Nativism in Linguistics:

Empirical and Theoretical Issues

Taki Flevaris

Abstract:

The debate concerning linguistic nativism has been ongoing for years. For that reason, few have bothered to clarify the initial intricacies of the argument, to fully understand why it is being claimed. This has lead to serious misunderstandings on both sides. What is needed is a unified exposition of the nativist argument; that is what follows.

The debate has lead to the formation of two general camps, which must be identified before this discussion can be fruitful. Specific nativist theories vary, but the crux of the position is that children have at least *some* linguistically specific knowledge innately. Data-driven theorists hold that children have no linguistically specific knowledge at birth. The nativist does not deny that children may *also* learn from their input, while the data-driven theorist holds that children learn *only* from their input.

In this paper I hope to shed light on some of the disputes that have transpired. First, I will ground the argument for nativism and show why it exists. From this it will become clear exactly what needs to be shown to discount it. Second, I will examine whether the critics of nativism have made significant steps toward establishing a data-driven account of language acquisition: I will investigate whether or not the critics have proven any of the necessary points outlined in the first section. Finally, I will provide supplementary evidence which bolsters the nativist theory, all of which is unexplainable by a data-driven account of language acquisition.

TABLE OF CONTENTS

1. Introduction

Any scientific theory is an attempt to explain a given property of the world, which is usually represented by a type of data (e.g., the theory of gravity serves to explain why objects drop toward the Earth's center). In the same vein, a scientific theory has emerged within the field of linguistics. Begging explanation is the acquisition of human language, and the challenge is explaining how nearly every speaker of a given language attains the same rich and structured linguistic knowledge by adulthood. The data is twofold: 1) the environmental input to the language learner, and 2) the linguistic knowledge that the language learner has attained at any given point in time. The problem lies in aligning the two. After a critical analysis of the general experience of language acquisition and all attendant circumstance, it becomes clear that the only way to align the data is the supposition of innate linguistic knowledge. This position has become known as *nativism*. The alternative to nativism is the theory of *data-driven learning*, which stipulates that children learn language solely from their environmental stimulus, and that they have no innate linguistic knowledge.

It is laudable that linguists are questioning nativism; critique is a necessary part of any scientific development. Even so, these critics have failed to provide any plausible alternatives, nor have they put a significant dent into the nativist argument itself. Nativism is not proven beyond any doubt, nor has it ever been claimed to be; rather, it is a necessary assumption given the available data and our current understanding of the world. Lasnik & Uriagereka (2002) points out that detractors of nativism must themselves supply a "scientific alternative." Otherwise, until a more convincing theory rears its head, nativism is an unavoidable conclusion.

2. The Nativist Argument

2.1 Establishing the Nativist Argument

One important presumption of nativists (and linguists by and large) is that adults generally converge on the same interpretation of what is grammatical and what is not. This is what it means to speak 'English' or 'Irish'; namely, that we have the same judgments of grammaticality as other speakers. In fact, this is what allows us to call anything a language; after all, there is no 'English' in the world to which we can point.

Of course, it must be made clear what "grammatical" means. In a broad sense, it encompasses whatever a native speaker considers possible for use in their linguistic communication with others. This concerns not logical possibility but pragmatic possibility (i.e., would you ever hear that sentence uttered?). There are exceptions: suppose I were goaded into uttering an ungrammatical statement; my declaration would be used in a normal conversation, but would be ungrammatical.

Barring such extraneous hypotheticals, the concept of linguistic grammaticality should be clear: what determines grammaticality is simply what the speakers in a given group think is grammatical. If grammaticality were defined stringently, it would have to be true for people who spoke different languages (with different structures and phonemes). In contrast, if it were defined only subjectively for each person, there would be no constancy with which communication could take place, making grammar meaningless. A language without rules common to its speakers is not a language at all. Thus grammar must be a constant among speakers of any given language, and the

majority of those speakers define it. In the words of Michael Tomasello, "To become a competent speaker of a natural language it is necessary to be conventional: to use language the way that other people use it." (Tomasello, 2000: 209) This is not to say that adults never disagree whether a given statement is grammatical; the point is instead that they disagree *almost* never. The structure and rules of English, then, are not to be found in *Merriam-Webster* or *American Heritage*, but in the minds of English speakers.

Although grammar is nearly a universal constant among adults, it is absent among newborns. It is clear that children begin with no knowledge of any specific languages. This is evidenced first by the gradual nature of the learning process, and second by the fact that children learn only the language(s) spoken around them during maturation. Without an initial understanding of their native language or its rules, children eventually all attain the same linguistic knowledge and communicate with one another regularly and with ease. The most straightforward way to explain this would be to refer to their observation of their native language. The idea would be that by listening to those around them, children would be able to abstract from the sentences they hear and discover the structural rules used to form them. It is obvious that at least some of the learning process is driven by the environment (e.g., the discerning of labels of specific objects). What is not obvious is whether the environmental input to nearly every child is sufficient to explain the acquisition of the complicated structural knowledge that, as adults, they all have. After all, the individual experiences of children clearly vary in many ways.

We are left with an inconsistency: children, who begin with no specific language and whose experiences with language are varied (and sometimes quite lacking), all

eventually have the same intuitions about grammaticality in adulthood. Only nativism can account for this progress to eventual unity. The argument is twofold:

1)Positive: Linguistic knowledge is attained for which there is little if any evidence in the environmental stimulus. This means that children cannot possibly learn all of their linguistic knowledge from their environment alone. Hence, they must have innate linguistic knowledge.

2)Negative: Humans create sentences they've never heard before; language acquisition is not merely imitative. For this reason, our language must abide by structural rules, and humans use these rules to govern their language production. However, for any given statement there is an infinite number of possible explanatory structures and rules from which the statement could have been created. Although the environmental stimulus lacks evidence of the *ungrammaticality* of any of these structures, children only entertain a limited number of possible hypotheses, usually only the grammatical ones. Since nothing in the environment guides them in this way, there must be something guiding them innately.

The lack of positive and negative evidence during children's language acquisition proves that they could not possibly begin tabula rasa.

Further details concerning children's experience preserve the clout of the nativist view. First, the universality of grammar entails that for any data to be seen as conducive to empirical learning it must be present universally. That is, it must be present for nearly every child, since it is the very fact that *every* child acquires this knowledge which we are

trying to account for. Fair enough, says the empiricist. Yet we do not stop here; it must

also have sufficient intensity. This is because the input children are presented with is not

perfect: adults make mistakes from time to time in their language production (e.g., slips

of the tongue, false starts, etc.), thus producing ungrammatical statements. Newport et al

(1977) found that one out of every 1500 utterances spoken to children is a true disfluency

(e.g., '*The boy Susan didn't do it'). This does not include any sentences that Newport et

al found to be unanalyzable ("utterances which were either incomplete and broken off in

midstream or were either partially or wholly unintelligible"), which would also confound

the children's input. Children are somehow able to filter out these ungrammatical

statements, as evidenced by their later knowledge that these statements are, in fact,

ungrammatical. The most likely way that children could distinguish such errors would be

by their infrequency. In that case, however, they would also discount any grammatical

statements made with the same level of infrequency. Of course, children wouldn't

discount single statements, since almost every statement is unique. Rather, they would

have to be able to identify the abstract structures behind given statements, and then

discount whatever structures appeared infrequently. This poses significant problems for a

data-driven account, in that children know the grammaticality of structures which appear

less frequently than relevant errors. It is also difficult to explain how children would

initially decipher which structures were used to form the given statements before

deciding what kind of statements are ungrammatical.

Children also make errors of their own. Interestingly, despite the variability of

each child's specific acquisition experience, children tend to make the same types of

errors. Additionally, the number and types of errors they make is very limited. While

children don't assume random hypotheses of grammatical structures, they do assume more hypotheses than those suggested by experience alone. This is inconsistent with a data-driven account. If children discount certain structures because of their infrequency, it would not hold to reason that they would assume any structures to be grammatical which are *never* upheld in the environment. This issue will be given more attention later, but for now suffice it to say that children are not merely reproducing the structures they hear.

We have now seen that there is a large set of circumstances which necessitate some sort of innate linguistic knowledge. This knowledge must designate certain structures as possibly grammatical and others as ungrammatical, and in the literature has been generally dubbed Universal Grammar (UG). It is important to note that the distinction made between the positive and negative side of the nativist argument is of great consequence, because either side is *sufficient* in proving the existence of UG. That is, in order to discount the necessity of UG, one must disprove *both* the negative and the positive side of the nativist argument.

