Theoria, 2020, 35(1), 29-43
https://doi.org/10.1387/theoria.20464

THEORIA

ISSN 0495-4548 – eISSN 2171-679X

Some theoretical and empirical background to Fodor's systematicity arguments

(Algo de trasfondo teórico y empírico para los argumentos de la sistematicidad de Fodor)

Kenneth AIZAWA*
Rutgers University

ABSTRACT: This paper aims to clarify certain features of the systematicity arguments by a review of some of the largely underexamined background in Chomsky's and Fodor's early work on transformational grammar.

KEYWORDS: Chomsky, Fodor, systematicity, productivity, transformational grammar.

RESUMEN: *Este artículo intenta clarificar ciertos aspectos de los argumentos de la sistematicidad revisando parte del trasfondo en gran medida infraexaminado en el trabajo temprano de Chomsky y de Fodor sobre gramática transformacional.*

PALABRAS CLAVE: Chomsky, Fodor, sistematicidad, productividad, gramática transformacional.

* ***Correspondence to:*** Kenneth Aizawa. Department of Philosophy. Rutgers University, Newark. 175 University Avenue. Newark NJ, 07102. USA – ken.aizawa@gmail.com– https://orcid.org/0000-0002-7143-4901

How to cite: Aizawa, Kenneth. (2020). «Some theoretical and empirical background to Fodor's systematicity arguments»; *Theoria. An International Journal for Theory, History and Foundations of Science,* 35(1), 29-43. (https://doi.org/10.1387/theoria.20464).

Received: 17 December, 2018; Final version: 8 July, 2019.

I must admit that I have only two intellectual heroes. One was Jerry Fodor; the other is Noam Chomsky. Despite my deep respect for Fodor and his work, I had only the briefest of interactions with him, limited to a few comments and questions at conferences and workshops. For more personal insights, I rely on others. On this score, Georges Rey tells the following story that is both amusing and, I think, illuminating:

> I once asked him, "Jerry, you probably know more scientific psychology than any other philosopher: why, when you give an example of a psychological law, do you take a trivial folk example, such as Eating potato chips can make you want to eat more, instead of any serious ones from actual psychology?" Without hesitation, he replied, "Citing the science would be vulgar". (Rey, 2018, p. 335)[1]

I cannot vouch for Fodor knowing more scientific psychology than any other philosopher, but Fodor's habit was very much in evidence when he explained that human cognition is systematic by claiming, "no native speaker comes to understand the form of words 'John loves Mary' except as he *also* comes to understand the form of words 'Mary loves John'." (Fodor, 1987, p. 150).

I note Fodor's stylistic idiosyncrasy because it is one, among many, reasons that can make the systematicity arguments difficult to follow and evaluate. Indeed, some have suggested that such simple, or simplistic, formulations as Fodor, and others following him, invoke are not scientific.[2] Others have observed that Fodor does not document the existence of systematicity through references to the scientific literature, so that the systematicity of thought is a myth or a hoax.[3]

This paper will try to compensate a bit for Fodor's idiosyncrasy by reviewing some of his and Chomsky's work in linguistics in the 1960s. This history will (I hope) provide a new angle on the systematicity arguments. It will, first and foremost, draw attention to part of the basis Fodor had for thinking that thought has systematic features. Second, it will help focus attention on a key feature of the systematicity arguments, namely, a role for some notion of what is ad hoc. Third, and finally, it will provide a reason to be skeptical of the hypothesis that the systematicity of thought derives from the systematicity of language.[4]

Section 1 will review the systematicity arguments highlighting three important concepts, namely, productivity, systematicity, and the idea of the ad hoc. Section 2 will review the discussion of the productivity of natural language in Fodor (1961), Katz et al. (1963), and Chomsky (1964). This review helps us see how Fodor adapted his early thinking about the productivity of natural language into a case for the existence of systematic features of thought and a compositional explanation of them that is not ad hoc. Section 3 will draw attention to Chomsky's idea that good explanations in linguistics require more than that a grammar generate all and only the sentences of a language. There are, in addition, considerations of what is simpler and not *ad hoc*. Such considerations underlie the sense that the

[1] This is not the only amusing or insightful bit of Rey's discussion of Fodor, but it is the one that best suits my present purposes.

[2] See Travieso et al. (2014, p. 377).

[3] See Chemero (2014, pp. 355-9).

[4] See Gomila et al. (2012); Travieso et al. (2014).

systematicity arguments are instances of inference to the best explanation wherein what makes for a better explanation is a kind of simplicity and avoidance of the *ad hoc*. Section 4 will try to show how the considerations developed in sections 2 continue to be theoretically and empirically relevant, as they speak to recent attempts to explain the systematicity of thought by appeal to the systematicity of language.

