Hamilton College
Russell Marcus

Class 27 - Linguistic Platonism
Katz, "The Unfinished Chomskyan Revolution"

I. Nominalism, Conceptualism, and Platonism

Katz's arguments against Chomsky parallel two arguments in the philosophy of mathematics.
The first argument concerns three philosophical positions on the ontology of mathematics.
The second argument, which will take us through a detour into infinite arithmetic, relies on an argument from Langendoen and Postal; I will discuss that argument in section VII, below.

Traditionally, there are three distinct positions in the philosophy of mathematics concerning ontology.
The following are rough characterizations of the three positions.
The mathematical realist, or platonist, claims that mathematical objects exist, independently of us, as abstract objects.
Though we lack sense experience of numbers and perfect geometric figures, we still have knowledge of those real objects.
Plato, Descartes, and Frege are all mathematical realists.
The mathematical conceptualist, or intuitionist, claims that mathematical objects are mental constructs.
Intuitionists are inspired by Kant, and many of them reject the law of the excluded middle and proofs by reductio ad absurdum.
The mathematical formalist, or nominalist, says that mathematics is just a formal game, with no ontology beyond that of mathematical inscriptions.
Mathematical axioms and theorems are rules for playing the game, rules for manipulating inscriptions.
No one really prominent defends formalism, though it is often attacked; Katz ascribes formalism to Hilbert, though he recognizes that Hilbert wasn't technically the formalist he caricatures.

We can discern three analogous positions concerning the ontology of language.
The linguistic platonist claims that languages are abstract objects.
The linguistic conceptualist claims that languages are mental constructs.
The linguistic nominalist claims that languages have no ontology beyond that of token inscriptions or utterances.

We have already seen that Frege held linguistic platonism, that senses of expressions and sentences were abstract objects.
Chomsky is a linguistic conceptualist, and he argues for his conceptualism, in part, on the basis of an argument against the linguistic nominalism he and Katz ascribe to Bloomfield.
You might take a moment to think about the ontology of language for various of the philosophers we have studied.
In particular, notice that Locke would be a conceptualist, holding that the meanings of our terms are ideas in our minds.
Also notice that behaviorist theories of language, like those of Wittgenstein, which claim that meaning is use, are naturally taken as linguistically nominalist theories.

Katz's argument in "The Unfinished Chomskyan Revolution" is that Chomsky's argument against linguistic nominalism also applies to linguistic conceptualism, and that therefore the only plausible position of the three is linguistic platonism.

II. The Type/Token Distinction

The argument that Katz uses against Chomsky relies on the type/token distinction.
We have spoken previously about types and tokens, so only a brief review may be required.
Katz cites Peirce, who may have first made the distinction, and a nice passage from Quine, p 274.
Here are some examples of linguistic tokens: inscriptions on a page, or blackboard, or computer screen; utterances; gestures with conventional meaning.
Here are some examples of linguistic types: words and sentences, in the sense that they can be repeated; propositions.
When one thinks of meaning as use, one naturally thinks of semantic theories as constructed out of tokens.
Bloomfield, according to Katz, presented a linguistic theory which also relied on tokens.

> Bloomfield construed linguistic reality nominalistically, as concrete acoustic phenomena... Bloomfield and his followers formulated a methodology for nominalistically constructing grammars as category structures which taxonomize all expressions of a natural language. The methodology consisted in segmentation and classification procedures which, working up from actual expression tokens in a nominalistically acceptable way, construct a categorical structure describing all expression tokens of the language, including the possible but not actual ones (275-6).

The central problem with constructing a linguistic theory from tokens is that there are not enough of them to account for the novel and indefinite number of sentences that speakers can form.
Bloomfield thus posited possible tokens, in addition to actual tokens.
Chomsky's argument against the nominalist Bloomfieldian linguistics centers on just this issue.

