

Conclusion

What Is Language?

“objects ... are constituted through
the artful creativity of scientists”
(Latour & Woolgar 1979.129)

0.0 Prologue

Having come all this way, it seems reasonable to expect some release. Such closure might be provided by a response to the question *Well, what is language?* (or even just what is *this* specific language, with a lower case *l*). If forced to respond to either, then I have to admit that I do not know.¹

If we are speaking of some named language, then I am pretty certain of some things that that language is *not* ... For example:

The totality of utterances that can be made in a speech-community is the *language* of that speech-community (Bloomfield 1926.154)

From now on I will consider a *language* to be a set (finite or infinite) of sentences, each finite in length and constructed out of a finite set of elements. (Chomsky 1957.13)

Nor is it a *construct-i-con* (Goldberg 2003.8):

What makes a theory that allows for constructions a “construction based” theory is tenet 7 [The totality of our knowledge of language is captured by a network of constructions: a “construct-i-con”]: the idea that the network of constructions captures our knowledge *in toto* i.e., it’s constructions all the way down.

It seems fairly certain that, for example, it is not correct to assert that “the totality of utterances that can be made in ... [the Kutenai] speech-community is the *language* of that speech-community [i.e., Kutenai].” Nor is it any more useful to assert that Kutenai “[is] ... a set (finite or infinite) of sentences, each finite in length and constructed out of a finite set of [Kutenai] elements.” Nor is “the totality of our knowledge of [Kutenai] ... captured by a network of

¹ But that does not mean that I do not have an answer.

Given the title of this work, *Syntax & Semantics*, it should also be clear, but it can be reemphasized, that the question *What is language?* is constrained to the portion of language where we find meaning and its grammatical expression. Hence, aspects of language like phonology, sociolinguistics, etc. are not addressed, nor is grammatical complexity, although any response suggested here must be of a sort that can host an answer that extends to these broader concerns.

[Kutenai] constructions ... a [Kutenai] “construct-i-con” ... [that] captures our knowledge [of Kutenai] *in toto*.” The neologism *constructicon*, with its suffix *-icon*, connotes, like *lexicon*, an enumeration (of constructions) which parallels the *totality* of Bloomfield and the *set* (or *unbounded system*. Cf. fn. 3) of Chomsky.² From 1926 to 1957 to 2003, until now, it seems, the sense of what a language is has not changed much.³

² Each of those characterizations identify something, but I think it is not Kutenai, nor any language.

³ At least in these representations.

Postal (2003.233) initiates his discussion of the *foundations of linguistics* with this characterization of an empirical linguistics, which he denominates the *received view* (234):

As with any intellectual domain, there exist foundational issues about the nature of linguistics, the kinds of objects it studies and, therefore, the kind of field it is. While all surely accept the banality that linguistics is about (natural) language (NL), most linguists currently seem to believe, under the influence of the voluminous writings of Noam Chomsky on the topic, that the proper interpretation of this means that linguistics is about human linguistic knowledge and its development, and is therefore about something mental, possibly then about something biological. Linguistics would then be an empirical science concerned with the development of NL in normal humans and the innate, biologically given mechanisms which determine/permit this development.

but which Postal then rejects because it is “deeply in error and involve[s], inter alia, a fundamental confusion between NL and knowledge of NL” (234). But if one intentionally asserts that knowledge of natural language is what natural language is, there can be no “confusion”, although there may be empirical error. To replace the *received view*, Postal advocates support for Jerrold J. Katz’ proposal (Cf. Postal 2003.234 fn 3 for bibliography) for a “core of linguistics [that] is a formal science, one of the same type as logic, mathematics, and theoretical computer science. This pursuit is a study of NL” Ultimately (237),

I believe ... that an NL is a set-theoretical object, a collection, in fact, a bit more precisely, a collection of sets, where each set is a complex object composed of syntactic, semantic, and expression objects. The traditional term for these sets is “sentence,” so that it is appropriate to say that an NL is a collection of sentences.

and we are back to *totalities*, *sets*, and *constructicons*.

Even one of the most recent efforts, the Minimalist Program, appears retrograde in this sense. The “minimalist program reconstructs syntactic theory around Merge as the central computational operation” (Jackendoff 2013.591), and Merge (Chomsky 2005.11):

An elementary fact about the language faculty is that it is a system of discrete infinity. Any such system is based on a primitive operation that takes *n* objects already constructed, and constructs from them a new object: in the simplest case, the set of these *n* objects. Call that operation Merge. Either Merge or some equivalent is a minimal requirement. With Merge available, we instantly have an unbounded system of hierarchically structured expressions.

Two observations. **First**, “the primitive operation” operates (Jackendoff 2013.599):

In the beginning, how one understands *language* will be strongly affected by how one first encounters it. The trivial observation that all we humans will normally engage in a variety of behaviors that we, in English, bundle together and call *language*⁴ has the effect of privileging our first brush with the phenomenon ... noise, talk, articulation ... and to think of language in those terms.⁵ But articulation alone cannot constitute language. The conclusion that meaning is present, but not contained in the physics of articulation, is expressed by the terms *duality of patterning* (Hockett 1960) and *double articulation* (Martinet 1957).⁶ The coexistence of expression and meaning as partners in the *language* enterprise was one of the easier, and nondisputed, discoveries.⁷

step by step: A is combined with B, then {A, B} is combined with C, and so on. This leads to an inherent bottom-up order to building structure.

and (Di Sciullo & Jenkins 2016.211):

Merge is the basic combinatorial operation capable of deriving the discrete infinity of language. It is necessarily a part of the computational procedure of the language faculty. Merge is a binary operation that takes two syntactic objects a and b and derives another syntactic object consisting of the two objects that have been merged.

This operation is identical to Zellig Harris' (1946.170) proposal of a bottom-up "from morpheme to utterance":

The procedure in assigning these raised numbers which indicate uni-directional substitutability is in essence as follows: we assign raised ¹ to each class symbol, say X [Even the 'X' of 'X-bar' comes from Harris, PWD], when it first appears. Next time the X appears in an equation, we assign it the same number ¹ if the equivalents of this X can be substituted for X¹ in every equation which has so far been written. If the new X cannot be substituted for all the preceding X¹ we number it X². If we later obtain an X which cannot be substituted for all the preceding X¹ or X², we will number it X³, and so on.

Merge recycles a seventy-five year old idea (without acknowledging the primary source). **Second**, also recycled is the conclusion that language is a *totality*, *collection*, *set*, or *unbounded system of hierarchically structured expressions* (Chomsky 2005.11). By this measure, the field of linguistics has made no conceptual advance in 95 years.

⁴ Martinet (1987.31ff) considers the effects that the different meanings in the linguist's native language may have on his/her understanding, e.g. *langue* and *langage* in French.

⁵ We do not, for example, first contemplate language from the perspective of neuroanatomy or cellular biology. The subtitle of Sapir's 1921 *Language* is *An introduction of the study of speech*.

⁶ *This sentence have four error.*

A minister, a priest, and a rabbit walk into a bar. The rabbit says "I'm a typo."

⁷ This is, after all, a continuation of Saussure's *sign*, constructed by the *mutual implication*

But how does one go about discovering, understanding, and then expressing the constancy that we recognize as a language? There can be several ways to approach this task, and they all must initially assume some “constancy that we recognize as a language.” Lacking that initial experience, there will be nothing to wonder about and nothing to say. The project recognizes its target phenomenon, and the initial engagement with it may yield some provisional answer(s).⁸ The process avoids circularity by constantly adding to a growing pool of content and by constantly evaluating (and increasing) the degree to which our understanding of that content is mutually supportive and satisfactory. The “bad” answers/ideas are discarded.

The following are three questions which provide a way to instrumentalize the task of responding to the question *What is language?*⁹ That is, they provide a practical guide for us to work through this mess.

First, what delimits language from everything else?

Second, where in the universe does this language reside?

Third, what is language like internally?

Each of the three has been pursued at some point in the history of linguistic practice, but no attempt(s) in any of the paths has yet to yield more than some partial success. In the following paragraphs, we summarize a selection of those efforts.

The First. Hockett’s (1960) *Origin of Speech* is an attempt at delimitation. Thirteen properties of communicative systems are identified: (1) vocal-auditory channel, (2) broadcast transmission and directional reception, (3) rapid fading, (4) interchangeability, (5) total feedback, (6) specialization, (7) semanticity, (8) arbitrariness, (9) discreteness, (10) displacement, (11) productivity, (12) traditional transmission, and (13) duality of patterning. The

between a *signifier* and a *signified*.

E.g.,

... most researchers distinguish between at least two major stages just as in the processing of individual words, one involving the recovery of the phonological information (sound structure) and the other involving access to lexical-semantic information (Hickok 2009.122).

⁸ Say, *naïve*. Over time, I have come to think that our practice should be modeled on that which structured the guild system. One begins as an apprentice, moves to being a yeoman, and finally becomes a master. One learns by watching the master, giving it a try and then improving. There are no rules, only practice, and there is at the end only some degree of connoisseurship.

⁹ There may well be others. but these will serve in the beginning.

list as a whole is intended to be sufficient to isolate human language,

There is solid empirical justification for the belief that all the languages of the world share everyone of them ... It should also be noted that the listing does not attempt to include all the features that might be discovered in the communicative behavior of this or that species, but only those that are clearly important for language.

(89, 91)

but the members on the list are not equally crucial. Thus, one might omit the first, the vocal-auditory channel, and still succeed in the delimitation.¹⁰ While mostly empirically accurate, the emphasis is on delimitation, and we are left with a view of language as a black box. We may have been instructed as to what language is not, and hence how to recognize its presence. We know that it is, but we are uninformed as to what it is.

Hauser, Chomsky & Fitch (2002) return to the same issue of distinguishing human language from non-human communication in this way (1571):

... we take as uncontroversial the existence of some biological capacity of humans that allows us (and not, for example, chimpanzees) to readily master any human language without explicit instruction.

While the focus remains on the distinction between human language and everything else, the answer is not now in terms of language itself, but on whatever it is that permits language.

In the varieties of modern linguistics that concern us here, the term “language” is used quite differently [from “informal usage”] to refer to an internal component of the mind/brain

A distinction is then drawn between a *Faculty of Language — Broad Sense (FLB)* and a *Faculty of Language — Narrow Sense (FLN)* (Hauser et al. 2002.1570 et passim) (Hauser et al. 2002.1571):

FLN is the abstract linguistic computational system alone¹¹, independent of the

¹⁰ Of course, the existence of sign languages makes number 1 moot. The fact that language can be implemented in a written form then makes *rapid fading* less important.

¹¹ Accepting language, i.e., FLN, to be a *computational system* is the source of the acceptance of language to be an *unbounded system of hierarchically structured expressions* (Chomsky 2005.11). Recall from above that “Merge ... [is now] the central computational

other systems with which it interacts and interfaces. FLN is a component of FLB, and the mechanisms underlying it are some subset of those underlying FLB.

and (1578)

...most if not all of FLB is shared with other species, whereas FLN may be unique to humans ... this represents a tentative, testable hypothesis in need of further empirical investigation.

The FLN functions as does Hockett's 13 design features in being the singular basis for distinguishing human language from everything else (Hauser et al. 2002.1569),

We hypothesize that FLN only includes recursion and is the only uniquely human component of the faculty of language.

and both have the same black box nature. No one is certain what is within.¹²

The Second. The second tact towards understanding language takes it to have physical properties. It's out there somewhere and within reach of our five senses. The consensus initial suggestion places language anatomically in the human brain, based upon assumptions such as this (Bogen & Bogen 1976.834):

Can a lesion of the cerebrum produce a deficit in language that is far in excess of the concomitant deficit in intelligence? Asked in this way, almost everyone would answer yes.

The neurology of language is made more precise by locating it primarily in the left hemisphere.¹³ Further articulations recognize the duality of patterning (or double articulation) of language. The portion of language that is the

operation" (Jackendoff 2013.591).

The notion that language is an *abstract computational system alone* is not a twenty-first century innovation. Louis Hjelmslev's 1943 (1961) proposed conceptual opposition between *system* and *structure* is among the earlier conjectures of a computational or generative interpretation of language. *System* is the locus of computation.

¹² "The internal architecture of FLN, so conceived, is a topic of much current research and debate" (Hauser et al. 2002.1571).

¹³ Bogen & Bogen (1976.834):

If you have a person who is definitely right-handed, and he has a cerebral lesion that produces a loss of language far out of proportion to the loss of intelligence, the odds are about 50 to 1 that the lesion is in the left hemisphere.

expressive side is identified with the portion of the brain called Broca's area (Tremblay & Dick 2016.61):¹⁴

Paul Broca ..., in 1861, based on observations of brain lesions and associated behavioral consequences, first described the posterior two thirds of the inferior frontal gyrus (IFG) as the seat of the ability to articulate language.