Let's take a closer look. What would be involved in discounting the positive side of nativism? One would have to prove that, in fact, there is no extraneous grammatical knowledge that is unaccounted for in the input to children. A problem would remain: although the data would hypothetically evidence all structures that are grammatical, it would still fail to show any structures to be *ungrammatical*. Unless children considered *only* the structures in the input to be possibly grammatical, without innate knowledge they could not be limited in the hypotheses they assumed. Children entertain many novel hypotheses, but are limited in doing so. One point of interest is that the types of

grammatical errors that children produce are usually analogous to structures of other languages; quite a coincidence if UG had not made it so. The point to be made is that even when the positive argument has been discounted, the negative argument still necessitates UG.

Now for the reciprocal: To discount the negative argument, one would have to establish that the input discards every ungrammatical structure, something I will later show to be impossible. For the sake of argument, however, let us assume that such a task were achieved. Even then, children would know that some structures are grammatical for which no evidence existed. Only UG could provide such knowledge if there is none in the input.

It seems fairly clear that both sides of the argument are sufficient for the existence of UG. The only way to cast doubt upon its existence, then, would be to establish either of the following:

(3) Children systematically uphold *only* whatever structures are definitively designated by their environmental stimulus to be grammatical

Or

(4a) There is enough environmental stimulus for knowledge of all grammatical structures

and

(4b) There is enough environmental stimulus to establish the ungrammaticality of all other structures

The two positions would both simply prove that the given input can somehow logically establish our final adult grammar. Proof of (3) would have to assume that the child has some sort of innate system for filtering the input, or that the input could be filtered by logic alone, while (4a) and (4b) would indicate that the child could begin tabula rasa.

Note that I have taken care not to claim that such arguments would *disprove* UG and nativism. While the establishment of such arguments would make a data-driven hypothesis more plausible, the nativist theory would remain a possibility. Although we express much of our linguistic knowledge regularly through dialogue, this does not automatically imply that children utilize it for learning, nor that it is by any means *necessary* for learning. We could purposefully instruct children with detailed linguistic analyses and informative examples of much of our language's complexity, viz. try to present all of the positive and negative evidence that nativists claim children don't have. Such a turn of events would not suddenly prove that children learned from this evidence, nor by any means would it expunge any innate knowledge if it does indeed exist. A further implication is that the scope of UG could (and probably does) extend beyond only those palpable cases in which evidence is lacking. Specifically, cases of lacking evidence which solidify UG probably signify a greater overarching system. Regardless of any evidence in experience, there remains one important matter that the data-driven accounts have difficulty explaining, namely that of language creation. Thankfully, our era of linguistic consciousness has been witness to the creation and development of an entirely new language; this issue will be discussed in detail later on.

We see now that the opponents of nativism can only seek to better establish the empirical hypothesis in the proof of (3), or of both (4a) and (4b). It seems to me that the most fortuitous route for doing so would be to prove (3), but the importance of this position has been largely (if not fully) ignored in the literature. In what follows, (4b) will be shown to be theoretically impossible to establish, and subsequently false. That being the case, (3) seems to be the *only* possible way to better establish the empirical hypothesis, but I will eventually show (3) to be unfounded as well. Surprisingly, the brunt of the debate has instead concerned the truth of (4a) (Tomasello 2000, Pullum & Scholz 2002, etc.), i.e., that children can attain their grammatical knowledge from the environment alone. (4b) has also been explored a bit (Marcus 1993, Saxton 1998, Strapp 1999, etc) but none have claimed to come close to establishing its truth.

To summarize: Under the nativist theory, children begin with UG, innate knowledge of what are possible grammatical linguistic structures. From wholly varied and impoverished experiences with no negative evidence and in most cases sparse positive evidence, children entertain a limited number of hypotheses and develop a unified final grammar in accordance with the positive evidence that *is* presented. Without UG, this positive evidence would not be sufficient to form a definitive grammar; in that case, adult grammars would vary significantly. We know that this is not the case.

2.2 Universal Grammar

Any exhaustive backing of nativism and UG should involve an assessment of what could possibly limit and guide children's hypotheses concerning linguistic

structures/rules. Such an investigation would be necessary for a data-driven account to even begin to explain how children could identify any structures to be ungrammatical (to respond to the negative argument). As I see it there are three rational principles by which a child could possibly be limited. The first is *simplicity*, by which the child would isolate the most basic possible interpretation to account for the data and assume said interpretation until it failed (was unable to correspond to subsequent statements of the same supposed structure). Assuming a data-driven account, this would most likely be the way in which children learn language. However, this is not the case. One relevant example is the formation of questions via auxiliary-movement, as in "Is he __ cold?" from the statement "He is cold." Here, the most basic rule to account for such question-formation would be something along the lines of 'move the first auxiliary to the front of the sentence'. This rule fails, however, as seen in more complex formations: "Is the boy who is dangerous __ in the building?" and not *"Is the boy who __ dangerous is in the building?" (from "The boy who is dangerous is in the building"). Crain & Nakayama (1987) tested 3- to 5-year old children on this type of question formation. In their own words:

> "It seems reasonable to suppose that children encounter the simpler forms first…
> one might suppose that at least *some* children entertain [the simplest hypothesis]
> during grammatical development, contrary to the claims of Universal
> Grammar… Our main finding was that *no* child produced an utterance in which a
> verbal element in a relative clause was moved."
>
> (Crain & Nakayama 1987: 526,530, emphasis added)

The finding that not a single child assumed a first-auxiliary hypothesis indicates that children do not abide by a principle of simplicity in limiting the hypotheses they consider. Moreover, considering that the number of possible hypotheses is infinite, it seems nearly impossible for any system to distinguish the most basic. After all, wouldn't fronting the last auxiliary be just as simple as fronting the first, and account for the same basic data?

The second possible principle is *probability*, which would involve monitoring which structures were produced most frequently and abiding by those, while being wary of structures which were produced infrequently. Unfortunately probability cannot substantially help the problem of choosing competing theories because there is almost always more than one structure which is possible given the entirety of the child's input (in other words, none of the possible structures is *most* probable). Beyond this, to abide by probability children would have to first know what structures produced the given statements, but this is the task probability is supposed to guide. Legate & Yang (2002) effectively establishes that children learn in clear defiance of probability:

"But children often learn their languages in ways that clearly defy such distributions. Take the subject drop phenomenon: while almost all adult English sentences contain a subject, children only acquire this aspect of the grammar at the much delayed age of 3:0. In contrast… the placement of finite verbs over negation/adverbs is acquired by French children at a far earlier age (1;6), while this form only occurs at a frequency of 7 percent in adult French. Thus, we have one grammatical pattern that is heard frequently but learned late, and the other that is heard rarely but learned early…"

The third and final principle by which language acquisition could abide is, quite simply, *Universal Grammar*. Under this interpretation, children possess innate knowledge of what hypotheses are possibly grammatical, and what hypotheses are hence ungrammatical. This accounts for the fact that children entertain a limited number of novel ungrammatical hypotheses, as well as for their knowledge of the grammaticality of certain structures not evidenced in their stimulus; UG is the only valid explanation for language acquisition currently available. This list is not necessarily exhaustive; other limiting principles could theoretically be possible, but such principles have never been proposed, nor can I conceive of any. It seems that no such principles could explain why children produce certain novel structures, nor why children know the grammaticality of structures for which there is little to no evidence -- neither of these sets of structures/rules derive from the input in any recognizable or logical way. In this regard, the list can be seen as *theoretically exhaustive*, as only UG (or some form of it) could account for these latter elements.

Given the two sides of the nativist argument, one can fashion a rudimentary idea of UG:

- UG must take the place of positive evidence: it must propose structures as being possibly grammatical.

- UG must take the place of negative evidence: it must rule out an infinite

number of structures as ungrammatical. In this way, UG implies that *only* what it

specifically allows can possibly be grammatical.

The most obvious course of action is to discover exactly what UG consists of. This is

precisely what nativists believe should guide much of today's linguistic investigation; the

past few years have seen much progress toward deciphering what the properties of UG

really are. The general sense is that UG is a constraint on what is possibly grammatical,

and serves as a guide to the language learner, all regardless of what their specific

language may be.

Structure clearly varies across languages, and many inter-linguistic structures are

even mutually exclusive. UG would have to allow for *any* of these structures, since a

child is apt to learning *any* language. Learning certain languages, then, would result in

the ungrammaticality of what is possible within UG. For example, in some languages a

head occurs before its complement, while in others the reverse is true (the semantics of

this example are not crucial, suffice it to say that two alternatives exist for a certain type

of construction). A native speaker of a language with either interpretation would

subsequently find the opposing interpretation ungrammatical. We have already stated

that both interpretations are possible in UG, and also that the input does not provide *any*

sufficient evidence for what (if anything) is ungrammatical. This is a seeming dilemma,

since children eventually find certain structures to be ungrammatical which they did not

initially.