1. Productivity, Systematicity, and the Ad Hoc

The productivity of a natural language, such as English, refers to the hypothesis that there is no finite bound on the length of a sentence. This unbounded character might be suggested by sequences such as the following

John loves Mary
It's not the case that John loves Mary
It's not the case that it's not the case that John loves Mary
It's not the case that it's not the case that it's not the case that John loves Mary.

This is the house that Jack built
This is the malt that lay in the house that Jack built
This is the rat that ate the malt that lay in the house that Jack built.
This is the cat that killed the rat that ate the malt that lay in the house that Jack built.

By a natural idealization, one has it that natural languages are productive.[5] There is no point where, say, adding another negation converts a grammatical sentence of English into an ungrammatical sentence. One gets the idea that *thought* is productive, if one is granted the assumption that speakers of a natural language can entertain the thoughts expressed in their natural language.[6]

Fodor (1987) and Fodor et al. (1988) propose, for the sake of argument, not to rely on the idealization to the infinite representational capacities invoked in the productivity argument. Instead, they suppose that there are only finitely many possible occurrent thoughts.[7] Consider, then, in Fodorian style, a simplified set of possible occurrent thoughts.

(1) John loves John
John loves Mary
Mary loves John
Mary loves Mary.

[5] Chomsky (1956, p. 15), Chomsky (1957, pp. 23-4) proposes that it is simplicity in the grammar that motivates the view that natural languages are productive.

[6] Cf. Fodor et al. (1988, p. 39); Fodor (1987, p. 151) where the systematicity of thought is inferred from the systematicity of natural language along with the assumption that understanding a sentence of one's natural language is a matter of thinking the thought the sentence expresses.

[7] The discussion that follows will assume only that it is occurrent thoughts—thoughts that "come to mind" so to speak—that are systematic. The reason is that dispositional thoughts might be stored in another format that is used to construct occurrent thoughts. McLaughlin (2009, p. 254) however, proposes that systematicity extends to all thoughts.

Let these be all and only the thoughts that might occur to some cognitive agent. Fodor and Pylyshyn claim that such a set has two features. First, some of the thoughts are "intrinsically connected." Second, some of the thoughts are semantically related.[8] For the sake of simplicity, I believe, Fodor and Pylyshyn focus on pairs of thoughts, such as John loves Mary and Mary loves John. We might understand the "intrinsic connection" between them as one of counterfactual dependence. The agent would not have the capacity to think that John loves Mary unless she had the capacity to think that Mary loves John, and *vice versa*.[9] To say that thoughts are semantically related is to say, to a first approximation, that they are about the same individuals in the same relation.

I take it that interests of simplicity are behind the focus on such pairs. First of all, the point is not about pairs *per se.* All the members of the following cluster of thoughts are putatively counterfactually dependent and semantically related:

(2) John introduced Mary to Bob
John introduced Bob to Mary
Mary introduced John to Bob
Mary introduced Bob to John
Bob introduced John to Mary
Bob introduced Mary to John[10]

Second, consider the intrinsic connections among the thoughts in (1). The thought that John loves Mary and the thought that John loves John are (presumably) not mutually counterfactually dependent, since one might have the capacity to have the thought that John loves John, without having the capacity to have the thought that John loves Mary. One might not have the concept of Mary. Third, consider semantic relatedness. The thought that John loves Mary and that John loves John does not have the same individuals in the same relations. The first thought is about Mary, where the second is not. So, even though Fodor evidently presupposes that the set of possible occurrent thoughts is something like (1), he does not try to provide a characterization of the "intrinsic connections" and semantic relations among all the thoughts in the set.

Note that sometimes Fodor, Pylyshyn, and McLaughlin propose something somewhat stronger than we have set out above.[11] They propose, roughly, that one does not find cognitive agents who can think that thought that aRb, unless they can think the thought that

8 Fodor et al. (1988) describe the first feature as "the systematicity of cognitive representations," (section 3.2) and the second as "the compositionality of representations." (section 3.3).

9 McLaughlin (2014, p. 33), Ramsey (2014, p. 255) accept this way of understanding systematicity. McLaughlin (2009) also treats these in terms of "systematicity laws". Fleshing out the idea of "intrinsic connection" in terms of counterfactual dependence may help dispel the sense that the explanandum presupposes that the thoughts have the same components, thus begging a question in favor of a syntactically and semantically compositional language of thought. (Cf. Matthews (1994, p. 355) regarding question begging. See McLaughlin (2009, p. 257f) for a reply.)

10 McLaughlin notes "The idea that thought abilities come in clusters can be captured by saying that they come in pairs" (McLaughlin, 2009, p. 253). I read him as saying that you can get the point of counterfactual dependence by looking to pairs; not that one must explicate counterfactual dependence by looking to pairs.

