

Deconstructing SYNTAX

Jean-Pierre Koenig 

University at Buffalo

Karin Michelson 

University at Buffalo

Proceedings of the 21st International Conference on
Head-Driven Phrase Structure Grammar


University at Buffalo

Stefan Müller (Editor)

2014

Stanford, CA: CSLI Publications

pages 114–134

Koenig, Jean-Pierre & Karin Michelson. 2014. Deconstructing SYNTAX. In Stefan Müller (ed.), *Proceedings of the 21st International Conference on Head-Driven Phrase Structure Grammar, University at Buffalo*, 114–134. Stanford, CA: CSLI Publications. DOI: 10.21248/hpsg.2014.7. 

Abstract

There are at least two distinct ways of conceiving of syntax: the set of rules that enable speakers and listeners to combine the meaning of expressions (compositional syntax), or the set of formal constraints on the combinations of expressions (formal syntax). The question that occupies us in this paper is whether all languages include a significant formal syntax component or whether there are languages in which most syntactic rules are exclusively compositional. Our claims are (1) that Oneida (Northern Iroquoian) has almost no formal syntax component and is very close to a language that includes only a compositional syntax component and (2) that the little formal syntax Oneida has does not require making reference to syntactic features.¹

There are at least two distinct ways of conceiving of syntax:

Definition 1. *Syntax is the set of rules that enable speakers and listeners to combine the meaning of two or more expressions (words or phrases) (hereafter, compositional syntax)*

Definition 2. *Syntax is the set of formal constraints on the combinations of two or more expressions (words or phrases) (hereafter, formal syntax)*

Syntactic rules in most languages partake of both conceptions of syntax: They are statements about how speakers can combine the meaning of expressions while at the same time restricting the form of the expressions they license the combination of. But only the first, i.e. compositional syntax, is a conceptual necessity. Whatever syntax does, it must at a minimum ensure that when two expressions of the right semantic kind combine, they combine semantically in the right way. This is what the syntax of natural and logical languages share. It seems impossible to imagine a natural language whose syntax would not provide recipes for combining the meanings of expressions that are part of well-formed constituents. But, because most syntactic rules in most languages also include a formal component, we tend to think of syntax in the second sense (what we call *formal syntax*) as syntax proper. The question that occupies us in this paper is whether *all* languages include a significant formal syntax component or whether there are languages in which most syntactic rules are exclusively compositional and do not restrict the form of the expressions that combine. Our claim is that Oneida (Northern Iroquoian) is such a language. Most of its syntactic rules or constructions are strictly compositional, and very few include a formal component and that formal component is very restricted. More precisely, we make the following two claims.

Claim 1. *Compared to most languages, Oneida has almost no formal syntax component and is very close to a language that includes only a compositional syntax component.*

¹The examples come from a compilation of recorded texts or stories (Michelson, Kennedy and Doxtator, to appear), and we are grateful to those who have contributed recordings; we would like especially to acknowledge Norma Kennedy and the late Mercy Doxtator for their collaboration. References to works on Iroquoian languages may be found in an annotated bibliography (Michelson 2011). Grammars of several other Iroquoian languages are presently underway (e.g. Chafe In press).

Claim 2. *The little formal syntax Oneida has does not require making reference to syntactic features: Words and phrases in Oneida do not carry any syntactic feature.*

Our paper is organized as follows. In Section 1, we discuss traditional kinds of evidence that can justify positing formal syntactic rules or constraints as well as evidence for positing syntactic features. Section 2 briefly reviews our previous work in Oneida that argues that the kind of evidence adduced for formal syntactic constraints and syntactic features is absent in Oneida. Section 3 is the core of our paper, as it discusses many of the constructions present in Oneida and demonstrates how one can do compositional syntax without formal syntax. Section 4 argues that the little bit of formal syntax you have in Oneida does not require syntactic features (at least, as long as you have constructions as first-class grammatical objects). The paper concludes with a brief discussion in Section 5.

1 What good are syntactic features?

Syntactic features are the mainstay of syntactic theories since at least the time of American structuralism. Within HPSG or SBCG three kinds of syntactic features can be distinguished:

1. *Selectional features*: ARG-STRUC, VAL, SPEC, MOD for bounded dependants and REL, EXTRA, SLASH and the like for unbounded dependants
2. *Categorial features*: CASE, VFORM, AUX, ...
3. *Other features*: ROOT, LEX, ...

One of our claims is that the grammar of Oneida does not require the introduction of any of these features on words or phrases. Since the category *other features* is heterogeneous and the need for these features less cogent than that of other features, we do not discuss them any further here and focus instead on selectional and categorial features. Both selectional and categorial features percolate from *syntactic* heads to mothers of local trees (except, of course, SPEC and MOD, which are introduced to model selection of heads by non-heads). Although percolation is not definitional of either kind of features, percolation is one of the main reasons you need syntactic features in the first place, and it is one of the consequences of syntactic selection.

1.1 What do you need categorial features for?

In most languages, categorial or part-of-speech features are needed to constrain the combination of expressions beyond semantic types. In other words, if one cannot predict what combinations are grammatical or ungrammatical on the basis of the semantic types of combining expressions, categorial features are needed. Taking well-known examples from English, nouns cannot combine with (nominative

or accusative) noun phrases, only with possessive noun phrases or PPs (see(1)); prepositions cannot combine with verbs (see (2); note that we analyse gerunds as verbal categories but not verbs following Malouf 2000); some verbs subcategorize for PPs (headed by particular prepositions) (see (3)), for particular kinds of clauses, or VPs whose main verb has a particular form (see (4)); singular count nouns require determiners (see (5)). The fact that our description of these well-known facts may not be the most appropriate or that some of the constraints may follow from more general principles is not important. What is critical is that the kinds of constraints illustrated in (1) through (5) provide the traditional motivation for positing categorial distinctions in English, for it is not clear how to *derive* these constraints from the semantic types of combining expressions no matter how semantically motivated some of these constraints may be.