This problem has been accounted for with the "Principles and Parameters" theory.

(Chomsky 1981) 'Parameters' would be constraints within UG that consist of alternative

grammatical possiblities. In the Head-Complement example, the parameter would consist of the two possible orders. Then, once enough positive evidence suggesting one of the orders was presented, that parameter would be set in accordance. This would make the opposite order ungrammatical. Ray Jackenfoff explains that with this theory, "Learning a language can then be thought of roughly as like customizing the settings in a software package." (Jackendoff, 75) Parameters wouldn't all necessarily function the same way. For example, a parameter could consist of two non-exclusive grammatical possibilities. After a given time, whatever positive evidence was presented in accordance with either of the possibilities would subsequently give them grammatical status, respectively. If such evidence was not presented, after a given time, either of the structures would become ungrammatical. This is in contrast to the types of structures which, even after a lack of evidence, are considered grammatical, as is witnessed by some of the novel (and locally ungrammatical) constructions of children, or their knowledge of grammatical structures for which no evidence exists. However, even some of *these* hypotheses -- namely the former -- are abandoned after a given period of time. Whatever the case may be, UG is a complicated assortment of linguistic knowledge and principles. Given the guidelines of scientific development previously mentioned, it is clear that for the time being the best theory for explaining language acquisition is UG. Further exploration into its makeup is clearly necessary in the linguistic field.

At this juncture it is helpful to imagine what language acquisition would be like *without* UG: without a limiting principle for language acquisition, we would be left with infinitely many possible hypotheses for any given construction. Hence, infinitely many random hypotheses would be considered and used in language production. This would

make eventual unified collaboration nearly impossible. For subsequent children, data would be more and more contradictory since new adults would be entertaining wholly varying and sometimes contradicting structures. One may argue that the principles of simplicity and probability would limit the number of possible hypotheses and hence limit the subsequent chaos. We know that children don't abide by such principles. However, even if they did, the number of considered hypotheses couldn't possibly be limited significantly, or there would at least be many more than there are today. In fact, it follows that there would be no constant grammar, meaning no grammar at all. This is not an accurate depiction of the world, which gives us all the more reason to uphold UG.

3. The Critics

3.1 Michael Tomasello and FCL

Nativism must first be discounted before any theory of data-driven learning can become plausible. This is overlooked by Tomasello (2000), which only seems to rationalize why a data-driven account is perhaps more plausible than nativists give credit. Tomasello refers to some unimpressive particulars concerning children's language acquisition which he sees as impossible to reconcile with a nativist view. One of the essential problems is that Tomasello fundamentally misinterprets nativism as an argument for, as he puts it, "that young children operate with adult-like linguistic competence." (Tomasello, 2000: 210) This is a mistake. Rather, nativism argues that

UG acts as a *guide*; this accounts for a gradual acquisition process, limited by factors like maturity and parameter-setting.

Tomasello thus sets out to disprove that children have full syntactic knowledge, but this is not an implicit conclusion of the nativist argument. If children do lack full syntactic competence, the nativist argument certainly remains valid. All of Tomasello's propositions and conclusions are fully compatible with nativism; in truth, some of his examples reinforce its necessity. Additionally, many of his misgivings refer to problems of both nativism *and* data-driven learning, which does not make either option more attractive. For example, he says that it is difficult to conceive of how children identify what a subject is, but the problem of subject-identification would be true of both intepretations of acquisition.

One of the findings that Tomasello focuses on is that children don't initially generalize rules of verb-formation in their production of different verbs. While they may show mastery of verb-formation rules for a few verbs, this will not necessarily carry over to their production of others; they will produce 'walked' from 'walk' and 'opened' from 'open' but not 'picked' from 'pick'. This, Tomasello claims, shows that they have no syntactic knowledge of something like 'verb' or of structural rules in general. First I must emphasize that this data is not inconsistent with nativism; language acquisition is undeniably gradual, partly for reasons aforementioned. Even so, Tomasello's interpretation is flawed in other ways.

The first problem with his conclusion is that it fails to distinguish between production and comprehension, which are quite different. Many suspect that children know much more than their production implies, and this has been shown to be the case.

While nativists will claim that children have innate linguistic knowledge, how this will translate into speech production is never made clear. In fact, testing for linguistic comprehension can be quite enlightening, as Hirsch-Pasek & Golinkoff (1996) explicates quite well:

> "In sum, studies of language comprehension have three advantages over studies using spontaneous production: they permit researchers to probe for structures that are not yet produced; they offer a new window onto the process by which the child acquires a particular structure before that structure emerges full-blown; and they permit a degree of methodological conrol not available from observing production."

>   (Hirsch-Pasek & Golinkoff 1996: 107)

Tomasello mentions some ways of testing syntactic knowledge, but one modus operandi for which he gives only a fleeting consideration is that of 'preferential-looking': children are presented with two different scenes, along with a linguistic (auditory) input which coincides with only one of them. Spelke (1979) provided the theoretical basis for such a procedure, in which it was established that children will attend to whatever event matches the auditory stimulus, and do so only if they actually *understand* what is being said. It is thus possible to force children into specific contextual choices, allowing the discovery of which interpretations and structures they may utilize, which are otherwise unobservable.

Along the same lines, Crain (1991) suggests that, much of the time, procedural complexity causes child error rather than syntactic incompetence. He refers to a few occasions in which procedural adjustment of a study in subtle ways lead to the emergence

of syntactic knowledge that had been formerly latent. Much of the time, children are unable to exhibit knowledge even if they possess it. Tomasello rebuts that because the children could produce the verb constructions for *some* verbs, performance factors were irrelevant. However, this is not necessarily the case. Children may have difficulty remembering to apply generalizations; perhaps their production is confounded by other factors like apprehension, contextual information, etc. In contrast, these problems would be absent in something like a preferential-looking test.

The argument for nativism does not necessitate that children have *any* syntactic knowledge whatsoever. The essential circumstance is that *whenever* given knowledge is attained, not enough positive or negative evidence has been part of the input up to that point. The earlier that any knowledge is attained, the greater the likelihood of a lack of evidence and subsequently the more attractive nativism becomes. This has driven researchers to demonstrate grammatical knowledge in children of very young ages (Lidz et al, in press; etc). However, one must realize that if knowledge is present at any point by which not enough evidence has been part of the environmental stimulus, the need for nativism to account for such an inconsistency is equally pressing. With this in mind, Tomasello's observation of a lack of syntactic knowledge in children is irrelevant if he cannot account for the verb-formation knowledge they *eventually do acquire*. In fact, many have claimed that linguistic knowledge is innate but not necessarily active at birth:

> "The term 'innate' also requires comment. For a first approximation, it means
> 'present at birth.' However, it is customarily used more broadly to denote a
> characteristic that appears automatically in the course of an organism's

development, whether before or after birth. For instance the number and

organization of human teeth, which develop after birth, can be said to be innate."

<div align="right">(Jackendoff, 72)</div>

The delayed application of verb-construction rules in children's speech is perfectly

coherent with the nativist position; hence, Tomasello falls substantially short of

challenging nativism.

Nativism assumes that children will attain language during a gradual process of

maturation and the setting of parameters. UG explains acquisition in the face of lacking

environmental evidence, which would otherwise be necessary for the child to reach adult

grammar. In this light, none of Tomasello's findings exclude UG as a linguistic guide,

nor do they shed any light on the nativist argument itself. The argument is still as

insistent as it has always been. Children's production and even their comprehension can

be considerably limited by factors like processing and memory capabilities, and not by

the absence of syntactic knowledge or UG.

Ironically, one of the studies which Tomasello cites to demonstrate of a lack of

children's syntactic knowledge solidifies the argument for UG. The issue at hand is the

production and acceptance of varying sentence formations, namely SVO, SOV, or VSO.

In English, the grammatical formation of sentences accords to SVO (subject, verb, object;

e.g., I ate the apple), while other languages abide by other formations such as SOV or

VSO. Tomasello's findings, as he stated them:

"…the children produced utterances that involved a different configuration of

SVO than was typical in almost all of the speech they had previously heard or

<div align="right">20</div>

produced… when they heard one of the non-canonical SOV or VSO forms, children behaved differently at different ages. Only 1 of the 12 children at 2;8 and 4 of the 12 children at 3;6 consistently 'corrected' the non-canonical adult word order patterns to a canonical English SVO pattern, whereas 8 of the 12 children at 4;4 did so."