11 Cf, e.g., Fodor et al. (1988, pp. 40, 44); Fodor et al. (1990, p. 202); McLaughlin (2009, p. 254).

bRa. This is a stronger claim than what we have set out above insofar as it implicitly assumes that the schematic letters "a" and "b" can take on any individuals and "R" any relations as instances. This stronger claim has, of course, invited various putative counterexamples.[12] It has also invited questions about just how pervasive these intrinsic connections are supposed to be.[13] Fodor and Pylyshyn, however, are aware of the "patchiness," we might call it, of the intrinsic connections and semantic relations: "It's uncertain exactly how compositional natural languages actually are (just as it's uncertain exactly how systematic they are)" (ibid. p. 42.) Fodor elsewhere suggests that a precise characterization is unnecessary. "Just a little systematicity of thought will do to make things hard for Aunty, since, as previously remarked, [rival conceptions] are compatible with there being no systematicity of thought at all" (Fodor, 1987, p. 153.) Fodor's point, in this last, reinforces what I take to be his intent in focusing on pairs of thought capacities in order to run his argument. The upshot, perhaps, is that the challenge of the systematicity arguments is not so much to explain why thought is systematic, but to explain why thought is as systematic as it is. In more detail, the challenge is not so much to explain why there are intrinsic connections/semantic relations between thoughts, but to explain why there are as many intrinsic connections/semantics relations between thoughts as there are. Why, one might ask critics, is this not a perfectly good challenge?

Fodor and Pylyshyn propose that cognitive representations have a combinatorial syntax and semantics. Cognitive representations constitute a language of thought. Oversimplifying, the meanings of some cognitive representations (i.e. molecular representations) are determined by the meanings of atomic representations and the way in which those atomic representations are put together. So, we might have {John, loves, Mary} as our syntactic atoms with the obvious meanings and a single cognitive-grammatical construct. Fodor and Pylyshyn implicitly presuppose that this syntactically and semantically combinatorial language of thought bests all comers as an account of the structure of cognitive representation, so that, for purposes of illuminating how the systematicity arguments work, we might consider a very simple rival hypothesis, namely, that all cognitive representations have an atomic semantic and syntactic structure.[14] Let the atomic symbols be {◆, ♥, ♣, ♠} with the meanings John loves John, John loves Mary, Mary loves John, and Mary loves Mary. Following (Fodor, 1987), let us call this very simple hypothesis "Intentional Realism."

Consider the language of thought explanation of the intrinsic connections among possible occurrent thoughts. According to the language of thought story, if our hypothetical cognitive agent were to lose the capacity to think that John loves Mary, this could be because the agent lost the concept of John, or of loves, or of Mary, or the grammatical construction that forms the thought that John loves Mary. But, if the agent were to lose one of those things, then the agent would thereby lose the capacity to form the thought that Mary loves John. Similarly, if the agent were to have the capacity for the thought that John loves Mary, then the agent must have the concept of John, the concept of loves, the concept of Mary, and the grammatical construct that forms the thought that John loves Mary. But, a cognitive agent that has all that would thereby have the capacity to form the thought

[12] Cf, e.g., Matthews (1997, p. 162), Johnson (2004), Travieso et al. (2014, p. 378). For a defense of this stronger claim about systematicity, see McLaughlin (2009).

[13] See Chemero (2014).

[14] Ramsey (2014, p. 258) concurs that Fodor and Pylyshyn mean to take on all rivals.

https://doi.org/10.1387/theoria.20464

that Mary loves John. Here we have a straightforward explanation of the counterfactual dependence between the thoughts. What, however, would be the Intentional Realist account? Why would losing the symbol ♥, meaning John loves Mary, have any implications for losing the symbol ♣, meaning Mary loves John? There is no account.

Consider the language of thought explanation of the semantic relatedness of possible occurrent thoughts. In theory, one could have an agent whose possible occurrent thoughts are {John loves Mary, bears sleep in the woods, Einstein was a physicist}. But, one does not find such agents. Why? Because the stock of possible occurrent thoughts is built up out of a stock of syntactic and semantic atoms that form a set of semantically related thoughts. What, however, would be the Intentional realist account? Why would the stock of symbols be semantically related, rather than not?

Notice that Fodor and Pylyshyn, on the one hand, devote one section of their paper to the intrinsic connections among thoughts and another to the semantic relatedness of thoughts. This suggests that they are supposed to be separate explananda. On the other hand, they claim that, "Compositionality is closely related to systematicity; perhaps they're best viewed as aspects of a single phenomenon." Then, later, they expand upon this,

> We now add that which sentences are systematically related is not arbitrary from a semantic point of view. For example, being able to understand 'John loves the girl' goes along with being able to understand 'the girl loves John', and there are correspondingly close semantic relations between these sentences: in order for the first to be true, John must bear to the girl the very same relation that the truth of the second requires the girl to bear to John. By contrast, there is no intrinsic connection between understanding either of the John/girl sentences and understanding semantically unrelated formulas like 'quarks are made of gluons' or 'the cat is on the mat' or '2 + 2 = 4'; it looks as though semantical relatedness and systematicity keep quite close company. (Fodor et al., 1988, pp. 41-2; cf., Fodor et al., 1990, p. 202)

