Á. Gallego & D. Ott (eds.). 2015. *50 Years Later: Reflections on Chomsky's* Aspects. Cambridge, MA: MITWPL.
Available for purchase at http://mitwpl.mit.edu

Complementarity in Language

Juan Uriagereka
University of Maryland

> There is surely no reason today for taking seriously a position that attributes a complex human achievement entirely to months (or at most years) of experience, rather than to millions of years of evolution or to principles of neural organization that may be even more deeply grounded in physical law.
>
> N. Chomsky, *Aspects of the Theory of Syntax*, ch. 1, p. 59

1 Introduction

The quote above reveals that Chomsky had minimalist thoughts already in 1964. In (2005) he attributed the so-called Strong Minimalist Thesis to "principles of structural architecture and developmental constraints that enter into canalization, organic form, and action". It is useful to reflect on what it would take to "ground in physical law" some of the properties of linguistic systems that go back in their initial formulation to *Aspects*.

2 Quantum Cognition

If contemporary physics is to help understand a linguistic computation, the apparatus of quantum theory, dealing with distributed systems, may well be a starting point. Generative linguists would not be the first to initiate such efforts: the field of quantum cognition, for staters, already exists, with reasonable questions and challenges. Their Wikipedia entry starts as follows:

> Quantum cognition [. . .] applies the mathematical formalism of quantum theory to model cognitive phenomena such as information processing by the human brain, decision making, human memory, concepts and conceptual reasoning, human judgment, and perception.

The field is not to be confused with any hypothesis that the brain itself involves quantum processes. While quantum cognition is compatible with such a possibility, the theory is more abstract. It is based on the idea, expressed in Aerts and Aerts 1994, Atmanspacher et al. 2002, Khrennikov 2006, that quantum probability theory can help understand information processing in minds, by factoring in probabilistic notions and contextual dependencies. More specifically, quantum cognition:

> . . . uses the mathematical formalism of quantum theory to inspire and formalize models of cognition [focusing] on modeling phenomena in cognitive science that have resisted traditional techniques or where traditional models seem to have reached a barrier (e.g., human memory), and modeling preferences in decision theory that seem paradoxical from a traditional rational point of view (e.g., preference reversals). [Wikipedia]

The implication is that neural networks produce effects that can be modeled by the interference of probability amplitudes (a wave which, upon measurement, gives rise to a probability).

To get a sense of that mouthfull non-technically, consider tuning a guitar. This involves settling on the high E (string #1), then proceeding to equalize the remaining strings in pitch at a higher fret. At stake is the frequency at which a string vibrates when plucked, as altered by tensing the tuning keys. One knows two strings have been equalized by successively listening to the notes each plays: the shorter one listens, the less clear the analysis. The question is why a very short (i.e. very precise) time of listening doesn't entail an equally precise analysis of the note. Waves can be modeled as sums of sines and cosines, per Fourier analysis. The variables at stake – the frequency at which a string vibrates and the time it takes to listen to this frequency – are "conjugate", as they cannot be simultaneously measured (in one formalism, their operators do not commute). The more precisely we measure one of the variables – say, timing – the less precisely we can express its conjugate, frequency. The famous Uncertainty Principle expresses a mathematical inequality that pertains to the fundamental imprecision of complementary conjugate variables. This is just a mathematical fact about measuring Fourier duals. A better device wouldn't do any good, since that reduces the measuring sample on one variable, at the expense of the measuring precision of the other – therein their complementarity.

What was remarkable for physicists a century ago was to translate that mathematical fact about measuring waves into a claim about reality. That's gutsy metaphysics: to propose that reality, when involving complementary variables, presents itself in one of two incompatible ways: existing as one variable or the other, not both. Do processes in mind present complementarity? Scientific practice suggests that the way to go about answering that is seeing what it would buy us. Of course, even if one were to find situations of complementarity in mind, one need not exhibit the metaphysical guts that quantum physicists had, to thereby declare that mind is, hence, "dual". That step in the reasoning is as drastic for mind as it has been for physical reality. Moreover, physically it is of course patent that brains propagate waves and their activities can and have been described in terms of vector fields. But that certainly doesn't entail that the minds that emerge from those brains should then, ipso facto, be wave-dependent. Whether they are rests on how the mind-body problem is resolved, which may not happen any time soon.

