future science of linguistics, to say nothing of the actual one, that external factors will justifiably be excluded from linguistic explanation.

A critical issue that has, I believe, made the role of external entities in linguistic theory difficult to treat, notwithstanding many efforts to do so, is the focus on E-language as the primary entity of concern. This choice has weighed the conclusions against the sensible construal of external linguistic entities. Even though I will argue that external linguistic entities are quite prevalent in linguistic explanation, it is in fact little surprise that we cannot find many uses for E-language per se in the details of specific elements of linguistic explanation. Because, in fact, we also do not have much use for I-language per se in the elements of a theory either. In the details of any theory, the system as a whole rarely figures into theoretical explanations, any more than cells as a whole figure into explanations of organelles, or bodies in explanations of the structures of anatomical parts. Rather, I will reframe the discussion of external factors in terms of the linguistic entities, internal and external, that are commonly used in linguistic explanations. The entities I will focus on in this case are words of various kinds, which I will roughly distinguish into I-words and E-words.

## **1. Varieties of words**

There is a small philosophical literature and a large linguistic literature on the nature of words. Philosophical treatments of the metaphysics of words have generally arisen in the context either of controversies over the type/token distinction or controversies over the modal status of T-sentences. A few proposals were made a number of years ago about the nature of words, including Devitt's suggestion in *Designation* that words are temporally extended and causally connected sets of physical tokens, Millikan's treatment

of words as reproductively established families, and Bromberger's account of words as linguistic types<sup>3</sup>. Somewhat more recently David Kaplan stimulated renewed discussion of the topic with an article proposing a theory of words in the interest of solving Kripke's puzzle about belief.<sup>4</sup> The linguistics literature has had different concerns entirely. The most sustained treatment of words is probably Di Sciullo and Williams' monograph *On the Definition of Word*,<sup>5</sup> but the topic is widely discussed in many subfields of linguistics. Naturally, theories of the lexicon are concerned with the nature of words, but also most books on morphology have at least a chapter on the nature of the word, and in phonology as well there has recently been a flurry of activity on interpretations of the "phonological word."

Most conspicuous in all this material is just how many different senses 'word' has. Here I want to start by clarifying the different kinds of word in these many contexts, but I want to do so with an eye to noticing and taking on the burden of explaining the relations among these as well. It is conceivable that a word like 'word' can come to refer to a number of different phenomena that are only superficially similar to one another, but it is also entirely reasonable to ask whether there is not a unifying explanation for that similarity.

#### *Philosophical treatments of words*

First, let me point out a couple of ways in which philosophers have taken 'word' to be ambiguous. Most obvious is the type-token ambiguity. In counting words, as Quine pointed out, we tend to count tokens in certain contexts and types in others.<sup>6</sup> We say that there are eight words in an inscription of the phrase, "A rose is a rose is a rose," but if we are counting the number of words in a child's vocabulary, we do not count 'rose' more

---

<sup>3</sup> Devitt (1981); Millikan (1984); Bromberger (1989).

<sup>4</sup> Kaplan (1990)

<sup>5</sup> Di Sciullo and Williams (1987)

<sup>66</sup> Quine (1987); Szabo (1999).

than once, regardless of the number of times the child utters or thinks it.

Another ambiguity brought out by the recent work on the metaphysics of words involves the role of semantic value in individuating words. Intuitively, homonyms like ‘bank’ and ‘bank’ are different words, though they have the same orthography and pronunciation. A simple way of explaining this intuition is to take a word’s semantic value to be an essential property of the word, so that homonyms are different words in virtue of having different semantic values. A different perspective on this “semantic essentialism,” however, has arisen from the literature on deflationary theories of truth. If the semantic value of a word is an essential property of that word, that seems to imply that T-sentences (e.g., “‘Snow is white’ is true if and only if snow is white.”) are necessary, rather than contingent. This is a consequence that many philosophers have disliked;<sup>7</sup> consequently, in recent discussions of words most philosophers have assumed that semantic essentialism is false.<sup>8</sup> There is thus a potential divergence between an intuitive sense of ‘word’, in which semantic value figures into their individuation, versus a sense in which ‘word’ refers only to the sign, independent of semantic value. This point remains in dispute.

Whichever metaphysical line is advanced in philosophical treatments of words, however, there is one conspicuous piece of common ground: the primary instances of words about which these theories are constructed are not mental tokens. Some theories, such as Kaplan’s, do include mental factors in their accounts of word-individuation, but mainly these are theories of utterances and inscriptions, i.e., entities that have tokens or stages that are physical and largely external to individuals. This is in contrast to the variety

---

<sup>7</sup> Cf. Field (1994); Soames (1995); Price (1997); Garcia-Carpintero (1997).

<sup>8</sup> This is explicitly argued by Herman Cappelen, in a recent attack on stage-continuant pictures of words (Cappelen (1999)), such as those advanced by Devitt, Millikan, Kaplan, and Mark Richard (Richard (1990)). But many of these continuant pictures of words also distance themselves from semantic essentialism. In “Words,” Kaplan intends to give an account of the difference between ‘bank’ and ‘bank’ as different “lexico-syntactic entities.” This also seems to be the approach taken by Szabo in his recent “representation” theory of words (Szabo (1999)).

of ways ‘word’ is treated in linguistics.

*Linguistic treatments of words*

For any given architecture of the language faculty, there are plausibly a number of senses of ‘word’. A major complication for disambiguating the senses of ‘word’ in linguistics is that there is far from a consensus on linguistic architecture, with each approach expanding or collapsing the distinctions that others make.<sup>9</sup> The best known distinction between kinds of words in linguistics is typically drawn between the listed items, which Di Sciullo and Williams call “listemes,” and the syntactic words, which they call “syntactic atoms.” These they understand to be the items that are opaque to syntactic manipulation, the items that can be placed in  $X^0$  positions in syntactic structures. In *Aspects of the Theory of Syntax*, Chomsky uses ‘lexicon’ to mean the repository of all the words the speaker knows, and he does not pay much attention to larger and smaller units that have to be stored.<sup>10</sup> But it is clear that the distinction needs to be made. Idioms suffice to illustrate this, since the meaning of idioms cannot be derived from their parts, and yet they are not opaque to syntactic rules. For example:

- (1) a. He dragged Michael’s name through the mud.  
b. Michael’s name was dragged through the mud by him.  
c. \*He Michael’s dragged name through the mud.

Di Sciullo and Williams argue that a further distinction needs to be made between not just syntactic words and listemes, but also syntactic words and morphological words, or what they call “morphological objects.” The distinction between these is more complicated, since the boundaries of

---

<sup>9</sup> In light of the continual flux of views on linguistic architectures, it is almost surprising that we can even purport to talk about “standard linguistic entities,” since there is no standard or consensus about what even the basic categories of entities there are. To be sure, linguistic theory constantly makes use of a wide variety of entities or constituents, many of which have been part of linguistic theorizing for generations. At the same time, however, what linguistic entities there are depends on how the divisions are made among components of the language faculty. This is a particular problem for words.

<sup>10</sup> Chomsky (1965); Cf. Jackendoff (2002), p. 153.

morphology are more in dispute, but some simple examples show this distinction too to be very plausible.

First, it has long been noticed that a distinction must be made between listemes and morphological objects. Halle (1973) notes that word generation is productive, as in:

(2) missile, anti-missile missile, anti-anti-missile missile missile.

To distinguish morphological words from syntactic words, Di Sciullo and Williams give the example of romance compounds such as *essuie-glace* ‘windshield wiper’. Many romance compounds are opaque to syntactic rules, and thus act as syntactic atoms. For instance:

- (3) a. un bon essuie-glace  
b. trois essuie-glace  
c. Ils ont fait l’amour comme des essuie-glace.  
d. \*essuie-bien glace  
e. \*essuie-quelques-glaces  
f. \*essuient-glace

The compound *essuie-glace* in this way is different from the compound *timbre-poste*, which seems to act more like an idiom or listed phrase, and is not syntactically opaque.<sup>11</sup>

Other examples of romance compounds that are syntactically opaque are:

- (4) a. trompe-l’oeil, fend-la-bise, coupe-la-soif  
b. boit-sans-soif, pince-sans-rire, monte-en-l’air  
c. bon-à-rien, haut-de-forme, juste-au-corps  
d. home-de-paille, boule-de-neige, arc-en-ciel  
e. hors-la-loi, sans-le-sous, hors-d’oeuvre

Although these are syntactic words (on the opacity interpretation), nonetheless there is good reason not to take them to be morphological words. As Di Sciullo and Williams say, “if morphology were enriched to generate these directly, not only would the generality of morphological principles be compromised but the fact that all the examples are compatible with the laws

---

<sup>11</sup> One difference between *essuie-glace* and *timbre-poste* is that the latter is head-initial.

of French syntactic form would be missed.”<sup>12</sup>

A fourth variety of word in linguistic theory is that of a phonological or prosodic word, or “pword”, defined within the phonology component (and denoted by  $\omega$ ). Phonological words are an element of the phonological hierarchy, typically understood potentially to include utterances, intonational phrases, phonological phrases, phonological words, feet, syllables, moras, segments, and features.<sup>13</sup>

There is a good deal of evidence that the phonological words are an explanatory kind in phonological theory, but there is controversy over what the criteria for pword boundaries are. T. Alan Hall has recently reviewed three sets of considerations arguing for the role of the pword in phonological generalizations: its playing a role in the domain of phonological rules, in phonotactic generalizations, and for minimality constraints.<sup>14</sup> Booij (1984) and Nespor and Vogel (1986), for instance, argue that the explanation of Hungarian vowel harmony applies only when the trigger and target occur in the same pword. For phonotactic constraints, Booij (1999) and Hall (1999b) argue that in German and Dutch, there are constraints barring some sound or sequence of sounds occurring at the beginning or end of a pword.

Nevertheless, the criteria for pword boundaries are still in flux. Aronoff and Sridhar (1983) argue that stress patterns are the main criterion for pwords, for instance, while Raffelsiefen (1999) argues that pword boundaries depend on other elements of the prosodic hierarchy, i.e., syllable, mora, and foot.

Many linguists see the pword as involved in the interface between morphology and phonology, and assume that there is a connection between pword boundaries and morphological boundaries. Nonetheless, the pword differs from both the morphological word and the syntactic word. Szpyra (1989), for instance, employs a distinction between two different kinds of

---

<sup>12</sup> Di Sciullo and Williams (1987), p. 83.

<sup>13</sup> Cf. Selkirk (1984).

<sup>14</sup> Hall (1999a)

affixes to distinguish those that are incorporated in the pword of the stem to which they attach from those that do not, such as:

- (5) a. im+polite → (impolite)<sub>ω</sub>  
b. un+believable → (un)<sub>ω</sub> (believable)<sub>ω</sub>

Nonetheless, there may be limitations to the mismatch. Nespor and Vogel (1986) have argued that pwords cannot span two grammatical words, but others have argued that they can. For instance, in the phrase:

- (6) I'll cross the street.