The meaningful side of language is associated with Wernicke's area (or *the Wernicke area*):

Carl Wernicke ..., in 1874, described two patients who had difficulty understanding spoken language, even though their articulation was fluent (Wernicke, 1874/1969). An autopsy conducted by Wernicke on these patients revealed lesions in the superior temporal gyrus, which led Wernicke to conclude that this region was crucial to language comprehension (61).

Yet the field still lacks consistent definition of either region, over 150 years after their initial introduction (63).

Presuming that the aspect of language affected lay at the site of the lesions, the conclusion was to pair the location of brain lesions with the language disfunctions they produced.¹⁵

In the 1940's-1950's, neurological inspection moves from autopsies to active surgical intervention. Most notably, Wilder Penfield and Lamar Roberts's 1959 *Speech and Brain-Mechanisms*.

Further along this path, technological innovations such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) permit observation of actual activation of specific portions of the brain in association with normal, not damaged, language functioning (Bookheimer 2002:152):

Functional brain imaging, particularly activation PET and functional magnetic resonance imaging (fMRI), rely on a very different fundamental approach to

¹⁴ This conclusion is not unambiguous: "... Broca's *area* was hypothesized to support some aspect of syntactic processing" (Hickok 2009.128).

¹⁵ There were the expected disagreements in executing this plan. Are the boundaries discrete or graded? What are the disabilities (i.e., aphasia) involved? Etc. Tremblay & Dick (2016.64, 66) conclude that "we should simply retire the labels ... A more interesting question, we believe, might be: How does the brain accomplish and integrate the various sub-functions that comprise human language, can we parse the network implementing these sub-functions into its constituent components, and can we identify the role specific patches of cortex (or subcortical nuclei or regions) play in the context of the broader system implementing language."

understanding brain organization. These techniques reveal brain areas involved in, though not necessarily essential to, the ongoing performance of a task.

Because the new tools allow a dynamic (not a static, inert) observation of place, the language correlate acquires an equivalent dynamism (“performance”), and the word *processing* becomes prominent:

“The results of this experiment suggest that overall sentence *processing* occurs in regions of the left perisylvian association cortex.” (Stromswald 1996. 452.)

“modern work has identified areas outside of the classical regions that are implicated in language *processing* ... there are cortical and subcortical regions that clearly contribute to normal language *processing* ... Paralleling the development of the hemodynamic imaging techniques was the emergence of multichannel EEG and the advent of MEG, both of which allowed researchers to chart the timecourse of neural events underpinning language *processing*.” (Hickok & Poeppel 2003.5, 6).

“... we analyzed 120 functional neuroimaging studies focusing on semantic *processing*.” (Binder & Desai 2009.2767)

“... Finally, we will discuss the neural basis of higher-order aspects of language *processing* ... The STS [superior temporal sulcus] has emerged as an important site for representing and/or *processing* phonological information ... Functional imaging studies have also implicated posterior middle temporal regions in lexical-semantic *processing* ... Anterior temporal lobe (ATL) regions ... have also been implicated both in lexical-semantic and sentence-level *processing* ... a number of different brain areas and circuits have been implicated in grammatical and sentence-level *processing*” (Hickok 2009.123, 124, 128, 137)

“To summarize this section, semantic *processing* of familiar auditory stimuli activates a distributed set of regions that surround the ventral, anterior, and posterior borders of the perisylvian regions supporting prelexical auditory speech *processing* ... This suggests that the left anterior superior temporal sulcus is involved in multimodal sentence *processing* ... Semantic *processing* of single words extends even further in the anterior, ventral, and posterior directions into the middle and inferior temporal cortex.” (Price 2010.68, 69, 75)

“Some of the imaging studies showing modality-specific activations during language *processing* are summarized in Figure 1.” (Binder & Desai 2011.528)

“A Re-examination of Neural Basis of Language *Processing*” (Dufau et al. 2014)

“Grounding language *processing* on basic neurophysiological principles “ (Friederici & Singer 2015)

In this paradigm, some language task is selected, and while the subject performs, the brain is monitored by PET or fMRI (Bookheimer 2002). The result is a neurological location paired with some (supposed) portion of language (Cabeza & Nyberg 1997.1-2):

The standard method of analyzing PET data involves comparing the pattern of activity associated with a “target” condition with that of a “reference” condition ... The target and reference tasks are designed to differ only in terms of the process of interest. In this way, the neural correlates of this process can be identified by subtracting the pattern of brain activity in the reference condition from that in the target condition. Regions showing a higher level of blood flow during the target task than during the reference task are often called “activations,” and those showing a lower level of activity during the target task than during the reference task are referred to as “deactivations.”

The specific locations where differences are observed are usually expressed as three-dimensional (x, y, z) coordinates in reference to the stereotaxic brain atlas of Talairach and Tournoux (1988). The use of a common metric for localization allows ready comparison of PET results across experiments and across laboratories, and is critical for the objective of mapping the human brain.

Occasionally, the techniques transcend mapping and inform the understanding of language (Jaeger et al. 1996.489, 490) :

It appears that when a linguistic phenomenon is extremely regular, simple, and productive, such as the regular past tense pattern in English, the most efficient way for the mind to deal with it is to store it as a rule of the grammar that can be implemented on-line in a rapid, error-free way. On the other hand, it appears that when the linguistic is inconsistent and contains a number of differing, nonproductive patterns, such as the irregular past tense patterns in English, the most efficient way for the mind to deal with it is to simply store the various forms and index their relationships to each other (in this case semantic, morphological, and phonological) as part of the lexicon ... we feel that our results demonstrate the value of PET methodology for assessing both linguistic theories and processing models.

The work introduced in the preceding paragraphs represents one of two threads united in name only as *biolinguistics*. They are the ii) in this formulation of Martins & Boeckx (2016.25):

The study of the biological foundations of language is sometimes called “biolinguistics”, a term that has gained considerable traction in recent years. Those who agree that something like a language faculty exists are inclined to use it as a means of emphasizing that their object of inquiry is ultimately a biological one. There are, however, two different factions that prominently adopted the term:

- i) those who use it as a rebranding of theoretical linguistics of the generativist persuasion
- ii) those who use it as a departure from the common practices of theoretical linguistics, firmly oriented towards biology.

The attitude in i) returns to FLN introduced in *The First* above (Chomsky 2005.2):¹⁶

[a] ... more basic question from the biological point of view is how much of language can be given a principled explanation, whether or not homologous elements can be found in other domains or organisms. The effort to sharpen these questions and to investigate them for language has come to be called the “Minimalist Program” in recent years, but the questions arise for any biological system and are independent of theoretical persuasion, in linguistics and elsewhere.

The Third. There are several notable threads that have addressed the character of language from the perspective of its internal substance. I shall discuss three of the more interesting: Information Structure, the work of Sandra Thompson, and that of Paul Hopper.

The approach now labeled Information Structure and, in its modern form, generally attributed to Wallace Chafe (1976) was discussed in Chapters 13 & 25, along with the typologies of FOCUS and TOPIC. It is the initial orientation to the problem that most reliably connects the practitioners of Information Structure, who do not necessarily agree in their conclusions. The criticisms provided, then, do not necessarily apply to all of the work in Information Structure. My strongest disapproval derives from¹⁷

- (1) The beginning point.
- (2) The selection of substances.
- (3) The ending point.

The *beginning point* (1) is to concentrate on a selection of semantics — those noted in Chafe’s lengthy 1976 title *Givenness, Contrastiveness, Definiteness, Subjects, Topics, and Point of View*. The orientation is properly away from the formal, computational model and toward language in its natural habitat — talk. Having begun in this manner, the discussion turns to one of the

¹⁶ Cf. also Boeckx & Grohmann 2007 and Bickerton 2016.

¹⁷ There may be more.

identified issues, say, Topic and Comment (or Focus).¹⁸ It is the way in which these are introduced that seems bothersome.¹⁹ The practice commonly is first to prescribe (or *define*) the semantics of Topic and Comment (or Focus). In place of being drawn as the last step in the investigation, their nature is taken as something that can be known initially. The direction takes what should be the conclusion and sets it prior to empirical inquiry. In this way, some conceptualization is proposed a priori for Topic and Comment, and then the actual investigation is to discover the degree of fit between the theoretical frame and language. In this way, Topic and Comment (or Focus) are frequently seen as complements, so that, taken together, they exhaust the substance of an utterance with no remainder. Gundel & Fretheim (2009.176):

Unless otherwise noted, we use the term FOCUS in this paper to refer roughly to the function described in Chao's notion of logical predicate, and we use the term TOPIC to refer to the complement of focus. Topic is that the sentence²⁰ is about, focus is what is predicated about the topic.

In addition to this tactic ending with an impoverished understanding of Topic and Comment/Focus, it is otherwise faulty.²¹ The languages examined

¹⁸ I will discuss only one here.

¹⁹ Following Deborah Schiffrin (1994), Paul Hopper (2004.239) recognizes this approach as the *etic* one:

Etic approaches begin with an assumption about a grammatical construction, generally identified first by its compatibility with one of a list of putative discourse functions, and invite the linguist to study how it is exemplified and used ... Etic approaches are top-down approaches.

The complementary

Emic analyses are bottom-up [I.e., they begin with supposed data and try to understand that. Cf. section 2.0 below. PWD]. They aim to infer the speaker's own repertoire of constructions by examining bodies of text.

²⁰ Notice that *the sentence* is taken as given. It is never questioned.

²¹ Gundel & Fretheim (2009.176, 177) distinguish between two kinds of given — new: the *referential* and the *relational*:

Referential givenness-newness involves a relation between a linguistic expression and a corresponding non-linguistic entity in the speaker/hearer's mind, the discourse (model), or some real or possible world, depending on where the referents or corresponding meanings of these linguistic expressions are assumed to reside ...

previously here, show that TOPIC and FOCUS are semantically independent (though obviously related). They are neither semantically nor grammatically complements.

The *selection of substances* (2) is equally troublesome. Givenness, for example, is a presence (or absence, i.e., New) that holds other substance together. That is, while there may be some grammar in a language that can be reasonably taken to manifest Topic, there is no grammar in a language that segregates/isolates Givenness. There is no Given in and of itself, just for the sake of being Given. It is always an aspect of something else, a repeated, distributed ingredient. E.g. Topics can be/are Given; Subjects (if we were to acknowledge them) can be Given, Pronouns can be Given, etc. Furthermore, across languages, the semantic task accomplished by Given will be in some language be done by other semantics.²² Probably, the most reliably, consistently Given are the Speech Act Participants *I* and *you*.²³

The *ending point* (3). While not expecting an answer to the question *What is language?*, nor even that it be addressed, Information Structure seems to propose a list of semantics that may be discoverable in discourse, but which seem otherwise disconnected. The language properties cited in Table 1 below cohere, and when taken together, their grammatical expression will very nearly exhaust the grammatical resources of the language.

In addition to Information Structure, a second thread in *The Third*, is represented in the work of Sandra Thompson (Hopper & Thompson 1980 and 1984, Li & Thompson 1976, Matthiessen & Thompson 1988, Ono &

Relational givenness-newness, in contrast, involves a portion of the semantic/conceptual representation of a sentence into two complementary parts X and Y, where X is what the sentence is about (the logical/psychological subject) and Y is what is predicated about X (the logical/psychological predicate). X is given in relation to Y ... Y is new in relation to X in that it is new information that is asserted, questioned, etc. about X ...

Topic and focus, as we use these terms here, are thus relationally given and new, respectively.

If the two categories are defined mutually (*relationally*, but also circularly — each is what the other is not) how is it we know which is which? Why is Y not the Topic? Introduction of “referential givenness-newness” may break the circle. And it is, surreptitiously perhaps, by saying “Y is *new* in relation to X in that it is *new* information ... about X.” The only way this avoids being meaningless is to take *new* as *referentially new*.

²² For example, if Givenness-Newness be taken as the basis of DETERMINACY, then it can have at best only partial success in that role. Cf. for example the discussion of Bella Coola DETERMINACY in Chapter 37 and of Yogad *yu* and *tu* in Chapter 38 and in Davis et al. 1998, section 3.3. Any sense of *given* - *new* is essentially irrelevant to these languages.

²³ Other pronominal persons and numbers are less tied to Given: *We'll see*, *You all come back!*, *He who laughs last, laughs best*, *They always say that*, etc.