3 Some Global Effects

A typical argument within quantum cognition works as follows. Suppose a subject can toss a coin to win $200 or lose $100. In Condition A of an experiment testing decisions made in subsequent tosses, subjects are allowed to find out about the toss outcome before playing a second time; in Condition B, they are not. In both instances they're given the option of tossing again. As Tversky and Shafir (1992) show for a version of this experiment, these results ensue:

(1) a. When believing they won the first round, a majority of subjects choose to play on the second.
b. When believing they lost the first round, a majority of subjects choose to play on the second.
c. When not given the results of the first round, the majority of subjects choose not to play again on the second.

Needless to say, (1c) is inconsistent with what happens in the majority of cases in (1a) and (1b). Quantum cognitive scientists account for this violation of the law of total probability as a quantum interference. Deviations from standard rational expectations can be understood by assuming that the overall conceptual landscape influences the subject's choice – the decision process is "context-sensitive", in a sense to be discussed shortly.

Another example stems from understanding concepts as prototypes. A classic conundrum was posed in Osherson and Smith 1981: the typicality of a concept like a pet-fish (say, a guppy) is not a function of the typicality of the conjoined terms in the compound: a typical pet is experimentally ascertained to be warm and furry, while a typical fish is experimentally ascertained to be relatively large; the conjoined typicality of the terms is not predictable from the typicality of the concept ensuing from combining the two. Of course, this is only one of many problematic aspects compounds pose. So while it may be good to have a theory that predicts peculiarities in compound "composition", it is important not to over-predict. Witness:

(2) I consider my fish a pet. My trout is definitely a pet-fish!

There doesn't seem to be anything non-compositional here. Moreover, it would seem as if the fish and pet concepts in these sentences are dull exemplars. The point is: whatever is going on with compounds, one does want syntactic composition in the normal relations between arguments and predicates that sentential syntax instantiates.

With that as background, examine instances of so-called "amelioration" in syntactically unacceptable sentences. One such case was first raised by Haj Ross:

(3) They want to hire someone who speaks a Balkan language, but I don't remember which (Balkan language) (?*they want to hire someone who speaks.)

While the example in full presents a violation of an "island" condition (extracting from inside a relative clause), the ellipsis (without the last parenthetical) is fine. A very different amelioration instance involves "that"-trace violations, as follows:

(4) Who do you think (*that) t died?

The trace of a Wh-extraction from subject position cannot follow an overt complementizer like *that* – though it is fine if the complementizer is dropped. Now witness:

(5) a. Who do you think that, **after years and years of cheating death**, t finally died? (Example by Barss and Deprez).
b. Who did Leslie said that, **to KIM**, t had given the money? (Culicover)
c. Who do you think that, **as we're aware**, t loves silly books? (Ackema)
d. Who does John doubt whether, **and Bill suspect that**, t cheated on the exam? (De Chene)
e. Who do you think that t WROTE Barriers (as opposed to edited it)? (Drury)

None of their authors claim that these sentences are perfect, but they all indicate that the sharp violation in (4) ameliorates within them. The examples in (5) would seem to have the same structure as (4), with *that* as an overt complementizer and *who* as the extracted subject. In these instances there is "more discourse context", although of various sorts: a "long" adverb in (5a), a displaced focalized phrase in (5b), an "as" interjection in (5c), right-node raising in (5d), contrastive focalization in (5e)… One final amelioration example involves the PRO-gate effect discovered by Jim Higginbotham:

(6) Who did PRO/*his kissing his mother upset?

The difference at stake here is rather trivial: pronunciation of the first *his*. When the pronoun is overt, though, the Wh-extraction leads to a Crossover effect; if the pronoun is dropped (replaced by PRO), the result improves.