This later passage obviously suggests an intimate connection between the two explananda. Here is an interpretation of what is going on. Fodor and Pylyshyn seem to have in mind three distinct psychological features to be explained. First, there are intrinsic connections between some possible occurrent thoughts. Second, there are semantic relations among some possible occurrent thoughts. And, third, the possible occurrent thoughts that are intrinsically connected are also semantically related. John loves Mary and Mary loves John are intrinsically connected and semantically related. Why? There is no conceptual or logical connection between intrinsic connectedness and semantic relatedness. In principle, one could have a cognitive agent in which there are thoughts that are intrinsically connected, but not semantically related. Such a cognitive agent might be such that its capacity to think that John loves Mary and its capacity to think that the cat is on the mat are intrinsically connected, but they are clearly not semantically related in the relevant sense. Moreover, in principle one could have a cognitive agent in which there are thoughts that are semantically related, but which are not intrinsically connected. Such an agent might have the capacities to think that John loves Mary and Mary loves John, but is such that neither thought is counterfactually dependent on the other. But, given that intrinsic connectedness and semantic relatedness do go hand in hand, a cognitive scientist should have an explanation of this. But, the language of thought story given in the last paragraph gives an answer to this. The language of thought explanation of the intrinsic connection between thoughts yields an explanation of their semantic relatedness "for free" one might say. Once one has in place the language of thought account of the intrinsic connections between possible occurrent thoughts, one needs no additional hypotheses to explain the semantic relatedness of possible occurrent thoughts. There is a simple, non-*ad hoc* account of why intrinsically connected thoughts are semantically related. This is what seems to underlie the sense in which the language of thought hypothesis is supposed to be the best explanation of the systematic features of thought.[15]

This pass through the systematicity arguments has no references to any background linguistics literature on systematicity. But, even a casual reader of Fodor and Pylyshyn's discussion will notice the references to Chomskyan ideas in generative linguistics. In what remains of this paper, I will try to provide what seem to be some of the salient features of this background in order to elaborate on some of the finer points of the systematicity arguments. I do not mean to claim that Fodor borrows the systematicity arguments from Chomsky or from theories of transformational grammar. Instead, Fodor draws on some of the implicit ideas in linguistics to develop the systematicity arguments.

2. *Why Believe Language and Thought have Systematic Features?*

While some have suggested that the intrinsic connections and semantic relations among thoughts is a myth, I propose to flesh out Fodor's claim that the principal ideas of the systematic arguments may be found in the earlier discussions of the productivity of natural language. Consider Chomsky's following contention:

> The central fact to which any significant linguistic theory must address itself is this: a mature speaker can produce a new sentence of his language on the appropriate occasion, and other speakers can understand it immediately, though it is equally new to them. Most of our linguistic experience, both as speakers and hearers, is with new sentences. (Chomsky, 1964, p. 50)

Notice that this contention does not presuppose that language is productive. It could be that speakers/hearers are able to cope with novel sentences, because there are infinitely many of them. Or, it could be because there is a large, but finite number of them. Notice what Chomsky's "central fact" implies. Given experience with a finite sample of sentences from a natural language, speakers develop the capacity to produce/understand many—perhaps infinitely many—novel sentences. Given experience with a sample of English sentences, children will "extrapolate" from their sample to many, many new sentences. Thus, there appears to be a counterfactual dependence between the production/understanding of some sentences and the production/understanding of others.

There is little doubt that Fodor was aware of this putative "central fact," as Chomsky's paper appeared in the (Fodor et al., 1964) anthology. Moreover, Katz et al. (1963) (reprinted in the anthology along with Chomsky's paper) expand on Chomsky's point in the following:

> A fluent speaker's mastery of his language exhibits itself in his ability to produce and understand the sentences of his language, INCLUDING INDEFINITELY MANY THAT ARE WHOLLY NOVEL TO HIM ... The emphasis upon novel sentences is important. The most

[15] Cf., Fodor (1987, p. 149); Aizawa (1997, p. 117); Aizawa (2003, Chapter 2).

> characteristic feature of language is its ability to make available an infinity of sentences from which the speaker can select appropriate and novel ones to use as the need arises. That is to say, what qualifies one as a fluent speaker is not the ability to imitate previously heard sentences but rather the ability to produce and understand sentences never before encountered. The striking fact about the use of language is the absence of repetition—almost every sentence uttered is uttered to the first time. This can be substantiated by checking texts for the number of times a sentence is repeated. It is exceedingly unlikely that even a single repetition of a sentence of reasonable length will be encountered. … Since a fluent speaker is able to use and understand any sentence drawn from the *infinite* set of sentences of his language, and since, at any given time, he has only encountered a *finite* set of sentences, it follows that the speaker's knowledge of his language takes the form of rules which project the finite set of sentences he has fortuitously encountered to the infinite set of sentences of the language. (Katz et al., 1964, pp. 481-2)

Here Katz and Fodor embraced what was the going conclusion of the day, namely, that the set of sentences of a natural language is infinite, but, as noted above, that is an inessential feature of the argument for a compositional basis to natural language. The principal point is that language acquisition involves taking a finite set of sentences, and then extrapolating to a larger set.