Di Sciullo and Williams argue that 'I'll' is a phonological word, but it is not a word in any other sense. It is not a syntactic atom, since clearly the position of 'I'll' in the sentence is determined by the syntactic distribution of I and will, and hence is composed of syntactically accessible parts.<sup>15</sup>

Clearly, all these distinctions are controversial, and in this discussion I do not intend to evaluate or endorse these particular distinctions or ways of characterizing them. In fact, when we step back and consider them as a group, there are some peculiarities about how they are specified. Morphological words seem to be determined through the characteristics of the interface between morphology and other components, while phonological words are simply entities that play a role in one layer of the phonology component, and have an unclear relation to the interfaces with other units. Syntactic words also seem to be defined in terms of a role, but are defined through a single theoretical criterion, i.e. opacity to syntactic rules. It would not be inconsistent for all these characterizations to hold, but we can surmise that we have not yet arrived at a stable ontology in the theories of every one of these components.

A natural reaction is that given all this ambiguity, we might ask whether it even makes sense to speak of words in the first place, as opposed simply to introducing several different entities playing different roles. Indeed, these different uses of word need to be disambiguated, and kept track of. Moreover, even apart from the issue of the many kinds of words, some

---

<sup>15</sup> Di Sciullo and Williams (1987), p. 107.

linguists are skeptical that it is even coherent to speak of words within a linguistic component. A number of theories of syntax weaken or deny the distinction between syntax and morphology, and take morphemes to be the primary unit of syntactic operations. Marit Julien, for instance, has recently argued that words are not grammatical entities at all, but rather are a kind of linguistic epiphenomenon arising from distributional regularities in expressions constructed of morphemes.<sup>16</sup>

In linguistics we thus have two grounds for being skeptical of a single coherent notion of “word”: the ambiguity of the notion across linguistic components, and the difficulty of delineating it even when limited to a single component.

I will suggest that these two issues are related to one another, and that we can illuminate the nature of words within a component by considering their role in the language faculty altogether, that is, how the kinds of word-like entities in one linguistic component are related to those in another. Before we examine whether or how there can be unity in their differences, let us start with a way of keeping track of entities in linguistic components. We will see that there actually are far more than the four or six varieties I have mentioned.

## **2. Keeping track of linguistic entities**

Ramsey-Lewis sentences are a useful device for characterizing and keeping track of the differences among linguistic kinds.<sup>17</sup> As I will explain, I will use the sentences simply as characterizations of terms referring to these kinds.

To apply these sentences to linguistic kinds, consider the phonological word. While there may be an economical explicit definition of the phonological word, it is helpful to characterize it by means of a theory of the phonological component. A full theory of the universals of the phonological component will have clauses involving all the elements of the phonological

---

<sup>16</sup> Julien (2007).

<sup>17</sup> Cf. Lewis (1970); Lewis (1972); Lewis (1997).

hierarchy. For reasons I will talk about shortly, I will use the notation  $\overline{B}$  to denote the components as typically treated in linguistics, and  $\overline{T}$  to denote the typical theories (B and T without the bars will show up below). For the universal aspect of the phonology component, as typically treated by linguistics, I will use the notation  $\overline{B}_U^P$ , and for a good theory of  $\overline{B}_U^P$  I will use the notation  $\overline{T}_U^P$ .

Let us take as a simplified theory of phonology the following fragment, given by Jackendoff (2002) based on Selkirk (1984):

- (7) A version of  $\overline{T}_U^P$ :
- a. An utterance consists of a series of one or more concatenated intonational phrases forming a flat structure.
  - b. Each intonational phrase is a sequence of phonological words.
  - c. Preferably, the intonational phrases are of equal length.
  - d. Preferably, the longest intonational phrase is at the end.
  - e. Possibly, there is a maximum duration for intonational phrases
  - f. Etc.

Suppose that there are n theoretical terms in  $\overline{T}_U^P$ , and the term ‘phonological word’ is the  $i^{\text{th}}$  term among those n terms. Then the Ramsey-Lewis definition of the  $i^{\text{th}}$  term is given by the description:

$$(8) \quad D_i(T) =_{\text{def}} \neg y_i \exists y_1 \dots y_{i-1} y_{i+1} \dots y_n \forall x_1 x_n (T[x_1, \dots, x_n] \leftrightarrow y_1 = x_1 \wedge \dots \wedge y_n = x_n)$$

where  $T[ ]$  is the sentence obtained by conjoining the clauses of  $\overline{T}_U^P$  and replacing the n theoretical terms in  $\overline{T}_U^P$  with free variables  $x_1, \dots, x_n$ .<sup>18</sup>

---

<sup>18</sup> Lewis (1970), p. 88.

Ramsey-Lewis descriptions are most familiar in now-outdated functionalist theories of mind, so let me say a couple of words about my use of these descriptions. When a philosopher of mind characterizes a functional property with a Ramsey-Lewis description, she will often intend to suggest that the description exhausts the property, that the property is strictly functional and can be realized by any physical substrate satisfying the function. That is not the suggestion here.

The problem we have in speaking of linguistic entities is that even if we are realists about these entities, it is a big task simply to designate which entities we are speaking of. A theory of a highly complex system, such as of the language faculty, will involve terms referring to the “natural” kinds or entities in that system. But it is an illusion to suppose that there is a sharp distinction, in any special science, between the “natural” kinds and the non-natural ones. Rather, in linguistics as in other special sciences, we construct theories based on the most privileged or “quasi-natural” entities and kinds, those that give rise to the phenomena of concern, and that are involved in relevant explanations, in the science.

In a domain in which there are many candidate entities, as in the candidates for “words” in the language faculty, we therefore need a device for designating which quasi-natural one we are referring to with a particular term. This is where Ramsey-Lewis descriptions are useful. These descriptions allow us to designate entities using a theory as a whole, i.e., to take an implicitly described entity and construct an explicit description, presuming that a good theory employs just the most natural or privileged entities. That does not, however, imply any particular view on the metaphysics of linguistic entities.<sup>19</sup> To be sure, I do not employ Ramsey-Lewis descriptions to suggest

---

<sup>19</sup> If we are realists about linguistic kinds, then there will be true theories of the components in which the kinds play a role. A realist might say, with early David Lewis, that the terms are implicitly defined in the theory, and we hope that a true theory we construct will be uniquely realized, so that there is one actual kind that satisfies the definition. A different approach a realist might take is that the description does not define the term ‘phonological word’, but is intended to be a reference-fixer, so that if the theory is uniquely realized it fixes the reference of the term to the actual kind. If we regard linguistic entities as structures, then a realist would say that the

that linguistic entities are functional, but only as a device for being clear as to which entities we are referring to by a given term.

$\overline{T}_U^P$ , of course, is not the only theory in which “words” show up. Each linguistic component has its own theory, and there are also theories that consist of the conjunction of the theories of individual components. What is the relation among the “words” across these theories, except that we happen to use the word ‘word’ to refer to some kind in many of them?

Below, I will speak of some pressures in the design of multi-component systems for what I will call their “ontological matching” across components. For the moment, I will leave the idea intuitive, that among the kinds playing a role in a component of the language faculty, some are “wordlike,” either structurally or extensionally closely matching the intuitive notion of a word.<sup>20</sup> For notational simplicity, I will intuitively define the function  $W$  on a theory as one that chooses this word-like kind from those entities playing a role in the theory. If we take linguistic kinds to be defined implicitly in a theory, then the  $W$  function is simply (8) with an appropriate choice of the index  $i$ .

---

structural description given in a true theory gives the definition of the theoretical term. An anti-realist might agree that the structural description gives the definition of the theoretical term, but deny that there is a true theory. Even among the realists, this approach also leaves it open what the metaphysics of linguistic entities is, and in particular whether the entities are psychological or the objects of a nonpsychological but linguistic ontology.

My own view is that a quasi-realist picture of these entities is the right direction to go in, that is that linguistic entities are real kinds to the extent that any entities in the special sciences are. Some are individuated strictly in terms of internal factors, some strictly externally, but most have a mix of individuating factors. This is the case for most special sciences, and in this respect the entities of linguistics are not particularly exceptional. All that is neither here nor there, for the purposes of the discussion in this paper.

<sup>20</sup> For any given component, there may be more than one wordlike kind, and there clearly may also be components that do not make use of any wordlike kinds. For our purposes, however, we will suppose there not to be multiple wordlike kinds in a given component, but rather just a single one that plays a dominant role in it, and from the discussion above, at least for the components under consideration, each does have at least one wordlike kind playing a role in it.

Thus in the case of the theory of phonological universals,  $W(\overline{T}_U^P)$  is the phonological word. There will be similar characterizations given theories of the syntax and morphology components, and potentially the lexicon.

Let us now consider a range of theories of the idiolect of a particular speaker, call her ‘Martha’. When we construct theory  $\overline{T}_U^P$  for Martha, it is meant to be a theory of that particular speaker’s phonological component. As long as we are defining the linguistic kinds determined by Martha’s idiolect, however, we should also notice that the entities playing a role in theories of the *universals* of Martha’s idiolect are not the only, or even the most intuitive, sorts of phonological, syntactic, and other entities that play a role in her idiolect altogether.

When we speak of linguistic entities, we tend to focus on the structures that figure into theories of linguistic universals. In constructing a linguistic theory of Martha’s phonological component, for instance, our typical intent is to construct a theory of the generic phonological component, of which Martha’s particular phonological component, with her parameters set as they are, is an instance. But the most natural systematization of Martha’s own idiolect is the one that does not abstract away Martha’s parameter settings, but rather gives a theory of her full-blooded idiolect, parameters set as they are. In developing theories of the structures of universal grammar, of course, we assume that we can reasonably hypothesize about the genus, of which each of our full-blooded idiolects is a species, and use the species data to infer the constraints on the structure of the genus. Still, we should note that a true theory of a full-blooded idiolect will involve its own theoretical terms, which will be implicitly designated in more detailed theories. There thus is more than one theory even of Martha’s phonological component, morphological component, and so on, corresponding to which there is more than one kind of phonological word, morphological word, and so on. The reason I mention this is that to make sense of the claims of matches and mismatches of words across components, it will not only be the universals that are relevant, but the matches once the parameters are set as well. In fact, it will be the latter that I

will focus the bulk of my attention on.