- (1) Bill's daughter/*Bill daughter/The book of Job/*The book Job
- (2) Bob dreams of getting a new car/*get a new car
- (3) John laughed *at* the idea
- (4) I want for him to be happy no matter what he ends up doing
- (5) Milk/A student/*student/students

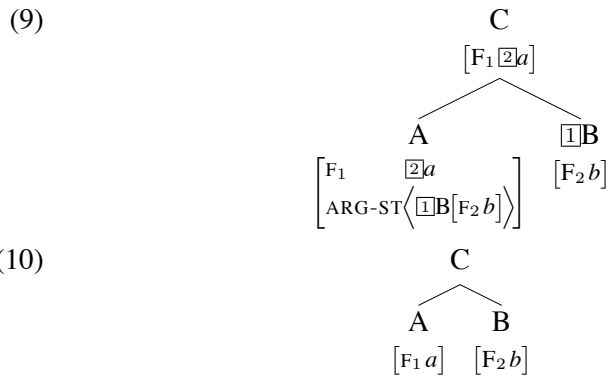
1.2 What is the evidence for syntactic selection?

Most syntactic thinking since at least Adjuckiewicz (1935) assumes that a good portion of the syntax of natural languages can be characterized through a relation of selection between an expression and other dependant expressions. The evidence is multi-varied, and we only mention some of the basic kinds of evidence: Some dependants of heads are "obligatory" (see (6)); the order of dependants and heads is (relatively or partially) "fixed" (see (6)); verbs undergo valence alternations (or their phrase-structural, movement-driven equivalents) (see (7)); dependants (and heads) can enter into binding relations (including WH-dependencies) (see (8)). Scare quotes around some of the terms are meant to suggest that various analyses are possible. Again, what matters is not whether the descriptions of the facts illustrated in (6)-(8) are the most appropriate ones; rather what matters is the existence of these kinds of facts since they motivate positing relations of selection between, say, heads and dependants. Note that in many cases selection goes together with an ordering of the dependants selected by heads, as the binding of the English reflexives in (8) illustrates (see Bickel 2011 for a discussion of the kinds of evidence for (ordered) grammatical relations).

- (6) *(Mary) loves *(John).
- (7) John is not loved by his students.
- (8) Mary loves herself.

1.3 Different kinds of formal syntax

Although we have discussed models of the facts illustrated in (6) through (8) in terms of selection, the facts we discussed in Sections 1.1 and 1.2 can be modelled in two distinct ways. We can use *selectional syntactic constraints* to ensure that the presence of one daughter appropriately restricts the presence (or absence) of other daughters, their position, and their *form*. We can alternatively use *constructional syntactic constraints*, i.e. posit one-off structural patterns which may restrict the form or formal properties of daughters but do not require one daughter to select other daughters. Of course, one and the same approach to syntax can adopt both selectional and constructional constraints. In HPSG (Pollard and Sag 1994), syntax is mostly projectionist but some is constructional; in SBCG (Boas and Sag 2012), a little more syntax is constructional but selection still plays an important role. We illustrate selectional and constructional syntax in (9) and (10), respectively.



The choice of a projectionist vs. constructional syntax is not orthogonal to the issue of selection. Selecting expressions are most often heads and the mother node records whether selected expressions were realized locally. There is little doubt that syntactic selection and projectionist approaches to selection have been very successful. But, irrespective of whether one is more inclined to projectionists or constructional solutions, models of syntactic knowledge include a good bit of formal syntax and that formal syntax requires the use of syntactic features, typically for the kinds of reasons we alluded to in this section. The question is whether this is part of the design of natural languages or is merely an overwhelmingly frequent aspect of natural languages syntax. In the next section, we suggest that the latter answer is the correct one. There are languages like Oneida where syntactic features are dispensable and formal syntax is minimal: In such languages, most syntactic rules are examples of combinatorial syntax.

2 Formal syntax in Oneida

We have argued in past work that Oneida syntax is not based on syntactic selection or ordering of syntactic dependants (subjects and objects) (see Koenig and Michel-

son 2012; Koenig and Michelson 2014). We cannot recapitulate our argument in detail here. Its logic is simple: None of the evidence typically adduced in favor of syntactic selection or ordering of syntactic dependants is present in Oneida, and in the absence of such evidence the simplest model of the grammar of Oneida does not make use of these notions.

2.1 There is not much formal syntax in Oneida

In this section, we list the phenomena whose presence justifies syntactic selection or syntactic part-of-speech information and whose absence in Oneida suggests its syntax is of a different kind than that of most languages (see Koenig and Michelson 2014 for much more detail).

No syntactic selection

1. There is no requirement by words based on noun or verb stems that they have dependants, as a comparison of the Oneida discourse in (11) and its translation shows.²
2. There is no necessary co-indexing relation between a word's semantic arguments and expressions that further specify these arguments. (12) shows that an external referential phrase can denote a subset of the entities referenced by the pronominal prefix on the verb while (13) shows that an external referential phrase can denote a superset of the entities referenced by the pronominal prefix on the verb.
3. There is no restriction on the order of dependants headed by words based on noun or verb stems (see (14) vs. (15)). As argued for by Mithun (1987) for Mohawk, the concept of default word order has no application in Iroquoian.
4. There is no restriction on the *form* of dependants headed by words based on noun or verb stems.
5. There are no valence alternations. All (derivational) morphological operations that have a reflex in pronominal prefixes on verb stems are morphosemantic in Ackerman's (1992) sense and the morphological effect follows from the fact that they alter the meaning of the stem they apply to.