Thompson 1995, Ford, Fox & Thompson 2013, Thompson 2019, and Thompson & Couper-Kuhlen 2005). The internal substance that is central in this approach derives from the transactional use to which language is put (Thompson 2019.255):

... humans do not have the goal of building grammar as they go about their daily interactions. Rather, they are interacting to meet their basic needs for food, shelter, sex, and social contact and manipulation, and grammar emerges as they use and re-use language to meet these needs.

The pivotal term seems to be *grammar*:

At the heart of ‘linguistic structure’ is what linguists call ‘grammar’, by which we mean regular patterns at the level of sounds, words, and larger units such as phrases, clauses, and sentences. (Thompson & Couper-Kuhlen 2005.482).

and *grammar* centers about *clause*:

Linguists generally assume ‘clause’ to be a basic unit in the analysis of grammatical structure (Thompson 2019.254)

and *clauses*:

... the distinction between predicating and non-predicating elements is one which “reflects [a] universal distinction that every language makes”. In this paper, I adopt this distinction: I take ‘clauses’ as being built around a ‘predicate’²⁴ (Thompson 2019.257-258)

Although *clause* is ill-defined grammatically²⁵, it remains central to the discussion because

²⁴ Unfortunately, having introduced *clause* — a term of expression — as central to the discussion, we now pursue the issue of just what a *clause* is, do all languages have them, in what way, etc. The answers are provided with reference to form: “NP arguments are routinely expressed” (258), “an ‘elaborated predicate’ ... where NP arguments are typically inferred” (258), “we could speak of ‘clauses’ as typically consisting entirely of predicates” (259).

²⁵ *Clauses* as natural entities or *units* exist only as artificial constructs (Thompson 2019.255):

... as linguists, we also need ways of talking about the phenomena we study ... and it is in this sense than an emergent structure such as ‘clause’ can be viewed as a ‘unit’.

Clauses are in fact not entities but are emergent in the sense of Hopper (cf. below).

clauses are the primary way in which humans do social actions with talk, such actions as requesting, assessing, inviting, informing, assessing[?], etc. (260-261)

... we cannot expect to understand the patterning in the way grammar works in any language unless we understand its profoundly social underpinnings,²⁶ and the organizational patterns of interactional encounters ... [and] the primary vehicle for carrying social action is the predicate, which, in languages like English, together with expressed or inferred arguments, forms a clause. (260)

The emphasis on *social action* is complementary with the semantics that we have concentrated upon (Table 1), and it is unlikely that study restricted to *interactional encounters* will/could discover those semantics and their grammar.²⁷

Perhaps the best way to begin a consideration of Paul Hopper's proposal of *emergent grammar* is to reconsider this century-old passage from Edward Sapir's *Language* (1921.21):

The habitual association of radical elements, grammatical elements, words, and sentences with concepts or groups of concepts related into wholes is the fact itself of language. It is important to note that there is in all languages a certain randomness of association. Thus, the idea of "hide" may be also expressed by the word "conceal," the notion of "three times" also by "thrice." The multiple expression of a single concept is universally felt as a source of linguistic strength and variety, not as a needless extravagance. More irksome is a random correspondence between idea and linguistic expression in the field of abstract and relational concepts, particularly when the concept is embodied in a grammatical element. Thus, the randomness of the expression of plurality in such words as books, oxen, sheep, and geese is felt to be rather more, I fancy, an unavoidable and traditional predicament than a welcome luxuriance. It is obvious that a language cannot go beyond a certain point in this randomness. Many languages go incredibly far in this respect, it is true, but linguistic history shows conclusively that sooner or later the less frequently occurring associations are ironed out at the expense of the more vital ones. In other words, all languages have an inherent tendency to economy of expression. Were this tendency entirely inoperative, there would be no grammar. The fact of grammar, a universal trait of

²⁶ This is a bit stronger statement than the earlier 2005 assertion (Thompson & Couper-Kuhlen 2005.481):

... the study of interaction can contribute to an understanding of linguistic 'structure'

²⁷ Such a study might ultimately, somehow end at this point, but the route would certainly be circuitous. I think that the criticism, just above, of *The Ending Point* of Information Structure applies here. The semantics of social interaction, though certainly present in language, are not the meat of the morphosyntax of language.

language, is simply a generalized expression of the feeling that analogous concepts and relations are most conveniently symbolized in analogous forms. Were a language ever completely “grammatical,” it would be a perfect engine of conceptual expression. Unfortunately, or luckily, no language is tyrannically consistent. All grammars leak.

Grammar, here, is not the grammar that is the product of the linguist’s work, but exists priorly as the object of his/her study. The nature of the *leak* lies in the dimension of *generalized expression* and *analogous forms*. Thus, the degree to which a *randomness* is present, that is, the degree to which *economy of expression* is depressed is a measure of the degree of *leak*. Irregular, idiosyncratic forms are certainly present in language, and their existence is what prevents language from being *a perfect engine of conceptual expression*, and thus to *leak*.²⁸ In contrast, the problem we seem to have encountered is that language has not yet achieved the status of *engine*, perfect or imperfect.

Paul Hopper (1987, 1988, 1995, 1998, 2002, 2004 and Bybee & Hopper 2001) has, I think, a similar perspective on this condition, which he expresses as *emergent grammar*, yet there are significant differences between Hopper’s conclusions and the ones drawn here. Hopper’s (1987.141) statement is:

I believe ... [that] grammar, which like speech itself must be viewed as a real-time, social phenomenon, and therefore is temporal; its structure is always deferred, always in a process but never arriving, and therefore emergent

(Hopper 2002.3-4):

The term “emergent” is to be sharply distinguished from “emerging”, though the two are often confused. “Emerging” means “in the course of development toward completion”; “emergent” suggests a perpetual process in which movement *toward* a complete structure of some kind is constant but where completion is constantly deferred. Linguistic structure is intrinsically incomplete, a work in progress, a site under construction.

(Hopper 2004.256):

It [the English pseudocleft] never achieves closure as a fixed schema. It is, to repeat Haj Ross’s formulation, a construction that doesn’t quite make it. I suggest that we will always encounter this openness and structural indeterminacy when we

²⁸ Note the Constructionists’ acknowledgement of this condition some decades later by their introduction of the terms *core* and *periphery*. Some discussion can be found in Chapter 36, section 4.2.

examine constructions one by one from a discourse perspective. For, like politics, all grammar is local.

All this appears to suggest that even while *emergent*, there is still a *structure* or *schema* that may be contemplated and anticipated, although *always deferred* and *never arriving*. As for Sapir, language here is never *a perfect engine of conceptual expression*. Sapir and Hopper appear to share a common attitude in their emergent and/or leaking grammars.

It is the absence expressed by *perpetual process* and *constantly deferred* that the view presented in this chapter shares with Hopper's ... with this difference: There is no *complete structure*, no *fixed schema* coming. It is our history that prompts us to want/need/expect its presence, but the wholeness and completeness suggested by the use of *structure* does not exist and hence, neither can *emergent* exist.²⁹ Because there is no "complete structure" in fact, nor in principle, there can be no "movement toward" it. There is nothing out there to be *emergent*.³⁰ "All grammar is [indeed] local",³¹ as Hopper proposes, and because of this there is no contrast between *fragment* and *totality*. It is all the same, *fragment* or *totality* as one wants.

This conclusion creates a conundrum. If all there is, is the *local*, there must be some additional dimension to house what we expect to be language. It cannot be all here and all now. Hopper (2004.239) proposes:

Previously heard utterances [habitually combined ... fragments], rather than grammatical rules, form the basis of current and future utterances.

²⁹ Hopper's nihilism is somewhat undercut by this assumption (Hopper 2004.239):

The following are some of the assumptions that underlie the present paper. A language is an inventory of constructions that its speakers use.

Inventory of constructions is too close for comfort to Bloomfield's *totality of utterances* and Chomsky's *set (finite or infinite) of sentences* and Goldberg's *network of constructions*. But now the backing off begins:

In casual spoken discourse constructions appear not as neatly bounded sentences or clauses but as unstructured fragments that are habitually combined with other fragments to make utterances.

Fragments are the essence of language, but *fragments* are *fragments* only if there is a *totality*, a *construction*, to contrast them with. Without the contrast, they are indifferently *fragment* or *totality*. It is here that *emergent* emerges.

³⁰ We are waiting for Godot.

³¹ Hence, static and without *process*, *movement*, or *deference*.

We no longer have *language*. We have a *basis*, a memory of *previously heard utterances*, an “inventory of constructions that speakers use” (Hopper 2004.239).³²

The **first** difference between Hopper’s attitude and the one here was the acceptance of a *complete structure* or *construction* (although one constantly deferred), thus creating emergent grammar and the contrast between a *totality* and *fragments*. The *basis* that Hopper suggests in the grammar of *previously heard utterances* is replaced here by the semantics of Table 1. And that is the **second** contrast. The essence/basis of language is meaning, not form (structure, form, construction, clause, etc.).

I believe that I do not know, cannot define or describe, any entity that is Kutenai (or any of the remaining 7,116 languages) because such an entity does not exist. There is no *thing* that is Kutenai.³³ Nor is there a *thing* that is *language*. Analogous to the discussion of *proposition* in Chapter 36, the

ASSERTION

FOCUS

TOPIC

VOICE, EVENT-PARTICIPANT ROLES & PROPOSITIONAL ROLES

EVENT and PARTICIPANTS

DETERMINACY, i.e., the relation of PARTICIPANTS to our experience independently of the first six above.

Table 1: *Some Semantics of Language.*

impression that there exists an encapsulation uniting the semantics of Table 1 (and certainly other additional semantics) into *propositions* and further into a *totality* (Sapir 1930) or *set* or *something* that is *language* is a product of the fact that those semantics work ... because evolution has fashioned them perfectly to the need.³⁴

They are the evolutionary result of human intelligence having been placed into the flow of experience. Each of these semantics has its own evolutionary motivation and history. Having that much in common, they remain independent shards, cobbled together, entangled, and cooperating in the

³² For Thompson, the *basis* exists in the “routinization of social actions in everyday interactions ... entities ‘doing their own thing’ in real time” (Thompson 2019.255).

³³ Nor verb.

³⁴ In the manner of Voltaire’s pre-Darwinian Dr. Pangloss (*Candide*), *This is the best of all possible worlds.*

successful nesting of human intelligence into its context.³⁵

In this effort, I have assumed the semantics repeated in Table 1. Examining the current population of 56 languages one by one (to some arbitrary degree), the goal was to discover whether each of the semantics of Table 1 was recognizable in each of the languages, and how, both semantically and grammatically. The answer *yes* or *no* was supported by what appeared to be a convergence. Although there were initial suspicions and expectations, there were no arbitrary initial definitions. Patterns attributable to the semantics of Table 1 were discovered (repeatedly), not imposed, and the result was itself not a definition, but an understanding that everything in Table 1 has some sensible meaning and that there are reasonable interactions among them. Without providing some absolute number, it remains true that the vast bulk of a language's grammatical resources are devoted to the expression of these (Table 1) semantics. There is no grammar but this grammar.³⁶

In the remainder of this chapter,³⁷ I will consider some of the implications of this suggestion, which I believe is not negative, just not the expected one. As usual, I begin by considering a piece of one language, paying attention to my interaction with the speaker and to the outcome of that experience. That is section 1.0. Section 1.0 provides some basis for a consideration in Section 2.0 of the implications of *data*. Section 3.0 returns to *language* as a way of

³⁵ Hopper (2004.256) suggests that

prefabricated [?] fragments of discourse [the expressive correlate to our shards, PWD] are acquired for their usefulness in managing effective discourse [what would language be if not this? PWD]

but

planned and written modes have normativized these bits and pieces into longer and more rule-governed syntactic constructions

Note that this stands somewhat at odds to a statement elsewhere, which downplays the presence of *rules*:

Previously heard utterances [habitually combined ... fragments], rather than grammatical rules, form the basis of current and future utterances.

³⁶ That is certainly too strongly stated. But while drawing things together, the claim simultaneously permits elaboration. For example, PARTICIPANTS have multiple parameters. They may or may not exist without explicit grammatical marking, e.g. the Bella Coola prefixes *ti-*, *ta-*, etc. (Chapter 37 and Davis & Saunders 1997a). They may exist by degrees as content moves between EVENT and PARTICIPANT, e.g., Unbounded (*in school* vs. *in the/a school*)— Mass (*air*) — Collective (*faculty* vs. *faculties*) — Count (*wife*). They may be complicated by Gender (Romance), Classes (Bantu), Classifiers (Mandarin), etc.