(Tomasello, 2000: 222)

These children are producing structures which they never hear in their environmental stimulus – this is a fundamental part of the nativist argument. It may be objected that the children heard these structures immediately prior to their production of said errors, however these examples are more infrequent than many other forms of error that the children could -- but don't -- consider. We know language acquisition is not merely imitative. Most (if not all) of these children have *never* heard such constructions. For some reason, however, they are considering SOV and VSO to be grammatical. This concurs with the existence of UG, which allows these structures to be grammatical (as is witnessed by their use in certain languages).

Additionally, Tomasello's findings fit very nicely with a gradual account of parameter setting. The supposition would be that children have a parameter of varying structural forms like SVO, SOV, and VSO. These structures would each be possibly grammatical according to UG (but not necessarily mutually exclusive), hence it would follow that only over time would the parameters of SOV and VSO be set as ungrammatical while establishing SVO as a grammatical form. This is exactly the kind of progression that is evident in Tomasello's findings.

In spite of this, Tomasello objects to the validity of parameter-setting; he relies on Mazuka (1995) to formulate the objection:

"Setting a Head Direction parameter by analyzing the syntactic structure of the input involves a paradox. The Head Direction parameter is supposed to determine the order in which the head and complement should appear in the language the child is acquiring. But, for a child to set this parameter, she must first find out which units are the heads and the complements in the sentence she hears. If her linguistic skills are sophisticated enough to know which are heads and complements, she will also know which order they came in. If she already knows which order the head and the complements come in a sentence, there is no need to set the parameter."

(Mazuka, 1995: 24-25)

This argument is misguided. First, the idea is not that parameters determine, but rather *are determined*. As a child discovers the specific linguistic constructions of his/her language, parameters are set accordingly. Mazuka does not appreciate the gradual aspect of the whole process. The child will initially have knowledge that either order is possibly grammatical; with *enough* positive evidence for one of the possibilities, over time the parameter will be set. The gradual attainment of this knowledge *is* the setting of the parameter, but Mazuka confuses the two as wholly different.

The further objection concerning the difficulty of deciphering what is the head and what is the complement is an argument concerning language acquisition generally, not the proposition of nativism. If the child learned solely from input, the problem of deciphering said input would remain (in fact, it would only be more difficult).

Briefly, it seems that a child is somehow able to utilize a collaboration of syntactic, contextual, and functional information, perhaps an interpretation of what a speaker is trying to communicate, and in the end is able to distinguish the language's varying parts of speech and syntactic structure. This basically sums up Tomasello's proposal for a data-driven alternative, but by no means does its truth preclude nativism. His theory is a form of data-driven learning he calls "Functional and Cognitive Linguistics" (FCL). FCL involves a conception of language acquisition by means of an understanding of linguistic functionality and an ability to abstract linguistic input, resulting in grammatical knowledge (i.e., adult grammar). Regrettably, FCL does not account for the problems and inconsistencies of language acquisition that make nativism a logical necessity. Nor does it account for the types of novel structural hypotheses that children entertain.

Function fails to fully explain language acquisition. First, a functional account would have to abide by principles of simplicity and probability, but we know that children don't follow these types of systematic data-analysis, nor are the structures of adult grammar simple. If language were guided by function, children wouldn't hypothesize novel linguistic structures, especially not in areas in which alternative structures already serve a given purpose (they do this, as we saw in Tomasello's study concerning SVO and VSO constructions). Even more questionable is how a functional theory could explain the constancy of adult grammar in the first place. Each individual may have certain linguistic functions that are highly unique. We have a great deal of linguistic knowledge that is never relevant to our lives or daily communication; if

language were developed from function, it would be highly unlikely that we would all have such extensive grammatical knowledge.

Tomasello ultimately ignores nativist objections to data-driven learning. In his own words:

> "This general approach is **usage-based** in the sense that all linguistic knowledge -– however abstract it may ultimately become – derives in the first instance from the comprehension and production of specific utterances on specific occasions of use."
>
> (Tomasello, 2000: 237-238)

This theory has the same pitfalls of every other theory of data-driven learning thus far. If language is only created from specific use and then developed via subsequent abstraction, how and why do children hypothesize new structures which can't be logically abstracted from the input? Even more so, why do they hypothesize *those* hypotheses and not others? On its own, FCL cannot discount the nativist argument. To do this Tomasello would have to address specific inconsistencies of language acquisition, and show that they are explainable by FCL. This he simply fails to do. However, such a task *has* been attempted (although without much success), as we will see in the next section.

3.2 Pullum & Scholz and the Positive Argument

Athough Pullum & Scholz (2002) avoids the mistake of ignoring the nativist argument completely, it is nonetheless shortsighted in its undertakings. Pullum and

Scholz are wise in the modesty of their claims: "The argument of this article…does not challenge linguistic nativism directly, or claim that it is false." (P&S, 2000: 10) What then *does* the argument accomplish? Pullum & Scholz (2002)'s answer:

> "Linguistic nativism is the view…that human infants have at least some linguistically specific innate knowledge. We argue that linguists have not achieved what they are widely thought to have achieved. The empirical work that would be needed to support their claims has not been carried out… Certainly, humans are endowed with some sort of predisposition toward language learning. The substantive issue is whether a full description of that predisposition incorporates anything that entails specific contingent facts about natural languages… We are concerned not with whether empiricist claims about language learning are true, but we are concerned with whether linguists have established their falsity."
>
> (Pullum & Scholz, 2002; 10)

Pullum and Scholz have elected to argue against nativism directly, to decipher whether or not nativist claims have proven the falsity of solely data-driven learning.

Unfortunately, their project is bound to failure. Once again, the complexity of the nativist argument has not been fully appreciated: the nativist argument sufficiently falsifies any possiblity of data-driven learning *unless* (3), or both (4a) and (4b), are proven. Restated, the arguments to be proven are:

> (3) Children systematically uphold *only* whatever structures are definitively designated by their stimulus to be grammatical

Or

(4a) There is enough environmental stimulus for knowledge of all grammatical

structures

and

(4b) There is enough environmental stimulus to establish the ungrammaticality of

all other structures

The implications here are noteworthy. Pullum and Scholz have elected to discover

whether or not data-driven claims have been falsified, and we know that they certainly

are falsified, barring the proof of one of the two aforementioned positions. In spite of

this, Pullum and Scholz relegate themselves only to the issue of (4a). This, they say, is

where one finds "the strongest and most productive [argument] for nativism." (P&S,

2002: 14) What of (4b)? Pullum and Scholz claim that "Addressing this sort of case

would take us away from the task we have set for ourselves in this article." (P&S, 2002:

16) Yet (4b) involves an equally strong and productive argument, as I have shown.

Hence, to accomplish the task they have set out for themselves, viz. to show that nativism

has not proven the falsity of data-driven learning, Pullum and Scholz need to address the

negative argument as well. Alas, they are limited only to further deciphering the scope of

UG (if anything), and this is precisely the task that nativists have offered the linguistic

field.

Before their primary investigation, Pullum and Scholz present a supposed

theoretical inconsistency of nativism, originally put forward by Sampson (1989):

"Sampson raises this question: How does Angela know that F is a fact about L?
If the answer to this involves giving evidence from expressions of L, then Angela
has conceded that such evidence is available, which means that S could in
principle have learned F from that evidence. That contradicts what was to be
shown. If, on the other hand, Angela knows L natively, and claims to know F in
virtue of having come to know it during her own first-language acquisition period
with the aid of innate linguistically-specific information, then Angela has
presupposed nativism in an argument for nativism. That is viciously circular."

(Pullum & Scholz, 2002: 15)

This argument claims that if the linguist has knowledge because of any given amount of
evidence, the child, *in principle*, can learn from the same evidence. The problem is that a
linguist has access to much more evidence than any child ever could (even 'in principle').
First, the linguist can elicit knowledge and evidence that would otherwise not have been
presented (and is *not* presented in the normal environment). Second, the linguist is able
to gather examples from various and far-reaching settings, an impossible task for
children. What must once again be emphasized is that for the evidence to be relevant, it
must be universal. In other words, although a single linguist may be able to find an
instance of relevant evidence, this evidence would not be present in many children's
input. Finally, the linguist is an adult. This means that he/she will most likely have the
pertinent knowledge regardless of the examination of *any* evidence. This is not an
assumption of nativism, but of grammar generally. After all, the evidence he/she would

gather would be from ordinary speakers of English, who have no more authority than the linguist, also a speaker of English.