²Abbreviations are A agent, CAUS causative, CSL cislocative, COIN coincident, DISTR distributive, DL dualic, DP dual-plural or non-singular, DU dual, EX exclusive, FACT factual, FI feminine-indefinite, FZ feminine-zoic, FUT future, HAB habitual, JN joiner vowel, LOC locative, M masculine, NEG negative, OPT optative, P patient, PART partitive, PL plural, PNC punctual, PRES present, PROG progressive, REP repetitive, SG singular, STV stative, TRL translocative. Z/N zoic/neuter. The symbol > indicates a proto-agent acting on a proto-patient; for example, 3M.SG>1SG should be understood as 3rd person masculine singular acting on 1st person singular. Not all clitics or particles are glossed, as appropriate word glosses do not always exist. A set of robust phonological modifications occur at the end of "utterances." A pervasive utterance-final process is the devoicing of a word-final vowel or syllable, indicated by underlining.

6. There is no syntactic *ordering* (subject, object, ...) of dependants of words based on noun or verb stems. Phenomena that typically justify ordering dependants are absent in Oneida:

- (a) Principle A is irrelevant (reflexive/reciprocal marking is strictly morphological)
- (b) There is no VP ellipsis/conjunction (*pace* Baker 1996)
- (c) There is no clear evidence supporting the claim that Oneida shows weak cross-over effects (*pace* Baker 1996)
- (d) Principle C is not operative in Oneida (*pace* Baker 1996), as (16) and (17) show.³

- (11) Né=s wí né n tshiwahu níse? lon-u?wéskwani-he?
 so it's a long time ago 3M.DP.P-enjoy-HAB
 a-hati-yat-a-kó n-a? ká',
 OPT-3M.PL.A-wood-JN-go.somewhere.to.harvest-PNC see
 tahnú=s kwí kwahotoká u tsi? wa-hu-nakla:kó tho
 and=habitually just for real that FACT-3M.PL.A-move.away:PNC there
 y-a-hu-náklat-e? tsi? nú ye-hoti-yo?tá-st-a?
 TRL-FACT-3M.PL.A-settle-PNC where TRL-3M.DP.P-work-CAUS-HAB
 'A long time ago they used to like to go cut wood, and so they would move away and they would settle over there, where they were working.' (Mercy Doxtator, *Some Woodcutters Get a Visitor*, recorded 1996)
- (12) N_A kwí wa-hy-atlihwísa-ne? ka?iká ló-nhah-se?,
 so then FACT-3M.DU.A-agree-PNC this 3M.SG>3M.SG-hire-HAB
 so then he (my father) and his boss agreed/planned, (Norma Kennedy, *A Haunted Car*, recorded 2010)
- (13) tsi? náhte? wa-h-atkátho-? uky-atyóha.
 that what FACT-3M.SG.A-see-PNC 1DU.P-brother-in-law
 '(This is my favorite story,) what my brother-in-law saw.' (Mildred Cutcut, *The Hunter*, recorded 1982)
- (14) n_A kwí úska útlatste? thiká Tsyó khále? í yakn-í tlu-?,
 so then one time that Joe and me 1EX.DU.A-be.at.home-STV
 'so then one time Joe and I were home,' ((Clifford Cornelius, *A Lifetime Working*, recorded 1994)

³Both these examples were elicited to test for Principle C. They (and similar sentences) are based on situations that occurred in the recorded texts. Our consultant considered the sentences perfectly okay. However, in certain respects they are somewhat atypical, which is perhaps not unexpected for elicited sentences.

- (15) né· tsiʔ ni-ho-naʔkhwá·-u thiká lakeʔnihká,
 because PART-3M.SG.P-get.mad-STV that my late father
 ‘because my father was so mad,’ (Clifford Cornelius, *A Lifetime Working*,
 recorded 1994)
- (16) Waʔ-utat-hlo·lí kwí tsiʔ yako-yo·té· aknulhá·
 FACT-3FI.SG>3FI.SG-tell-PNC that 3FI.P-work:STV my mother
 aʔé· Heinz Factory.
 way over there Heinz Factory
 ‘My mother told me that she was working way over at the Heinz Factory.’
- (17) Wa-hak-hlo·lí tsiʔ wa-huwá-hsle-ʔ lakeʔníha
 FACT-3M.SG>1SG-tell:PNC that FACT-3>3M.SG-chase-PNC my father
 nΛ ka-list-aʔké-shuʔ te-ho-tawΛlye-háti-ʔ.
 when 3Z/N.SG.A-iron-LOC-DISTR DL-3M.SG.P-travel-PROG-PRES
 ‘My father told me that she chased him when he was going along on the
 railway tracks.’

No syntactic parts of speech. Oneida has a robust notion of stem classes. Several derivational and inflectional processes allow us to distinguish between four kinds of Oneida stems: noun stems, verb stems, uninflected stems, and kin stems (see Koenig and Michelson 2010 on Oneida kin stems). But, the same is not true of Oneida words. To illustrate the difference between nominal stems and what would be putative NPs (and their N heads), we counted in the naturally produced discourses in Michelson, Kennedy, and Doxator (*To appear*) all referring expressions headed by words based on the four kinds of stems. Table 1 summarizes the relevant part of this corpus study: Over 60% of referring expressions (what typically would be encoded by NPs in English) are headed by words based on stems with *no* nominal morphology.

Table 1: Proportions of referring expressions according to morphology

	REs headed by words with exclusive nominal morphology	REs headed by words with some nominal morphology	REs headed by words with no nominal morphology	Total
Count	575	686	2027	3288
As % of REs	17.5%	20.8%	61.7%	100%
As % of Wds	1.9%	2.2%	6.5%	10.6%
As % of clauses				39.93%

The data in Table 1 do not provide conclusive evidence that there is no need for syntactic part-of-speech information in Oneida. In fact, there can only be negative evidence: No syntactic rule/constraint makes reference to part-of-speech, as we show in the next section. But the low percentage of referential expressions headed

by words based on noun stems (less than 40%) will, hopefully, suggest to readers how different Oneida is from many better-studied languages. The irrelevance of part-of-speech distinctions is, of course, not surprising since without syntactic selection there is little need for part-of-speech information. Similarly the absence of syntactic headedness in Oneida is not surprising since without syntactic selection syntactic headedness is hardly useful.