³⁷ ... which is a revised version of Davis 1995b.

understanding those *data*, and section 4.0 finally ends (or stops) the work.

1.0 Introduction: On Ilokano

Ilokano is a Western Austronesian language spoken in the northwestern portion of Luzon in the Philippines.³⁸ It is typical of the Philippine languages in that Ilokano is verb initial and endowed with a complex system of VOICE. The Agent ROLE is implemented with several VOICE affixes which are appropriate to that ROLE alone (*nag-*, *mai-*, *=um=*, *ag-*, and *mai-*); the same is true of the Patient and Instrument ROLES. The purpose of the introduction is to address the puzzle presented by the multiplicity of AGENT affixes, to produce some statement which seems to resolve the puzzle, and then to reflect self-consciously upon the process which has led to that statement and its implication for our conceptualization of language.

What I call *Ilokano* are conclusions drawn from content recorded in conversation with Dr. Layus. In discussing Ilokano, I appeal not only to the Ilokano utterances and their glosses, but also to possible circumstances to which the utterances might or might not be suitable. My exchange with the speaker of Ilokano consisted of asking how some English utterance might be said in Ilokano and of whether some Ilokano utterance which I composed might actually be used by a speaker of Ilokano. After these issues were settled to our satisfaction, I would ask whether some context which the Ilokano utterance (and its gloss) might suggest would be appropriate. In this, I did not ask the speaker to introspect and to tell me what the utterance meant, but to react to the utterance in some context. Frequently, several such contexts were collected for each Ilokano utterance, especially when two Ilokano utterances shared an English gloss. Questions of concerning the match of utterances to circumstances are frequently much easier for a speaker to respond to than are questions which require settling upon some English equivalent for an Ilokano utterance, and the speaker will frequently volunteer contrasting contexts. Cf. (10) and (11) below, concerning bulls and bus riding, respectively. Responses of this sort are 'robust' in that they are easily verifiable and analogous responses will be obtained for analogous contexts. When terms such as 'meaning', 'semantics', etc. appear below, they refer to the English glosses

³⁸ The variety of Ilokano represented here is that spoken in the state of Ilokos Norte. I would like to thank the Rev. Dominador Layus, from Batác (Ilokos Norte), for his patient help in providing the information which is the basis of the description given here. The spelling of the examples which is used in this paper is a phonemic one (except where an English word intrudes), and will differ in certain respects from the popular orthography. The 'equals' mark (=) is used to surround an infix. AF is the traditional abbreviation for Agent Voice, and PF (below) labels Patient Voice. Those terms are retained here to label the appropriate affixes. The forms *ti*, *iti*, and *ni*, which appear in these examples, are determiners. For a description of their semantics, see Baker 1991 and Baker 1994.

and also to the penumbra of contextual content which is tapped by the Ilokano sentences and the relation which the Ilokano sentences have to these contexts.

An alternative to this mode of inquiry might be a more passive one which simply enticed the speaker to speak (somehow) so that I could experience Ilokano without interfering with its production. I would just witness and record. The motivation for such an alternative is that “elicited” responses are somehow tainted by the process of elicitation and are not “real” Ilokano.³⁹ Certainly, unelicited pieces of Ilokano are relevant, but a restriction of the inquiry to only such examples, for the sake of some kind of purity, strikes me as unnecessary.⁴⁰ Whatever conclusions one reaches, based on interactions with Ilokano, from whatever source will ... if matters are done well ... ultimately cohere. If not, then it is probably not the Ilokano that is the problem. It is the linguist. The more information representing the language that one can accumulate, the better.

Turning now to Ilokano itself, sentences (1) and (2) demonstrate that semantics other than AGENT ROLE is also present in the VOICE affixes which are our concern:

- (1) (a) Ag-bú'uy ti ubíŋ ti táwa
 [AF-break child window]
 'The child is about to break the window'
- (b) Ag-bú'uy ti táwa
 [AF-break window]
 'The window might crack'
- (2) (a) B=um=ú'uy ti ubíŋ ti táwa
 [break=AF=break child window]
 'The child might break the window'
- (b) B=um=ú'uy ti táwa
 [break=AF=break window]
 'The window will crack'

³⁹ Such a limitation, or even privileging, connotes that the linguist is claiming to know what is “correct” or even “true”. “This is language; that is not.”

⁴⁰ It recalls the once-upon-a-time prohibition of *mixing levels* in phonology. I recall a plenary lecture given at the 1960 LSA summer school in Austin, Texas. I believe the speaker was Bob Stockwell, and his lax attitude toward such mixing was too much for George Trager, who was in the audience. Prof. Trager stood and left the lecture, striding determinedly for the exit.

The breaking modulated by *ag-* in (1) may be attributed to some source outside the object that breaks. It is an external circumstance, e.g., the heat of a summer's afternoon, which is the cause of the breaking. In (2), there is a similar attribution of cause, and again it is to some circumstance; but now it lies within the object itself which has broken. Here, the window may have been incorrectly tempered, and the cause is carried within the window itself. The semantics of 'internal' and 'external' is repeated in (3):

- (3) (a) Ag-púdut ti útu
 [AF-heat auto]
 'The car overheats'
- (b) P=um=údut ti útu
 [heat=AF=heat auto]
 'The car heats'

In (3a) the heating is not normal; there is an outside influence to the heating of the car, e.g., a broken hose, and the result is overheating. But in (3b) the process is done by the car as it warms up through use; the source is internal and the result is the normal heating. A similar pair is found in (4):

- (4) (a) Ag-maéstru ni Hwan
 [AF-teacher Juan]
 'Juan is becoming a teacher'
- (b) M=um=aéstru ni Hwan
 [teacher=AF=teacher Juan]
 'Juan is coming to look like a teacher'

In (4a), Juan may never have tried teaching, but he has been trained and is qualified as a teacher because of that training. In (4b), Juan was not trained, but he has the ability, which may have become apparent in his performance in the classroom. In the first, the capacity is an overlay upon Juan, and in the latter, it comes from within. An *ag=kankánta* (based on the root *kánta* 'sing') may be a talentless person trained to sing, but a *k=um=akánta* is a talented singer, trained or not.

A second dimension to the difference between *ag-* and *=um=* seems to be 'spontaneity.' It is the *=um=* which denotes the more spontaneous occasion of an EVENT:

- (5) (a) Ag-pápas ni Hwan ti agbyáhe
 [AF-extend Juan travel]
 'Juan is extending his travel'
- (b) P=um=ápas ni Hwan ti agbyáhe
 'Juan is extending his travel'
- (6) (a) Ag-pakní ák
 [AF-bring.to.a.safe.place I]
 'I'll shelter myself in a safe place'
- (b) P=um=akni ak
 'I'll pull over'
- (7) (a) Ag-saɲpét ak ti baláy-ku
 [AF-arrive I house-I]
 'I'm coming home'
- (b) S=um=aɲpét ak ti balay-ku
 'I'm coming home'

and

- (8) (a) Ag-saɲpét ak ti baláy-ku alaséys
 [AF-arrive I house-I at.six]
 'I'm coming home at six'
- (b) S=um=aɲpét ak ti baláy-ku alaséys
 'I'm coming home at six'

In (5a), John may have planned to attend a convention in New York City; and before leaving home, he decided to extend the trip to include Toronto since he knew that he would be close. In (5b), the same extension occurs; but here it happens on the spur of the moment. Realizing that Toronto is so close, John decides to postpone his return in order to take advantage of the opportunity. In (5a), the travel is deliberate, and in (5b), spontaneous. Sentence (6a) implies search for a shelter for some longer term than does (6b), as when a boat takes cover to protect itself from an approaching storm. The latter implies, perhaps, pulling over to the shoulder of the road in order to check a sudden noise in the car. The specific gloss of (6b) is determined by that context, and any

circumstance which occasions an unanticipated movement to a safe place (or getting out of the way) will be appropriate to (6b). In (7a), the arrival is scheduled and normal; it reports a usual activity or an activity at a definite time. But in (7b), the arrival is unusual and noteworthy. Perhaps it reports the arrival of a student who has been away at school. In (8a), the speaker is reporting what time she arrives home; it is the expected time of arrival. But in (8b), the six o'clock hour is not the usual time. And with more stative EVENTS the contrast is between a process (=um=) versus an achieved state.

- (9) (a) Ag-kápuy ni Ben
[AF-weak Ben]
'Ben has gotten weak'
- (b) K=um=ápuy ni Ben
'Ben is getting/becoming weak'

One generalization from these examples is that =um= attributes to the Agent *the motile spark which is the occasion of the EVENT*, and in that capaci-

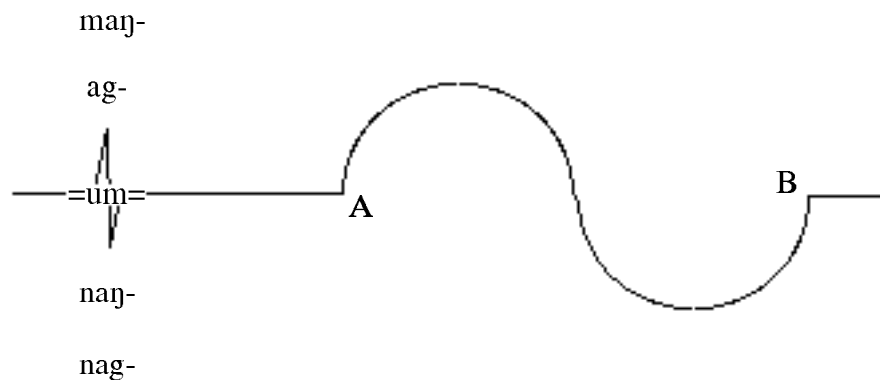


Figure 1: AGENT VOICE in Ilokano.

ty it is *internal* to the AGENT and not 'externalized' as when *ag-* appears. And the motile impulse is also more *spontaneous*, a true unpredictable spark of initiation. Figure 1 gives expression to this contrast and also to the relation of three other Agent VOICES. The curved line of Figure 1 depicts the history of the EVENT from its discernable inception at **A** to its completion at **B**. Placing =um= within the line before **A** visualizes the position of =um= in the history of the EVENT, a history in which the impulse for the EVENT arises from the AGENT. Notice the iconicity of the infixation of =um=.

Let us now consider the affixes *naŋ-* and *nag-* in more detail, beginning

with the following two utterances from Ilokano:

- (10) (a) Naŋ-kámat ti túru kadakwáda
 [AF-chase bull them]
 'The bull chased them'
- (b) Nag-kámat ti túru kadakwáda
 [AF-chase bull them]
 'The bull chased them'

The English glosses are the same, but the way in which the people are chased differs between the two. In (10a), the bull was not really angry and only halfheartedly chased the intruders from his pasture, perhaps part way to the fence; but in (10b), the bull was truly angry and chased them as far as he could. The formal difference between (10a) and (10b) lies in the prefixes *naŋ-* and *nag-*. Now let us add the following pairs:⁴¹

- (11) (a) Naŋ-tugáw ni Ben ti trak
 [AF-seat Ben bus]
 'Ben sat on the bus'
- (b) Nag-tugáw ni Ben ti trak
 [AF-seat Ben bus]
 'Ben sat on the bus'
- (12) (a) Naŋ-i-sukát ak ti bumbílya
 [AF-I-replace I lightbulb]
 'I replaced the lightbulb'
- (b) Nag-i-sukát ak ti bumbílya
 [AF-I-replace I lightbulb]
 'I replaced the lightbulbs'

To understand the difference between (11a) and (11b), it is necessary to know that in the Philippines, bus tickets may be purchased either in the bus station or from the conductor on the bus. Practically, what happens is that some individuals will occupy a seat on the bus, thus reserving the seat, but with no intention of making the trip. When the bus has filled, they then give up their

⁴¹ The affix *-i-* is glossed as itself, 'I.' It is not relevant to the discussion here, and its place in the VOICE system of Ilokano is detailed in Davis 1991 and in Davis Ms.b.

seat (for payment) to a person who is really intending to travel. Bus tickets are in that way scalped. Sentence (11a) is appropriate to the person who is selling his seat and (11b) for the person is making the trip. It is the prefix *nag-* which describes the intent to complete the trip; *nay-* describes the case in which the person is only a ‘pretend’ traveler. In (12a), one bulb is replaced, but in (12b) more than one is replaced as when one examines a string of Christmas tree lights before installing them. The semantic ‘thoroughness’ and ‘honest intent’ of *nag-* in (10b) and in (11b) is manifest in (12b) in multiple occurrences; and the semantic halfheartedness of (10a) and (11a) is manifest as the one time occurrence of (12a). The theme is played out as ‘partial’ versus ‘whole’ in these pairs:

- (13) (a) *Naŋ-i-bútu* *ti* *hispánik* *ken ni* *Bush* *idi napalábas*
 [AF-I-vote Hispanic for Bush last time]
 ‘The Hispanics voted for Bush last time’
- (b) *Nag-i-bútu* *ti* *hispánik* *ken ni* *Bush* *idi napalábas*
 [AF-I-vote Hispanic for Bush last time]
 ‘The Hispanics voted for Bush last time’

In (13b), the Hispanics voted as a block; but in (13a), although the Hispanics voted for Bush, the vote is split.