III. Chomsky's Rejection of Nominalism

Instead of studying language from the bottom-up, from actual tokens to categories of possible tokens, Chomsky proposed studying generative grammars.
Generative grammars, modeled on formal systems, like those of mathematics and logic, have compositionality built-in.
Given a lexicon and principles which guide the composition of sentences, Chomsky allows for the construction of the indefinite number of sentences that we desire.
But, generative grammars may not be nominalistically acceptable, because they allow for too many linguistic objects.
So, the only remaining positions from our list of three are conceptualism and realism.

> The nominalist interpretation of grammars as descriptions of acoustic phenomena was thus replaced with a conceptualist interpretation of grammars as theories of the linguistic aspect of human psychology, or, as he...now expresses it, a 'language organ' in the human brain (277).

Recall, Chomsky argued that language is grown in humans rather than learned behaviorally/inductively.
Like the mathematical intuitionist, the ontology which underlies our linguistic knowledge is mental.
Of course, Chomsky takes the mind to be the brain (or near enough).
Still, the conceptualist holds that the generative grammars, and even the concepts underlying the lexicon, are actually built into the brain.

We can see how literally Chomsky intends this claim by noticing how Chomsky's theory may be best confirmed.
A linguistic theory which accords with a native speaker's intuitions about language (including, for example, phenomena like analyticity and redundancy) would be some evidence in favor of Chomsky's theory.
Transformation rules which map natural languages to UG would be even greater evidence.
But, the best evidence of all would be the mapping of the language centers of the brain in such a way that we could see how UG were built in.
That is, Chomsky is presenting a theory which admits of empirical, neurological evidence.

IV. The Boomerang

Katz's central claim, then, is that Chomsky's argument against the Bloomfield nominalist applies to Chomsky's conceptualism itself.
The argument against establishing a semantic theory on the basis of linguistic tokens was that there just are not enough of them.
We would need a denumerable infinity of linguistic objects.
But, there are not enough neural tokens, either.
If we take a language to consist of an indefinite number of expressions, there will not be enough of any kind of token for the theory.

> Infinity in linguistics cannot be squared with concretism in the foundations of linguistics (279).

If, on the other hand, we take a language to be an abstract object, we have plenty of room.

> The source of the problem is taking linguistic reality to be concrete. Taking expressions to be acoustic objects is just one way of taking linguistic reality to be concrete. Another way is taking expressions to be mental/neural objects... The fundamental problem of...the abstractness of language was not solved in the Chomskyan revolution, but swept under the rug (277-8).

Once we accept the consequences of the initial observation that linguistic theories are theories of linguistic types, we have adopted a platonist ontology in language.
All attempts to naturalize this platonism, to conceptualize abstract objects, are doomed to fail.
The only way to do justice to the conclusion is to accept that linguistics is a formal theory, of abstract objects, rather than a physical theory, of concrete objects.

V. Ontology and Generative Grammar

Chomsky was drawn to conceptualism by his desire to make linguistics compatible with natural science.
He agreed with Bloomfield that linguistics is an empirical theory, with empirically testable results.
His claim against Bloomfield, and structural linguistics generally, was that it could not account for indefinitely many novel sentences.
The Chomskyan revolution in linguistics was really centered on the move from taxonomies of utterances to generative grammar.
But, Chomsky placed the generative grammar in the brain.
Linguistics became a branch of empirical, cognitive psychology.

I mentioned at the end of my notes on Chomsky that his is an odd position, marrying apriorism about linguistic facts (like those governing analyticity) with empiricism about languages themselves.
Katz, in contrast, argues that any discipline, and its methods, must correspond to the objects it studies.
Along with Chomsky's generative grammar came the abstract objects of linguistic types.
(Actually, abstract objects were present in pre-Chomskyan linguistics, even if they went unnoticed.)
Since linguistics studies abstract objects, Katz argues, it must be a formal, not an empirical, science.
There are empirical questions we can ask about a formal science.
We can ask about our knowledge of mathematics, and of language.
We can trace our learning processes.
But, when we study the language itself, we are no longer engaged in empirical research.
We have entered into a formal, rather than empirical, science.