But, what are these extrapolations like? Again, simplifying in Fodorian fashion, we might have at one time the set of producible/understandable sentences

(3) "John loves John"
"John loves Mary"
"Mary loves John"
"Mary loves Mary".

But, then, the agent learns the word "likes," from "John likes Mary," so that the set of producible/understandable sentences becomes

(4)	"John loves John"	"John likes John"
	"John loves Mary"	"John likes Mary"
	"Mary loves John"	"Mary likes John"
	"Mary loves Mary".	"Mary likes Mary"

But, then, at a later stage, the agent learns, say, the passive grammatical construction, so that the set of producible/understandable sentences is

(5)	"John loves John"	"John likes John"
	"John loves Mary"	"John likes Mary"
	"Mary loves John"	"Mary likes John"
	"Mary loves Mary".	"Mary likes Mary"
	"John is loved by John"	"John is liked by John"
	"John is loved by Mary"	"John is liked by Mary"
	"Mary is loved by John"	"Mary is liked by John"
	"Mary is loved by Mary"	"Mary is like by Mary."

At each stage, there will be pairs of sentences that are intrinsically connected and semantically related. So, in (5), there will be

"John loves Mary" and "Mary loves John,"
"John likes Mary" and "Mary likes John,"

"John is loved by Mary" and "Mary is loved by John", and
"John is liked by Mary" and "Mary is liked by John"[16]

Notice that that there are two types of counterfactual dependencies in play here. One is diachronic; the other synchronic. So, there is a diachronic dependence between an agent's encounter with the sentence, say, "John likes Mary" at time t_0 and the possession of all the "like sentences" in (4) at a later time t_1. The agent would not have all the members in the set (4) at t_1 if the agent had not encountered "John likes Mary" at t_0. The intrinsic connection of interest in the systematicity arguments, however, is synchronic. The idea is that the agent would not have the capacity to produce/understand the sentence "John loves Mary" at time t_0 if the agent did not have the capacity to produce/understand the sentence "Mary loves John" at time t_0, and *vice versa*.[17]

But, what mechanism enables this pattern of extrapolation? Katz and Fodor propose that, "This problem requires for its solution a rule which projects the infinite set of sentences in a way which mirrors the way speakers understand novel sentences. In encountering a novel sentence, the speaker is not encountering novel elements but only a novel combination of familiar elements." (Katz et al., 1963, p. 171). The elements here are the words or morphemes of the language and the grammatical structures underlying them. So, one can see in this argument that there will be counterfactual dependencies and semantic relations among some of the pairs of sentences a speaker can understand. Indeed, one can see that, on this account, the sentences that are counterfactually dependent are also semantically related.

At the risk of belaboring Fodor's familiarity with this argument, we might note the following from one of Fodor's earliest papers:

> Nevertheless, it seems that some aspects of the ability to understand and produce novel sentences may be characterized in terms of our present knowledge about the systematic relations of sentences in natural languages. To give one example: it is clearly a necessary condition for understanding a new sentence that one should be implicitly capable of giving a grammatical analysis of the sentence. This involves at least the ability to decide correctly upon those substitutions for various components of the sentence which preserve grammaticality. The set of such substitutions, however, partially determines the relation between the sentence in question and other sentences in the language. Thus, understanding
> 1. The boy went to the store
> involves knowing such facts as that
> 2. The girl went to the store
> is a grammatical sentence in English while
> 3. The gives went to the store
> is not. (Fodor 1961, 73)

[16] Travieso et al. (2014) note that "John loves Mary" and "Mary is loved by John" are not counterfactually dependent, since speakers typically acquire the former years before the latter. That is, of course, right, but misinterprets which pairs Fodor supposed to be both semantically related and counterfactually dependent.

[17] NB. The diachronic counterfactual dependence is asymmetric, where the synchronic counterfactual dependence is symmetric.

It is striking that Fodor uses the word "systematic" to describe these relations. This use seems to hark back to Fodor's talk of "intrinsic connections" between thoughts as in (1) above, but not the idea of counterfactual dependence between thoughts as found in this, and many other passages like it:

> What does it mean to say that thought is systematic? Well, just as you don't find people who can understand the sentence 'John loves the girl' but not the sentence 'the girl loves John,' so too you don't find people who can think the thought that John loves the girl but can't think the thought that the girl loves John. Indeed, in the case of verbal organisms the systematicity of thought follows from the systematicity of language if you assume-as most psychologists do-that understanding a sentence involves entertaining the thought that it expresses; on that assumption, nobody could understand both the sentences about John and the girl unless he were able to think both the thoughts about John and the girl. (Fodor et al., 1988, p. 39; cf., Fodor & Pylyshyn, 1988, pp. 41-2, quoted above)

It might have been helpful had Fodor and Pylyshyn provided a reference to Fodor (1961), Katz et al. (1963), or Chomsky (1964). Nevertheless, if one were to take Fodor and Pylyshyn at their word that there were productivity arguments for a combinatorial syntax and semantic in natural language, one could track them down.