The opposition between *nay-* and *nag-* in (10) and (11) is not in terms of percentages or fractions as it is in (12) and (13); it is more continuous and in terms of ‘pretense’ versus ‘ernestness’. The contrast involving the association of ‘ernestness’, ‘completeness’, and ‘totality’ is repeated in these:

- (14) (a) *Naŋ-paténga* *ti* *preparasyún*
 [AF-finish preparation]
 ‘They stopped preparation’
- (b) *Nag-paténga* *ti* *preparasyún*
 [AF-finish preparation]
 ‘They finished preparation’
- (15) (a) *Naŋ-paténga* *ák* *ti* *libru*
 [AF-finish I book]
 ‘I finished the book’
- (b) *Nag-paténga* *ák* *ti* *libru*

[AF-finish I book]
 'I finished the book'

In (14a) the workers have reach the end of a middle phase in the completion of a job. Perhaps in the construction of a building, they have just finished topping off the frame; but in (14b) the building is ready to be occupied. It is totally prepared. In (15a), part of the book has been read, while in (15b) the book is read from cover to cover. Or in (15b), the reader has finished getting from the book what was wanted (even if only part was read): 'I finished with the book'.

In these pairs:

(16) (a) Naŋ-i-tarús ak ti gayyém-ku idyáy hutél
 [AF-I-stay I friend-I this hotel]
 'I put my friend up in this hotel'

(b) Nag-i-tarús ak ti gayyém-ku idyáy hutél
 [AF-I-stay I friend-I this hotel]
 'I used to put my friend up in this hotel'

the speaker of (16a) has put up the friend just this once. But if *nag-* is employed, the sense is that there has been a series of such lodgings. Consistent with (16b), we find that we can say 'I put my friend up in this hotel last Friday' with *naŋ-* (cf. [17a]), but not with *nag-* (cf. [17b]). The less precise *iti vyernes* 'one Friday' (as opposed to *idi vyernes* 'last Friday') is acceptable with *nag-* (cf. [18]):

(17) (a) Naŋ-i-tarús ak ti gayyém-ku idyáy hutél idi byénes
 [AF-I-stay I friend-I this hotel last Friday]
 'I put my friend up in this hotel last Friday'

(b) *Nag-i-tarús ak ti gayyém-ku idyáy hutél idi byénes

(18) Nag-i-tarús ak ti gayyém-ku idyáy hutél iti byénes
 [AF-I-stay I friend-I this hotel one Friday]
 'I put my friend in this hotel one Friday'

And these:

(19) (a) Naŋ-pídut ni Ben ti dúlyar

[AF-pick.up Ben money]
 ‘Ben picked up the money’

- (b) Nag-pídut ni Ben ti dúlyar
 [AF-pick.up Ben money]
 ‘Ben picked up the money’

In (19b), Ben may have sat down at a restaurant table before it had been cleared by the busboy; he found the tip left for the waiter and took it all. In (19a), the circumstance is similar except that Ben only took part of the tip.

In these, the senses of *naj-* and *nag-* differ more markedly from earlier examples:

- (20) (a) Naj-swítik ni Ben ti eksámen
 [AF-cheat Ben exam]
 ‘Ben cheated on this examination’

- (b) Nag-swítik ni Ben ti eksámen
 [AF-cheat Ben exam]
 ‘Ben cheated on some examination’

In (20a), the examination is immediate, but in (20b), it is some examination in the past; the precise one is not an issue. Sentence (20a) is appropriate to say about Ben as we leave the examination room; (20b) is not.

- (21) (a) Naj-i-náyun ak ti asúkar
 [AF-I-add I sugar]
 ‘I added some sugar’

- (b) Nag-i-náyun ak ti asúkar
 [AF-I-add I sugar]
 ‘I added some sugar’

- (22) (a) Naj-pugsú ti radyadúr
 [AF-spew radiator]
 ‘My radiator spewed up’

- (b) Nag-pugsú ti radyadúr
 [AF-spew radiator]
 ‘My radiator spewed up’

In (21b), the speaker added sugar, perhaps to the tea of someone without being noticed and without the knowledge of the person concerned. This implies that the sugar could not have been added to the speaker's own tea in (21b); and, indeed, (21b) is not appropriate when that is so. But in (21a), the addition is noticed, and it can now apply to the speaker's tea as well. In (22b), the sense is that whenever the radiator gets hot, it boils over and puts out a geyser of steam and hot water. But in (22a), the idea is that this happens because the speaker was careless and left the cap loose so that someone was injured. The unacceptability of (23) about the Philippine volcano Pinatubo seems consistent with the acceptability and sense of (22a):

- (23) **Naŋ-betták* *ti* *Pinatubu*
 [AF-burst *Pinatubo*]

Sentence (24), however, with *nag-* is appropriate:

- (24) *Nag-betták* *ti* *Pinatubu* *ket* *nag-i-pugsú*
 [AF-burst *Pinatubo* and AF-I-erupt]
 ‘Pinatubo burst open and erupted’

The sentences (25) - (27) continue in this vein:

- (25) *Naŋ-byág* *ni* *De Bakey* *ti* *pasyénte*
 [AF-life *De Bakey* patient]
 ‘De Bakey gave the patient life’
- (26) **Nag-byág* *ni* *De Bakey* *ti* *pasyénte*
 [AF-life *De Bakey* patient]
- (27) *Nag-byág* *ti* *pasyénte*
 [AF-life patient]
 ‘The patient came to life’

The sense of (25) is that the patient is facing certain death and De Bakey, by performing a bypass operation, saves the person's life. But (26) is not possible; however, (27) is possible in a medio-passive sense, e.g., a patient who has been given up for hopeless by the doctors, but who nevertheless recovers without their aid. In (27) as in (24), *nag-* is appropriate to natural occurrences in which no impulse is imputed, but *naŋ-* imposes such a causing force. And in (23), applied to Pinatubo, such imposition is factually

inappropriate as it is in (28):

- (28) *Naŋ-byag ti pasyénté

In the sense ‘The patient came to life’, (28) is unacceptable.⁴²

Another contrast between *naŋ-* and *nag-* appears in these:

- (29) (a) Naŋ-layús ti karayán
[AF-flood river]
‘The river flooded’
- (b) Nag-layús ti karayán
[AF-flood river]
‘The river flooded’

In (29b), the flood occurred so that no one was affected; no houses or property used were destroyed. But in (29a), there was some effect, say, some agricultural land was covered so that the farmers complained. The effect is not necessarily a negative one; for example, when the annual beneficial flood of the Nile occurs, (29a) could also be used in expressing approval of the event.

With *naŋ-*, when the EVENT is grammatically intransitive, some externalization of cause (e.g. [22a]) or of effect (e.g. [29a]) is implied. When the EVENT is inherently stative and admits no performance, no occurrence of *naŋ-* is possible:

⁴² In another Philippine language, Tagbanwa, spoken on the island of Palawan, the contrast between *nag-* and *naN-* is that *nag-* implies a ‘purposive element’ (Green 1979.70):

- (i) nag-buwat aku nat sakayan mu
[AF-make I now the.boat your]
‘I have made your boat now’

In (i), “the person for whom the boat was made had previously ordered a boat made for him.” The *nag-* both in Tagbanwa and in Ilokano are similar in that they can imply that the impulse for the EVENT originates outside the involved parties. In Ilokano, this sense emerges when there is only one PARTICIPANT, i.e., in intransitive usages, while in Tagbanwa, it pervades the grammatically transitive occurrences as well. The affix *naN-* appears in (ii):

- (ii) nan-leteg aku kagayna it sakayan duʔut laud
[AF-see I a.while.ago a boat there.in.the distance]
‘I saw a boat there in the distance a while ago’

Sentence (ii) “could denote either durative action or the fact that the boat was only one of the things the actor saw as he was distributing his act of looking over a number of things on the water” (Green 1979.74).

- (30) *Naŋ-baknaŋ
[AF-rich]
- (31) *Naŋ-sakit ni Ben
[AF-sick Ben]
- (32) *Naŋ-lúkmeg
[AF-fat]
- (33) *Naŋ-uŋar
[AF-risen]
- (34) *Naŋ-dakkel
[AF-big]

But when the causative *pa-* is also present, then *naŋ-* is again compatible with the stem:

- (35) Naŋ-pa-sakit ni Ben kanyák
[AF-PA-sick Ben me]
'Ben made me sick'
['That's physical', e.g. chicken pox.]
- (36) Naŋ-pa-lukmeg ni Ben ti báka
[AF-PA-fat Ben cow]
'Ben fattened the cow'
- (37) Naŋ-pa-dakkel
[AF-PA-big]
'He let it get big'

We now have several ways in which *nag-* and *naŋ-* may differ from each other (Figure 2). At first glance, the senses associated with the respective affixes in Figure 2 appear to be without pattern, but reference to Figure 1 may permit some clarification. With *nag-*, the spark which initiates and impels the event is at its greatest remove from its origin in the EVENT, i.e., at its greatest distance from =um=. This implies that its effect may have *suffused* throughout the EVENT implying a *repeated* performance or a *completed* performance or a *true* one; but at the same time, this greatest remoteness may be seen to imply

| <i>nag-</i> | <i>naɪ-</i> |
|-------------|--------------|
| ‘complete’ | ‘partial’ |
| ‘repeated’ | ‘once/first’ |
| ‘true’ | ‘pretend’ |
| ‘natural’ | ‘caused’ |
| ‘no effect’ | ‘effect’ |

Figure 2: *More AGENT VOICE in Ilokano.*

the *weakening* of the motile spark and the possibility of natural and *uncaused* occurrences, e.g., weather and geological phenomena. When this remoteness is lessened, the diffuseness is made more *focussed*; but at the same time, because it is more focussed, it is *fragmented* and less pervasive. Thus, we see the *partial* senses, the initiation or *first time* senses as well as the clearer presence of *some operant source* of the EVENT and its effect (its *inapplicability* to spontaneous natural occurrences). Perceived in this way, we may detect a contrast of REMOTE — IMMEDIATE (*nag-* versus *naɪ-*). As we move from the EVENT-internal =*um*= to the externalized localization of the motile origin of the EVENT, we progress first through the more IMMEDIATE realization of that causation until finally we reach the most REMOTE separation of the motile impulse from the EVENT.