I will use the following notation for good theories of Martha’s phonological universals, morphological universals, syntactic universals, and so on.<sup>21</sup>

$$(9a) \quad \overline{T}_U^P, \overline{T}_U^M, \overline{T}_U^S, \dots$$

A full theory of Martha’s idiolect, as I mentioned, will consist of more than a theory of her universals, but includes her parameter settings as well. Presuming some variety of the principle-parameters approach is correct, then for any component *b* of her language faculty, there are values to which Martha’s *b*-parameters are set, the theory of which I will notate as:

$$(9b) \quad \overline{T}_V^b$$

Thus any component has a theory combining the principles and parameters, which I’ll notate as:

$$(9c) \quad \overline{T}_I^b = \overline{T}_U^b \cup \overline{T}_V^b$$

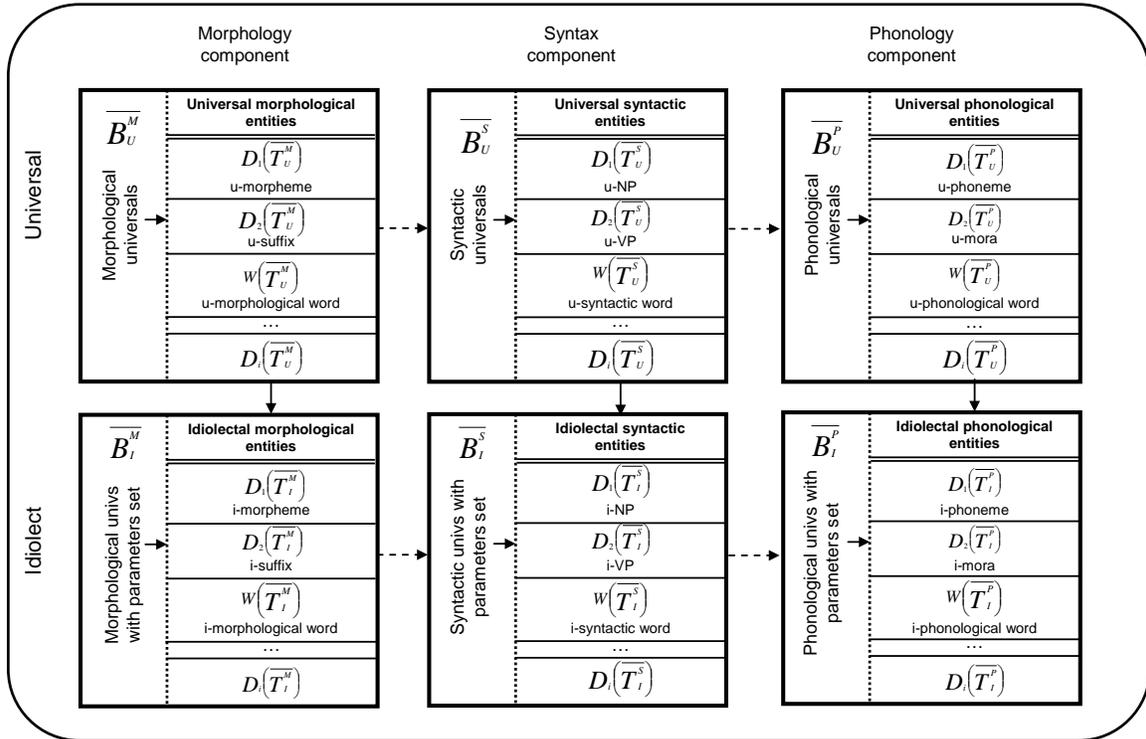
Figure 1 illustrates a simplified architecture of Martha’s language faculty, with the top row of boxes representing the universal structures of her morphology, syntax, and phonology components, and the bottom row of boxes represent the components including both the universals and their parameter settings.

The entities implicitly characterized in good theories of these components, i.e., those obtained by constructing Ramsey-Lewis sentences with the theories listed in (9a) and (9c), are the ones labeled as “universal” and “idiolectal” entities, within the respective components.

---

<sup>21</sup> I.e., true theories that relatively completely capture the functional characteristics of the components. For these purposes, I will assume that existing linguistic theories are close to good ones in this sense.

Fig. 1

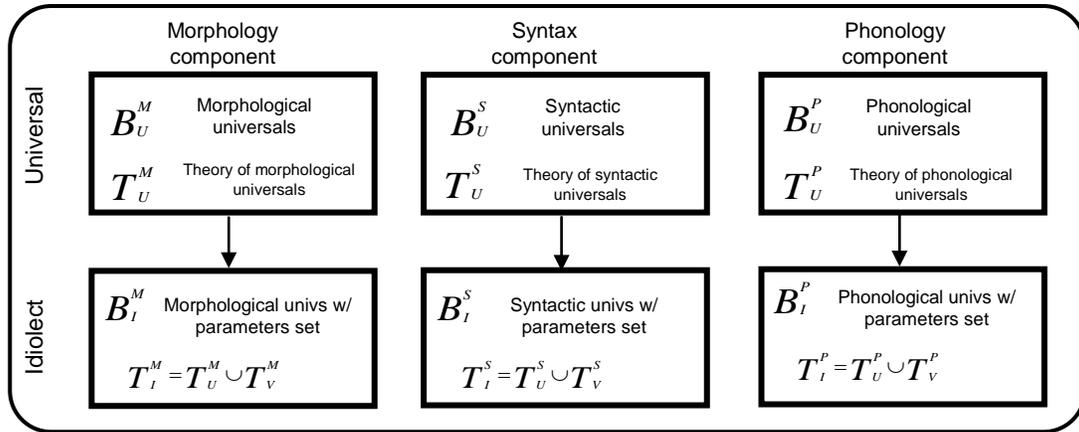


### System-constrained entities

One of the key complexities in designating linguistic entities arises from the fact that the language faculty, as depicted in figure 1, consists of components that are connected to one another. This means that an adequate theory of some component is not really a theory of the component itself. Rather, it is a theory of the component as constrained by the components that figure into the inputs to that component. This is a general characteristic of theories of modules in multi-part systems, and appreciating it helps clarify what exactly the quasi-natural entities are that we refer to in linguistic theorizing. It becomes particularly illuminating when we apply it to multi-part systems when we already have a theory of a component as isolated from the system. The “internal/external” linguistic system will be such an example.

To point this out, let us consider three different ways we can depict the components of Martha's language faculty. First is to consider the simplified language architecture depicted in figure 1, but without the connections among the components. That is, let us imagine that the various components were treated as isolated from one another. If we constructed theories of the isolated components, then we would only theorize about the characteristics and implied entities within each component.

(Fig. 2a) Isolated components



The component in isolation, i.e., with properties resulting from its internal structure rather than including the systematic constraints imposed by other elements of the system, I denote with B (absent a bar). Likewise a theory of a component in isolation is T without a bar. In the diagram, a division is again made between the universal and idiolectal aspects of the components, depicted in the top and bottom rows. For each module, the idiolectal component is the universal component with parameters set.

When the components are combined into a system, however, there is a change in a good theory of the syntax and phonology components. The theory of the universal syntactic component in isolation,  $T_U^S$ , is not the same as the syntactic component as it occurs in Martha's language faculty, i.e., as constrained by the inputs from the morphology component.

(Fig. 2b) Components constrained by system

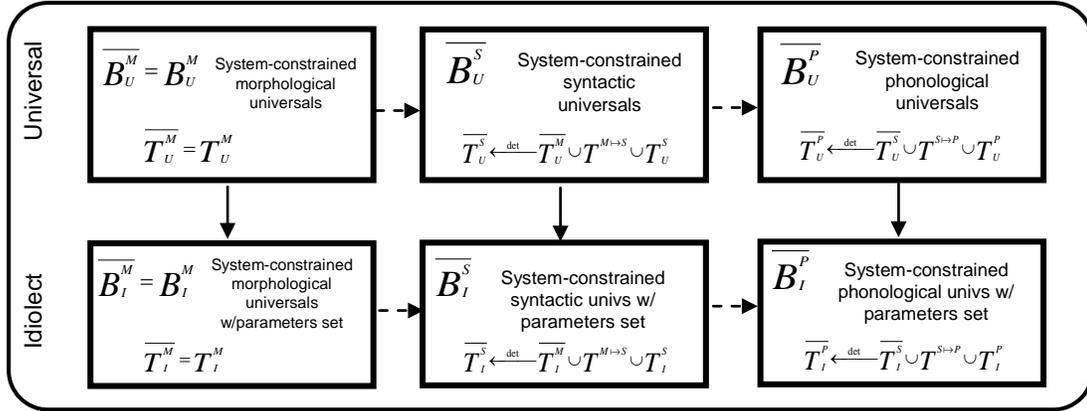


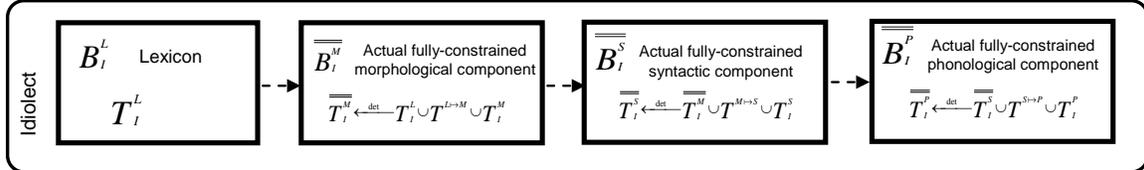
Figure 2b, which corresponds to the depiction in figure 1, incorporates the constraints on components imposed by the other components in the language faculty. Given a theory  $T_U^S$  of the syntactic component in isolation, together with the theory  $T_U^M$  of the morphological component and the theory  $T_U^{M \rightarrow S}$  of the connection between  $\overline{B}_U^M$  and  $\overline{B}_U^S$ , these suffice to determine the theory of the constrained universal syntactic component. This I have notated in the diagram as:  $\overline{T}_U^S \leftarrow_{\text{det}} \overline{T}_U^M \cup T_U^{M \rightarrow S} \cup T_U^S$ .<sup>22</sup> Likewise, the structures that will occur in the phonology component will be constrained by the structures of the syntax component and the syntax-phonology interface.

In a moment, I will point out why the typical linguistic theory is a theory of the constrained component  $\overline{B}$ , rather than the isolated component  $B$ . But let us also consider a third way of describing Martha's language faculty. This is actually the most natural way to treat describe it, but interestingly it is not the focus of linguistic theories. If we suppose that Martha is a competent

<sup>22</sup> It may be useful to think of a different modular system as an analogy, such as a mill and a bakery, or a CD player and a graphic equalizer. A good theory of a particular bakery, and the entities and kinds within it, will depend on whether it is treated in isolation, or whether it is connected to a corn mill or a rye mill or a wheat mill.

speaker, then the natural system for describing Martha’s language faculty is to incorporate the constraints imposed by her *actual* lexicon on the other components:

(Fig. 2c) Components constrained by actual lexicon



For each of the components, there is a set of quasi-natural entities that occur in that component, and that are designated in theories of the components. Figure 2c is plausibly the most intuitive theory of Martha’s idiolect, because the most obvious kinds of linguistic entities are just the kinds of actual items in Martha’s idiolect, fully constrained by her lexicon and everything else in her language faculty. That is, the obvious kinds of linguistic entities in Martha’s idiolect are her actual morphemes, actual phonemes, actual phonological words, and so on. By the “actual” ones, I do not mean the tokens that occur in performance, but rather those that simply are the morphemes, phonological words, and so on, in her actual idiolect as it is, with not only the parameters set, but her actual lexicon in place. These are the entities constrained by Martha’s actual lexicon, as well as all the constraints imposed by the structures of the components that provide input to her morphology, syntax, and phonology components. They are the sorts of entities we refer to most commonly when saying, for instance, that Martha has a forty thousand word vocabulary. I will call these entities her “idiolectal actual” (IA) entities.