Theorists have analyzed these situations in various ways. For example, Hornstein et al. 2007 accounts for the ellipsis amelioration by presenting a theory of islands that makes them sensitive to their phonetic shape. Hornstein and Kiguchi 2003 sees the PRO-gate amelioration as an argument that PRO may be the trace of A-movement, based on the fact that Crossover effects generally disappear with A-movement. Stories for the "that"-trace ameliorations have varied; the most recent is proposed in Sato and Dobashi 2013, which bases the account on the analysis of what counts as a prosodic phrase. These are three explanations for the three ameliorations.

Now recall the quantum interference approach to (1). Players that haven't been informed of the output of their first coin toss can globally imagine *both* next steps, unlike if they have been informed of the fate of their first throw (which decides on heads or tails). In that global state, subjects are not considering a fate that ensues after they have "collapsed" the choice to either heads or tails; rather, *both options are "live"*, and therefore subjects make the "calculated" move not to play the second time around. Observe the possibilities:

(7) a. won \$200; (i) won \$200 = won \$400 (.25 chances)
(ii) lost \$100 = won \$100 (.25 chances)
b. lost \$100; (i) won \$200 = won \$100 (.25 chances)
(ii) lost \$100 = lost \$200 (.25 chances)

Given these odds, one can debate whether taking a second toss pays off. But according to this approach, what gives subjects pause is *not knowing* what happened in the intermediate toss, which keeps all options for a further toss open, good or bad.

We may think of that state in terms of Coleridge's famous "suspension of disbelief", which keeps a subject engaged in a task for its duration, *regardless of its immediate plausibility*. With that in mind, note that, for grammaticality judgments, a pattern holds across the ameliorations above. It can be abstractly represented as in (8):

(8) a. [Wh-phrase… [… t …]…]
↑_ _ _?_ _ ***R***_ /

b. [Wh-phrase … [PRO…] upset t]
↑ ↑___***R***___/|
|_ _ _ _ _ _?_ _ _ _ _ _|

c. [Wh- phrase …[that [X Y Z] [t…]…]
↑ _***R***_↑ |
|_ _ _ _?_ _ _ _|

R in these examples denotes a context-sensitive relation of grammar that is *independent of the one that yields the ungrammaticality*. R in (8a) is ellipsis, it is control in (8b), and a variety of relations in the complementizer-trace effects, all complex. The question mark ? in (8) denotes the offending relation, which *ameliorates when its context of application overlaps with R's context of application.* This "suspension of ungrammaticality" is abstractly similar to the suspension of disbelief effect alluded to above, in that it is clearly a global process that "takes another look" at a bad example; so one may wonder whether this, also, could be seen as a quantum interference of some sort.

There are two distinct, though related, components to the possible approach being entertained. One is the very globality of the analysis. Of course, somewhat global conditions in grammar arise the moment one thinks of obviation effects, for example:

(9) She feared (people thought…) [the FBI would kill Marilyn].

Familiarly, *she* and *Marilyn* cannot co-refer, regardless of how far apart they are. However, there is a further twist to these ameliorations: the "suspension" effect appears to involve the computation of equally complex relations (*R* in these instances), which somehow puts things in a different light for speakers. Other ameliorations seem to take on that shape too, for instance "connectivity effects" discussed in Kayne (1984):

(10) a. ?*What do you wonder why I bought?
b. Who wondered why I bought what?

(11) a. ?*What did who see?
b. What did who see where?

The logical forms for the multiple questions in (10b) and (11b) are very similar to the ones involved in the simpler, albeit ungrammatical, examples (a Wh-island in (10a) and a Superiority effect in (11a)). Yet whatever causes the ungrammaticality in the relevant relation, as the new dependency *R* is introduced that overlaps with it, the results improve. It is, thus, the joint action of a certain globality in the computation, together with the overlap in context-sensitive relations, that yields the suspension of grammaticality. If there is interference here, in the quantum sense, it ought to involve both of those factors.