The dimension of REMOTE — IMMEDIATE may be implemented in the realis mode, but the same dimension may extended into other modes as well. The Agent affixes *ag-* and *maɪ-* present a contrast similar to *naɪ-* and *nag-* but in the direction of irrealis (upward in Figure 1):

- (38) (a) S=*um*=aksí ak
[witness=AF=witness I]
‘I’ll witness’
- (b) Ag-saksí ak
‘I witness for somebody’
- (c) Maɪ-saksí ak
‘I’ll testify’

In (38a), the infix =*um*= repeats its earlier semantics of ‘internalized source’ in connoting that the witness is testifying voluntarily, while in (38b), the witness has been suppoenaed to appear. It may also be that in (38b) that the

witness is determined to testify. In this case the witness will try any means possible to appear on the stand. In (38c), the witness is merely available to testify. The following triplet with *tíkag* ‘to be dry’ parallels this:

- (39) (a) T=um=iíkag
[be.dry=AF=be.dry]
‘It will be dry’
- (b) Ag-tíkag
[AF-be.dry]
‘It will be dry’
- (c) Maŋ-tíkag
[AF-be.dry]
‘It will be dry’

The prediction of dryness in (39a) is based on the belief that at certain points in the lunar cycle, there will be no rain, while in (39b), there has already been a series of dry days and the conjecture is now, because of that, ‘It will [now] be dry’. That is, summer (the dry season) is here, and the dryness is extended and made more pervasive. Finally, in (39c) the dryness is projected upon the basis of the climatological information rather than from immediate experience, and that declaration would be appropriate to such areas as Arizona. The present dryness is extended still further. As we move from the immediate source of the dry spell in the lunar cycle, to its broader origin in the season for drought and, lastly, to the still more diffuse distribution of dryness in the climate of the geography, we move from =um=, to *ag-* and then to *maŋ-*. A parallel progression occurs with the form *uŋét* ‘to be angry’:

- (40) (a) Um-uŋét ak
[AF-be.angry I]
‘Now I’m angry’
- (b) Ag-uŋét ak
[AF-be.angry I]
‘I’m getting angry’
- (c) Maŋ-uŋét ak
[AF-be.angry I]
‘I’ll get angry’

The prefix *um-* directs attention to the point of transition, the last straw. In (40a), the speaker was not angry and was putting up with the messes the children were making until this last time. It was one time too many, and at that point, anger erupted. In (40b), anger is slowly growing (e.g., a slow burn) as the children ignore commands to be quiet. And in (40c), the speaker arrives home to find a mess, but no culprit. There is no one at whom the anger can be directed and hence no point to it; the anger is postponed until there will be a target for it. The event *sérbi* ‘to serve’ amplifies this progression:

- (41) (a) S=um=erbí ak iti armáda
 [serve=AF=serve I army]
 ‘I’ll serve in the army’
- (b) Ag-serbí ak iti armáda
 [AF-serve I army]
 ‘I’ll serve in the army’
- (c) Maṇ-serbí ak iti armada
 [AF-serve I army]
 ‘I’ll serve in the army’

In (41a), the speaker is volunteering to serve, but with (41b), the speaker is more compelled to enter into service, perhaps from patriotism (in which case the service may be either positively or negatively thought of ... looked forward to or not). In (41c), the speaker will serve in some capacity if not actually as a combatant, perhaps in the capacity of a recruiter ... and if not in this way, then in some other indeterminate way. This indeterminacy is reflected numerically in this pair:

- (42) (a) Ag-lukát ka ti rikép
 [AF-open you closure]
 ‘Open the window shutters’
- (b) Maṇ-lukát ka ti rikép
 [AF-open you closure]
 ‘Open some shutters’
 [‘Maybe one or two ... a few’]

In (42a), all the shutters will be opened, but in (42b), there is an indeterminacy and therefore only some of the shutters will be opened and not all of them. As

a command, then, (42a) is perceived as less polite. Peremptorily, it demands more; (42b) is more polite and more likely to be used to a spouse or parent, while (42a) may be used to a child. In the following examples, number also appears to be a criterion in the distinction between *ag-* and *maŋ-*:

- (43) (a) *Ag-i-pakní* *ak ti maŋgá*
 [AF-I-put.in.safe.place I mango]
 'I'll put the mangoes in a safe place'
- (b) *Maŋ-i-pakní* *ak ti maŋga*
 [AF-I-put.in.safe.place I mango]
 'I'll put the mango in a safe place'

In (43a)/(43b), the contexts are these. There are several mangoes on the table and one of them has begun to get so ripe that it is required to put it into the refrigerator so that it will last, hence, the 'just one'. In (43a), it is *all* mangoes which are put together into the refrigerator. The Patient *maŋga* is only obliquely affected with *maŋ-*, but for its accomplishment, this obliqueness requires that the one mango be set off against the background of a collection of other unaffected ones. That is, *there are others* which are not put away. The plurality is obliquely affected, and that is 'just one'.⁴³ The contrast between absolute and partial recurs in a non-numerical, analog way in this pair:

- (44) (a) *Ag-búsur* *ni Turner ti monorail*
 [AF-oppose Turner monorail]
 'Turner opposes monorail'
- (b) *Maŋ-busur* *ni Turner ti monorail*
 'Turner opposes some aspects of monorail'
 ['Not totally opposed to']

In (44a), Turner is said to oppose monorail completely and repeatedly, i.e., he is campaigning against it. But in (44b), he is only opposed to some aspects of monorail and approves it in principle. The focus of determinacy associated with *ag-* appears in (45b), in which it is the last chapter of the book which is finished:

⁴³ There can be more than one mango affected in (43b), but then they must be in a container so that the unity is preserved. It is the *bag* of mangoes which is put away, while others are not.

- (45) (a) P=um=atenga ák ti kapitulu
[finish=AF=finish I chapter]
'I'll just finish the chapter'
- (b) Ag-patenga ák ti kapitulu madamdamá
[AF-finish I chapter later]
'I'll finish the chapter later'
- (c) Maŋ-patenga ák ti kapitulu madamdamá
'I'll finish the chapter later'

In (45a), the speaker is in the middle of the chapter when called to do some task. In order not to lose the thread, the speaker responds saying that she will just finish the chapter and then be right there. In (45c), it is *some* chapter which is finished, but its indeterminacy is mirrored by its not being the last chapter. This indeterminacy/partiality is extended to include an uncertainty:

- (46) (a) S=um=wi'tik ak ti eksámen
[cheat=AF=cheat I exam]
'I'll cheat on the exam'
- (b) Ag-swi'tik ni Ben ti eksámen
[AF-cheat Ben exam]
'Ben will cheat on the exam'
- (c) Maŋ-swi'tik ni Ben ti eksámen
[AF-cheat Ben exam]
'Ben will cheat on the exam'

In (46a), the cheating arises spontaneously when it is discovered that the exam is too difficult unless the exam taker is dishonest. In (46b) and (46c), the difference turns in part on the knowledge of *how* Ben will cheat. In (46b), the speaker knows the method Ben will use, but in (46c), the speaker does not know his method, only that Ben will cheat in all probability. He is not that close to the occurrence:

- (47) (a) Ag-i-switik ák ti tulágan
[AF-I-cheat I contract]
'I'll cheat on the contract'

- (b) Maṇ-i-switik ák ti tulágan
 ‘I’ll cheat on the contract’

In (47a) and (47b), the difference is manifest as one of imminence and a more remote, relaxed performance, respectively. There seems to be a greater intensity in the use of *ag-*, whether it is realized by ‘determinacy’ (‘all’), by ‘compulsion’ (‘patriotism’ or ‘by hook or by crook’) or by ‘imminence’. The use of *ag-* is also supported by a context in which the activity is preceded by previous occurrences of the same event or contexts in which the present occurrence is another in a series. This contrasts then with *maṇ-*, which appears in a context unsupported by prior occurrences; the sense then can be that of a change in habit or of a one time spontaneous occurrence:

- (48) (a) Ag-i-saṇpét ak ti úbas
 [AF-I-arrive I grape]
 ‘I’ll bring grapes’
- (b) Maṇ-i-saṇpét ak ti úbas
 [AF-I-bring I grape]
 ‘I’m bringing home some grapes’

In (48a), it is part of the routine to bring grapes home, but in (48b), it is not. The speaker may have seen the grapes on sale in the store and impetuously decided to take advantage of the low price. The contrast between *ag-* and *maṇ-* is summarized in Figure 3.

| <i>ag-</i> | <i>maṇ-</i> |
|-------------------------------|-----------------------|
| ‘season’ | ‘climate’ |
| ‘by hook or crook’ | ‘availability’ |
| ‘all/last one’ | ‘some/any’ |
| ‘next stage in a progression’ | ‘postponed’ |
| ‘certainty’ | ‘suspicion’ |
| ‘continuation of series’ | ‘one time occurrence’ |
| ‘less polite’ | ‘more polite’ |

Figure 3: *Even More AGENT VOICE in Ilokano.*

Semantic ‘determinacy’ (*ag-*) in opposition to ‘indeterminacy’ (*maṇ-*) may summarize the contrasts of ‘season’ opposed to ‘climate’, ‘by hook or by

crook' opposed to 'availability', 'all' or 'last (one)' versus 'some/any', 'next stage in a progression' versus 'postponed', 'certainty' versus 'suspicion', 'continuation of a series (habit)' versus 'a one time occurrence'. The 'determinacy' or more focused 'intensity' of *ag-* can also be felt as 'less polite' in comparison with the greater 'indeterminacy' or increasing 'dilution' (lesser 'intensity') of *mai-*. As the spark of Agency is externalized from *=um=* along the parameter of irrealis, it is first 'intense' in a variety of implementations; but then its 'intensity' is diminished as it moves away from its originating source. And that is the dimension defined by *ag-* and *mai-*.

Dimensions of the EVENT: IMMEDIATE — REMOTE & REALIS — IRREALIS.

In each of the occurrences of *ag-* and *mai-*, the specific EVENT has yet to be realized. When *ag-* is glossed as a progressive in English, the reference is to the preceding context which motivates its semantic 'intensity' and not to the initiation of the immediate EVENT, which has yet to be begun. In this series

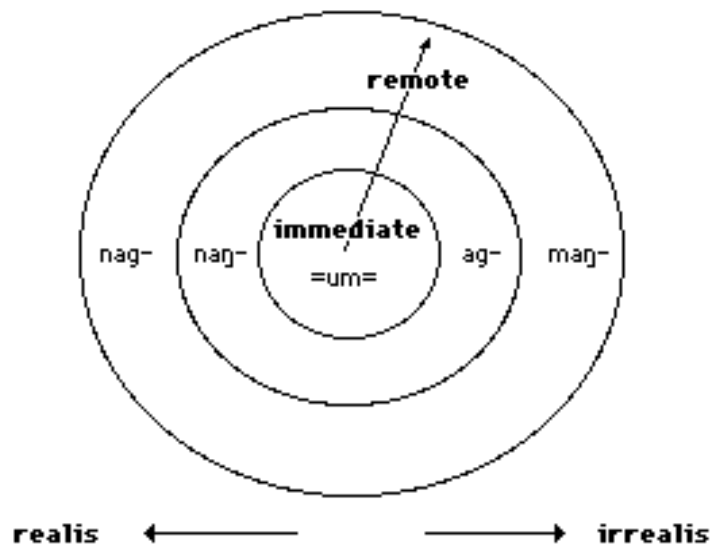


Figure 4: *Summation of AGENT VOICE in Ilokano.*

of *=um=*, *ag-*, and *mai-*, *=um=* again acts as the spring from which the EVENT erupts. As the EVENT is established, it acquires a maturity signalled by *ag-*, which may be the intensity of a pattern (e.g. [48a]), of determination (e.g. [41b]), of a determinate way (e.g. [46b]), of completion (e.g. [45b]), etc. As the cycle of the EVENT continues, the intensity of its mature condition diminishes. Signalled by *mai-*, this is recognized in its being a one time occurrence/out of pattern (e.g. [48b]), in its being done in an indeterminate

way (e.g. [46c] and [41c]), in its being partial/incomplete (e.g. [45c]). Figure 4 combines this as the dimension of IMMEDIATE — REMOTE semantics with the IRREALIS semantics of =um=, ag-, and *mai-* and with the REALIS semantics of the series of =um=, *mai-*, and *nag-*. The infix =um= denotes the spontaneous, momentaneous, emergent quality of the semantic agency. In (38a), (40a), (45a), (46a), for example, the speaker is *volunteering* to witness, has experienced the *straw that broke the camel's back*, will *just* finish reading the chapter, and has *just this moment* discovered that cheating is necessary. These combine with the earlier examples to place =um= at the center, at the cusp of the semantic agency from which the EVENT emerges. One last example of the contrast of =um=:

- (49) (a) S=um=ákit ti ŋípín
 [pain=AF=pain tooth]
 ‘My tooth hurts’
- (b) Ag-sákit ti ŋípín
 [AF-pain tooth]
 ‘The tooth is hurting’

In (49a), there is just the symptom of a bad tooth, but in (49b), there is a prior experience of pain and there is the impression of greater discomfort.

2.0 On ‘Data’.

As linguists, we consider ourselves students of language, and in order to study language, we must know where to find it.⁴⁴ We are not, after all, botanists. But while agreeing that we are not botanists, we still may not agree on where to seek language. Prescriptions (and proscriptions) as to where we should look (or where we are entitled to look) are well known.⁴⁵ We may examine certain noises which some individual (who is thereby a ‘speaker’) produces, i.e. ‘utterances’. The ways we proceed in the inspection can be varied. We may confine our attention to short noises (‘words’), or we may

⁴⁴ The ‘we’ is variably applied here and only means that (probably) somewhere in the field of linguistics someone will hold (or will have held) the position identified. It does not mean that I adhere to all of these beliefs.

⁴⁵ This is one of the early pre- and proscriptions (Bloomfield 1933:32):

... the linguist deals only with the speech-signal (r · · · s); he is not competent to deal with problems of physiology or psychology.

consider longer noises ('sentences'), or we may examine still longer noises which are cooperatively produced by one, or more than one speaker ('discourse'). We may extend our interest along another dimension and bring into consideration behavior which is not noise, or noise producing, but which occurs in conjunction with noise, i.e. 'nonverbal behavior'. And we may add a third dimension and also consider verbal behavior about verbal behavior. This last may include the common 'intuition'. Or it may include responses to such inquiries as 'What did that [the noise the speaker made] mean?' and 'Can I say this [the linguist makes noise] in your language?' Or it may be an extended discussion of some utterance or discourse (piece).