Another of Sampson's implications is that if one claims to know something natively, one is guilty of faulty logic. This is simply not the case. If Angela has "come to know [X] during her own first-language acquisition period with the help of innate linguistically-specific information" then it only follows that she does indeed know X natively. This is something like "if A, then A," one of the most fundamental truths of elementary logic. Of course Angela would still have to have some way of *knowing* that her first-language acquisition was as such, which I believe is the genuine argument. One way Angela could prove that her knowledge was innate would be to show that it could not have possibly been attained any other way. Such is the form of the nativist argument, to which Pullum and Scholz don't reply. The logic is something along the lines of: "I know that X is the case. Reasonably, X could *only* have been caused by either A or B. It is found that B could not have caused X. Hence, A must have caused X." More simply, the idea is that if either A or B must be true, and B (data-driven learning) is false, then A (nativism) must be true. With that said, it seems pretty clear that the argument for nativism is consistent indeed.

Ultimately, Pullum and Scholz dont confer enough consideration to these thoughts for them to be considered significant. Their attention is instead directed at four examples of the positive nativist argument, in which they attempt to show that there is enough evidence available to formulate the pertinent knowledge. For reasons of brevity, I will only analyze their last example; once it is obvious that the case for nativism is legitimate, the former arguments would only serve to amend the scope of positive

evidence. Suffice it to say that I have serious misgivings about the other examples as well. Our main task here, however, is simply to preserve the authenticity of the positive argument.

The most convincing case of the positive argument for nativism has already been mentioned, that of auxiliary movement in question formation, and Pullum and Scholz save the best for last. Not surprisingly, there are two elements of the nativist argument concerning auxiliary movement:

1) Positive: The rule that governs our question formation is movement of the main auxiliary. However, there is no evidence for this in the input. Regardless, children all attain the correct structural/procedural knowledge.

2) Negative: There is no evidence in the stimulus that the infinite number of other hypotheses which are possible are ungrammatical, but children all attain this knowledge, entertaining only the correct structural/procedural knowledge.

As already mentioned, Pullum and Scholz do not contend with the negative argument of nativism. In fact, they point out some very telling details concerning auxiliary movement which wonderfully support its necessity:

"[Structurally complex examples of auxiliary movement] actually only support the claim that sentences formed under the main-clause-auxiliary hypothesis are grammatical. They do not show that sentences formed under the first-auxiliary hypothesis are *un*grammatical. [Such a sentence is] compatible with the idea that

the correct principle allows *any* auxiliary to be fronted. We ignore this

interesting point (which is due to Freidin 1991), because our concern here is to

assess whether the inaccessiblity claim is true."

<div align="right">(Pullum & Scholz, 2002: 38)</div>

There is no negative evidence concerning the ungrammaticality of *any* hypotheses of

auxiliary movement. Even so, we all seem to converge on the same single interpretation.

Crain and Nakayama (1987) found not even a single child that showed signs of a

structure-independent hypothesis.

Pullum and Scholz fail to realize that these facts also aid the positive argument, in

that they establish the absence of any evidence for the *preference* of main-auxiliary

movement. The positive argument merely claims that there is no evidence which makes

the adult structure *more* attractive than any others. The correct interpretation will be

consistent with *any* given data – that is what makes it the correct interpretation in the first

place. The more substantive question is whether or not the given data provides a child

any good reasons to entertain such a hypothesis over any others. In the case of auxiliary

questions, the answer seems to be no. If children have no good reason to entertain main-

auxiliary movement, given no other guiding principles they would vary greatly in

whichever interpretations they decided to hypothesize. We know, however, that children

eventually uphold the same interpretation.

An analysis of the relevant data is in order. The great majority of auxiliary

questions are quite basic and involve only one auxiliary (e.g. - 'Is he cold?'). With such

examples, a great number of hypotheses are possible, most obviously (and simply) 'front

the first auxiliary,' or 'front the last auxiliary.' If only these types of sentences are

prevalent in children's input, there is a lack of any substantial positive evidence in support of the main-auxiliary movement hypothesis.

In response to this fact, Pullum and Scholz call attention to some empirical examples from the Wall Street Journal which involve multiple auxiliaries (e.g. – 'Is a young professional who lives in a bachelor condo __ as much a part of the middle class as a family in the suburbs?'). Although these questions could not be formulated with first-auxiliary movement, an infinite number of other hypotheses remain equally plausible. In fact, *all* of the complex examples that Pullum and Scholz cite are compatible with the hypothesis "front the last auxiliary," which is ostensibly simpler than the correct adult interpretation. This hypothesis eventually fails by allowing sentences like *"Is the boy who is studious is reading a book that __ very large?" as opposed to "Is the boy who is studious __ reading a book that is very large?" However, such sentences could not possibly be universal in children's input (nor is it clear how their presence would help matters).

The types of sentences that Pullum and Scholz mention are also dubious in regards to their universality. Once again, this criterion (of universality) must be stressed. Since it is nearly every child that eventually attains the same grammatical interpretation, the experience of nearly every child must be accounted for in a data-driven explanation. Hence, Pullum and Scholz commit a serious miscalculation in assuming that "a construction inaccessible to infants during the language acquisition process must be rare enough that it will be almost entirely absent from corpora of text quite generally," although they agree that "this assumption could be wrong." (P&S, 2002: 21)

Upon further investigation, it is found that the assumption *is* wrong. The language that children hear is by no means necessarily analogous to texts by and large. Language use varies quite significantly in different contexts. Pullum and Scholz indicate that certain linguistic properties are constant *among texts*, but how relevant and extensive these properties are is unknown. Additionally, constancy among texts does not necessarily extend to children's environmental stimulus nor to spoken language generally. In fact, a nationwide survey of parents with children as old as 3-years by The Commonwealth Fund found that "sixteen percent of parents do not read to their children at all." (Young, Davis, Schoen, 1996)

If anything, we can be certain that the actual occurrences Pullum and Scholz mention have never been directly present in any child's input – 2-year olds simply don't spend Sunday mornings over coffee and the Wall Street Journal. For such sentences to reach children's input, adults would have to consistently make such structurally complex productions in the children's presence. However, it seems contrary to reason that such complex questions would be routinely produced around children. Would a child normally hear: "Is a young professional who lives in a bachelor condo as much a part of middle class as a family in the suburbs?" Probably not.

By contrast, Legate and Yang (2002) analyzes two corpora of child input, which are clearly closer representations of the environmental stimulus than any unrelated texts. The findings are quite informative: in the first corpora, questions with multiple auxiliaries occurred only .068% of the time, in the second, .045%, both of which are "low enough to be considered negligible, that is, not reliably available for every human child." (Legate & Yang, 2002: 158) Additionally, *none* of these sentences discounted the 'front the last

auxiliary' hypothesis, leaving the dilemma unresolved. Without any reason to entertain the correct hypothesis, children must be guided in other ways. It is obvious that there is insufficient positive evidence with which to decipher the prevailing grammatical structure.

More generally, Pullum and Scholz object to nativism by proposing the very disconcerting notion that as a result of any environmental deficiency, adults may in fact possess different grammars. "That is, there could be speakers, interacting with us undetected, who have actually learned the wrong generalization." (P&S, 2002: 35) Yet this would seem to contradict our fundamental assumptions about grammar. Since it is composed of the generalization of a majority of speakers in a given language, there is no way for most of us to have a 'wrong' generalization. Until Pullum and Scholz can present us with bona fide examples of large groups of adults who are opposed to the main-auxiliary hypothesis, the argument should not be given credence.

We can see merely from the literature that the argument is false; if our grammatical interpretations did vary concerning such underepresented interpretations, they would vary dramatically and on a large scale. Our personal structures concerning auxiliary questions, for example, would be selected from among an infinite number of random possiblities. Hence, it would be easy to find disagreement among many speakers (what would be the chances that we all picked the *same* hypothesis from such a palette?). If *that* were the case, an objection over which form of auxiliary movement is grammatical would have been made in the literature, if not encountered by linguists many times over in their investigations.

Instead, Pullum and Scholz go on to call many structures 'ungrammatical' or 'wrong', even while considering the possible dilemma. What could possibly give them the authority to call anything ungrammatical if opinion on grammar could differ so freely? Clearly, Pullum and Scholz are (reasonably) assuming the ubiquitousness of an unstandardized grammar. Even more indicative of its certainty are the examples they find in an attempt to show the presence of positive evidence. While their findings are somewhat irrelevant in the realm of child input, they do point to the constancy of such grammatical knowledge among adults. The WSJ does *not* contain sentences like *"Is the boy who is studious is reading a book that ___ very large?" This shows that at the very least, *most* of us share the same views on grammar, and that is just the proposition we began with.