"Suspensions" (of disbelief or grammaticality) are strange, and the quantum cognition approach only gives us a way to talk about them. If the mind is acting in global terms in certain conditions, it is plausible to address these puzzles in quantum-like terms: by simultaneously considering the sum of representations being derived in parallel. Such an argument is strong only if alternative explanations aren't and only as elegant as the predictive power of its underlying theory. To date there are no systematic treatments of amelioration, so this line of research is potentially interesting, and should be tested with examples constructed to fit the bill (ungrammatical context-sensitive dependencies that ameliorate in the context of "interfering", equally complex, unrelated relations). Abstractly similar issues arise with grammatical preference, which typically involve situations as in (12):

(12) A boy saw every girl.
a. $\exists$x, boy x [$\forall$y, girl y [x saw y]]
b. $\forall$y, girl y [$\exists$x, boy x [x saw y]]

The structural ambiguity in (12) is well known. But despite the fact that both of those representations are grammatical, speakers feel that *(12a) is preferred over (12b).* To account for this, we need a comparison set and a procedure to establish the appropriate choice. If we find such a procedure within this framework, that would speak to its predictive power. In searching for this, we need to consider what sorts of grammatical conditions a quantum theory of syntactic categories and relations ought to involve.

4 Spooky Grammatical Paradigms

Consider whether familiar properties of grammars can, or should, be related to quantum considerations. Take simple projection, illustrated as follows:

(13) Chewing gum strengthens your jaws.

Of course *Chewing gum* can be a projection of *chewing* or a projection of *gum* instead. Projecting each lexical term means having one or the other determine the labeling of the ensuing combination. How does that process happen? Why is there not a "mixed" projection that has properties *of both* terms, instead? One way to actually predict this is by understanding the (*chewing*, *gum*) relation as involving, in some sense, conjugate variables, such that the verbal one is a dynamical variable tracking the spatial nominal one. Familiarly, one can think of a chewing as an event lasting while there is gum (left to chew), which can be modeled as the derivative over time of the mass term *gum*. If this is the right way to understand verb-theme relations, and this intuition is taken seriously in terms that have the dynamical (verbal) variable defined over the static (nominal) variable, it is plausible to treat the verb-noun relation as involving Fourier duals, with this consequence:

(14) *Complementarity in head-complement relations*
For (head, complement) relations treated as Fourier duals, corresponding variables may be mathematically complementary.

This *complementarity of complements* entails that the complement variable can be specified, or the head variable – not both. That certainly tracks projection. For example, "collapsing" the relation in terms of the static (nominal) variable yields a nominal projection; concentrating on the dynamical (verbal) variable yields a verbal projection, which continues combining.

If (14) is the general approach to head-complement relations, the *asymmetry of projection* follows, as well as the fact that projections must settle on either a head-centered expression (head-complement relations, allowing further composition), or a non-head-centered expression (which terminates in that particular collapse, yielding predication). The formalism itself has as a consequence facts that otherwise need to be stipulated, e.g. as modes of semantic composition. The approach makes interesting predictions for other head-complement relations, for example predicting dynamical hypotactic (complementizer-centered) vs. static paratactic (claused-centered) relations for complement clauses, or dynamical strong (determiner-centered) vs. static weak (noun centered) relations for determiner-noun relations.

Once complementarity is assumed for projection, similar questions can be posed for what is, in essence, a super-projection: the formation of movement chains. Can a chain also be considered a dual object, with a dynamical and a static expression, as it were? That is possible if all the steps of movement (the chain occurrences) are seen as superposed states of a dynamical object spanning over relevant syntactic contexts. The corresponding static variable is the one and only position in which such chain occurrences manifest themselves in interpretation, at PF and LF. A version of the Principle of Complementarity when applying to chains looks as follows:

(15) *Principle of Complementarity in Chains (PC_C)*
Within derivation D, it is impossible to ascertain the scope σ of a given chain occurrence C while at the same time ascertaining the reach ρ of the superposed set of chain occurrences of C in D. Therefore, observation of derivation D with reach ρ collapses D to ρ's conjugate variable σ.

Where ρ is the set of contexts that the chain occurrences involve, σ is the corresponding scopes of those contexts. Assuming the variables are conjugate, the effect that the PC_C has on PF representations is direct, where σ manifests itself as linearization. Consider in this regard Nunes's 2004 analysis of the linearization of chain links, as in (16):

(16)
a. Gum was chewed t.
b. **Gum** was chewed **gum**.
c. Gum was chewed ~~chewed~~.