Linguistic platonism is the name for the claim that linguistics is the study of abstract types, rather than the concrete utterances (which are studied by the pre-Chomskyan structural linguists) or mental states (which are studied by the Chomskyan conceptualist).

Once we accept linguistic platonism, we have to wonder whether generative grammars are sufficient to yield the requisite abstract objects.
Katz considers an argument from Langendoen and Postal that generative grammars fail to yield enough sentences.
The basic idea of the Langendoen and Postal claim is that there are more than denumerably many sentences of English, but generative grammars can only produce denumerably many sentences.
The argument is based on Cantor's diagonal argument which yields that there are different levels of infinity.
For example, there are more real numbers than there are integers, even though there are the same number, of even integers as there are integers.
To better understand the Langendoen and Postal claim, we will take a detour into infinity.

VI. Cantor's Diagonal Argument and the Different Sizes of Infinity

To get in the mood for infinite numbers, consider the infinite hotel.
All the rooms in the infinite hotel are booked.
But another person comes knocking at the door, looking for a room.
We can add the new guest by shifting everyone from room n to room n+1.
There are enough rooms for everyone.

Next, a busload of forty new guests arrives.
"It's no problem," says the manager.
"We'll move every current guest from room n to room n+40, and slide the new guests into the first forty rooms."
Again, there's room for everyone.

Soon, an even larger busload of new guests arrives: an infinite busload of new guests.
Again, the manager keeps his cool.
"We can add the infinite busload of guests, by shifting every one already in the hotel from room n to room 2n.
Then, we put the new guests in the odd-numbered rooms."

After all the new guests are settled, a new challenge arises: an infinite number of infinite busloads.
Still, we can accommodate the new arrivals.
We shift every one already in the hotel from room n to room 2^n.
We then place the people on the first bus in room numbers 3^n, the people in the second bus in rooms 5^n, the people in the third bus to rooms 7^n, and so on for each (prime number)^n.
Since there are an infinite number of prime numbers, there will be an infinite number of infinite such sequences.
And, we'll have lots of empty rooms left over.

The splitting headache which may arise from thinking about infinite numbers may correspond to a split between two ways to think about cardinal numbers.
We use them to measure size.
But, we also use one-one correspondence to characterize cardinal numbers.
In the finite realm, these two approaches converge.
But with transfinite numbers, as with the infinite hotel, the two concepts diverge.
The size of the integers seems to be bigger than the size of the even numbers.
But, they can be put into one-one correspondence with each other.

We call the claim that two sets are the same size iff they can be put into one-one correspondence with each other Hume's principle.
Hume's principle is the guiding force behind infinite arithmetic.
Cantor relied on Hume's principle to generate the different transfinite numbers, the different levels of infinity.
When we list the members of something, we are putting them into one-one correspondence with the natural numbers.
Cardinal numbers are the sizes of sets, the number to which we count when we put the set in one-one correspondence with the natural numbers.
Any set that can be put into one-one correspondence with the natural numbers, any set whose members can be listed, is called a denumerable set.

It turns out, according to Cantor's diagonal argument, that we can not make certain lists.
Some sets are non-denumerable.
For example, we can not list the real numbers.

Our interest in the diagonal argument is in its application to linguistic ontology.
To see how Langendoen and Postal use the diagonal argument, we have to look a bit more closely at transfinite arithmetic, and the set theory which underlies mathematics.
There are two different kinds of numbers, corresponding to their different uses.
Cardinal numbers measure size, ordinal numbers measure rank.
There is a long debate in the history of the philosophy of mathematics about whether the ordinals or the cardinals are primary.
Frege preferred to start with the cardinal numbers.
Most modern set theorists take the ordinals as primary, defining the cardinals in terms of them.
In fact, it is customary to think about numbers as particular kinds of well-ordered sets.
A set is well-ordered if, basically, we can find an ordering relation on the set, and it has a first element.
All of mathematics can be reduced to logic and set theory, as Whitehead and Russell showed.
The consequences of such reductions for mathematics are disputed, and beyond the scope of our course.
Our interest is in Cantor's diagonal argument.