2.2 The little formal syntax Oneida has


The previous section showed that the typical evidence in favor of formal syntax is nowhere to be found in Oneida. This is the sense in which Oneida's syntax is, as we show in more detail in the next section, almost exclusively compositional. But, not quite. There are several syntactic constraints that are formal, i.e. that make reference to the form, ordering, or lexical identity of daughters within a constituent. We mention a few here. First, some clauses begin with the word *tsi?*, e.g. argument clauses with realis interpretations. Second, some words must co-occur with other words: *ok* must follow the word *ukha?* for an “indefinite person” meaning to be encoded. Third, question words must occur first in a clause, while argument clauses must follow the verb whose propositional argument they further specify. Crucially, constraints such as these are very restricted. They all involve particular semantic types or particular words (or classes of words) and therefore do not require the projection of syntactic features (including the projection of part-of-speech information or categorial features). Linear order constraints need only mention the semantic type of daughters or their lexical identity. So, not only does Oneida have little formal syntax, the little formal syntax it has does not require the introduction of syntactic features. In other words, there is no need for a SYN attribute in Oneida. In the next section, we show what an almost exclusively compositional syntax and one that does not include syntactic features looks like.




3 The constructions (the goods)!

A few preliminary remarks are in order. First, we leave out a couple of constructions for reasons of space. Second, the list of constructions we have identified is most probably incomplete. It has been compiled over the last few years; more recently we added to it by, together, going through a few pages of texts on a regular basis and accounting for all the constructions speakers made use of. While the list is probably incomplete, we are fairly confident that it includes the bulk of Oneida constructions and that variants of these constructions or other constructions would not significantly alter our claims. Third, if syntactic phrases are not built through syntactic selection, semantic composition must be done constructionally (including variable identification) (see Bach 1976; Klein and Sag 1985). Fourth, since there is only “existential” quantification in Oneida (no quantifiers of type $< 1, 1 >$ in the sense of Peters and Westerstahl 2006; see Koenig and Michelson 2012), we

can dispense with (generalized) quantifiers altogether, have only free variables, and get the existential force of variables from the anchoring of atomic formulas (à la Kamp 1981; Kamp and Reyle 1993), as shown in (18). Fifth, we adapt (quite liberally) a semantic underspecification approach called *Lexical Resource Semantics* (see Richter and Sailer 2004). Semantic underspecification is quite useful when all semantic combinatorics is constructional, although we do not know if it is truly needed.

- (18) $\llbracket P(x_1, \dots, x_n) \rrbracket^M = 1$ iff
 there is an anchoring g such that $\langle g(x_1), \dots, g(x_n) \rangle$ that is in the denotation of P .

The following sections present the list of Oneida constructions. Our goal is two-fold. First, illustrate what compositional syntax looks like; second, demonstrate that we can model Oneida syntax without the use of syntactic features (SYN in HPSG parlance). For space reasons, we give an example of each construction with its English translation and the meaning of the mother node, but do not discuss the example nor provide interlinear glosses; parts of the sentences that exemplify the construction are in bold. We use the terms *entity expression* and *situation expression* for expressions denoting or describing entities and situations, respectively. Our analysis of Oneida syntactic constructions thus relies on a fundamental distinction between two semantic types. We leave a justification of these two particular semantic types to another venue. Finally, to increase readability, we indicate graphically the semantic import of constructions on the semantic translation of the (relevant portion of) examples.  indicates which of the daughters' index is the

index of the entire expression;  indicates identification of indices across daughters;  indicates that a predication was added by the construction itself; finally,  indicates the output of “previous” semantic composition.

3.1 Entity expression apposition

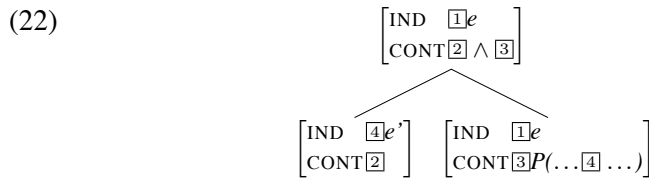
This construction states (1) that two entity expressions can, generally, co-occur in either order, (2) that the meaning of the whole bears the index of both daughters (which must be the same) and (3) that the meaning of the whole is the conjunction of the contents of the daughters. It is represented in (19) and an example is provided in (20a). Note that this construction applies more widely than apposition in English, as Oneida demonstratives are fully referential entity expressions that can occur in apposition to another fully referential entity expression, as shown in the example in (21).



- (20) a. **Lake?káha Leo**, né· kAs né· wahatkátho? thiká
 ‘My brother Leo saw it’
 (Rose Antone, What My Brother Saw, recorded 2011)
- b. brother’(‘I’, \boxed{x}) \wedge Leo’(\boxed{x})
- (21) nA ki? ok kwí· wa?ekwe?talukó· **ka?iká kaná talok**,
 ‘and then she cut into chunks **this bread** ,’ (Norma Kennedy, The Bird,
 recorded 2008)

3.2 Entity expression adjunction

This construction states (1) that two entity expressions can co-occur in either order (which is simplifying somewhat for reasons of space), (2) that the meaning of the whole bears the index of one daughter (the semantic head), (3) that the index of the other daughter is an argument of the semantic head’s content, and (4) that the meaning of the whole is the conjunction of the daughters’ contents. A representation of the construction is provided in (22) and an example in (23a). This construction illustrates the importance of INDEX selection for semantic composition in a language where syntactic selection and functional composition do not ensure the proper matching of variables and argument positions. The construction must specify that the index of one of the daughters is the index of the entire expression so that hearers can determine upon hearing (23a) who died, the person referred to via *aknulhá·* or the person referred to via *onulha?ká·*.