But though they are less obvious, the quasi-natural entities of the components depicted in 2b are theoretically more useful. (I will call these components, which I have notated as  $\overline{B}$ , “system-constrained” components.) These entities are most properly what we seem to mean when we speak of the “possible words,” or “possible intonational phrases,” etc., of an idiolect, or of human language altogether. The phonological words that are “possible” for human language, for instance, consist of those that satisfy the universals of

the phonology component as well as the universals of the constraints on the phonology component imposed by the other components. And the phonological words that are possible for Martha include those that do not violate her universals and parameter settings for her phonology component and its input constraints, even though they do not appear in her lexicon.<sup>23</sup> I will call these entities the “universal system-constrained” (UC) entities, and the “idiolectal system-constrained” (IC) entities.

To illustrate the distinction between actual and possible words most clearly, it might have been preferable to use a slightly less simplified model of the language faculty, which captures the constraints imposed by phonological structures on derivational morphology. In a model that captured this, then the IC-words are the ones we speak of when we say that jabberwocky, brillig, slithy, and toves are possible English words, while bbrjwckaeoy is not. The violation, ruling these out, is a phonological one pertaining to the “sonority hierarchy,” and so even if they are not ruled out by morphological structures, the inputs to the derivational morphology component rule them out as possible outputs. For an idiolect, the IC-words are thus intuitively the “possible” words in a component, and as for the ones that might appropriately be counted as possible in any human language, it is the UC-words that seem to be the right ones.<sup>24</sup>

The standalone entities are just the entities that occur in the components in isolation, unconstrained by the other components in the linguistic architecture as a whole. These I will call the “universal standalone” and “idiolectal standalone” (IS) entities. In a weak sense, an IS-phonological-word, for instance, is a “possible” word, given the structure of the phonological component alone. Nonetheless, it is a broader kind than what we would normally consider the “possible” ones. There are many tokens that

---

<sup>23</sup> This account is somewhat more complicated if the lexicon is structured.

<sup>24</sup> In the first component in a chain, there is no distinction between standalone and system-constrained entities, which is the reason that in the simplified architecture in figure 2, the US and IS entities in the morphology component are also UC and IC entities.

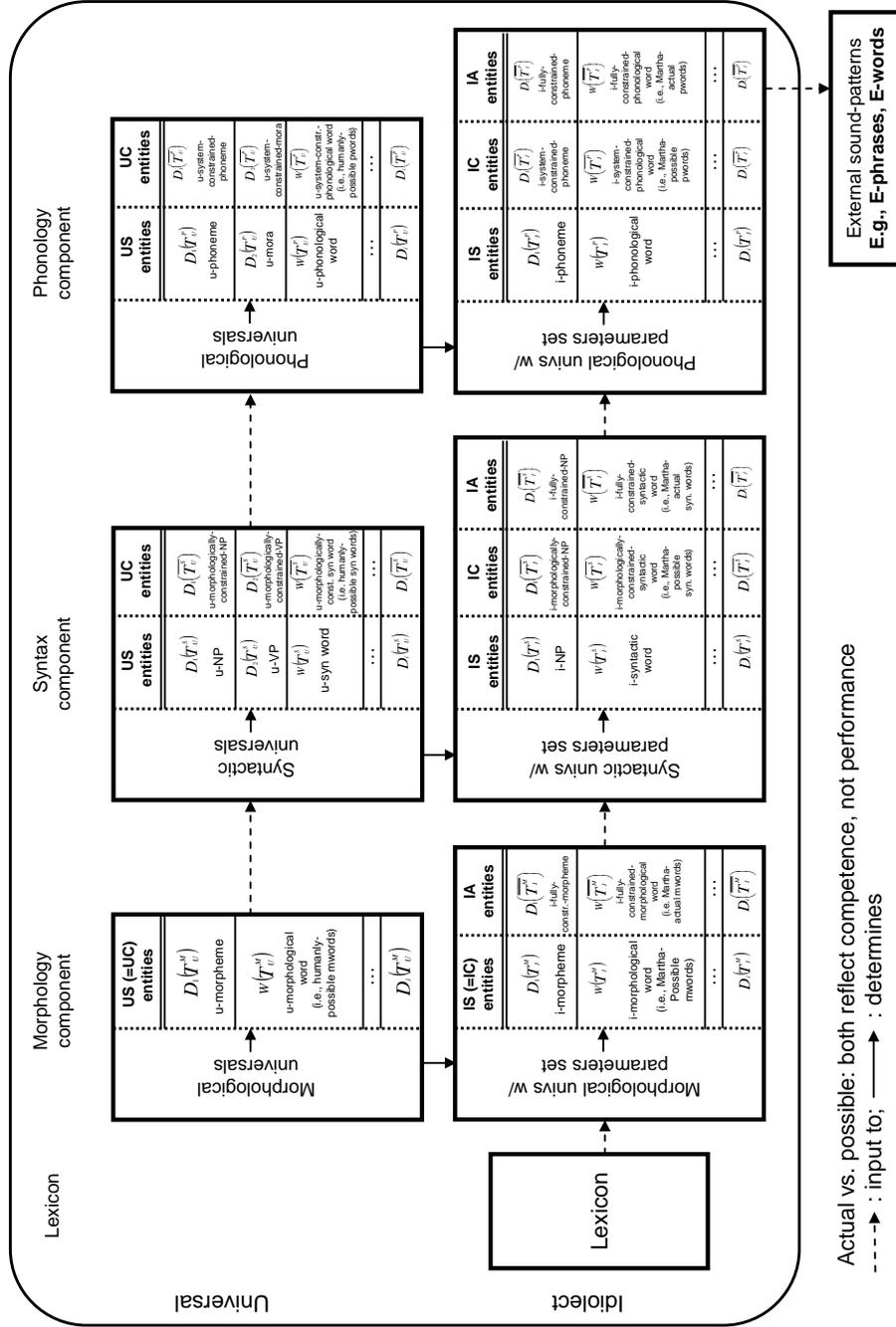
would satisfy the conditions given by the characterization of  $W(\mathbf{T}_i^P)$  that nonetheless could not be part of the idiolect; not because they do not happen to be lexical items, but because they are ruled out by the structures of the preceding components. Thus in order to pick out the intuitively “possible” phonological words in Martha’s idiolect, we ought to choose the idiolectal system-constrained (IC) phonological words. These are different from the actual (IA) phonological words, constrained by Martha’s lexicon, but are more constrained than just what is determined by the structure of her phonological component alone.

For any component of the language faculty, such as morphology or phonology or syntax, there are five kinds of entities: universal standalone (Type US); universal system-constrained (Type UC); idiolectal standalone (Type IS); idiolectal system-constrained (Type IC); and idiolectal actual (Type IA). Figure 3 is a depiction of the simplified linguistic architecture from figure 1, but now including all the entities that are implicitly determined by their being constrained by the elements the system that feed into them.<sup>25</sup>

---

<sup>25</sup> For simplicity, I have left out the theories of the connections between the components. A more sophisticated diagram would include bidirectional arrows, and a less simplified architecture for the system as a whole.

Fig. 3



With this, we have at least a baker's dozen theoretically characterized kinds that might count as different varieties of words. This is the thing about quasi-natural entities in the special sciences: there are lots of them. If we wish to take an anti-realist metaphysics of these, then there will be no trouble postulating a plenum of potential such entities, but then we have the task of accounting for which ones are salient to linguistic theory and why, and as far as I can tell, this is parallel to the realist's task of accounting for what the properties of the sparser (but still numerous) set of real entities is, in a particular domain.

### **3. The idea of ontological matching across components**

Simply introducing an operator to pick out the most wordlike kinds in different components does not, of course, mean that the kinds selected have anything to do with one another. If we were arbitrarily to choose a few separate domains of inquiry, construct theories of them, and then choose the most "wordlike" entities among the domains, then we would not expect there to be much resemblance among these entities. With linguistic components, however, the phenomenon is quite the opposite. In the first section I concentrated on pointing out that the word-kinds in each of the various components are not the same: the syntactic word is almost but not exactly the same as the morphological word, and both are almost but not exactly the same as the phonological word. That the kinds are mismatched is striking. But it shouldn't be overlooked that to the largest extent, they do match across components.

A few paragraphs back, for instance, in introducing IC-words with the jabberwocky example, I glossed over whether I was speaking of morphological or syntactic or phonological IC-words. The reason is that here there was no need to distinguish among them. *Bbrjwckaeoy* is not a possible word in any of these components, even though I argued above that the word-kinds across components are not extensionally equivalent to one another. The possibility of mismatch among the word-kinds only highlights the fact that there is a nearly perfect match of these kinds across the components, in spite

of the fact that the match is not entirely perfect. How and why are entities or kinds in one component coordinated with those in another? It is conceivable that it is simply an accident, but this seems unlikely. This is what we might call the problem of “ontological matching” among linguistic components.

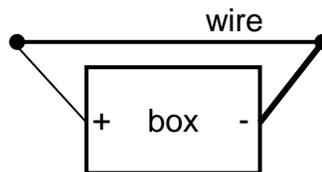
Interestingly, the phenomenon of ontological matching it is far from unique to linguistics, and in fact is a distinctive but common characteristic of multi-component systems. In considering how inter-component coordination can occur, I hope to shed some light on the interaction between other kinds of coordinations, such as between internal and external entities, and thus on the role of external entities in certain linguistic explanations.

#### *The box-wire system*

To clarify the idea of ontological matching, let us consider a simplified system. Below I will use this system as a model for discussing internal and external linguistic entities, but first let us consider how ontological matching can arise through the design of a pair of components to perform a function in coordination with one another.

Consider a box, capable of producing and receiving electrical signals, and having some internal structure. The box generates and accepts electrical signals, and let us suppose that the medium through which the signals are transmitted is just a wire consisting of three different lengths having different gauges. The wire we connect both to the output and the input of the box.

Fig. 4



Suppose we design the box and arrange the wires so that the box can send signals to itself and detect them as rapidly as possible. But suppose that the wires are unfortunately a bit thin and sensitive for the signal the box sends through it. The box is capable of sending, say, four different levels of current, from 50V to 200V, as well as sending no current at all, at a time. If

the voltages get too high, though, or if the voltages vary too rapidly from low voltages to high ones, the wire heats up, and its signal transmission becomes unreliable. And if a 200V signal is sent at all, the wire burns out entirely.