It is probably worthwhile to note three points where the ideas and arguments of Chomsky's and Fodor's early writings differ from what appears in the systematicity arguments. First, and most obviously, the early arguments were concerned with the productivity of natural language, hence the combinatorial syntax and semantics of natural language, whereas the later arguments are concerned with the productivity and systematicity of thought. Second, in the early arguments, Fodor and Chomsky relied on the novelty of sentences encountered. As noted above, sentences could generally be novel whether there were finitely or infinitely many of them. And this would seem to form a perfectly sound basis for a systematicity argument. Yet, it is probably only a manifestation of Fodor's idiosyncratic style to push the envelope so as not to rely on the novelty of encountered sentences at all. Instead, he could make do with homely examples, such as "John loves the girl" and "The girl loves John," and that anyone who understands the sentence "John loves the girl," will *ipso facto* extrapolate to an understanding of "The girl loves John." Third, in the passage from Fodor (1961) cited above, Fodor took the ability to give a grammatical analysis of a sentence to be constitutive of the ability to understand that sentence. Roughly speaking, if you cannot assign a parse tree to a sentence, you do not understand that sentence. So, a tourist who can produce a sound string does not thereby have a fluent speaker's understanding of a corresponding sentence.[18] One need not embrace this, however, in order to appreciate the force of the systematicity arguments.

3. *Inference to the Best Explanation in Generative Linguistics*

Turn now to another bit of the background to Fodor's presuppositions in the systematicity arguments. Chomsky's "Three models for the description of language" (Chomsky, 1956),

[18] Cf., Fodor, 1987, p. 149.

and *Syntactic Structures* (Chomsky, 1957), are among the seminal contributions to the cognitive revolution. In his introduction to the second edition, in 2002, for example, David Lightfoot described *Syntactic Structures* as "the snowball which began the avalanche of the modern 'cognitive revolution'" (Lightfoot, 2002, v). These works were surely well understood by Fodor. They were, for example, part of the background to Katz et al. (1963). Moreover, Fodor (1961) cites *Syntactic Structures* as part of the theoretical backdrop to his paper.

Although Chomsky allows that a language may either be finite or infinite, he maintains that human natural languages are unbounded. The human mind being finite, however, there must be finite means for generating and understanding each of the sentences of the language. The first model Chomsky examines for the description of natural languages is that they are given by so-called "finite state models." (Cf., Chomsky, 1956, section 2; Chomsky, 1957, chapter 3.) Chomsky argues against the adequacy of such models on the grounds that they do not generate many of a wide range of types of sentences found in English. To borrow some later terminology, Chomsky would say that finite state grammars are *observationally inadequate* (Cf., Chomsky, 1964, p. 53).

In the initial formulation of the argument against finite state models, Chomsky presupposes that there is no finite bound on the length of a sentence, in other words, that English is productive. Later, he entertains the hypothesis that there is a finite bound on the length of an English sentence. So, at some point, one would find that, say, "Carl and Fred and Larry and Tom ate the cheese" is an English sentence, but "Carl and Fred and Larry and Tom and Ken ate the cheese" is not. Such a move is comparable to Fodor's setting aside the productivity of thought in favor of considerations of intrinsic connectedness and semantic relatedness. Of the possibility of a finite bound on sentence length, Chomsky writes,

> Such arbitrary limitations serve no useful purpose, however. If these processes have no finite limit, we can prove the literal inapplicability of the [finite state] theory. If the processes have a limit, then the construction of a finite state grammar will not be literally out of the question, since it will be possible to list the sentences, and a list is essentially a trivial finite state grammar. But this grammar will be so complex that it will be of little use or interest.

If there is no finite limit on the length of sentences—if it is admitted that natural language is productive, then the model of language is "literally inapplicable" per the argument above. It is observationally inadequate. If, however, there is a finite bound on sentence length, then the grammar would be of the sort postulated by the Intentional Realist in section 1 above, namely, the well-formed strings of the language are just {◆, ♥, ♣, ♠, ...}. Such a grammar, used for a genuinely significant fraction of English "would serve no useful purpose" and "will be so complex as to be of little use or interest." Chomsky, thus, thinks that there is more to having a good grammar—a good system of representation—over and above mere generative capacity. There is more to having a good account of a system of representation than merely providing for observational adequacy. A better account is one that is, in some sense, simpler.