Even if data-driven learning were not directly disproven by nativism, it would still have to be shown to be a plausible scientific alternative. Pullum and Scholz don't take it upon themselves to present any such theory. Yet isn't the exploration of whether or not nativism disproves data-driven learning irrelevant if data-driven learning can't act as a valid scientific theory? Vaguely, Pullum and Scholz only imply that children can generalize from their environmental stimulus. First, this disregards all the pressing intricacies of the argument for nativism, which Pullum and Scholz simply mull over. Additionally, such a data-driven explanation cannot possibly account for children producing wholly novel structures, nor can it explain language creation. Pullum and Scholz ignore too much of the nativist argument to make any significant progress in questioning its authority. They even fail to substantially influence the positive side of the

debate. Much of their arguments lack any overarching persuasiveness, and what is still clear is that there is some grammatical knowledge for which the data cannot account.

3.3 The Negative Argument

The negative argument has been largely ignored by the combatants of nativist theory. All the undertakings which *have* investigated the argument have been attempts to discover the presence of error corrections in the stimulus. In other words, researchers have sought to discover if parents provide their children with evidence that the errors they make are ungrammatical. This could take the form of direct reprimand (No!, Wrong!), recasting of the ungrammatical sentence in its grammatical form ('The man *saw* the bird!' in response to *'the man seed the bird'), or other subtler forms. Of course, none of these researchers have been able to reach any definitive conclusions, much less provide a basis for the rejection of the negative argument.

The task of independently disproving the negative argument for nativism (i.e., proving (4b)) cannot possibly be successful. The amount of negative evidence necessary to prevent children from hypothesizing a random assortment of possible structures is incredibly daunting. Such a set of negative evidence would have to come in one of two forms:

1) A general law defining *everything* as ungrammatical except for a clearly

defined set of grammatical structures/rules, which I will call Set X.

2) Evidence identifying every single possible ungrammatical structure or rule,

and stating that it is just that, ungrammatical. The amount of possible

ungrammatical structures is, of course, infinite. Everything *not* identified would

be grammatical, i.e. would comprise Set X.

In other words, the evidence would have to exclusively define all that is grammatical or

all that is ungrammatical. It is clear that children are not given either array in experience.

No exhaustive set is ever defined as what is solely grammatical, no law made clear, and

they could never possibly be given an exhaustive account of what is ungrammatical (due

to there being an infinite number of ungrammatical structures).

The data-driven theorist is not lost yet. After all, there is no way to deny that

humans are *somehow* predisposed to language acquisition. The data-driven theorist only

objects to innate knowledge that is *linguistically specific*. Hence, the possibility remains

for children to have innate (non-linguistic) knowledge of what defines Set X, and that

only Set X is grammatical. The only entity that could define Set X would be some or all

of the environmental stimulus, since nothing else can provide linguistically specific

information. The relevant input could be the environmental stimulus in its entirety, or a

subset which was isolated by non-linguistic means (e.g. -- only the statements parents

produce, only the statements which appear more than once, etc.). As we will see, this is

the type of complicated proposition necessary to begin proving (3). It is odd, however,

that many who are hesitant of linguistically specific knowledge because of its detail have

no qualms with innate awareness of such a specific filter for acquiring that knowledge.

Now we are back to square one. This supposition (Set X as defined by the

environmental stimulus) cannot explain how children could logically filter out errors

from their input, while maintaining other grammatical structures which are, for example, less frequent. Nor does it account for children producing structures and rules absent from the input.

Hence, even if researchers stumble upon evidence of error correction in the stimulus, the negative argument will remain compelling. This will be true, in fact, even if *every* error were corrected upon its perpetration. The rate of error correction, however, has been found to be substantially low. Studies have repeatedly found that mothers recast their children's ungrammatical statements no more than 34% of the time (Hirsh-Pasek, et al., 1984; Demetras, et al., 1986; Farrar, 1990, 1992). That is not to say that every instance of recast is effective. In fact, Farrar (1992) found that children amend their ungrammatical interpretations in response to recast only 12% of the time. These are not promising numbers.

Another important detail in regards to error correction is that all the negative evidence is in response *only* to hypotheses that children actually produce. This means that error correction cannot possibly account for why children don't entertain any number of random hypotheses in the first place. By a strictly data-driven account, even perfect error correction would allow every hypothesis not yet produced to be grammatical. Also, error correction only deals with specific statements children make rather than with general structures. The *veiled error* would be left to interpretation, as well as how the alternative grammatical statement should be formed.

Most importantly, children would have to be able to identify any given error correction as a mutually exclusive alternative to their original production. How the child would know when the input involved error correction instead of additional grammatical

alternatives is beyond reason. For example, a child may produce a given statement using the active construction. It is quite possible that in discussing the point, or perhaps repeating it, the parent may use passive construction. The child eventually accepts *both* constructions as grammatical, meaning that he/she sees the parent's passive production as a grammatical alternative rather than an instance of error correction. Yet there doesn't seem to be any way for the child to reliably decipher the two possibilities without prior intuition of what is and is not possibly grammatical. In this sense, error correction could *only* work under the pretense of UG.

4. The Possible Alternative

The lingering possibility for the data-driven theory is to show that (3) is feasible, so this certainly deserves some attention. The general idea would be that children are able to examine the statements they hear and conclude which structures/rules were used to produce them. With knowledge of these rules, children could then produce statements of their own. Once more, the proposal would be:

(3) Children systematically uphold *only* whatever structures are definitively

designated by their environmental stimulus to be grammatical

This proposition contains quite a bit of complexity, hence it is essential to depict exactly what would be involved if it were true.

Assuming (3) were true, children would be *systematic* in two ways. First, children would abide by some sort of principle which would allow them -- given a particular set of statements -- to single out the proper structure from all logical possibilities. I will call this the *possibility choosing principle* (PCP). We saw the reason for PCP in our example of auxiliary questions: children need to be able to specify main-auxiliary movement as the correct structure from among any of the other possibilities. Of course, PCP would not be a principle based on the linguistically specific properties of sentences, since (3) is an attempt to show how children learn *without* linguistically specific knowledge. Instead, PCP must be a principle based on non-linguistic factors, such as simplicity or probability (sound familiar?). PCP would have to be flexible; with more input, it would reevaluate the supposed hypothesis. Only with enough evidence would PCP single out the correct (grammatical) hypothesis. That is to say, PCP would be ineffective with *no* evidence, but requires much less than if the child had begun tabula rasa.

Second, children would have to discount any ungrammatical statements which were part of the input. Such statements directly contradict the correct structure, so children could not decipher the correct structure (even with PCP) if they accepted such statements as being grammatical. This filtering would thus have to exclude *all* ungrammatical statements from consideration, and could very well exclude some grammatical statements as well. I will call this the *stimulus exclusion filter* (SEF). Just like PCP, SEF would not be based on the linguistic specifics of these statements, since that would imply that children already have knowledge of what is ungrammatical, which is exactly what is being discounted.

Instead, SEF would have to scrap ungrammatical statements via non-linguistic means. As a result of SEF, a portion of the input is thus excluded from data-driven learning, and from PCP's consideration.

After the input is filtered by SEF, the stimulus that remains would have to allow for the correct grammatical structures to be selected by PCP. In this sense, the input must be able to *definitively designate* grammatical structures via PCP and SEF. The SEF must be constricting, as to exclude any ungrammatical statements, but must not exclude any input which is necessary for PCP to determine a grammatical structure or rule.

Finally, beyond being systematic, children would deem *only* the structures selected by PCP to be grammatical. In this sense, any structure would be considered ungrammatical *unless* definitively designated by the input. By contrast, every grammatical structure would be inferrable by PCP from the filtered input alone. Even more importantly, *only* those structures would be inferred; to be a plausible way that children acquire language, PCP and SEF cannot lead to the selection of an ungrammatical structure.

We can see now why this position would discount both arguments for nativism: First, it limits the amount and extent of positive evidence needed for any child to arrive at the correct hypothesis, since they have PCP to guide them. Second, it does away with the need for negative evidence, since children would designate any structures not chosen by PCP as ungrammatical by default.

Moreover, unless (4a) and (4b) are proven, the aforementioned theory is the *only* way to account for language acquisition; if there is not enough positive evidence for

children to logically decipher the correct grammatical hypotheses, only with PCP and SEF can they do so. We know (4b) to be impossible, so the task for any data-driven theorist is now to decipher what PCP and SEF could be. No one has yet proposed any possibilities, but I will briefly discuss the issue.

As I said, PCP cannot be simplicity or probability. What else is possible? Let's take the example of main-auxiliary movement once again: Given sentences like 'Is he __ cold?', children would have to be able to form the rule of moving the main auxiliary. It is clear that no principle (linguistic or not) could isolate this hypothesis. Even if children were presented with sentences like 'Is the boy who is cold __ in the house?', I can think of no non-linguistic principle with which they could deduce main-auxiliary movement (and *only* main-auxiliary movement), nor do I believe that one exists. Given the complexity of our structures and rules, it seems that the only way to decipher them would be from extensively complex examples and the use of logic. This is, in fact, just the way that linguists decipher what our rules are. However, linguists can incite these examples, while children can only work with what they are given. PCP, then, is nothing but a pipe dream.