Cantor's diagonal argument applies both to numbers and to sets.
Langendoen and Postal invoke the set-theoretic version of the argument.
But the number-theoretic version is easier to understand.

Transfinite numbers share some properties of finite numbers, but they have some properties of their own.
For all cardinal numbers a, b, and c, whether finite or transfinite, the following hold:

A1. $a+b=b+a$
A2. $ab=ba$
A3. $a + (b + c) = (a + b) + c$
A4. $a \bullet (b \bullet c) = (a \bullet b) \bullet c$
A5. $a \bullet (b + c) = ab + ac$
A6. $a^{(b+c)} = a^b \bullet a^c$
A7. $(ab)^c = a^c \bullet b^c$
A8. $(a^b)^c = a^{bc}$
A9. $2^a > a$

But, the following properties hold of transfinite numbers, but do not hold of finite numbers:

T1. $a+1=a$
T2. $2a=a$
T3. $a \cdot a=a$

We demonstrated T1-T3 in considering the infinite hotel.
Technically, we showed that there was a bijective mapping from one set to the other.

Let's return to A9, which is the key claim for which the diagonal argument is used.
In set-theoretic terms, A9 says that $\wp(a) > a$.
'$\wp(a)$' refers to the power set of a, the set of all subsets of a set a.
Consider a set $A = \{2, 4, 6\}$
Then $\wp(A) = \{\{2\}, \{4\}, \{6\}, \{2, 4\}, \{2, 6\}, \{4, 6\}, \{2, 4, 6\}, \varnothing\}$
In general the power set of a set with n elements will have 2^n elements.
Sets with n members are the same size as sets with n+1 members, or with 2n members, or with n^2 members, for infinite n, as in T1-T3.
We might think that sets with n members are the same size as sets with 2^n members, too, and that A9 does not hold for transfinite numbers.
For, with infinite numbers, it is not always clear that what we think of as a larger set is in fact larger.

The claim that $\wp(a) > a$ is called Cantor's theorem.
The proof of the theorem is a set-theoretic version of the diagonalization argument.
It shows that the cardinal number C of the power set of a set is strictly larger than the cardinal number of the set itself (i.e. $C(\wp(A)) > C(A)$).
It suffices to show that there is no function which maps A one-one and onto its power set.
A function is called one-one if it every element of the domain maps to a different element of the range.
A function maps a set A onto another set B if the range of the function is the entire set B, i.e. if no elements of B are left out of the mapping.

Here is the meat of the proof; it is a bit technical:

> Consider any set A, and assume that there is a function f: A $\Rightarrow \wp(A)$
> Consider the set B = $\{x \mid x \in A \bullet x \notin f(x)\}$
> B is a subset of A, since it consists only of members of A.
> So, B is an element of $\wp(A)$, by definition of the power set.
> That means that B itself is in the range of f.
> Since, by assumption, f is one-one and onto, there must be an element of A, b, such that f(b) is B itself.
> Is $b \in B$?
> If it is, then there is a contradiction, since B is defined only to include sets which are not members of their images.
> If it is not, then there is a contradiction, since B should include all elements which are not members of their images.
> Either way, we have a contradiction.
> So, our assumption fails, and there must be no such function.
> $\wp(A) > A$

QED

The size of the natural numbers is called $\aleph_0$.
$\aleph_0$ is the first transfinite cardinal.
We call its corresponding ordinal number ω (omega).
The real numbers are the size of the power set of the natural numbers, $2^{\aleph_0}$.
We can proceed to generate larger and larger cardinals through the power set process.
Moreover, set theorists, by various ingenious methods, including addition of axioms which do not contradict the given axioms, generate even larger cardinals.
You can google 'inaccessible cardinals'.