In the discussion of Section 1, we began with the encounter (as linguist) with the person who speaks Ilokano. The whole enterprise was driven by the belief that Dr. Layus knew something which I did not and that that knowledge was what enabled him to act as my partner in this type of conversation. The experience of the encounter was indexed by audio recording and by written record. Then, the written record was examined, and from that examination, questions arose which were discussed in later meetings. A series of summations was produced and these are written in Figures 2 and 3. Although language may only be partially present as a distorted trace, it is somehow present in those data. Language has been taken 'to inscribe its own signature' (Latour & Woolgar 1979:63), however elaborately and however subtly; and our reading of it is manifest in this parsing. When we accept this perspective, there *will* then be better or worse places to look for language because we believe that somewhere in that array of sources (which still may not be exhaustive), language has left its track. It is somehow in there.

The belief which drives the alternative view of language which I suggest here is that *there are no data in the way data are commonly taken*. Latour & Woolgar (1979) describe the activity of a laboratory studying problems in neuroendocrinology. In one sequence, data appear to be first recognizable as a liquid withdrawn from rats with a syringe. But then

... over a period of several days, tubes are arranged in rows, other liquids are added, the mixtures shaken and eventually removed for refrigeration.

Periodically, the routine of manipulation and rearrangement of tubes is interrupted. The samples extracted from rats are put into one of the pieces of apparatus and undergo a radical transformation: instead of modifying or labeling the samples, the machine produces a sheet of figures ... the same tubes which had been carefully handled for a week, which had cost time and effort to the tune of several hundred dollars, were now regarded as worthless. The focus of attention shifted to a sheet of figures. The sheet of figures, taken to be the end result of a long assay, was used as the input to a computer ... After a short time, the computer printed out a data sheet and it was this, rather than the original sheet of figures,

which was regarded as the important product of the operation ... a technician work[ed] on several data sheets produced by the computer ... [to produce] a single elegant curve carefully drawn on graph paper. Once again, the focus of attention shifted ... and it was the peaks and slopes of the curve which excited comment from the participants in their offices.

The whole series of transformations, between the rats from which samples are initially extracted and the curve which finally appears in publications, involves an enormous quantity of sophisticated apparatus ... It is clear, then, that particular significance can be attached to the operation of apparatus which provides some kind of written output ... items of apparatus, which we shall call 'inscription devices,' transform pieces of matter into written documents ...

An important consequence of this notion of inscription device is that inscriptions are regarded as having a direct relationship to the original substance. The final diagram or curve thus provides the focus of discussion about properties of the substance ... The process of writing articles about the substance thus takes the end diagram as a starting point. (Latour & Woolgar 1979.49-51).

The term 'data' appears in Latour and Woolgar's description only at the point at which the computer has processed figures fed into it, for this is the first point in the series of transformations at which *a human searches for pattern*. But this operation is immaterial for the concept of data since all stages are functionally equivalent; all inscriptions are equivalent in that they have "a direct relationship to the original substance". The content of the final curve is there somehow in the original liquid taken from the rat (if read correctly). The point of relevance to us as linguists is not whether we lack inscription devices of the sort available to the neuroendocrinologist, but that the relationship between whatever we may call data (perhaps 'fieldnotes' or the 'transcription' of an audio or video recording⁴⁶) and the concept of language is *not direct* (in the sense of Latour and Woolgar), and could not be, even if the relevant inscription devices existed. *Nothing* which we manipulate in our work bears "a direct relationship to the original substance".

Language exists somewhat in the manner of the rainbow in Owen Barfield's discussion of western consciousness (Barfield 1965.15):

Look at a rainbow. While it lasts, it is, or appears to be, a great arc of many colours occupying a position out there in space. It touches the horizon between that chimney and that tree; a line drawn from the sun behind you and passing through your head would pierce the centre of the circle of which it is part. And now, before it fades, recollect all you have ever been told about the rainbow and

⁴⁶ Givón (1991.88-89), in describing linguistic activity, identifies a sequence of transformations similar to that observed in the neuroendocrinologists' laboratory, in which oral 'descriptions of what they [each speaker] saw' become 'pause-marked texts,' which 'constituted the input to the quantitative analysis.'

its causes, and ask your self the question *Is it really there?*

In another context, Michael Polanyi (Polanyi & Prosch 1975.29) responds:

Take any question to which you want to know the answer. For example, having planted some primroses today, you would like to know whether they will bear blossoms next spring. This question is not answered by a list of atomic positions and velocities at some future moment on May 1 of next year. *Primroses* [and rainbows], as such, are lost in the topography of *all* the atoms. Your question can be answered only in terms of primroses.

The rainbow, the primroses, and language are not there in themselves, but in our interaction with some aspect of our experience. Language escapes us when we confine ourselves to isolated words (e.g., Saussure), to isolated sentences (e.g., much of American structuralism including transformational grammar and GB), or even to texts of whatever length and complexity (e.g., many functional approaches, text generation, discourse analysis, etc.). Language does not exist as a detached and objective entity. This may be a familiar assertion to describe the circumstance in which we ourselves are speakers of a language, who ‘negotiate’ our way through what is then recognized post hoc as a conversation. But the same assertion that language is not objectively presented for examination is also true for the linguist, who is trying to grasp what coherence may be present there.

The problem is *not* that language is there and that we cannot see it. The difficulty is that *we create language by looking for it*. The fact the linguist’s knowledge of language originates in that interaction imbues language with a personal quality which Polanyi calls *tacit knowledge* (Polanyi & Prosch 1975.34, 35 and 38):

Consider the act of viewing a pair of stereoscopic pictures in the usual way, with one eye on each of the pictures. Their joint images might be regarded as a whole, composed of the two pictures as parts. But we can get closer to understanding what is going on here if we note that, when looking through a stereo viewer, we see a stereo image at the focus of our attention and are also aware of the two stereo pictures in some peculiar nonfocal way. We seem to look at them only as guides to the image on which we focus our attention. We can describe this relationship of the two pictures to the stereo image by saying that the two pictures function as *subsidiaries* to our seeing their *joint* image, which is their joint meaning. This is the typical structure of tacit knowing ... A characteristic aspect of from-to [tacit] knowledge is exemplified by the change of appearance which occurs when the viewing of a pair of stereo pictures transforms them into a stereo image. A stereo image has a marked depth and also shows firmly shaped ‘solid’ objects not present as such in the original pair. It therefore involves us in a novel

sensory experience, which has obviously been created by tacit knowing. Such *phenomenal formation* is a characteristic feature of from-to knowing. In this manner the coherence we see in nature has an actually new *sensory* quality not possessed by the sense perceptions from which it is tacitly created ... The relation of subsidiary to focus is formed by the *act of a person* who integrates one to the other. The from-to relation lasts only so long as a person, the knower, sustains this integration.

A rainbow, a primrose, and language are what we see when we peer through the stereopticon; but they are not there in the two pictures which prompt this response. There is value added. In this sense, there are no prior data; there are no givens.

Language is created in an artful way, in a way which is distinct from others (e.g., non-linguists) who have had different experiences, yet in a way which is enlightening. Although 'scientific knowledge' is fundamentally the same as 'everyday knowledge', there is an identifiable difference (Polanyi 1946.24):

... the capacity of scientists to guess the presence of shapes as tokens of reality differs from the capacity of our ordinary perception, only by the fact that it can integrate shapes presented to it in terms which the perception of ordinary people cannot readily handle.

And that difference lies in the 'skills of connoisseurship' (Polanyi & Prosch 1975.32) of the scientists. Latour and Woolgar (1979.128-29) express the dependence of the phenomenon upon the observer as follows:

... in emphasising the process whereby substances are *constructed*, we have tried to avoid descriptions of the bioassays which take as unproblematic relationships between signs and things signified. Despite the fact that our scientists [the ones in the laboratory which Latour and Woolgar are observing] held the belief that the inscriptions could be representations or indicators of some entity with an independent existence 'out there,' we have argued that such entities were constituted solely through the use of these inscriptions. It is not simply that differences between curves indicate the presence of a substance; rather the substance is identical with perceived differences between curves. In order to stress this point, we have eschewed the use of expressions such as 'the substance was discovered by using a bioassay' or 'the object was found as a result of identifying differences between two peaks.' To employ such expressions would be to convey the misleading impression that the presence of certain objects was a *pregiven* and that such objects merely awaited the timely revelation of their existence by scientists. By contrast, we do not conceive of scientists using various strategies as pulling back the curtain on *pregiven*, but hitherto concealed truths. Rather, objects (in this case substances) are constituted through the artful creativity of scientists.

Latour and Woolgar have in mind that the assertion of the scientist is not initially a 'truth' or a 'fact', and that that status is achieved by a history of social construction. Only retrospectively, does it appear that 'truths' were discovered, whereas in actuality they are negotiated and created. It is the sociology of science which validates (or not) these 'artful creations' as truth and as fact. Polanyi reaches a similar conclusion that the components of knowledge (scientific or not) are the artful constructions of the perceiver (Polanyi & Prosch 1975.30-31):

People miss the point when they speak of the exact predictions made by the mathematical sciences ... You might think that Newton's laws could predict the exact position of the planets at any future moment of time. But this they can never do. Astronomers can merely compute from one set of numbers, which they identify with the position of a planet at a particular time, another set of numbers, which will represent its position at a future moment of time. But no formulas can foretell the actual readings on our instruments. These readings will rarely, if ever, coincide with the predicted numbers as computed from Newton's laws, and there is no rule — *and can be no rule* — on which we can rely for deciding whether the discrepancies between theory and observation should be shrugged aside as observational errors or be recognized, on the contrary, as actual deviations from the theory. The assessment in each case is a personal judgment ... We may conclude quite generally that no science can predict observed facts except by relying with confidence upon an art: the art of establishing by the trained delicacy of eye, ear, and touch a correspondence between the explicit predictions of science and the actual experience of our senses to which these predictions shall apply.

Polanyi (Polanyi & Prosch 1975.63) concludes then that

Scientific inquiry is accordingly a dynamic exercise of the imagination and is rooted in commitments and beliefs about the nature of things. It is a fiduciary act. It is far from any skepticism in itself. It depends upon firm beliefs. Nor should it ever give rise to skepticism. Its ideal is the discovery of coherence and meaning in that which we believe exists; it is not the reduction of everything to a meaningless jumble of atoms or accidentally achieved equilibrations of forces ... It rests no less than our other ways of achieving meaning, upon various commitments which we personally share. We make use of these in science in creative and imaginative ways involving our very person.⁴⁷

⁴⁷ Such views of science as those advocated by Polanyi, Latour and Woolgar, and others (e.g. Kuhn 1962 and 1970) have not gone unchallenged. They, in fact, represent the minority position, which some take to be a 'danger' (Scheffler 1967.12) to the 'standard view' (Scheffler 1967.7) represented by the work of Lewis, Cohen, Nagel, Popper, and others. In the standard position, science differs from common sense and everyday knowledge in that science is objective and is directed towards discovery of a preexisting truth or coherence/pattern (Scheffler 1967.8):

Something enables the the behavior of speakers and secondarily, our record of that behavior. Language is the creation of the linguist in response to such experiences/records and is intended to soothe the irritation we feel that something is going on here and that we ought to be able to understand it. It is important to do so. Yet what is responsible for our experiences is other than the inscriptions we manufacture of them. The aim of such an alternative approach is not an ‘explanation’ of language constituted by “any merely *formal* subsumption of a natural law under a more general law” (Polanyi & Prosch 1975.55), but a “relief from puzzlement” (Polanyi & Prosch 1975.53):

Suppose we are puzzled by the way an intricate part of machinery is constructed and the way it works or by the layout of a building in which we keep losing our way. What we are seeking here is an understanding of the machine or the building — an insight into them, but not an explanation. Such insight is a particular type of tacit integration that has not yet been mentioned. Its subsidiary items are the particulars of the complex entity — the machine, or the rooms in the building; and when we integrate these particulars and thus bring out their joint meaning, their puzzling aspect is transformed into a lucid image. Our puzzlement in these cases is relieved by an insight which is itself simply our own meaning integration of the parts of the complex entity.