For the box to transmit information through the wire successfully at all, to say nothing of doing it well, the wire medium and its arrangement impose many constraints on the structure of the box. From the box's perspective, it would use all available current levels, switching them on and off as rapidly as it could, to maximize the amount of information transmitted. But the structure of the wire dictates that the highest state never be used, and also that there is a bias for using the lower states, as well as a bias against changing among the states too often. Thus for the system to be able to transmit information at all, there are constraints imposed on the design of the box, and if there are pressures for it be designed with the function of transmitting information efficiently, there are competing considerations that need to be balanced.

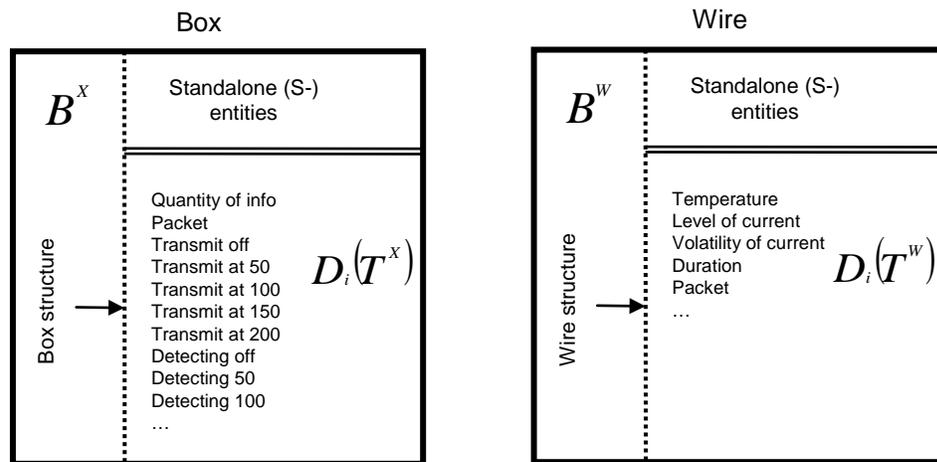
So let us suppose that a clever designer has assembled the internals of the box and the arrangement of the wire, so as to do a good job at optimizing how the box sends signals to itself through the wire. The wire component (call it  $B^W$ ), let us suppose, is arranged in order as heavy to medium to fine gauge. And the box component ( $B^X$ ) has a complex internal structure.

First consider the components designed as they are, treated in isolation from one another. For each, we can construct a theory of the isolated component,  $T^W$  and  $T^X$  respectively. These respective theories will each refer to different quasi-natural kinds and entities. In the theory of the wire we might refer to the gauge and composition of the wires, and also the generic electrical properties of the wire component, relating such variables as the level, duration, and volatility of current flowing through the wire to the temperature of the wire. One quasi-natural kind referred to in  $T^W$  might be a generic burst or signal or packet of current flowing through the wire. This could for instance be characterized in the theory as: a relatively steady voltage with some fixed duration, and having a substantial change in voltage to a relatively steady level before and after.

Theorizing about the box in isolation, we would have a very different set of quasi-natural kinds and entities. The box might have switches and counters, inputs and outputs, levels of transmission, and levels of detection. In theorizing about the electrical properties of the box in isolation, it might also have its own quasi-natural burst or signal or packet corresponding to the natural discrete bursts of current that flowing through the box.

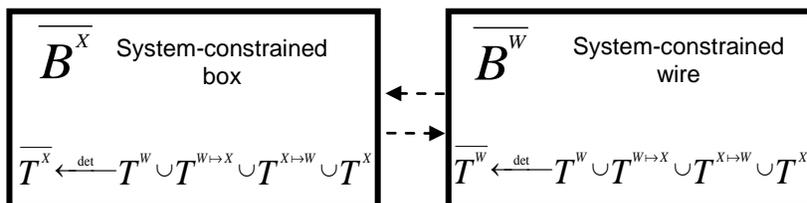
As with the linguistic case, we can thus depict the components of the box-wire system as isolated from one another, and list the quasi-natural entities referred to in good theories of these components:

Fig. 5



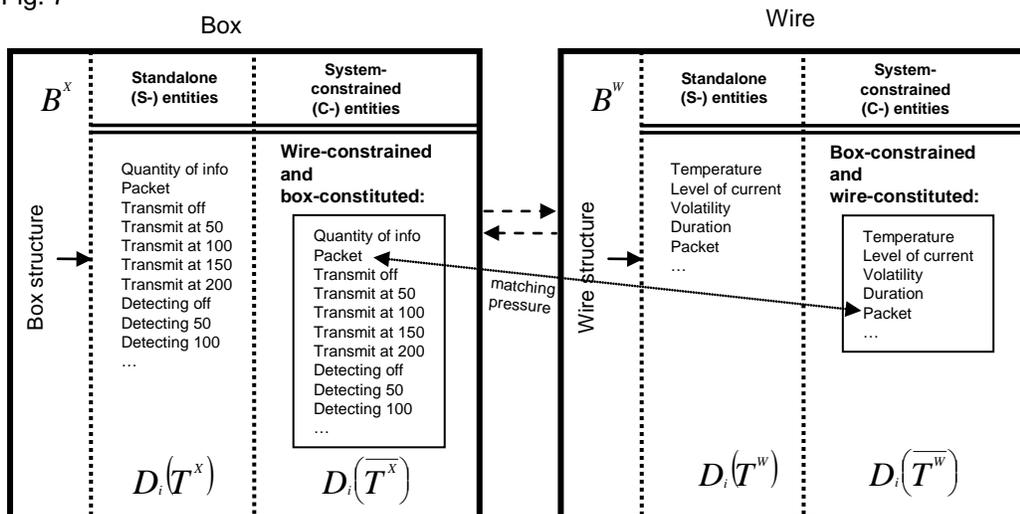
The entities referred to in  $T^w$  and  $T^x$  are S-entities, i.e. determined strictly by the characteristics of the wire and box taken as isolated components. As with the linguistic system, however, when we consider the system connected together, a different set of entities will be the natural ones that we will discern in the wire and in the box.

Fig. 6



In the combined system, the theories of the system-constrained box and the system-constrained wire are each determined by the theories of the box and wire in isolation, together with the bidirectional interfaces between the two. The entities in the box (i.e., the C-entities) will still be the ones that are materially constituted strictly within the box, and correspondingly for the wire. But the C-entities that are the quasi-natural ones in the box will not be determined by the characteristics of the box alone. And likewise for the C-entities in the wire. When we observe the system as a whole in action, it is the C-entities that we observe. A theory of the current flowing through the wire, as connected to the system, will thus look very different from a generic theory of current flowing through an identical wire.

Fig. 7



In the course of the design of the box-wire system, there is implicit

pressure for certain entities across components to match up with one another. The system is designed so that it maximizes the amount of information the box can transmit to itself. One thing this means is that the signals that the box outputs need to be received by the inputs of the box. So they cannot disappear when sent through the wire, because of noise or burnout or some other reason. That is to say, all the packets sent through and from and to the box should correspond to packets sent through the wire. And all the packets sent through the wire should correspond to ones sent through and from and to the box. This pressure is for a correspondence not between the ontologies of the box and the wire in isolation, but only for correspondence between their ontologies as constrained by the other, in the system in which they are connected. That is to say, there is pressure for the wire-C-packets and the box-C-packets to be ontologically coordinated.

The pressure on the design of the box, in other words, is not for it to be structured so as to involve a box-entity that matches an entity that is entirely natural to wire ontology. What is required, instead, is that the box be structured so as to involve an entity that matches one that can be entirely constituted by structures in the wire when constrained by the system as a whole. It does not matter that the C-packets in the wire seem, from the perspective of the theory of wires alone, to be highly gerrymandered entities. It also doesn't matter that there may be mismatches between the C-entities on occasion between the two components. We will find examples of box packets that do not count as wire packets, and conversely wire packets that do not count as box packets. This may be an artifact, for instance, of the flakiness of the wire, so that noise in the wire sometimes may drown out what voltages are output by the box. Or it may be an artifact of a constraint on the design of the box to match what count as packets for it to what count as packets when considered within the medium of the wire. Nonetheless, it is efficient to design the box so that when it wants a packet sent from its output to arrive at its input, that that packet will be tailored so as to conform to what can be a C-packet in the wire.

In the box-wire system, coordination facilitates the efficient functioning

of the system. For the box to ensure that it received the signals it sent, it is reasonable to design it so that the box-C-packets match up with the wire-C-packets. But what the quasi-natural kinds are in the wire in isolation – that is the wire-S-entities – does not particularly matter. Intuitively, the reason that these are the entities coordinated is that the ones that actually show up in the components are the C-entities, not the S-entities. We only need coordination in the cases that might actually occur, so it is not only a waste of effort, but is less efficient, if we coordinate more generic entities between components than the ones that show up given the constraints of the system.

Likewise, to explain the near-matching of entities across linguistic components, it is not the US- or IS-entities that we should expect to be coordinated, but the UC-, IC- and/or IA-entities. The linguistic data on varieties of words considered above, such as that *essuie-glace* is a morphological word in French and not a syntactic word, could be understood as a claim about morphological and syntactic words of type (IA), i.e., the words in Martha's (or perhaps Pierre's) actual idiolect, with parameter settings and lexicon fixed. In addition, the data also seems to be evidence about the mismatch of the possible or IC-words, i.e., those determined strictly by the universals and parameter settings of components of an idiolect. And it can be understood as bearing on the coordination of the universal C-words, the constrained universal structures of morphology, syntax, and phonology. It may be either the constrained or the actual entities that there is pressure to coordinate;<sup>26</sup> but inasmuch as there is pressure on the standalone entities, that is a consequence of the pressures on the actual and constrained ones. If one doesn't distinguish the constrained entities from the standalone ones, then there is a strong risk she will overlook the fact of ontological matching altogether. As I will discuss, this is a key source of confusion in discerning the role of external words in linguistic explanation.

---

<sup>26</sup> There may in fact be reason to expect that it is the IA entities that are coordinated between components, not the IC entities. That's something that we can empirically test: that there is better IA matching than IC matching. The way we do that is to compare the matching soon after acquisition, or words acquired earlier and later.

#### 4. Ontological matching across linguistic components

There are interesting similarities between the notion of ontological matching across linguistic components and the idea of “edge alignment” in recent work on the interface between morphology and phonology. The approach I will suggest, however, is quite different from that implicitly advocated in current treatments of edge alignment. In current theories, alignment is treated as an active constraint that acts across components. I will suggest, in contrast, that the coordination of linguistic components can be understood as arising from mutual pressures in the initial setting of parameters, rather than having to postulate ongoing coordination in the active generation of grammatical structures.