Finite state grammars were only the first of three models Chomsky considered. In addition, there are phrase structure grammars and transformational grammars. Chomsky prefaces his case for transformational grammars over simpler phrase structure grammars with the following:

> The strongest possible proof of the inadequacy of a linguistic theory is to show that it literally cannot apply to some natural language. A weaker, but perfectly sufficient demonstration of inadequacy would be to show that the theory can apply only clumsily; that is, to show that any grammar that can be constructed in terms of this theory will be extremely complex, *ad hoc*, and 'unrevealing'. (Chomsky, 1957, p. 34)

So, there are (at least) two grounds upon which a representational system can be faulted. First, it might be faulted for not being able to represent all that needs to be represented. This was the problem facing "grammars" that are mere lists. One might say that such a grammar does not provide an explanation of why certain strings are members of a language. Second, a representational can be faulted for being *ad hoc* or for being excessively complex. This was the problem facing phrase structure grammars versus transformational grammars. And, we might understand an *ad hoc* complex explanation to be worse than a non-*ad hoc* simple explanation. An *ad hoc* complex explanation is not one we would infer through an inference to the best explanation.

Return now to the way in which Fodor is implicitly relying, not on observational adequacy, but on some standard of better explanation. Recall the claim that

> No doubt it is possible ... to wire a network so that it supports a vector that represents aRb if and only if it supports a vector that represents bRa; ... The trouble is that, although the architecture permits this, it equally permits Smolensky to wire a network so that it supports a vector that represents aRb if and only if it supports a vector that represents zSq. (Fodor et al., 1990, p. 202)[19]

The idea is that one might get a system that can represent both aRb and bRa, hence be observationally adequate and have thoughts that are semantically related. And, one might also have it be the case that these representations are counterfactually dependent. But, in addition, one wants some explanation of why it is that the thoughts that are semantically related are also counterfactually dependent. Why do these two features of thought go together? On the language of thought hypothesis, the claim is that once one has on board what is needed in order to explain the counterfactual dependencies in thought, one has on board what one needs in order to explain the semantic relatedness of possible thoughts (and *vice versa*). One needs no "extra assumptions". The language of thought explanation is, in that sense, simpler. It is not, in that sense, *ad hoc*.[20]

4. The "Central Fact" about Language is Still Worthy of Attention

Consider, now, how Chomsky's "central fact" about language might bear on more recent attempts to deal with the systematicity arguments. Gomila et al. (2012, pp. 102-3) and Travieso et al. (2014, pp. 373-4) propose that the systematicity challenge is to provide a genetic explanation of why thought is systematic and that the correct explanation is that the systematicity of thought is induced by the systematicity of natural language. In more de-

---

[19] Cf., Fodor et al., 1988, p. 50; Fodor et al., 1990, p. 202; Ramsey, 2014, pp. 255-6.

[20] Aizawa, 1997, 2003, try to provide examples from the history of science that illustrate the kind of reliance on *ad hoc* hypotheses that are implicitly taken to make for inferior explanations.

tail, they seem to think that thoughts get their contents from language, so that the patterns among the contents of possible thoughts arise from the patterns among contents of sentences in the agent's natural language.[21] So, an agent comes to have the set of possible occurrent thoughts {John loves John, John loves Mary, Mary loves John, Mary loves Mary} because the agent has, say, the set of possible English sentences {"John loves John," "John loves Mary," "Mary loves John," "Mary loves Mary"}. In other words, they propose to explain the content relatedness of possible occurrent thoughts in terms of the content relatedness of possible sentences of a natural language.

Grant them this for a moment. This still leaves the counterfactual dependence among thought unexplained. Why is it that there is a counterfactual dependence between the John loves Mary thought and the Mary loves John thought? Travieso, Gomila, and Lobo provide no account. Nor do they provide an account of the co-occurrence of counterfactual dependence and semantic relatedness.

Now take back what was granted in the last paragraph. Of course, Gomila, Travieso, and Lobo can claim that the capacity to extrapolate from "John loves Mary" to "Mary loves John" in English also induces the capacity to extrapolate from the thought that John loves Mary to the thought that Mary loves John in "mentalese." But, that claim can't be all there is to the story. Why, as things stand, is this not a mere *ad hoc* stipulation? Why does the linguistic extrapolation enable the comparable extrapolation in thought? Why, for example, doesn't the linguistic extrapolation outstrip the extrapolation in thought? Suppose that English has a compositional syntax and semantics, but that thought does not. Why, then, is it not the case that agents have the capacity to parse the sentence "John loves Mary" and therefrom extrapolate to a parse tree for "Mary loves John," but still not understand (think the thought) that Mary loves John. To put the matter another way, in the case of English, one would have a combinatorial explosion in the number of possible sentences as the lexicon increases in size, but why, in the case of thought, would there be a similar combinatorial explosion, unless there were in fact combinations to explode in number? Again, if only language has a combinatorial syntax and semantics, then it would seem that the representational resources of language would easily outstrip the representational resources of thought. By contrast, according to the language of thought account, the explosive growth in both language and thought is possible because of the combinatorics of both. But, if thought has a combinatorial syntax and semantics, then thought would have a language of thought.