If movement steps (copies boldfaced in (16b)) are identical occurrences, in chain-formation the upper and lower links of given chains are in the paradoxical situation of having to precede and follow intervening elements. Nunes concludes that the system resorts to 'turning off' all copies but one, the linearizing element (as in (16c)). Suppose the set {{*gum*, *T'*}, {*gum*, *chewed*}} (including the immediate syntactic contexts for each chain occurrence) is the derivational object with a certain reach ρ. If ρ is the canonical conjugate of some σ, observing {{*gum*, *T'*}, {*gum*, *chewed*}} yields some value for σ in a PF linearization. This doesn't tell us what the value for σ should be – e.g. *gum* preceding. But it is a significant start, as the conclusion that we need a definite value for σ follows from the PC_C. In a nutshell: *Interpreting syntax is observing it*.

Identical considerations arise for the LF component, yielding what Hornstein (1995) calls conditions of "uniqueness" – the stipulation that multiple occurrences at LF are interpreted in a single location. This is more puzzling than it may seem at first, if the occurrences as copies are "real enough" to trigger reconstruction effects:

(17) a. Which pictures of **myself** [did John say t [Mary believes t [**I** took t]]]
b. Which pictures of **herself** [did John say t [**Mary** believes t [I took t]]]
c. Which pictures of **himself** [did **John** say t [Mary believes t [I took t]]]

The standard way to treat these sorts of effects is as in (18):

(18) a. [Which ~~pictures of myself~~] [did John say ~~[Which pictures of myself]~~ [Mary believes ~~[Which pictures of myself]~~ [I took [~~Which~~ pictures of myself]]]]
b. [Which ~~pictures of herself~~] [did John say ~~[Which pictures of herself]~~ [Mary believes [~~Which pictures~~ of herself] [I took [~~Which~~ pictures ~~of herself~~]]]]
c. [Which ~~pictures of himself~~] [did John say [~~Which pictures~~ of himself] [Mary believes ~~[Which pictures of himself]~~ [I took [~~Which~~ pictures ~~of himself~~]]]]

If the anaphor is local to a valid antecedent, different LFs arise, showing that, so long as the phrase *which* has matrix scope and the noun *pictures* establishes the variable position for the theme theta-role associated to *took*, where the restrictive material *pictures of …self* is interpreted is up for grabs. That said, consider the ungrammatical (19):

(19) * Pictures of themselves seemed to no photographer to appear to no model to have been retouched.

Notice: a sentence like *No man shows no woman bad pictures of themselves!* demonstrates that anaphors like *themselves* can take split quantificational antecedents. In turn, *Which picture of themselves would no man say no woman should destroy?* shows that such split antecedents can be picked in successive-cyclic fashion. So then why can't (19) be analyzed as in (20)?

(20) [[pictures ~~of themselves~~] seemed to no photographer
[[~~pictures~~ of themselves] to appear to no model
[[~~pictures~~ of themselves] to have been retouched]]]

The PC_C prevents the LF in (20), inasmuch as key to this object is the occurrence of *themselves* in two different intermediate sites. Uniqueness for LF occurrences impedes such an analysis – but it need not be stipulated: Uniqueness ensues as an LF consequence of a collapsed chain not being able to reach over two different scopes, as the PC_C predicts.

The uniqueness of occurrences may seem to trivially follow from the fact that a single lexical element is involved, which "re-merges" at the appropriate (displaced) syntactic contexts. The problem with that approach is that it presupposes a notion of syntactic token that has no understood reality. Any syntactic theory must distinguish two identical token words within a sentence (e.g. the word *word* when we say *this word precedes this other word*) from various (syntactically arising) occurrences of these elements as above, and those two from the corresponding lexical type, say, *word*. The lexicon doesn't contain each and every token of the word *word* that has been used in this paragraph, but just a basic (concept, expression) pair that linguistically encodes that particular notion. The issue is how to activate that lexical type multiple times in a sentence, in some instances letting those activations survive to interpretation, while in others, whatever activation takes place within the derivation (as occurrences) dies out as the sentence collapses into an interpretation within the reach of that activation. The quantum approach takes the chain to be the sum of the probability of each occurrence – a copy at a given context – and has the uniqueness of tokens follow from the fact that interpreting one such dynamical object is interfering-with/observing it, which results in the collapse. A different token is a separate wave, which starts at the formation of the lexical array from the lexicon.