To take one example of cross-component alignment, it has long been observed that the prosodic structure of a language has an influence on the admissible morphological forms in that language. A variety of theories have been proposed to account for the influences of natural prosodic breaking points with word boundaries.<sup>27</sup> McCarthy and Prince have proposed a very generic sort of constraint, which they call “generalized alignment,” that can operate within or between components. In optimality theory, constraints are prioritized, and act as violable successive filters on which candidate structures are determined to be grammatical.<sup>28</sup> Generalized alignment constraints are any constraints of the form *Align(Category 1, Edge 1, Category 2, Edge 2)*, that is, constraints that specify that either the right or left edge of some sort of string is aligned with the right or left edge of another sort of string. A particular instance of this might be that the right edge of a suffix is aligned with the right edge of a pword. In the generalized alignment approach, there does not have to be simultaneous alignment of both sides of an entity: just one edge of an entity can be aligned with an edge of another entity.

---

<sup>27</sup> Cf. McCarthy and Prince (1993), pp. 80-83.

<sup>28</sup> McCarthy and Prince (1993); McCarthy and Prince (1999).

For our purposes, the details are not essential.<sup>29</sup> What is relevant is the approach to understanding how pressures are applied to the generation of grammatical forms that are aligned across components. Inasmuch as alignment constraints are descriptive of the structure of expression generation, they operate as active constraints operating across categories, involving the simultaneous application of ranked constraints within both components.<sup>30</sup> McCarthy and Prince propose that generalized alignment is the only meeting point between morphology and phonology. Nonetheless, it is a powerful meeting point. The output of the two components is actively coordinated between the two, in narrowing down grammatical outputs. Thus the outputs of the components in their picture (again, so long as the description is taken to be of a generative system, rather than simply descriptive adequacy of the outputs) are not generated modularly, from one component passed on to another, and it effectively takes the components to be a single module.

Rather than being an active constraint on structure formation in the language faculty, I will suggest that ontological matching can arise through pressures exerted between components in the setting of parameters or rankings in an idiolect. As a characteristic of the interface between components, alignment is a violation of modularity. But if we understand the

---

<sup>29</sup> One point that may have significant bearing on the issues here, however, is the alignment of one edge of an entity without aligning the other edge. McCarthy and Prince provide a number of examples that suggest that alignment phenomena can indeed occur with edges independent of one another. This may have an important effect on how we consider what the pressures are for aligning entities across components. I will not focus on this issue, but only notice that one pressure is for word alignment across components, whether or not this pressure should be understood as two edge-alignment pressures.

<sup>30</sup> In much of optimality theory, it is unclear whether the constraints at work in the filtering of candidate forms are meant to describe an overall mathematical structure for systematizing grammatical outputs, or whether they are meant to correspond to generative systems within the language faculty. When we consider a cross-component constraint, this issue becomes critical, since it raises the issue as to whether linguistic components can be treated as independently operating modules, or must be seen as acting in coordination in the generation of grammatical structures.

setting of parameters or rankings to take place in coordination with common pressures, the same outcome can arise through alignment in acquisition, rather than through alignment in production.

Another difference between the treatment of alignment in optimality theory and the approach suggested by examining ontological matching relates to what is understood as the relevant kind of coordination across components. Generalized alignment involves the coordination of the edges of strings with one another. The notion of “ontological matching,” however, at least as it appears from the characteristics of words, involves the coincidence of entities across components, where the entities are not necessarily strings and the coordination does not necessarily involve only edges.<sup>31</sup>

#### *The coordination of entities across components through parameter setting*

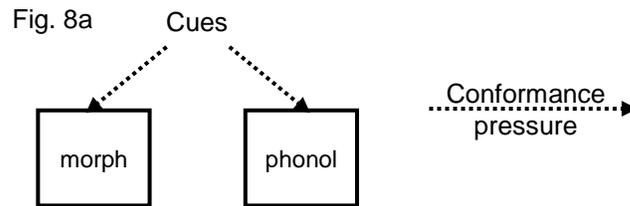
The treatment of edge-alignment with interface rules has the advantage that they are general and globally applicable. However, it has the disadvantage that it is at odds with the modularity of linguistic components. Inasmuch as such rules are intended for grammatical description, rather than generativity, this is not necessarily a fault. But if generativity is a

---

<sup>31</sup> An intuitive understanding of ontological matching is simply as coinciding extensions: the phonological and morphological and syntactic words in an idiolect match when each consists of the same actual words. There are two obvious problems with this. One is that we can only say loosely that the extension of the phonological word, for instance, matches the extension of the syntactic word. Ivan Sag has rightly mocked the image of linguistic entities passing from component to component like salesmen carrying locked suitcases. Jackendoff has emphasized a similar point, in noting that the way syntactic trees are typically drawn, with lexical items as the leaves, is misleading, as opposed to being structures with slots for lexical items (Jackendoff (2002), pp. 121-122). This is an issue that can be cleared up with a proper understanding of the kinds of entities occurring in multi-component systems. The other problem, though, is that even if there is a sense in which a word may be understood as a cross-component entity, it is not clear how their extensions are to be understood as even potentially matching across components. In this discussion, I will simply leave it at this intuitive level. In what follows, I will presume that we can make sense of inter-component matching, and focus on the question of the matching of IC-entities, i.e. what explains why the morphological and syntactic and phonological words that we intuitively consider Martha’s possible words nearly match with one another.

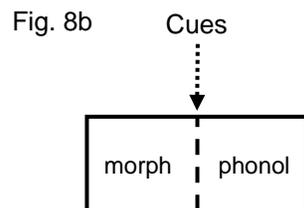
consideration in evaluating descriptive approaches, and if linguistic components generate grammatical forms modularly, then a constraint governing interfaces is a shortcoming. I will suggest instead that modularity can be reconciled with cross-module alignment, through the coordination in learning, or the setting of parameters, rather than in generation.

The generic issue in the case of setting parameters, as is depicted in figure 8a, is that there are cues received by the language acquirer, which exert pressures on how the parameters are set in each linguistic component. (The line with small dots represents design or conformance pressure exerted by some factor on a component.) The question is how, in light of the setting of the parameters in conformance to the cueing pressures, the resulting entities end up being ontologically aligned with one another. Again, the entities I mean to consider here are the IC-entities across the components. Martha's possible morphological words, for instance, are aligned with her possible syntactic ones, and so with her possible phonological words.



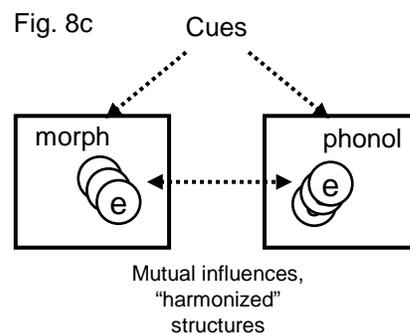
There are several sorts of explanation for the alignment. One is that it is simply accidental, which seems implausible. This is the solution that would be depicted by 8a, if there were no coordination at all between the modules.

A different solution is to take the cues that set Martha's parameters or constraint rankings to involve morphophonological alignment rules directly. This is depicted in 8b:



As I mentioned in connection with generalized alignment, however, this compromises the modularity of the morphology and phonology components.

A third solution is that the alignment between modules can be explained through mutual pressures or coordination mechanisms in the setting of parameters or rankings, without needing to postulate global alignment rules or constraints. Figure 8c depicts the components being mutually influenced or coordinated, as a result of influences interacting between them. Here there are pressures leading to the structures of the components being coordinated, so that entities that are quasi-natural kinds across those structures are aligned. Both of these involve mutual coordination, due to interactive conformance pressures of some sort.



If we look at the literature on language acquisition with an eye to observing the coordination of components in parameter-setting, we find that it is thick with examples. Slobin (1985), for instance, presents a variety of cases of the development of morphological paradigms on phonological grounds. In acquiring gender morphology, children across a wide variety of languages employ phonological cues to work out inflectional structure.<sup>32</sup>

Evaluation the coordination of parameter-settings across modules involves a variety of empirical issues. The picture requires both that linguistic components be modular, so that there is at least reason to doubt that expression generation is governed by overarching principles or constraints. It

---

<sup>32</sup> Slobin (1985), pp. 1212-1219. See also, Levy (1983); Pinker (1984); MacWhinney (1985).

also requires that the way that the morphological parameters are set influences and is influenced by the way that the phonological ones are, and likewise for the syntactic ones. And moreover, that there are influences in parameter setting between components suggests but does not entail that those influences involve ontological coordination.

Still, for there to be an allocation of responsibility and a coordination of ontologies between two components, there needs only to be a minimal source of pressure exerted between the components. In all the models, there is pressure exerted by the cues to have the parameter settings somehow affected by them. Yet it is reasonable to assume that conformance to the demands of the cues, together with the universal structures of the component, are not the only constraint on the setting of the parameters, but that there are also overall demands, such as to enhance overall pronounceability, augment the efficiency of the system as a whole, or reduce energy consumption, or increase processing speed. These crude pressures can account for the generation of complex sorts of coordination between components. This issue has been explored in most detail in economic theory, both in the theory of games and in the theory of the firm.<sup>33</sup> Investigation of similar phenomena as applied to cognitive structure is a potentially fertile topic for inquiry.

#### *The coordination of universals*

In the preceding discussion, I have suggested that a plausible explanation for the ontological matching of idiolectal entities, most straightforwardly the ones I've called IC-entities, is provided by coordination in the setup of the components, rather than by happenstance or in virtue simply of the active coordination between the components in the generation of grammatical forms. An obvious next step is to ask the same about whether or not there is matching of entities determined in the universal structures of various

---

<sup>33</sup> A generic form of the mutual pressures that affect the components in tandem with one another has a relatively straightforward mathematical description, which helps make it explicit how they are interdetermined. This is a topic that has been addressed in the study of the pressures influencing the coordination of firms in economic theory, notably in the literature sparked by Grossman and Hart (1986).

components and seek an explanation for what would account for such coordination as well. The obvious source of potential coordination among these is natural selection, acting as a pressure on the design of the universal structures of various components in tandem with one another, as opposed to exerting pressure on each component entirely independent of one another. The coordination of multi-component systems by natural selection is again a common but as far as I know relatively unexplored phenomenon.

The viability of this explanation, of course, depends on the question of the evolution of the language faculty altogether. The sorts of linguistic data on words I presented above can, I believe, serve as evidence for the ontological matching of universals across components as well as for the coordination of parameter-setting across components. Furthermore, it may be reasonable even to take these phenomena as some evidence for a selection account of linguistic universals, as I mentioned.