SEF runs into similar problems. The sentences which SEF allows are taken in by PCP, which leads to a given structure that coincides with all of the selected input, or so the theory goes. The sentences SEF does not allow are tossed aside as ungrammatical. What non-linguistic characteristics uniquely distinguish ungrammatical statements? It is quite difficult to say unless one already knows that they are ungrammatical. We know that their frequency is an unlikely candidate. In fact, an additional problem to what was already mentioned is that children would not yet know exactly which structures were

used to make any given sentence (for that, PCP would have to have analyzed the *already-filtered* input). Hence, they would not know which structures appeared more often than others.

What, besides frequency, could constitute SEF? One possibility would be for something non-linguistic to accompany every ungrammatical statement made. This is highly unlikely, although it is plausible that many ungrammatical statements are accompanied by negative reactions such as 'I mean, ___' or 'whoops!' along with some type of body language or facial expression. This is the type of functional learning that Tomasello espouses, in that children would realize the adults made a mistake by witnessing, for example, their frustration. In that case, SEF would filter out any statements accompanied by such frustration or body language. Perhaps this could limit the quantity of the remaining ungrammatical statements to a small number, but that still leaves a big problem for PCP. After all, as seen in Tomasello (2000), children will suppose a structure to be grammatical that they have heard only *once*. It seems that a unique reaction would have to accompany *every* ungrammatical statement for the theory to be plausible. What about false starts, in which a person begins a sentence but does not finish it? The problems begin to…

Although PCP and SEF seem to be impossibilities, suppose them to be plausibly explained. What are the implications of these two entities, and (3) generally? First, children would never consider any structures to be grammatical that were not designated as such by PCP. Second, children would never be able to create language since their learning would be based solely on input filtered by SEF. As we'll see in the next section, neither of these inferences is empirically true.

5. Evidence for Nativism

There are two facts pertaining to language acquisition that no data-driven theory can account for. First, that children believe certain ungrammatical structures to be grammatical that are *never* evidenced in their input, and second, that children can create a language and/or add to its complexity in significant ways. Once it is noted that instances of both cases have been substantiated, it is difficult -- if not impossible -- to maintain a data-driven account.

One of the most telling examples of children hypothesizing a completely novel structure is the case of English-speaking infants upholding the 'medial-wh' phenomenon (Thornton, 1990; McDaniel et al, 1995; Crain & Thornton, 1998; Crain & Pietroski, 2002). In this case, children adhere to a surprisingly detailed structural hypothesis *never* evidenced in adult English. Thornton (1990) involved the elicitation of long-distance questions from 3- to 4-year old English-speaking children, and found that about one third of the children inserted an extra wh-word into their productions:

Who did I see *who* went into the store?

What do you think *what* I have in my hand?

This structure is nowhere to be found in adult English, nor is there an analogue from which one could deduce its grammaticality; this structure is wholly new. However, production does not necessarily equate with comprehension. In fact, some studies have

shown that children produce statements that they will judge to be ungrammatical (Hiramatsu & Lillo-Martin, 1998, etc). Fortunately, McDaniel et al (1995) tested children's judgements concerning the 'medial-wh' construction and found that over a third of children also *judged* the statements to be grammatical; children not only produce such statements, they also consider their underlying structure to be grammatically correct. This lies in fundamental contradiction with data-driven learning, which asserts that children will develop their knowledge of grammaticality based solely on environmental input; here, there is none.

Crain & Pietroski (2002) extends the investigation even further by comparing these productions to a similar construction in German:

"This 'error' by English speaking children is presumably not a response to the children's environment, since medial-wh constructions are not part of the primary linguistic data for children in English-speaking environments. However, [such structures] are attested in dialects of German… Further investigation shows that the similarity of Child-English to [German] runs deep. For both adult Germans and American children, lexical (full) wh-phrases cannot be repeated in the medial position… Finally, children never used a medial-wh when extracting from infinitival clauses… Nor is this permissible in languages that allow the medial-wh."

(Crain & Pietroski, 2002)

Such distinct and specific similarities in usage are striking, especially considering that the children are not given the basic structure in the first place. For example, while allowing

statements like "Who do you think who will win?," neither these children nor speakers of German will allow '*Who do you want who to win?" because of the presence of the infinitive.

Children's medial-wh constructions are noteworthy for a couple of reasons. First, it is remarkable that a full third of children are producing such a structure. No data-driven account could explain why so many children are attesting to a hypothesis which is never proposed -- let alone definitively designated as grammatical -- by the input. This directly disproves (3) because children are not considering *only* what is designated by their input to be grammatical. Second, the striking similarity of the children's medial-wh construction to a similar construction in German seems beyond chance. Not only are children producing the same structure, they are using it in the same specific contexts in which the German construction also applies. This is significant evidence for UG's guidance in language acquisition.

The need for children to propose wholly novel hypotheses becomes even more interesting when considering language creation; in such a case, children's novel constructions may be *necessary*. After all, how would a data-driven account explain language creation? If children learn only from their data, language would have to first be created in order to be learned. In other words, adults would have to consciously select the rules and structures of the language as it developed. For example, the rule of main-auxiliary movement would have been randomly selected by an adult faced with many possibilities of question formation.

It seems as though, without some sort of linguistic guidance, different adults would choose different hypotheses, causing much confusion and difficulty. One theory

may be that one person or group of people were put to the specific task of creating a language and deciding its rules. This seems like a nearly impossible task without language in the first place. Given the complexity of every language (which the entire field of linguistics has not nearly exhausted), it is more likely that rules and structures must be developed over time by subsequent generations. It is quite questionable whether such adults could effectively institute their conscious linguistic decisions upon other adults, which would be necessary for the constructions to be present in nearly every child's experience. Adults have quite a difficult time learning any language, so it seems that only during childhood are humans quite apt for learning their native tongue, but also for creating it.

Unfortunately we do not have a detailed historical account of the creation of English grammar, so we can only guess how it developed. However, history *has* granted us a recent instance of language creation: Nicaraguan Sign Language (NSL) was created in the 1980s and has slowly developed over time into a syntactically complex language. The pertinent question is how such development occurred, and by whom. The language appeared among a community of young students who, without any instruction, developed a rudimentary sign language with marked rules and structures. Over the years, the language has developed syntactically and become closer and closer to other established sign languages.

Senghas & Coppola (2001) investigated the development of NSL, and from where such development was emerging. What was specifically observed was the amount and use of spatial modulation (a part of every established sign language), which involves signing in non-neutral locations to perform various functions. NSL speakers were

divided into two groups based on age and time of acquisition, and the findings were quite significant. First, spatial modulations were found to be more frequent in the younger group, which implied that the younger group was not merely reproducing the language of their input. Second, spatial functions were applied to a more specific function in the younger group – that of shared reference -- which allowed for long-distance grammatical relationships of words. Such is the use of spatial modulations in many established sign languages. Finally, the younger group was also more fluent in the language, meaning that limitation of spatial modulations was not a result of limited input or a partial learning of the language. Senghas and Coppola conclude:

"Drawn together, these results show that the youngest members of the second cohort, as children, surpassed their input, taking a partially developed language and systematizing it in a specific way. This finding is especially striking considering that the second cohort had had fewer years of experience with the language."

(Senghas & Coppola, 2001; 327)

It is obvious from the findings that adults are not the cause of the syntactic development of NSL. This gives us a good idea of how every other language developed; namely, in the minds of each generation of children.

Some would argue that the evidence cited is much like the ungrammatical medial-wh questions of English-speaking children, in that the NSL-speaking children are temporarily upholding ungrammatical structures. This would imply that they will eventually espouse the adult interpretation of more general spatial modulations.

However, Senghas and Coppola point out that at the time of testing, none of the participants were still children:

> "These data show a permanent, creative impact of children's learning capacities on their language… the participants were no longer young children at the time of testing… these regularizations represent changes that have taken hold and persist in the language of today's adolescents. Nicaraguan Sign Language had evidently not yet stabilized in the mid-1980s, when the second-cohort learners arrived. Otherwise, this cohort would have shown the normal pattern of 'unlearning' their 'errors' and eventually adopting adultlike productions."
>
> (Senghas & Coppola, 2001; 327)

Such creative and spontaneous language creation poses substantial problems for a data-driven theory of language acquisition. The proposition that adults consciously institute language creation and development towards complexity seems implausible. Pollom and Scholz referred to the Wall Street Journal in an attempt to demonstrate relevant positive evidence for today's English-speaking children. However, it is unclear how the writers of the WSJ who produced such evidence initially learned of its grammaticality. Perhaps the previous writers of the WSJ produced equally relevant evidence. These authors would have to have learned from previous evidence, however, and the line of reasoning takes us on a neverending search for more positive evidence. It is obvious that *at the very least,* during the creation and development of a given

structure/rule, no positive evidence existed. This being the case, a data-driven

account of acquisition is simply insufficient.