- (23) a. Tahnú· **aknulhá· onulha?ká** tshahanáklate? Bill ne? thó·ne? né· t-yakaw-ahé·yú. ‘And **my mother’s mother** died when Bill was born.’
 (Olive Elm, Visits to my Auntie’s, recorded 1993)
- b. mother’(‘I’, \boxed{x}) \wedge late.mother’(\boxed{x} , \boxed{y})

3.3 Clausal constructions

Oneida has several constructions that build clauses out of a situation expression and various other kinds of expressions. Figure 1 summarizes the hierarchy of clausal constructions we mention in this paper.

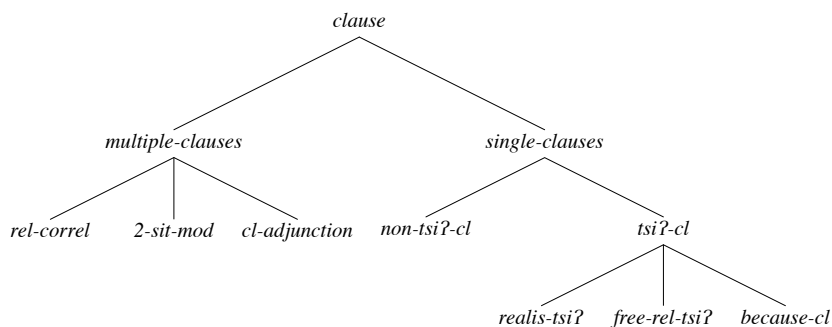


Figure 1: A hierarchy of Oneida clausal constructions

Single clauses. This is the basic construction Oneida uses to build clauses. Its formulation is complex because our analysis of clauses in Oneida is “flatter” than the analysis of clauses in many approaches, and as a consequence expressions of distinct semantic types are sisters to the semantic head. Flatter VPs have been posited for a long time in HPSG (see Bouma, Malouf, and Sag 2001 or Kim and Sag 2002, among others), but our analysis of Oneida clauses is even flatter. We have two main reasons for adopting such a flat structure. First, we know of no evidence to posit more structure; in the absence of such evidence positing additional structure would be imposing onto Oneida what is relevant to other languages. Second, the order of expressions of distinct semantic types can vary and the number of possible orderings is quite large (see Section 4). We could, of course, make use of domains (see Kathol 1999 among others), but in the absence of evidence for more hierarchical structure the introduction of this rather heavy mechanism would be *ad hoc*. Furthermore, the motivation for distinguishing linearization issues from “structural” combinations, which is at the root of linearization-based approaches, is absent in Oneida if, as we argue, there is no syntactic selection. Oneida thus lacks the very motivation for positing a level of representation in which functors and arguments combine that is distinct from their linear order. The net effect of a flat clausal structure and the absence of functional application (or its derivatives) is that semantic composition is case-based for this construction, as shown in (24). The construction is represented in (25) and an example is given in (26a). Note that semantic underspecification makes it relatively easy to have a case-based definition of semantic composition.

- (24) A situation expression can consist of a situation-describing word (the semantic head) preceded by zero or more expressions and followed by zero or more expressions. The index of the whole is that of the situation-describing word and the semantic content of the whole is determined as follows:
1. If a non-head daughter is an entity expression, its index must be included in the content of the semantic head (co-indexed with one of the head’s argument or an argument of an argument ... of the head), and the content of the whole includes the conjunction of the content

of the non-head daughter and the content of the head

2. If a non-head daughter is a situation expression, its content must be included in the content of the head
3. If a non-head daughter is a time or location expression, it takes the index of the semantic head as argument, and the content of the whole includes the conjunction of the content of the non-head daughter and the head
4. If a non-head daughter is a scopal operator (e.g. negation) its argument must include the content of all expressions to its right that are scope sensitive

$$(25) \left[\begin{array}{c} \text{IND } \boxed{1}^s \\ \text{CONT } \alpha \end{array} \right] \rightarrow \left(\begin{array}{c} \text{IND } \textit{none} \\ \text{CONT } \textit{Op}(\beta) \end{array} \right)^* , \left(\begin{array}{c} \text{IND } \boxed{2}^e \\ \text{CONT } \boxed{3} \end{array} \right)^* , \left[\begin{array}{c} \text{IND } \boxed{1} \\ \text{CONT } \textit{P}(\boxed{1}, \dots, \boxed{5}) \end{array} \right], \\ \left(\begin{array}{c} \text{IND } \textit{l} \\ \text{CONT } \textit{P}'(\boxed{1}, \dots) \end{array} \right)^* , \left(\begin{array}{c} \text{IND } \textit{s}' \\ \text{CONT } \boxed{5} \end{array} \right)$$

$\boxed{2} \triangleleft \boxed{4}$; $(\boxed{4} \wedge \boxed{6}) \triangleleft \alpha$; $(\boxed{3} \wedge \boxed{4}) \triangleleft \alpha$; $\alpha = \textit{leftmost}(\textit{Op}(\beta))$ (where *leftmost* selects the semantic content of the leftmost daughter that contributes a scope-sensitive operator)

- (26) a. **Lake?káha Leo**, né k_{AS} né **wahatkátho?** thiká
 ‘My brother Leo saw it’ (Rose Antone, What My Brother Saw, recorded 2011)

b. $\boxed{\text{brother}'('I', \boxed{x}) \wedge \text{Leo}'(\boxed{x})} \wedge \text{see}'(\boxed{s}, \boxed{x}, y)$

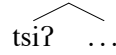
The statement of the construction in (24) and (25) stipulates that scope-sensitive expressions follow a left-to-right order so that expressions on the left take the semantic content of expressions on the right as arguments. Determining which expressions are scope-sensitive is a difficult issue we cannot go into in this paper and our assumption that scopal relations follows a left-to-right order should be considered provisional. Sentences (27) and (28) illustrate the semantic effect of inverting the order of the quantificational expression *akwekú* and the negative particle *yah*, which partially motivates our tentative hypothesis.