Nevertheless, rather than examining universals I want to turn instead to the coordination of internal and external entities; that is, of certain I-word kinds with certain E-word kinds. Again, I am less concerned with proposing a metaphysics of these entities than with bringing into relief their role in linguistic explanation, so as to counter the claim that they are irrelevant to linguistics. To do so, I want to “flatten out” the distinction between the multiple components of the language faculty and the physiological apparatus for the production and reception of language and the external manifestations of language, including sounds, inscriptions, and the media in which they occur.

## **5. External entities in linguistic explanation**

Ludlow points out in *Philosophical Issues in Generative Linguistics* that “relational sciences,” i.e., sciences involving the relation between individuals and the environment, are not entirely irrelevant to building theories in “individualistic sciences”:

The general phenomenon of individualistic and relational sciences informing each other is fairly common. One example I drew on [in previous work] was the contrast between primate physiology and primate ecology, in which the former draws upon individualistic facts about (for example the skeletal structure of the primate) and the latter addresses relations between the primate and its environment. Despite the apparently distinct properties under investigation, it is possible for facts about primate physiology to support a particular claim about the relation of the primate to its environment and vice versa...My favorite illustration of the general phenomenon remains Webster and Webster (1988), in which it is observed that individualistic anatomical structure can place constraints on the types of (relational) environmental functions that are possible, and vice versa.<sup>34</sup>

Chomsky (1995) likewise affirms that relational and individualistic sciences can inform each other. A study of the human breathing cycle, for instance, may involve the exchange of oxygen with the atmosphere.

Both of these, however, are raised mainly to draw the distinction between the investigation of internal from environmental factors, rather than to explore a role for the environment in linguistic investigation. It is left unclear, for instance, whether there is then a relational science that informs linguistics, as primate ecology does for primate physiology. Moreover, these points are not accompanied by an acknowledgement of the potential relevance or even coherence of external linguistic entities.

Applying the relations between words in various internal components of the language faculty to the internal/external divide, we can make significant headway in clearing up the candidates for what we are referring to in speaking of external words as well. Clarifying the similarity between the external candidates and the internal ones makes it clear that it is rash to deny

---

<sup>34</sup> Ludlow (2007), Ch. 4.

the coherence of external linguistic entities outright, and also to deny their role in linguistic explanation.

Once again, here I am not considering the design of linguistic universals in response to external constraints. To vindicate the role of external entities in linguistic explanation we can stick to the setting of parameters with cues.

*Matching internal and external words*

In the discussion of the box-wire system, I treated both box and wire as analogous to components of the language faculty. However, it is straightforward to see that the wire is even more closely analogous to the acoustic environment in which verbal signals are communicated from a speaker to herself and between speakers.

Fig. 9

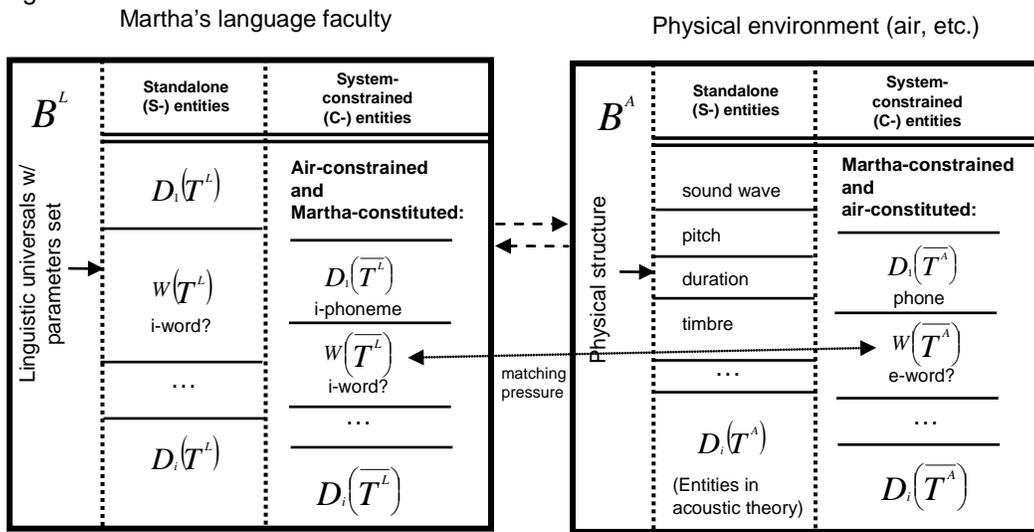


Figure 9 depicts the system of Martha's language faculty as a whole connected to the external physical environment, as the box was connected to the wire. Martha's language faculty is here taken as a unit, and in isolation is denoted by ' $B^L$ ', with  $T^L$  being a good theory of  $B^L$ .  $B^A$  is the relevant external environment – principally, the physical characteristics of the air through which vocalizations are transmitted – and  $T^A$  is a good theory of  $B^A$  in isolation.

$T^A$  in isolation is the theory of acoustics. Acoustics has as its subject matter the transmission of sounds in the air in general, and there are many quasi-natural entities in acoustics, arising from the sound-transmission properties of the air and physical structures alone. A good theory of acoustics will involve terms for such kinds and properties as sound waves, pitch, duration, amplitude, etc. On the other hand, from the perspective of acoustic theory, *words* are not likely to be particularly natural entities at all, just as *the timbre of a Stradivarius* is not quasi-natural to acoustic theory alone, but only to acoustic theory in combination with Stradivarii.

In considering the physical environment from the perspective of linguistics, however, we are not principally concerned with S-entities of the physical environment. Rather, the relevant quasi-natural kinds are the C-entities of the environment in combination with the language faculties of individuals. It should be no surprise that acoustic theorists have a difficult time generating strictly acoustic criteria for word-hood. From the perspective of acoustic theory – that is, a theory of  $B^A$  alone – words are highly gerrymandered entities. But in a theory of the system – that is a theory of  $\{B^L, B^A, B^{L \rightarrow A}, B^{A \rightarrow L}\}$  – they will be among the most natural of entities.

As in the other multi-component systems, presuming at least some design pressures on the system as a whole, we should expect that any matching across the components should be between the C-entities. It is not difficult to see, as in the case of the box, what the functional utility is of a match between Martha's C-words (i.e.  $W(\overline{T^L})$ ) and the external C-words (i.e.  $W(\overline{T^A})$ ). Even if we consider only Martha's babbling to herself, rather than her communicating with a linguistic community, she babbles efficiently to herself just in case the tokens of the C-words in the environment occur when she has tokens of C-words in her head.

For this coordination to take place, it is not necessary to adapt both the physical environment and her language faculty in the design process. Rather, we can take the physical environment for granted, and take her language faculty to have been designed so as to accommodate those fixed

characteristics.<sup>35</sup> In other words, the component  $B^L$ , and the resulting quasi-natural entities  $D_i(T^L)$ , were designed, in conformance with the strictures imposed by  $B^L$ 's being in an environment having structure  $B^A$ . However, that does not imply that  $B^A$  understood as isolated (or the entities  $D_i(T^A)$ ) are what is salient to the explanation of the design of  $B^L$ . If we are to understand the problem of the design of  $B^L$  as an optimization problem, it is the coordination of  $W(\overline{T^L})$  with  $W(\overline{T^A})$  that is relevant. That coordination problem can only be expressed by considering the ontology of the system as a unit, rather than considering the language faculty to be directed by the characteristics of acoustic entities treated in isolation. Only once the component  $B^L$  has been designed, through pressures coordinating the C-entities of the system, then the internal “anatomy” of  $B^L$  is determined, and we can theorize about  $B^L$  in isolation.

This is what we described with the box-wire design. In the box-wire case, it was mostly the box that was designed, rather than the wire. The designer of the system was able to modify the arrangement of the segments of wire, but mostly the wire presented a fixed set of characteristics that the design of the box had to accommodate. This does not, however, imply that the only wire-entities about which to theorize are its S-entities. In constructing a theory of the wire, it is true that we are naturally inclined to consider it in isolation. But from the perspective of the box-wire system together, the wire-S-entities are not the natural external ones. Instead, it is the wire-C-entities, i.e., the entities that are determined by the system as a whole, albeit whose tokens are materially constituted entirely in the wire, that are the quasi-natural external ones in box-wire theory.

Even in the study of the *anatomy* of the box or of the language faculty, it may not be possible or desirable to ignore the C-entities. If we had perfect visibility into the internals of the box, we might be able to characterize its structure, i.e., to construct a good theory  $T^X$ , simply from the observation of

---

<sup>35</sup> Remember that here we are taking the “design” of her language faculty to include the setting of her parameters, not just the adaptation of her biological structures.

the box alone independent of its environment. But if there are any limitations at all to the data we have about the structure of the box, or even if the box is reasonably complex, then considering  $\overline{T^X}$  may be indispensable for generating an acceptable  $T^X$ . After all, at least part of the explanation for why the box-S-entities being what they are is derivative on the characteristics of the box-C-entities. If there are, for instance, simplicity or elegance or nonredundancy considerations applied to generating hypotheses about the structure of S-entities, then what these are properly applied to is the simplicity or elegance or nonredundancy of various box-internal solutions to the ontological matching problem between cross-component C-entities.

Given the almost complete opacity of the language faculty to direct inquiry, these sorts of considerations play an enormous role in linguistic theorizing. It is thus very implausible that there are such informative data-sources on the structure of the language faculty as the external factors that figure into the way parameters are set, and yet that these external factors will be outside the scope of linguistic explanation.

#### *What E-words are*

Although there are many possible candidates for what the term ‘word’ can refer to, that is an argument for disambiguation, not for the denial of the coherence of one notion or another. E-words are no more problematic from this perspective than the various candidates for I-words, or I-morphemes or I-phonemes. In every domain, we can identify the most “word-like” entities, but it must be admitted that the entities picked out will not be particularly natural in certain domains, such as acoustic theory taking in isolation.<sup>36</sup> Nonetheless, there are at least two excellent candidates for E-words identifiable in the Martha-physical system. One is the physical-C-words,

---

<sup>36</sup> Also, in domains where it is not particularly natural, there are likely to be many candidates that are basically equally good. It is as important a condition for being the clear referent of a theoretical term that it be the unique near-realizer of the Ramsey-Lewis description as that it be quasi-natural, though these two conditions are interrelated.

which are entirely materially constituted by sound waves, but are constrained by the structure of Martha's language faculty. These entities do have physical regularities, but not ones that the science of acoustics will easily discern, in the absence of considering the speakers as well. Given that they are materially constituted entirely externally, we can characterize these things in strictly external terms. That is, we can construct a theory of physical-C-words in entirely acoustic terms. But we should expect that this theory will look highly gerrymandered from the perspective of general acoustic theory.