A functional view like Tomasello's FCL cannot explain such language

creation either. Although many hold that language developed only from the

necessity for communication, this does not explain how or why children develop

the *same* hypotheses (for example, why a *third* of English-speaking children use

medial-wh questions and not other forms). In the end, grammatical collaboration

is still left unexplained even by a functional approach. Only with the assumption

of UG can we explain children's collaborative creation of language.

6. Conclusion

In the search for a scientific theory to explain language acquisition, we have

found no alternative to nativism. The argument is initially convincing: how can children

learn so much from so little? Fundamentally, however, nativism is based upon the

disproof of its contrary: it is not convincing in itself, but only in light of the impossibility

of a solely data-driven account of language acquisition. It is clear that children's

grammatical knowledge can not be derived solely from the input they receive; they are

somehow guided in their knowledge of what structures are possibly grammatical.

Thus far, the arguments against nativism have fallen significantly short of their

goal; much of the literature has focused on discussion which is probably self-defeating.

After evaluating the argument and the debate surrounding it, it is clear that nativism is the

only plausible assumption, for a few reasons. First, none have been able to give a viable

scientific alternative to nativism. What is needed is an unambiguous account of how children can learn from the input, and an examination of all the relevant inconsistencies within such an account. Data-driven theorists cannot simply assert that children can learn from their input or deny the fact that our grammatical interpretations are somehow all in synchrony. The more important point is that *until* such an undertaking is accomplished, nativism must be accepted.

Second, most have concerned themselves with specific contentions which have no impact on the nativist argument generally. In other words, they can only serve to discover the specific nature of the nativist argument rather than to discount it. It is the position of most nativists that linguists should concern themselves with the discovery of the specifics of UG. Nativism is – at least for now – the most explanatory theory of language acquisition. That being the case, the science of linguistics – *at least for now* – should take it upon itself to explore its intricacies. By the same token, linguists should presume the truth of nativism in their investigations.

Third, some of the inconsistencies which nativists point out have placed the possibility of grounding a data-driven account of language acquisition far out of reach of any of today's theorists. In other words, it is useless to uphold a theory of solely data-driven learning until our knowledge about the world changes. For example, children's completely novel productions cannot, as far as we know, be explained by a data-driven theory. In fact, even if data-driven theorists did show the possibility of data-driven learning from today's environmental stimulus, they would still be at a loss for the explanation of language creation.

Were data-driven learning possible, we would have no way of knowing if children learned by their environmental stimulus alone or were guided by some principle like UG. However, children's environments are wrought with ambiguity and error. Hence, we are in a unique position to assert that children do *not* learn from their input alone, but are also guided by linguistically specific knowledge. For all the controversy and debate surrounding nativism, it is quite clear that it is a legitimate scientific position.

REFERENCES:

Chomsky, N. (1981). Principles and parameters in syntactic theory. In Hornstein, N. & Lightfoot, D., *Explanation in Linguistics,* London: Longman.

Chomsky, N. (1975). *Reflections on Language*, New York: Pantheon Books.

Clark, E.V., Barron, B.J.S. (1988). A Thrower-Button or a Button Thrower? Children's Judgments of Grammatical and Ungrammatical Compound Nouns. *Linguistics, 26*, 3-19.

Crain, S. & Pietroski, P. (2002). Why language acquisition is a snap. *The Linguistic Review, 19*, 163-183.

Crain, S., & Thornton, R. (1998). *Investigations in Universal Grammar: A Guide to Experiments on the Acquisition of Syntax and Semantics*. Cambridge: The MIT Press.

Crain, S. (1991). Language Acquisition in the Absence of Experience. *Brain and Behavioral Sciences, 14*, 597-650.

Crain, S., & Nakayama, M. (1987). Structure Dependence in Grammar Formation. *Language, 63*, 522-543.

Demetras, M.J., Post, K.N., & Snow, C.E. (1986). Feedback to first language learners: the role of repetitions and clarification questions. *Journal of Child Language, 13*, 275-292.

Farrar, M.J. (1992). Negative evidence and grammatical morpheme acquisition. *Developmental Psychology, 28*, 91-99

Farrar, M.J. (1990). Discourse and the acquisition of grammatical morphemes. *Journal of Child Language, 17*, 607-624.

Hiramatsu, K., & Lillo-Martin, D., (1998). Children Who Judge Ungrammaticl What They Produce. *Proceedings of the Annual Boston University Conference on Language Development, 22*, 337-347.

Hirsh-Pasek, K. & Golinkoff, R.M. (1996). The Intermodal Preferential Looking

Paradigm: A Window onto Emerging Language Comprehension. In McDaniel, D., McKee, C., & Cairns, H.S., *Methods for Assessing Children's Syntax*, Cambridge: The MIT Press.

Hirsh-Pasek, K., Treiman, R., & Schneiderman, M. (1984). Brown and Hanlon revisited: mothers' sensitivity to ungrammatical forms. *Journal of Child Language, 11*, 81-88.

Hornstein, N., & Lightfoot, D. (1981). *Explanation in Linguistics*. London: Longman.

Jackendoff, R. (2002) *Foundations of Language*. New York: Oxford University Press.

Lasnik, H., & Uriagereka, J. (2002). On the poverty of the challenge. *The Linguistic Review, 19*, 147-150.

Legate, J.A., & Yang, C.D. (2002). Empirical re-assessment of stimulus poverty arguments. *The Linguistic Review, 19*, 151-162.

Lidz, J., Waxman, S., & Freedman, J. (2002). What Infants Know about Syntax but Couldn't Have Learned: Experimental Evidence for Syntactic Structure at 18-months. In press.

Lightfoot, D. (1982). *The Language Lottery: Toward a Biology of Grammars*. Cambridge: The MIT Press.

Marcus, G.F. (1993). Negative evidence in language acquisition. *Cognition, 46*, 53-85.

Mazuka, R. (1995). Can a grammatical parameter be set before the first word? In Morgan, J., & Demuth, K., *From signal to syntax,* Hillsdale, NJ: Erlbaum.

McDaniel, D., Chiu, B., & Maxfield, T.L. (1995). Paramaters for Wh-Movement Types: Evidence from Child English. *Natural Language and Linguistic Theory, 13*, 709-753.

Musolino, J., & Lidz, J. (2002) Why children aren't universally successful with quantification, (submitted).

Musolino, J., Crain, S., & Thornton, R. (2000). Navigating negative quantificational space. *Linguistics, 38*, 1-32.

Newport, E.L., Gleitman, H., & Gleitman, L. (1977). Mother, I'd Rather Do it Myself: Some effects and non-effects of maternal speech style. In Snow, C.E., & Ferguson, C.A. (eds), *Talking to Children: Language input and acquisition.* Cambridge: Cambridge University Press.

Pullum, G.K., & Scholz, B.C. (2002). Empirical assessment of stimulus poverty arguments. *The Linguistic Review, 19*, 9-50.

Sampson, G. (1989). Language acquisition: growth or learning? *Philosophical Papers, 18*, 203-240

Saxton, M., Kulcsar, B., Marshall, G., & Rupra, M. (1998). Longer-term effects of corrective input: an experimental approach. *Journal of Child Language, 25*, 701-721.

Saxton, M. (1997). The Contrast Theory of negative input. *Journal of Child Language, 24*, 139-161.

Senghas, A., & Coppola, M. (2001). Children Creating Language: How Nicaraguan Sign Language Acquired a Spatial Grammar. *Psychological Science, 12*, 323-328.

Spelke, E. (1979). Perceiving bimodally specified events in infancy. *Developmental Psychology,* 15, 626-636.

Strapp, C.M. (1999). Mothers', fathers', and siblings' responses to children's language

errors: comparing sources of negative evidence. *Journal of Child Language, 26,* 373-391.

Thornton, R. (1990). Adventures in long-distance moving: the acquisition of complex Wh-questions, unpublished Ph.D. dissertation, university of Connecticut.

Tomasello, M. (2000). Do young children have adult syntactic competence?. *Cognition, 74*, 209-253.

Young, K.T., Davis, K., & Schoen, C. (1996) *The Commonwealth Fund Survey of Parents with Children,* New York: The Commonwealth Fund.