Let's start counting, thinking about the ordinals, now.
We are thinking about the order of the numbers, which is really the order of certain kinds of sets.
For any number, the next number in the order is called its successor.
7 is the successor of 6; ω+1 is the successor of ω.
Ordinals generated in this way (i.e. by adding one) are called successor ordinals.
In transfinite set theory, there are also certain sets (which stand for certain infinite numbers) which are called limit ordinals.
We generate limit ordinals by combining (taking the union of) all the members of a set.
For example, if we combine all the sets that correspond to the finite ordinals into a whole, we get another well-ordered set.
This will be a new ordinal, and it will be larger than all of the ordinals in it.
So, there are two kinds of ordinals: successor ordinals and limit ordinals.

Here are some ordinal numbers, in order:

$1, 2, 3, \ldots\ \omega$
$\omega+1, \omega+2, \omega+3 \ldots 2\omega$
$2\omega+1, 2\omega+2, 2\omega+3 \ldots 3\omega \ldots$
$4\omega, \ldots, 5\omega, \ldots, 6\omega, \ldots. \omega^{2,,,}$
$\omega^3, \ldots, \omega^4, \ldots, \omega^\omega \ldots$
$(\omega^\omega)^\omega, \ldots, ((\omega^\omega)^\omega)^\omega, \ldots, \varepsilon^0$

The limit ordinals are found after the ellipses on each line.
That's enough infinite mathematics for now.

VII. Langendoen and Postal's Argument Against Generative Grammar

Chomsky's generative grammars have two key features that leave them open to Katz's criticisms.
First, they are constructed like systems of logical inference.
From a finite stock of lexical items, we can derive or generate new sentences.
We can derive in this fashion a denumerable number of new sentences.
But, there is no way to construct a limit ordinal in this way, no way to get to a second-level of infinity.

The second key feature that leaves Chomsky's conceptualism susceptible to Katz's criticisms is that the generative grammar is a facet of the brain.
Thus, any attempt by the Chomsky to expand the tools available to the language user is limited by the finite mind.
If Langendoen and Postal are correct that there are non-denumerably many sentences of English, then the possibility of a mentalistic linguistic ontology, a theory of language which is native to the physical brain, seems doomed.

To show that there are non-denumerably many sentences of English, Langendoen and Postal derive a version of Cantor's diagonal argument.
Specifically, they use the set-theoretic argument that the power set of a set is always strictly larger than the set itself.
Consider the sequence of sentences that Katz calls E:

I know that I like cheese.
I know that I know that I like cheese.
I know that I know that I know that I like cheese.
I know that I know that I know that I know that I like cheese.
…

Consider any conjunction of any pair of sentences in E; it will be a grammatical sentence, by principles of compositionality.
Even an infinite conjunction (the conjunction of infinitely many sentences in this sequence) will be grammatical.
Generative grammars will be able to construct the infinite sequence E.
But, now consider the power set of E, $\wp(E)$.
Each element of $\wp(E)$ can be turned into a grammatical sentence, by conjoining its elements.

But, since ℘(E)>E, there will be non-denumerably grammatical sentences of English.
In fact, there will be non-denumerably sentences of English which discuss only my knowledge of my taste for cheese!

The conceptualist worries about the existence of ℘(E), and rightly so.
There is no way that the generative grammar of the brain could support so many sentences.
But, the linguistic platonist can construct such sentences.

> On linguistic realism, the existence condition for a string type is the consistency of its specification. If a string type is consistently specifiable, it's a possibility, and if it's a possibility, it exists as a string type. In the case of abstract objects, there is no extensional difference between the possible and the actual (288).

VIII. Summing Up

Katz argues that Chomsky makes two errors.
The first is the claim that languages are essentially mental/biological.
The second is the claim that the best account of our knowledge of language relies on a native theory of generative grammar.
Instead, Katz argues that languages are abstract objects, and that our best account of our knowledge of language relies on the intuitions about grammaticality, and Katz's autonomous theory of sense, that we saw in our previous readings.