The consequences of the PC_C, thus, seem reasonable for chains. They actually neatly extend to a treatment of the preferences in (12), if analyzed as in Martin and Uriagereka 2008. In the present approach chains are comprised of several *simultaneous* derivational stages: In their dynamic state $\boldsymbol{\rho}$ they exist in the contexts the chain links span over. To interpret a chain with reach ρ is to *collapse* the chain in scopal site σ. The chains in (12) can collapse as in (21) (where φ signals a Case/agreement site, θ a θ-position):

(21) a. [$_{IP}$ a boy [$_{\underline{v}P}$ every girl [$_{VP}$ saw]]]
φ φ
b. [$_{IP}$ a boy [$_{\underline{v}P}$ [$_{VP}$ saw every girl]]]
φ θ
c. [$_{IP}$ [$_{\underline{v}P}$ a boy [$_{VP}$ saw every girl]]]
θ θ
d. [$_{IP}$ [$_{\underline{v}P}$ every girl [a boy [$_{VP}$ saw]]]]
φ θ

75% of these representations collapse in an LF where *some boy* has scope over *every girl*; only 25% collapse in the inverse reading. These rates may correspond to individual speaker preferences. That makes sense only if individual speakers simultaneously access all the chain states in (21), which the quantum approach allows. An interesting question of course emerges in terms of what the nature of the preference set is (*where* "scope interactions" are possible). This probably ought to be related to the nature of the global domains where amelioration has been shown to obtain, but we set the topic aside now.

After having motivated the PC_C and some of its consequences, we should ask also about notions that presuppose complementarity. Consider the following thought experiment. Imagine preparing two distinct syntactic categories-as-waves in such a way that they that start from "reduplicated" token waves within the same lexical array, of the sort arising in the emphatic *(these are) pictures pictures (not photocopies)*. At some later time in the derivation, after movement, we have one of the prepared tokens collapse in a given site, while we have the other prepared, identical, token collapse in a crucially different site:

(22) a. PF: Pictures of themselves pictures of themselves seemed to no photographer to appear to no model to have been retouched.
b. LF:
[[pictures ~~of themselves]~~ [pictures ~~of themselves]~~ seemed to no photographer
[[~~pictures~~ of themselves] ~~[pictures of themselves]~~ to appear to no model
[[~~pictures of themselves] [pictures~~ of themselves] to have been retouched]]]

Given the initial correlation between the prepared tokens (joined in the lexical array by their reduplicative identity) does this gambit not allow us to know *both* the reach *and* the scope in an object as in (22)? Of course that would violate the PC_C, if meant ontologically. If we manage to determine the reach for a chain-wave C within a sentence, we cannot ascertain the scope of one of its occurrences, *even in a roundabout fashion.* This is good, as (22) is ungrammatical.

But now imagine the possibility that identical chains in these "prepared" reduplicative conditions were capable of instantly transmitting information, so that (surprisingly) the more definite the scope of one of the chain-waves is upon its collapsing, the reach *of the other* would then be more indefinite, and vice versa. While that may not yield a viable result for the collapses in (22), it may actually provide a rather interesting account for an old chestnut:

(23) a. Lexical Array {tried, {to, leave, John, {John}}}
b. PF: John tried to leave.
c. LF: John tried [John to leave]