Following others (e.g., Polanyi 1964, Holton 1973, Latour & Woolgar 1979, Jones 1982), who have argued that scientific constructs are the socially negotiated product of their respective fields (e.g., TRF [Latour & Woolgar 1979], matter [Jones 1982]), I assume that language is also a concept which is a consequence of negotiation and not a prior given/datum to be elucidated.⁴⁸

3. On ‘Language’ Again: What it May Be Like.

Combining the themes from Sections 1.0 and 2.0, we may guess that none of the activity recorded and remembered between me and the speaker of Ilokano was itself ‘Ilokano’. What I experienced was behavior which was enabled by ‘Ilokano’, but which behavior is of a nature different from the language. Figure 4 is my attempt to gain some relief from the puzzlement of those encounters. Returning to the machinery/building metaphor of Polanyi mentioned above (Polanyi & Prosch 1975.53), my circumstance was something like this: I had discovered a house, but knowledge of its construction (Polanyi’s ‘layout’) provided no understanding of it. The

‘... this view affirms the objectivity of science; more specifically, it understands science to be a systematic public enterprise, controlled by logic and by empirical fact, whose purpose it is to formulate the truth about the natural world.’

⁴⁸ Forget it, Jake. It’s Chinatown.

character of the construction — the artifact — did not reflect directly the tools which had enabled the house nor the wit which drove the tools. In this perspective, it would not matter how large the house was nor how much of the house was examined. The relation between the building (or the experience with ‘Ilokano’) and the intelligence which enabled it (‘Ilokano’ itself) was too indirect to allow reasoning from the first to the second. Figure 4 is an attempt to imagine something of the dimensions of the wit which might have allowed the exchanges between me and the speaker to occur as they did. In gaining relief from the puzzlement engendered by these meetings, no one kind of experience/observation/encounter can be privileged over another. Each provides a partial and incomplete refraction of language. Because of this, we must be open to any and all pieces of information which appear to inform us of language. Whether we begin our pursuit of language with inspection of inscriptions taken from aural or visual recordings of conversations or extended monologues (‘narratives’), from directed exchanges with speakers (‘elicitations’), from found documents (‘texts’), from apparent errors, etc., none — nor even their sum — will inform us of what language is, because *in examining those inscriptions we are not examining language*. It is not a matter of not having enough to examine, ‘good enough’ data, nor of having incomplete records. There is no ‘final diagram or curve’ to be derived from these ‘data’ (no ‘underlying structure’ in any sense); and in none of our encounters have we touched, nor will we touch, language directly.

Not having immediate access to language and being unable to perceive it directly, we must infer (or ‘imagine’ it). We must wonder what its ontological status is and whether there is an ‘it’. Here, our condition as linguists shares something with the contemporary (this century) condition of physics. More and more, it has become apparent that the perception of the universe in the usual manner in which it is measured does not accord with the nature of the universe when it is not being recorded by the physicist. In the context of observation, physical phenomena seem to be particulate, yet there are hints that the particles are themselves in reality not discrete in the common sense manner. For example, an electron, which may be measured as if it were a particle, still betrays, in a common observation, that it (even each electron and not a collection) is continuous and not divisible. An electron reveals itself to

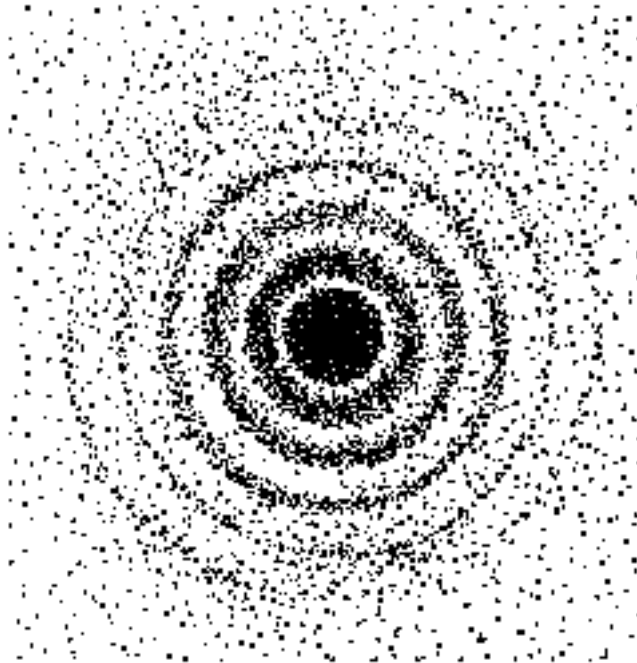


Figure 5: *An Airy Pattern.*

be a wave-like phenomenon in the Airy pattern produced by a succession of them as they pass from a source through two slits in a barrier to strike a phosphorus screen (Conveney & Highfield 1990.121-124). The Airy pattern (Herbert 1985.62) is typical of waves of any sort, but not of particles. Complementarily, in the 19th century, the supposed wave nature of light (which had prompted a search for luminiferous ether, its supposed medium) yielded instead to a particulate conception of light in order to explain certain phenomena (black-body radiation, the photoelectric effect, and the Compton effect. Cf. Herbert 1985.39.). This entanglement of continuous-and-discrete sorts itself out so that the ‘representation’ (Herbert 1985.98) of physical phenomena in quantum theory is unfailingly accurate; but quantum theory is a representation of discrete measurements.⁴⁹ It is known that ‘behind’ the measurements there exists an indeterminate and continuous universe in which ‘reality’ exists as a probability (Herbert 1985.73), without parts (Herbert 1985.84), with infinite attributes (Herbert 1985.104), and in a non-local way (Herbert 1985.223). Yet that is not what quantum theory describes; it describes the discrete measurements and indirectly represents an unfixed reality (Herbert 1985.111-12):

⁴⁹ Herbert (1985.94) writes that ‘Quantum theory by design only predicts the results of measurements.’

This theory deals with the world in a particularly indirect manner. It focuses strictly on measurement acts, not on how the world might behave between measurements; it does not describe single measurement events but only patterns of events, for which it gives statistical predictions.

The discontinuity between the unmeasured world and the world of quantum measurement (or conversely, their point of contact) can be placed at any spot along a chain of measurement opportunities; but “as far as final results are concerned, you can cut the chain and insert a collapse [find the discontinuity, PWD] anywhere you please” (Herbert 1985.147).⁵⁰ John von Neumann (mathematician, quantum theorist, and computer scientist) suggested this conception, as reported by Herbert (Herbert 1985.148):

Von Neumann could not find a natural place to locate his ‘miracle’ [the point in a laboratory context at which unbounded quantum reality appears as bounded quantum measurement, PWD]. Everything, after all, is made of atoms: there’s nothing holy about a measuring instrument. Following the von Neumann chain, driven by his own logic, in desperation von Neumann seized on its only peculiar link: the human mind. This is the only process in the whole von Neumann chain which is not mere molecules in motion. Von Neumann reluctantly came to the conclusion ... that human consciousness is the site of the wave function collapse.⁵¹

Language has been traditionally taken to be a discrete phonemon, both in reality and in its representation in description. The more continuous portion (perhaps *parole*, *performance*, and *usage*) has been of secondary importance. Yet analogous to the condition in physics, it is becoming more apparent that discreteness is not the reality of language. The existence of an ‘inventory’, whether of *phonemes* (Copeland 1991) or of *morphemes* (Lu 1991 and Davis 1993), can be questioned. The nature of ‘distribution’ (Fox 1991) similarly undermines the notion of well-defined morphemes. The dissolution of entities extends to *grammar* and *constructions*, which are ‘emergent but never present’ (Hopper 1987.148). The assumption of a small number of discrete *grammatical categories*, e.g. ROLES, can also be recognized to be an illusion. ROLES as well are ‘distributed’, created on demand (Davis 1994); and there

⁵⁰ In the example from neuroendocrinology above, the insertion of a mechanism of measurement corresponds to the introduction of the term ‘data’ (noted above from Latour and Woolgar 1979.49-51) at the point at which a computer prints out of a sheet of figures, i.e. collated measurements.

⁵¹ The sort of subjectivity, which this view forces, is ignored by philosophers of science such as Scheffler (above), who continue to maintain the ‘objectivity’ of science.

exists an unlimited number of them (within and across languages).⁵² Language is also beginning to appear to be an unformed wave-like phenomenon and to be discretely shaped only as we perceive it naively.

4.0 Conclusion.

Here, language (e.g. 'Ilokano') is assumed to be unformed; and it is the imposition of human observation upon the experience of language which gives the impression that language is discrete (and therefore that it can be modeled as a structure). The pattern of Figure 4 is (my impression) of the track left by some the principle(s) of 'Ilokano'. These principles (whose nature is different from the terms in which I encountered and in turn presented them) must necessarily be of a scope broader than Figure 4, broader than 'Ilokano', and present in all 'languages'.⁵³

I suggest that linguists have two conditions in common with the physicists. *First*, like them we *cannot directly seize the phenomenon* of interest. We can only obliquely imagine what it is like, while being nevertheless convinced of its existence. *Second*, like the physicists, the phenomenon of our interest is *not an object/thing* even though common sense experience with language suggests that 'thingness' is a basic property. But unlike physicists, we do not have a mathematics which can be used to represent discrete measurements/observations of the phenomenon of our interest.⁵⁴ To attempt to use intermediary discrete models must necessarily distort the phenomenon and guarantee that it be misunderstood. Unlike physicists, *we must attempt directly to imagine what that unbounded existence is like*. Herbert (1985.⁵⁷ and 67) comments on the physicists' quandary:

Since quantum theory fits the facts exactly, many physicists are sure that it bears some relationship to reality itself: such a perfect match between theory and fact is no accident ... Quantum theory doesn't show up *directly* in the quantum facts: it comes indirectly out of the quantum theory, which perfectly mirrors these facts ... However, the quantum facts give us not one description but two — each one separately inadequate, and both together contradictory. Moreover the knot that

⁵² This does *not* imply that there is no coherency (or 'universality') to ROLE. Davis (1994) argues that the limitation on ROLE lies in the operation of principles of intelligence which channel language into one or another motivated patterns. The semantics of word order, which can be shown to be varied and ill-defined in the received sense, can also be shown to be coherent in the same way (Huang & Davis 1991).

⁵³ Some specific suggestions concerning these principles are attempted in Davis 1994 and Ms.a., Davis & Hardy 1988, Davis & Saunders 1984 & 1989, and Huang & Davis 1991.

⁵⁴ Although it is common to try to mathematize the practice. Some are better, more interesting than others, e.g., Hockett 1967.

connects these two descriptions is the act of observation; leave out observation and neither description makes sense.

In arriving at a comfortable conceptualization of language, we must provide some means of accommodating the unshaped, unstructured (and counter-intuitive) nature of language while at the same time accommodating the perceived discrete form. But we must *not* force the one view (the latter) upon the other. One way of expressing our problem as linguists is to explain how a continuous, unformed experience can give us (as we live within it and as we attempt objectively to examine it as if from without) the impression that it 'consists' of 'pieces'

Physics is not infrequently cited as the most 'scientific' of sciences; and linguistics, not infrequently, has emulated science in attempts to acquire respectability. In the view outlined here, it is important to note that I am suggesting no emulation of physics. I suggest (in Section 2), rather, that linguistics and physics share many of the same problems concerning the reality of the respective 'objects' of our study. The position which I advocate is that there is a universe of intelligence which is prior to the mathematicizations and the forms (the 'objects') of the sciences. It is intelligence which enables 'reason', 'logic', 'mathematics', etc., but which is itself not constituted of those. No reason, no logic, no mathematics can reach intelligence. As suggested by Michael Polanyi, intelligence must be imagined. And language, as a configuration of intelligence, shares this same condition. Language is not presented to us as a given (Section 2) to be read or deciphered; it, too, must be imagined.⁵⁵

The view of the condition of our practice which I have outlined above does not require that we abandon inquiry into language. It is still possible to seek relief from puzzlement, and attempts to understand language will be much more effective if we are aware of the fundamental problems involved. Constant self-reflection may aid us to detect ourselves in the act of creating ourselves ... and perhaps, in that moment, to see beyond.

[Completed: May 20, 1994]
[Version: July 19, 2021]

⁵⁵ Because there is no linear series of manifestations initiated by 'language' (which we do not observe directly) and ending with 'data' (which we do observe), we must be sensitive to all aspects of our experience which we impute to language. We cannot allow any one of these experiences to become hegemonic, nor can we ignore any one of them. Put more simply, we cannot rely exclusively on the observations from large scale corpora (or discourse narrative, or speaker errors and corrections, or encounters of the sort described in Section 1.). We must welcome them all, because each enriches our ability to imagine the nature of what enables our experience of language.