- (27) n_Λ kyuni? wí· né· **akwekú yah** kánike? te?sh_Λ·né·se?.
 ‘I guess **all of them** (those named in the preceding sentence) are **not** around anymore.’ (Pearl Cornelius, Family and Friends, recorded 1993)

- (28) **Yah akwekú** tehone·ká·se? onu?úseli?.
 ‘**Not everyone** likes squash (but some do like it).’ (Elicited)

One subset of clausal constructions consist of constructions that begin with *tsi?*; they are all subtypes of the *tsi?-cl* construction whose definition (minus inherited properties, of course) is provided in (29). This is the first construction we have mentioned that includes a formal component, i.e. a constraint that restricts the form of one of the daughters.

(29) *tsi?*-cl \Rightarrow



3.4 Relative clause formation (type-shifting)

Oneida has three kinds of relative clauses, internally-headed relatives, free relatives, and relative-correlatives. We only discuss internally-headed clauses and relative-correlative constructions in this paper. Relative clauses in Oneida have a purely semantic effect, as the language does not require syntactic part-of-speech information: Relative clauses type-shift a situation expression into an entity expression, provided that this entity expression is an argument (sometimes a semantic adjunct) of the predicate associated with the situation expression. The construction states (1) that an entity expression can have as sole daughter a situation expression, (2) that the content of the whole is that of the daughter, (3) that its index is that of one argument of the content. A representation of the construction is provided in (30) and an example in (31a). Interestingly, type-shifting internally-headed relative clauses (or their lexicalized equivalents) is used to encode quantification in Oneida (see (32a) and Koenig and Michelson 2012)



(31) a. *yah né· té·yAlhe? a'yutekhu ní· kA? niyaká·*,
 ‘the little one doesn’t want to eat,’ (Olive Elm, Visits to My Auntie’s, recorded 1993)

b. *small'(s, \boxed{x})*

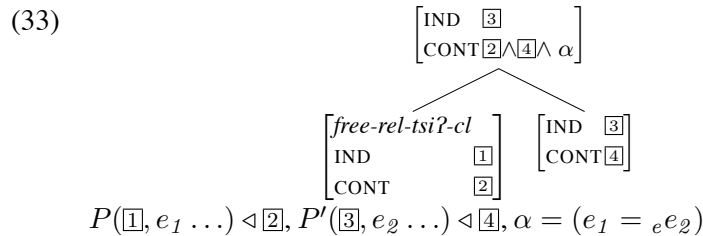
(32) a. *ÁhsA nikanláhtake Atésku?*
 ‘The [tobacco] leaves that amount to three, you are to hand them to me.’ [You are to hand me three leaves] (Olive Elm, Learning to Work in Tobacco, recorded 1998)

b. *leaf'(\boxed{x}) \wedge amount'(s, \boxed{x}), y) \wedge three'(y)*

3.5 Relative-correlative construction

Oneida has an interesting relative-correlative construction. The construction states (1) that a clause can consist of two clauses, a clause that describes a situation and a free-relative *tsi?* clause, (2) that the content of the clause is the conjunction of the contents of each clause together with an equality between the index of an entity expression within the situation expression clause and the index of the free relative,

(3) that the index of the whole is that of one entity in the clause that is not the free relative clause. A representation of the construction is provided in (33) and an example in (34a).



- (34) a. **tho yahunáklate? tsi? nú yehotiyo?tasta?**
 ‘they would settle over there, where they were working.’ (Mercy Doxtator, Some Woodcutters Get a Visitor, recorded 1996)
- b. settle’(s, x, l) \wedge work’(s’, x, l’) \wedge l=ol’

3.6 Modification

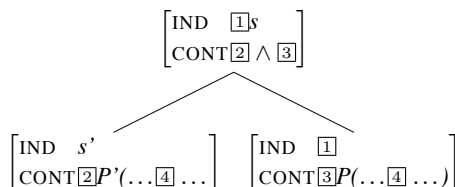
Oneida has several constructions encoding semantic modification in a broad sense of the term. We describe a few of them in this section.

Two situations modification. We begin with an interesting construction illustrated in (35a). This construction states (1) that two situation expressions (that belong to certain categories; e.g. one of them is a wearing situation) can combine, (2) that the meaning of the whole is the conjunction of the meaning of the parts, (3) that the index of the whole is the index of one of the expressions, and (4) that the content of each expression must include a shared argument. A representation of the construction is provided in (36). Readers might wonder why we do not analyze the text in (35a) as a sequence of two independent sentences. We have two reasons for tentatively assuming that the kind of sentences illustrated in (35a) exemplifies a stored pattern, i.e. a construction. First, the order of clauses would be pragmatically odd if the two clauses did not form a construction, since what is shiny is only introduced in the second clause (see the oddity of the English discourse *They_i are shiny. She is wearing shoes_i.*) Second, it seems this pattern is restricted to a couple of semantic classes of relations, in particular wearing relations (among perhaps a few others), a restriction that seems incompatible with the assumption that we are dealing with two independent clauses.

- (35) a. Kwahiká teyostaláthe? teyakohtáli?
 ‘She’s wearing shiny shoes.’ (lit. really it’s shiny she’s wearing shoes)
 (Georgina Nicholas, The Flirt, recorded 1980)

b. shiny'(s', [y]) ∧ shoes'(y) ∧ wear'([s], x, [y])

(36)

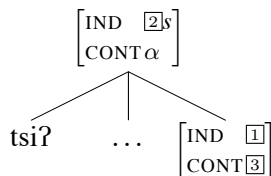


3.7 “Adjunct”-clauses

Until now, the semantic import of the constructions we have encountered was to identify variables in the semantic content of daughters and conjoin the semantic content of the daughters or embed the semantic content of one daughter into part of the content of another daughter (i.e., the semantic content of one daughter is an argument of the content of another daughter). We now consider a construction that adds a specific predication on top of the predications contributed by each of the daughters.

Adding the relation between situations. We define a *because-cl* and other kinds of “adjunct” clauses as an expression that describes a causal relation between two situations one of which is specified by the situation expression that is part of the *because-cl*. (37) represents the construction and sentence (38a) is an example of the construction.