A second candidate for what we might call an E-word is the quasi-natural word-like entity arising from the system as a whole, including both Martha and the environment. That is, the entity designated by  $W(T^L \cup T^A \cup T^{L \rightarrow A} \cup T^{A \rightarrow L})$ . This is an entity I have not highlighted in the various diagrams. Nonetheless it is reasonable to think that there will be a quasi-natural entity in the system as a whole, which closely matches both the Martha-C-words and the physical-C-words, yet is more quasi-natural than either when considering the system as a whole. Since the Martha-physical system includes Martha as a component, this entity might with equal justice be called an I-word as an E-word. Yet it is plausibly the intuitive entity that is instantiated both internally and externally, and hence that magically seems to cross boundaries between mind and world.

Notice that, however we understand E-words, they will likely not be perfectly matched with any of the dozen kinds of I-words in Martha's idiolect. Martha's parameters are set in part so as to match Martha-C-words with physical-C-words, but that design does not demand perfection. It is a fallacious argument against E-words to point out examples of sound-patterns that cannot be mapped one-to-one with I-words of some kind.

Finally, notice that the notion of an E-word does not depend so far on the publicity of language. The candidate E-words I have discussed only involve Martha and the environment, not a community of speakers. The approach I have presented, however, does indicate how we can begin to understand the individuation of public linguistic entities. If we consider a system of two people in an environment, then it is clear how they compose a system, with its

own sets of S- and C-entities, and a similar story can be constructed about the coordination of C-entities of all the components in that more complex system. The idea that we can speak of coherent E-words for a community depends on the question as to whether there is a tendency for various C-entities to match across the system as a whole. But our ability to communicate with one another strongly suggests that there is pressure for the ontological alignment of the C-entities within a community, where the community is considered as the system.

In considering the roles of various entities in linguistic explanation, it is helpful to worry about the internal/external distinction a little less, and instead notice that there are many ways of carving out multi-part systems along appropriately natural boundaries. Moreover, if there is any cunning to the history of scientific inquiry, it is to slap down those who arbitrarily stipulate that some domain of entities is not pertinent to explanations in a scientific field. This was the strong lesson that linguists have taken from the abortive exclusion of internal properties of the mind in early-century empiricism. It is equally clear, I believe, that in future linguistic investigation we are bound to find it an overreaction to have denied the reality of external linguistic entities or their pertinence to linguistic explanation.

All the discussion above is of course complicated by the fact that there are no perfectly definite boundaries between components of any kind. It has been repeatedly pointed out that the notion that the language faculty is a series of independent components, acting only through interfaces, is a rather gross simplification, whether the components are physically or even functionally delineated. Inasmuch as the functional structure of the language faculty can be treated as consisting of components, there are places where it is most plausible to regard them as “interpenetrated.” But the same is of course true of the boundaries between internal and external. As far as I can tell, this complication only breaks down further the pull to exclude the external from the direction in which linguistic theory is bound to be headed.

## REFERENCES

- Aronoff, M. and S. N. Sridhar. (1983). "Morphological Levels in English and Kannada; or, Atarizing Reagan." (In *Papers from the Paresession on Phonology, Morphology, and Syntax, 22-23 April 1983*. J. F. Richardson, M. Marks and A. Chukerman, Eds. (pp. 3-16). Chicago.)
- Booij, G. (1984). "Neutral Vowels and the Autosegmental Analysis of Hungarian Vowel Harmony." *Linguistics and Philosophy* 21: 249-280.
- Booij, G. (1999). "The Role of the Prosodic Word in Phonotactic Generalizations." (In *Studies on the Phonological Word*. T. A. Hall and U. Kleinhenz, Eds. (pp. 47-72). Amsterdam: John Benjamins.)
- Booij, G. (2005). *The Grammar of Words*. (Oxford: Oxford University Press.)
- Bromberger, S. (1989). "Types and Tokens in Linguistics." (In *Reflections on Chomsky*. A. George, Ed. (pp. 58-89). Oxford: Blackwell.)
- Cappelen, H. (1999). "Intentions in Words." *Nous* 33(1): 92-102.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. (Cambridge,,: M.I.T. Press.)
- Chomsky, N. (1995). *The Minimalist Program*. (Cambridge, Mass.: The MIT Press.)
- Chomsky, N. and P. Ludlow. (2003). Chomsky on Chomsky, Language, Mind, and Freedom. *The Stony Brook Interviews*.)
- Devitt, M. (1981). *Designation*. (New York: Columbia University Press.)
- Devitt, M. (2006). *Ignorance of Language*. (Oxford: Clarendon Press.)
- Di Sciullo, A.-M. and E. Williams. (1987). *On the Definition of Word*. (Cambridge: MIT Press.)
- Embick, D. and R. Noyer. (2007). "Distributed Morphology and the Syntax-Morphology Interface." (In *The Oxford Handbook of Linguistic Interfaces*. G. Ramchand and C. Reiss, Eds. (pp. 289-324). Oxford: Oxford University Press.)
- Field, H. (1994). "Deflationist Views of Meaning and Content." *Mind*

103(411): 249-285.

- Garcia-Carpintero, M. (1997). "On an Incorrect Understanding of Tarskian Truth Definitions." *Philosophical Issues* 8: 45-56.
- Grossman, S. J. and O. D. Hart. (1986). "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration." *Journal of Political Economy* 94(4): 691-719.
- Hall, T. A. (1999a). "The Phonological Word: A Review." (In *Studies on the Phonological Word*. T. A. Hall and U. Kleinhenz, Eds. (pp. 1-22). Amsterdam: John Benjamins.)
- Hall, T. A. (1999b). "Phonotactics and the Prosodic Structure of German Function Words." (In *Studies on the Phonological Word*. T. A. Hall and U. Kleinhenz, Eds. (pp. 99-132). Amsterdam: John Benjamins.)
- Halle, M. (1973). "Prolegomena to a Theory of Word-Formation." *Linguistic Inquiry* 4: 451-464.
- Haspelmath, M. (2002). *Understanding Morphology*. (London: Hodder Arnold.)
- Hauser, M. (1996). *The Evolution of Communication*. (Cambridge: MIT Press.)
- Ito, J. and R. A. Mester. (1999). "Realignment." (In *The Prosody-Morphology Interface*. R. Kager, H. van der Hulst and W. Zonneveld, Eds. (pp. 188-217). Cambridge: Cambridge University Press.)
- Jackendoff, R. (2002). *Foundations of language: Brain, Meaning, Grammar, Evolution*. (Oxford: Oxford University Press.)
- Julien, M. (2007). "The Relation Between Morphology and Syntax." (In *The Oxford Handbook of Linguistic Interfaces*. G. Ramchand and C. Reiss, Eds. (pp. 209-238). Oxford: Oxford University Press.)
- Kaplan, D. (1990). "Words." *Proceedings of the Aristotelian Society* 64: 93-119.
- Levy, Y. (1983). "It's Frogs all the Way Down." *Cognition* 15: 75-93.
- Lewis, D. K. (1970). "How to Define Theoretical Terms." *Journal of Philosophy* 67: 427-46.
- Lewis, D. K. (1972). "Psychophysical and Theoretical Identification."

- Australasian Journal of Philosophy* 50: 249-58.
- Lewis, D. K. (1997). "Naming the Colours." *Australasian Journal of Philosophy* 75(3): 325-42.
- Lieberman, P. (2006). *Toward an Evolutionary Biology of Language*. (Cambridge: Belknap.)
- Ludlow, P. (2007). "Philosophical Issues in Generative Linguistics," ms.
- MacWhinney, B. (1985). "Hungarian Language Acquisition as an Exemplification of a General Model of Grammatical Development." (In *The Crosslinguistic Study of Language Acquisition*. D. I. Slobin, Ed. 2 (pp. 1069-1155). Hillsdale: Lawrence Erlbaum Associates.)
- McCarthy, J. J. and A. S. Prince. (1993). "Generalized Alignment." (In *Yearbook of Morphology 1993*. G. Booij and J. van Marle, Eds. (pp. 79-153). Dordrecht: Kluwer.)
- McCarthy, J. J. and A. S. Prince. (1999). "Faithfulness and Identity in Prosodic Morphology." (In *The Prosody-Morphology Interface*. R. Kager, H. van der Hulst and W. Zonneveld, Eds. (pp. 218-309). Cambridge: Cambridge University Press.)
- Millikan, R. G. (1984). *Language, Thought, and Other Biological Categories: New Foundations for Realism*. (Cambridge: MIT Press.)
- Nespor, M. and I. Vogel. (1986). *Prosodic Phonology*. (Dordrecht: Foris.)
- Pinker, S. (1984). *Language Learnability and Language Development*. (Cambridge: Harvard University Press.)
- Price, H. (1997). "What Should a Deflationist About Truth Say About Meaning?" *Philosophical Issues* 8: 107-115.
- Quine, W. V. (1987). *Quiddities*. (Cambridge: Harvard University Press.)
- Raffelsiefen, R. (1999). "Diagnostics for Prosodic Words Revisited: The Case of Historically Prefixed Words in English." (In *Studies on the Phonological Word*. T. A. Hall and U. Kleinhenz, Eds. (pp. 133-202). Amsterdam: John Benjamins.)
- Rey, G. (2006a). "Conventions, Intuitions and Linguistic Inexistence: A Reply to Devitt." *Croatian Journal of Philosophy* VI.
- Rey, G. (2006b). "The Intentional Inexistence of Language -- But Not Cars."

- (In *Contemporary Debates in Cognitive Science*. R. Stainton, Ed. (pp. 237-55). Oxford: Blackwell Publishers.)
- Richard, M. (1990). *Propositional Attitudes*. (Cambridge: Cambridge University Press.)
- Ritchie, W. C. and T. K. Bhatia, Eds. (1999). *Handbook of Child Language Acquisition*. (San Diego: Academic Press.)
- Sandler, W. and D. Lillo-Martin. (2006). *Sign Language and Linguistic Universals*. (Cambridge: Cambridge University Press.)
- Selkirk, E. O. (1984). *Phonology and Syntax: The Relation Between Sound and Structure*. (Cambridge: MIT Press.)
- Slobin, D. I. (1985). "Crosslinguistic Evidence for the Language-Making Capacity." (In *The Crosslinguistic Study of Language Acquisition*. D. I. Slobin, Ed. 2 (pp. 1157-1256). Hillsdale: Lawrence Erlbaum Associates.)
- Soames, S. (1995). "T-Sentences." (In *Modality, Morality, and Belief*. W. Sinnott-Armstrong, D. Raffman and N. Asher, Eds. (pp. 250-270). Cambridge: Cambridge University Press.)
- Spencer, A. (1991). *Morphological Theory*. (Oxford: Blackwell.)
- Szabo, Z. G. (1999). "Expressions and their Representations." *Philosophical Quarterly* 49(195): 145-163.
- Szpyra, J. (1989). *The Phonology-Morphology Interface*. (London: Routledge.)
- Webster, D. B. and M. Webster. (1988). "Hypotheses Derived From Morphological Data: When and How They Are Useful." *American Zoologist* 28(1): 231-236.