This analysis of control provides an alternative to Hornstein (1995)'s analysis. The key to the chain-like dependency is to *start with a reduplicated pair* (John, John) in the lexical array, only to *use up one of the items in the first cycle*, thereby to yield *John to leave*. Subsequently, in the next cycle, the second *John* in the array is used. This of course presupposes that a lexical array need not be exhausted in a given cycle: it may be possible to use whatever part of the array yields a grammatical representation, passing on the remaining materials up to the next cycle. This procedure yields the LF representation in (23c), which appropriately tracks the semantics in terms of thematic relations: there is a "trier" and there is a "leaver". Of course, those two happen to be identical. But that aspect of the analysis is also captured if the two John's are actually *entangled*, in such a way that, at the point of determining the chain's reach, the system interprets the two John chains as a super-chain, with a unique derivational fate (signaled by the dotted line in (23)). Similarly, at the point of collapse the super-chain actually has a unique scope position, both at LF and at PF. While this super-chain does have chain characteristics, it does so *because of entanglement*, not because of complementarity (like regular chains).

Unlike complementarity, which operates on every local part of a construct, *entanglement acts globally.* So when we talk about entanglement on a sentence what we're really discussing is an abstract, global, structure that is distributed throughout (a phrase-marker). In effect this is "action at a distance", since the analysis does not commit us to the sorts of limitations (e.g., in terms of locality) that standard movement poses. Future research will need to determine the validity of such a consequence, as well as whether specific conditions limiting entanglement per se also obtain.

For example, the reason (22) cannot be salvaged as (23) is may be that in (22) there is *no configurational theta-role* to support the second *pictures of themselves* in a site separate from that of the first. In contrast, in (23) each *John* is in a valid theta configuration, as it enters the derivation. As a consequence of this analysis, one need not trivialize theta-roles as features (or otherwise giving up the good reasons why movement was traditionally considered different from control). While conditions of complementarity entail the standard chain behavior, conditions of entanglement that arise from complementarity allow for the chain-like analysis of instances that *start as separate tokens* in the lexical array. These notions (complementarity and entanglement) are part of the same machinery, but they are quite distinct. Complementarity is a property of waves under Fourier analysis; entanglement follows from an ontology that depends on taking complementarity seriously. Entanglement presupposes complementarity, not vice-versa.

The present approach provides us, also, with a tool to analyze "partial" control, where the entanglement between the co-referential items is only partial:

(24) John asked to dance (together).

(25) a. Lexical Array {asked, {(together), dance, John, [ϕ], {[ϕ]}}}
b. PF: John asked to dance (together)
c. LF: [John [ϕ]] asked [[ϕ] to dance (together)]

In this instance what is reduplicated is feature matrix [ϕ], essentially a pronominal clitic. While the first [ϕ] token in the array is used in the lowest cycle, satisfying the "dancer" theta-role, the second [ϕ] token, together with the category *John*, are passed on to the higher cycle – where they combine in clitic doubling fashion. Under these circumstances, the entanglement is between the clitic elements, so the co-reference has the weaker nature that pronominal dependencies exhibit more generally. Other chain-like phenomena – in terms of (PF, LF) interpretation, but not standard movement conditions – may be analyzed along the same lines – e.g. parasitic gaps, see Martin and Uriagereka 2014. Indeed, in principle *any* condition of identity of this "reduplicative" sort should be amenable to an entanglement approach. An intriguing such instance arises within ellipsis, which Chomsky 1995 speculates should involve "chain like" conditions. An entanglement approach may provide us with a way of illuminating the "parallel" effect in ellipsis, involving interpretive nuances that have never been explained.

5 Conclusions

This brief exploration of a syntax obeying complementarity is meant as promissory. The major theoretical move is bold, but scarce. Claiming that scope and reach within a syntactic chain are conjugate variables invites us to ask what in the nature of syntactic categories makes them be that way. It seems relevant that if movement ("internal merge") lives on scope-reach pairs, in parallel fashion (regular) merge should live on the static-dynamic duality that manifests itself within head-complement relations. This provides support for why syntactic composition should be bottom up and why (firstly merged) complements should be central in articulating internal thematic and quantificational relations. The question is why the complex variables (projecting heads, reaching occurrences) are defined over more elementary variables (complements, scope). That parallelism doesn't follow from anything understood, although it says something about how context-free computational relations constitute parts of context-sensitive ones.