(37)



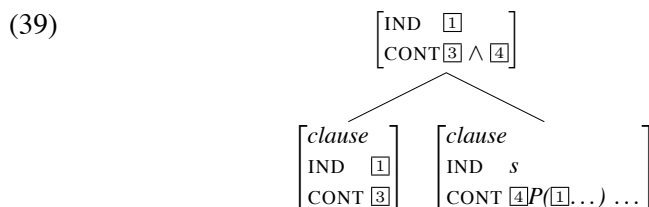
$\alpha = \text{causal-rel}([2], s', [1]) \wedge [3]$

(38) a. ya wét kyuhte wí yakotyánlustákhwa? ká slet, né tsi? tho lawΛheyú úhka? ok.
 ‘it was kind of like a haunted car, **because someone died** there.’ (Norma Kennedy, A Haunted Car, recorded 2010)

b. cause'([s], s₁, s₂) ∧ die'(s₂, x)

Clausal situation modifiers. The last construction we discuss in this paper puts together two clauses that describe situations. The construction states (1) that a situation expression can consist of two clauses that describe situation expressions, (2) that the content of the whole is the conjunction of the content of the two situation

expressions, (3) that the index of the whole is the index of one of them (let us call it *the main clause*), and (4) that the index of the main clause is an argument of part of the content of the other clause. The construction is represented in (39) and an example is provided in (40a).



- (40) a. ya wét kyuhthe wí yakotyánlustákhwa? ká'slet, né tsi? tho lawAheyú úhka? ok.

'it was kind of like a haunted car, because someone died there.'

(Norma Kennedy, A Haunted Car, recorded 2010)

- b. $\text{haunted}(\boxed{s_1}, y) \wedge \text{car}'(y) \wedge \text{cause}'(s, \boxed{s_1}, s_2) \wedge \text{die}'(s_2, x)$

4 The fun aspects of the compositional and formal syntax of Oneida

The previous section described the major Oneida constructions we know of; the constructions we omitted for reasons of space have similar characteristics. The crucial aspect of the descriptions of these constructions is that no syntactic features were needed to model all the constructions we listed and get the semantics right. The fragment of Oneida we described thus constitutes a good example of what compositional syntax looks like. Interestingly, the number of constructions we needed was not that numerous. The fact that in the absence of syntactic selection we could not rely on some very general constructions (Head-Complement; Head-Subject, ...) did not lead to a proliferation of constructions, as one might have feared. The constructions were also “minimal” and included nothing but the construction’s semantics, except in the case of *tsi?* clauses: Specifying the index and content of the combination of expressions, identifying variables across the contents of the combining expressions, adding a predication constructionally in a few cases was all that was needed. But Oneida syntax also requires some idiosyncratic ordering constraints, some a little odd. Some simple linear constraints are stated in (41)-(43). It should be rather obvious how such constraints can be stated within HPSG. But Oneida also includes a large array of linear order constraints associated with particles. Because particles and their orders are an important characteristic property of Oneida, the rest of this section is devoted to a brief overview of particles in Oneida.

- (41) A word that introduces a question variable (Gronendijk and Stockhoff 1997) must be clause initial.
- (42) An argument clause must follow the situation expression it further specifies an argument of.
- (43) In a *tsi?*-cl, *tsi?* must be initial.

Particles in Oneida are a *morphological class*, the class of uninflected words. In Michelson and Doxtator (2002), there are around 170 particles, 435 noun stems, and 2775 verb stems. Particles belong to various semantic types, including referring expressions (first person and second person pronouns, for example, are particles). Particles need not be utterance- or clause-initial, but they can “bunch together.” There are 165 distinct particles in Michelson, Kennedy, and Doxtator (To appear) and there are 2,059 *distinct* utterance-initial sequences of particles (it should be kept in mind that particles need not be utterance- or clause-initial, though). The order of particles is not arbitrary, but it is particularly complex. First, entity expressions and situation expressions can both occur between particles. Second, in the 2,059 distinct sequences of utterance-initial particles in Michelson, Kennedy, and Doxtator (To appear) there are on average for *each* particle, 10 particles that *only* occur before it, 10 particles that *only* occur after it, and 12 particles that occur *both* before and after it. Third, a very preliminary examination of particle orders suggests that scope follows a left-to-right linear order for particles whose meaning is scope-sensitive. This is what we model in (24). Now, given the number of possible combinations of order of particles and entity or situation expressions, this conclusion is provisional. But looking at the scope of all sequences of utterance-initial particles that include the negative particle *yah*, it appears correct. Fourth, “discourse particles” must be excluded from any order-sensitive scopal relations, as their semantic type should make them appear before, for example, particles that denote propositional operators, but they need not precede propositional operators (and sometimes are required to follow them). Fifth, some particles are complex (think compounds) in that the meaning they convey require *two* phonological words. These words occur in a strict order, but in some cases the two components need not be adjacent (see (44) for *úhka?* ... *ok* ‘someone’). A statement of the constraint needed to model the strict ordering required by *úhka?* ... *ok* (and other similar words) is provided in (45).

- (44) **úhka?** ki? **ok** uhte wí: luwa?áshΛ? thikÁ,
‘they stabbed **someone**,’ (i.e. someone got stabbed) (Norma Kennedy, A Haunted Car, recorded 2010)
- (45) If *úhka?* is a daughter and its content is that of an animate indefinite, *ok* must also be a daughter and must follow it.

5 Discussion

The goal of this paper was to distinguish two kinds of constraints syntactic rules can include, constraints on how to compose the meaning of two or more expressions and constraints on the form of combining expressions. Our current syntactic vocabulary and syntactic thinking is built around the second, formal, kind of constraints and particularly concepts of syntactic selection and percolation of categorial/selectional information. Without selection and category information, syntactic features are almost useless and syntax reduces (almost exclusively!) to what we call *compositional syntax*. The bulk of this paper described the broad outlines of the syntax of a language that comes close to being a strictly compositional syntax language and whose formal syntax does not require the introduction of syntactic features. In this paper's closing paragraphs, we want to place our work on Oneida in the larger context of recent work on the architecture of grammars.