Another way of putting the concern here is that Gomila et al. seem not to have come to grips with the apparent implications of what Chomsky and Fodor took to be the "central fact" of natural language. Why is it that, upon exposure to a finite number of sentences, a speaker is able to produce and understand (perhaps non-denumerably) many novel sentences? The Fodorian/Chomskyan answer is that natural language has a combinatorial syntax and semantics. But, if a speaker can understand all of these sentences of her natural language, then how it is that thought is able to "keep up" with language, if not with a syntactically and semantically combinatorial language of thought?

---

[21] They actually claim only an "isomorphy" between the sentences of natural language and of thoughts, but I assume they want synonymy. "John loves Mary" is isomorphic with "Mary loves John", and "Bob likes Alice", but is synonymous only with the first. They would seem to mean synonymy, since a set of semantically related sentences of a natural language could be isomorphic with a set of semantically unrelated sentence of mentalese.

5. *Conclusion*

My principal aim in this paper was to look at the systematicity arguments through the lens of some of Chomsky's and Fodor's early work in linguistics. This background is frequently mentioned in Fodor's writings on systematicity and productivity, but has not been given the kind of detailed references to the literature that would help a newcomer to the topic easily track this down. Nor has this background been reviewed to bring out its relevance. Nevertheless, it can provide us a better understanding of Fodor's thinking about systematicity, as well as a better understanding of the systematicity arguments for a language of thought and certain attempts to reply to those arguments.

REFERENCES

Aizawa, K. (1997). Explaining systematicity. *Mind & Language, 12*(2), 115-136.
Aizawa, K. (2003). *The systematicity arguments*. Boston, MA: Springer.
Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on Information Theory,* 2, 113-124.
Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
Chomsky, N. (1964). Current Issues in Linguistic Theory. In Jerry Fodor and Jerrold Jacob Katz, eds., *The structure of language: Readings in the philosophy of language* (pp. 50-118). Englewood Cliffs, NJ: Prentice-Hall.
Chomsky, N., & Lightfoot, D. (2002). *Syntactic Structures*. *2nd edition*. Berlin: Mouton de Gruyter.
Fodor, J. (1961). Projection and paraphrase in semantics. *Analysis* 21, 73-77.
Fodor, J. (1987). *Psychosemantics: The problem of meaning in the philosophy of mind*. Cambridge, MA: MIT Press.
Fodor, J., & Katz, J. J. (1964). *The structure of language: Readings in the philosophy of language*. Englewood Cliffs, NJ: Prentice-Hall.
Fodor, J., & McLaughlin, B. (1990). Connectionism and the problem of systematicity: Why Smolensky's solution doesn't work. *Cognition,* 35, 183-204.
Fodor, J. & Pylyshyn, Z. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition* 28, 3-71.
Johnson, K. (2004). On the systematicity of language and thought. *The Journal of Philosophy,* 101, 111-139.
Katz, J. J. & Fodor, J. (1963). The structure of a semantic theory. *Language,* 39, 170-210.
Matthews, R. (1994). Three-concept Monte: Explanation, implementation and systematicity. *Synthese,* 101, 347-363.
Matthews, R. (1997). Can connectionists explain systematicity? *Mind & Language*, 12, 154-177.
McLaughlin, B. (2009). Systematicity redux. *Synthese,* 170, 251-274.
McLaughlin, B. (2014). Can an ICS architecture meet the systematicity and productivity challenges? In Paco Calvo and John Symons, eds., *The architecture of cognition: Rethinking Fodor and Pylyshyn's systematicity challenge* (pp. 31-76). Cambridge, MA: MIT Press.
Ramsey, W. (2014). Systematicity and architectural pluralism. In Paco Calvo and John Symons, eds., *The architecture of cognition: Rethinking Fodor and Pylyshyn's systematicity challenge* (pp. 253-276). Cambridge, MA: MIT Press.
Rey, G. (2018). Remembering Jerry Fodor and his work. *Mind & Language* 33, 321-341.
Travieso, D., Gomila, A., & Lobo, L. (2014). From systematicity to interactive regularities: Grounding cognition at the sensorimotor level. In Paco Calvo and John Symons, eds., *The architecture of cognition: Rethinking Fodor and Pylyshyn's systematicity challenge* (pp. 371-396). Cambridge, MA: MIT Press.

Kenneth Aizawa received his PhD in History and Philosophy of Science from the University of Pittsburgh in 1989. He works primarily in the history and philosophy of psychology and is the author of *The Systematicity Arguments* and, with Fred Adams, *The Bounds of Cognition*. He has been Professor of Philosophy at Rutgers University, Newark, since fall 2013.

Address: Department of Philosophy. Rutgers University, Newark. 175 University Avenue. Newark NJ, 07102. USA. E-mail: ken.aizawa@gmail.com