Two notions of context are related here: the history of a derivation (or scope of a given occurrence) and context in the quantum sense that a dynamical system plays in determining the specific values of elementary duals. If syntactic context is physical context, one ought to ask whether other notions that are context-dependent, usually deployed in an information-theoretic sense, are another form of context that could be integrated. The question is then how syntactic context affects interpretive context, and "surface semantics" effects do suggest that a relation exists. It would be good to examine, also, whether complementarity effects arise at that supra-sentential level, e.g. in terms of "discourse-linking" entailing amelioration of island conditions.

Computational elegance holds within sentence grammar, not below or above. Infra-sentential matters are known not to be systematic, transparent or productive. Supra-sentential matters have been formally studied, but it is unclear whether speakers have robust judgments about them, or discourse coherence becomes less reliable with size and speaker-specific conditions. Neither extreme is surprising for the systems explored here. Quantum cognitive theorists have explored the lack of compositionality within compounds or similar conditions about textual (in)coherence in some areas – which may be related to whatever happens in ameliorations – to argue for the quantum nature of the system. But just as things get strange below words or above sentences, finding quantum effects within sentence is clearly difficult, as sentences are well-behaved computational entities. It is only when going into reconstructions, construal, ameliorations and other such strange notions that a quantum system may be showing its face.

How pretty that face is depends on whether, in the process of assuming this architecture, one also wants to predict why the computation's guts live in projected structures, displaced tokens, chain-like long-distance dependencies, and other such creatures. The debt owed to Noam Chomsky for having found such creatures and, more importantly, having provided a coherent analysis of their essential computational characteristics in *Aspects*, is only comparable to the most profound scientific breakthroughs of the 20th century, such as the quantum revolution.

Acknowlegments
This owes much to discussions with Massimo Piattelli-Palmarini, Doug Saddy, Ángel Gallego, Bill Idsardi and especially Roger Martin (who's helped develop the ideas) and Michael Jarret (who's helping to formalize them). Misunderstandings are the author's responsibility. The debt to Noam Chomsky is impossible to express – one hopes these thoughts are not too unorthodox for his taste.

References

Aerts, D. and A. Stein. 1994. Applications of quantum statistics in psychological studies of decision processes. *Foundations of Science* 1: 85-97.

Atmanspacher, H., H. Römer and H. Walach. 2002. Weak quantum theory: Complementarity and entanglement in physics and beyond. *Foundations of Physics* 32-3: 379-406.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. 2005. Three Factors in Language Design. *Linguistic Inquiry* 36: 1-22.

Hornstein, N. 1995. *Logical Form: From GB to Minimalism*. London: Blackwell.

Hornstein, N., H. Lasnik and J. Uriagereka. 2007. The dynamics of *islands*: Speculations on the locality of movement. *Linguistic Analysis* 33: 149-175.

Hornstein, N. and H. Kiguchi. 2003. PRO Gate and Movement. University of Pennsylvania Working Papers in Linguistics 8: 101-114.

Kayne, R. S. 1984. *Connectedness and Binary Branching.* Dordrect: Foris.

Khrennikov, A. 2006. Quantum-like brain: "Interference of minds". *Biosystems* 84-3: 225-241.

Martin, R. and J. Uriagereka. 2008. Competence for preferences. In *Festschrift for Patxi Goenaga.* X. Artiagoitia and J. A. Lakarra (eds.), University of the Basque Country.

Martin, R. and J. Uriagereka. 2014. Chains in Minimalism. In *Minimalism and Beyond: Radicalizing the Interfaces*, P. Kosta, S. Franks, T. Radeva-Bork and L. Schürcks (eds.). Amsterdam: John Benjamins.

Nunes, J. 2004. *Linearization of Chains and Sideward movement.* Cambridge, MA: MIT Press.

Osherson, D. N. and E.E. Smith. 1981. On the adequacy of prototype theory as a theory of concepts. *Cognition* 9: 35-58.

Sato, Y. and Y. Dobashi. 2013. Prosodic phrasing and the *that-trace effect*. Ms., National University of Singapore and. Niigata University.

Tversky, A. and E. Shafir. 1992. The disjunction effect in choice under uncertainty. *Psychological Science 3*: 305-309.