Jackendoff and Wittenberg (2014) describes grammars without syntax of increasing complexities. In their terminology, Oneida would exemplify a language with a recursive phrase grammar, but that is not a 'fully complex' language, as it does not have syntax in their sense (although it is quite complex morphologically, see Lounsbury 1953 and Koenig and Michelson 2015). From their perspective, the sketch of Oneida we present constitutes a demonstration of how much you can do 'without syntax.' To some extent, the difference between our way of describing Oneida and how they would describe it is terminological: What we call *compositional syntax* corresponds to their semantic structure cum interface rules between phonology and meaning, and what we call *formal syntax* corresponds to syntax proper in their approach. But slightly more than terminology is at stake here.

Jackendoff and Wittenberg's approach is implicitly and at times explicitly, teleological: Their grammatical hierarchy is meant to correspond partly to the evolution of language, as their discussion of creoles, homesigning, and young children's grammars makes clear. There is little doubt that languages that consist only of one-word utterances or two-word utterances are simpler than English. But we do not believe that the difference in complexity between compositional syntax and formal syntax is on a par with differences in complexity between two-word grammars and recursive grammars. If we define grammatical expressiveness as the range of semantic combinations licensed through grammatical means, Oneida grammatical expressiveness is roughly on a par with English. It does not leave much more to pragmatic enrichment than English does. Oneida's lack of formal syntax merely means that it lacks the crud that formal syntax has grafted on compositional syntax by vagaries of history. For us, what Oneida syntax illustrates is not a less complex grammatical system (one without syntax), but rather that what linguists typically think of as syntax is actually not the essential role of syntax in human languages. Compositional syntax is what is essential to human linguistic abilities; syntactic features are not, even though they are needed in the overwhelming majority of cases and an important aspect of syntactic theory.

References

- Ackerman, F. (1992). Complex predicates and morpholexical relatedness: Locative alternation in Hungarian. In I. Sag and A. Szabolcsi (Eds.), *Lexical Matters*, pp. 55–83. Stanford: CSLI Publications.
- Bach, E. (1976). An extension of classical transformational grammar. In *Problems of Linguistic Metatheory (Proceedings of the 1976 Conference)*, Michigan State University, pp. 183–224.
- Baker, M. (1996). *The Polysynthesis Parameter*. Oxford: Oxford University Press.
- Bickel, B. (2011). Grammatical relations typology. In J. J. Song (Ed.), *The Oxford Handbook of Language Typology*, pp. 399–444. Oxford: Oxford University Press.
- Boas, H. and I. Sag (Eds.) (2012). *Sign-Based Construction Grammar*. Stanford, CA: CSLI Publications.
- Chafe, W. (In press). *A Grammar of the Seneca Language*. Berkeley: University of California Press.
- Groenendijk, J. and M. Stokhof (1997). Questions. In J. Van Benthem and A. Ter Meulen (Eds.), *Handbook of Logic and Language*, pp. 1055–1124. Cambridge, MA: MIT Press.
- Jackendoff, R. and E. Wittenberg (2014). What you can say without syntax: A hierarchy of grammatical complexity. Tufts University.
- Kamp, H. (1981). A theory of truth and semantic representation. In A. G. Groenendijk, T. M. Janssen, and M. B. Stockhof (Eds.), *Formal Methods in the Study of Language*, pp. 1–41. Amsterdam: Foris.
- Kamp, H. and U. Reyle (1993). *From Discourse to Logic*. Dordrecht: Kluwer.
- Kathol, A. (1999). *Linearization-based Syntax*. Oxford: Oxford University Press.
- Klein, E. and I. Sag (1985). Type-driven translation. *Linguistics and Philosophy* 8, 163–201.
- Koenig, J.-P. and K. Michelson (2010). Argument structure of Oneida kinship terms. *International Journal of American Linguistics* 76, 169–205.
- Koenig, J.-P. and K. Michelson (2012). The (non)universality of syntactic selection and functional application. In C. Piñón (Ed.), *Empirical Studies in Syntax and Semantics*, Volume 9, pp. 185–205. Paris: Centre National de la Recherche Scientifique.

- Koenig, J.-P. and K. Michelson (2014). Invariance in argument realization: The case of Iroquoian. *Language*.
- Koenig, J.-P. and K. Michelson (2015). Morphological complexity à la Oneida. In M. Baerman and G. Corbett (Eds.), *Understanding and Measuring Morphological Complexity*. Oxford: Oxford University Press.
- Lounsbury, F. (1953). *Oneida Verb Morphology*. *Yale University Publications in Anthropology* 48. New Haven, CT: Yale University Press.
- Michelson, K. (2011). Iroquoian languages. In M. Aronoff (Ed.), *Oxford Bibliographies Online: Linguistics*. New York: Oxford University Press.
- Michelson, K. and M. Doxtator (2002). *Oneida-English/English-Oneida Dictionary*. Toronto: University of Toronto Press.
- Michelson, K., N. Kennedy, and M. Doxtator. (To appear). *Glimpses of Oneida Life*. Toronto: University of Toronto Press.
- Mithun, M. (1987). Is basic word order universal? Grounding and coherence in discourse. In R. Tomlin (Ed.), *Typological Studies in Language*, Volume 11, pp. 281–328. Amsterdam: John Benjamins.
- Peters, S. and D. Westerståhl (2006). *Quantifiers in Language and Logic*. Oxford: Clarendon Press.
- Pollard, C. and I. Sag (1994). *Head-Driven Phrase-Structure Grammar*. Chicago, IL: Chicago University Press.
- Richter, F. and M. Sailer (2004). Basic concepts of lexical resource semantics. In *ESSLLI 2003 Course Material*, Volume 5 of *Collegium Logicum*. Kurt Gödel Society Wien.