

12

The measurement of semantic complexity: how to get by if your language lacks generalized quantifiers

LISA MATTHEWSON

12.1 Introduction

Do languages vary in their semantic complexity? And if they do, does Universal Grammar place limits on that variation? These questions are not easy to answer. While there is a substantial literature investigating syntactic complexity in human language (see the other chapters in this volume, and references therein), semantic complexity has received much less attention, especially in the formal literature. We do not yet have an accepted definition of semantic complexity, nor do we know how to measure the semantic complexity of individual constructions, let alone entire languages. Empirically, we also have a problem: We know relatively little about how languages vary in their semantics in the first place.¹

Given this state of the art, my goals in this chapter are relatively modest. My first aim is to illustrate how one goes about detecting semantic variation. I do this by means of a case study on quantification in St'át'imcets (Salish). Based on the St'át'imcets data, I argue (following Davis 2010, but contra Matthewson 1998) that languages differ significantly in the semantics of their quantifier-like elements. To be specific, some languages lack generalized quantifiers. The moral of this section is that semantic variation exists, but is not surface-visible or easy to detect.

¹ Thanks to St'át'imcets consultants Carl Alexander, ~~Gertrude Ned, Laura Thevarge, the late Beverley Frank~~ and the late Rose Agnes Whitley. Thanks to Henry Davis and participants at the workshop on Formal Linguistics and the Measurement of Grammatical Complexity for helpful discussion, and to Fritz Newmeyer and Laurel Preston for comments on an earlier version of this chapter. This research was supported in part by SSHRC grants #410-2007-1046 and #410-2011-0431.

My second goal is to ascertain whether the attested variation in quantifier semantics involves a difference in semantic complexity. While it might seem obvious that a language which lacks generalized quantifiers is less complex than a language which possesses them, I will argue that this is not necessarily the case. A range of factors need to be taken into account, including formal complexity of the relevant lexical entries, paradigm complexity, and expressive complexity. I will conclude that the St'át'imcets quantificational system is in fact less complex than that of English, but not for the reasons one might have thought.

Finally, I present a brief language-wide comparison between English and St'át'imcets, assessing (as far as is possible with current knowledge and tools) the two languages' relative semantic complexity in a range of areas, including tense, aspect, evidentials and so on. Overall, the competition is pretty much a tie, with each language balancing paradigm complexity in some areas with simplicity in others. Is this balancing a coincidence? Could some languages be, overall, semantically more complex than others? While balancing makes functional sense, I argue that Universal Grammar has no means to prevent languages from differing in overall complexity. Universal Grammar places formal and substantive constraints on meaning, but not on semantic complexity per se. Universal Grammar does not even prevent differences in effability; thus, some languages are able to express meanings which other languages cannot (von Stechow and Matthewson 2008; Deal 2011). This claim is supported by the English/St'át'imcets comparison presented here, as the languages display at least one effability difference.

12.2 How do languages vary in their semantics?

Before we can ask whether languages vary in their semantic complexity, we need to ascertain how languages vary in their semantic properties, period. This is an area about which surprisingly little is known. Although cross-linguistic formal semantic research has been increasing rapidly since the 1990s, there is no accepted theory of what is universal and what varies in semantics. There is not even any active debate about what limits are placed on semantic variation. Instead, we have on the one hand isolated formal analyses of particular phenomena in different languages (often only a handful of languages), and on the other hand, large-scale typological research which relies on sources whose data were not gathered for the purpose of testing empirical hypotheses about semantics (see for example Dryer and Haspelmath 2011). As argued in [Matthewson *et al.* \(2012\)](#), the latter type of research usually fails to provide useful results for semanticists, as it offers only incomplete information about meaning. [Matthewson *et al.*](#) argue that what is needed is formal cross-linguistic research on as many languages as possible, testing concrete hypotheses, leading eventually to a formal semantic typology.

Exciting work is already going on in this vein, which will eventually lead us to an informed picture of universality and variation in semantics.² For now, we can say that there is restricted cross-linguistic variation in the inventory and lexical semantics of functional morphemes such as aspects, tenses, evidentials, modals, quantifiers, and determiners. The variation may be non-trivial, but it is not without limit or constraint. Semantic variation seems to be usually (but not always) non-parametric, and often consists of differences in which of a set of core building blocks of meaning are utilized, and/or in how finely languages lexically divide up the same semantic space (see von Stechow and Matthewson 2008 for discussion).

Let's turn to our case study on quantifiers.

12.3 Quantifiers

In this section I present a case study in semantic variation, that of quantifiers. Quantification is relatively well-studied from a formal, cross-linguistic perspective: See, for example, the collections in Bach *et al.* (1995), Matthewson (2008), and Keenan and Paperno (2012). Here I do not attempt to overview the state of the art, but concentrate on one specific empirical question, namely, whether some languages lack generalized quantifiers (GQs), and I look at just two languages, English and St'át'imcets. The question is of relevance to complexity because a language which lacks GQs may be less complex than a language which possesses them.

The answer to the empirical question (following Davis 2010) is yes, some languages do lack generalized quantifiers. A bonus moral of the discussion will be that one cannot establish the absence of a certain kind of quantifier by looking superficially, or by relying on translations or morphological make-up (cf. Everett 2005 on Pirahã). As for the complexity question, I will argue that a language which lacks GQs does *not* necessarily have a simpler quantification system than one which possesses them. In the case of St'át'imcets, however, there happens to be a lack of paradigm complexity in the quantificational system.

First, we need a definition of generalized quantifiers, so we know what to test for. A GQ is a noun phrase which enforces a particular relation between two sets, usually provided by a common noun and a VP. In (1a), for example, we see that *every girl* enforces a subset relation. *Every girl dances* asserts that the set of girls is a subset of those who dance. *Most girls* in (1b) enforces a majority relation. *Most girls dance* asserts that more girls dance than don't dance.

- (1) a. $[[\textit{every girl dances}]] = 1$ iff $[[\textit{girl}]] \subseteq [[\textit{dance}]]$
 b. $[[\textit{most girls dance}]] = 1$ iff $|[[\textit{girl}]] \cap [[\textit{dance}]]| \geq |[[\textit{girl}]] - [[\textit{dance}]]|$

² Recent examples include Cover (2011), Deal (2011), Tonhauser (2011), McKenzie (2012).

Noun phrases of this type, which enforce a relation between two sets, are assumed not to be analyzable as denoting either individuals or predicates. Thus, a noun phrase of the form *every N* can neither be analyzed as being of type *e*, nor as being of type $\langle e, t \rangle$ (see e.g. Heim and Kratzer 1998). Instead, *every N* must be of type $\langle \langle e, t \rangle, t \rangle$, denoting a function from sets/one-place predicates to truth values. We call noun phrases of type $\langle \langle e, t \rangle, t \rangle$ ‘generalized quantifiers’ (see Barwise and Cooper 1981, Szabolcsi 2010, among many others).

When asking whether some languages lack GQs, what matters is whether the language has noun phrases which *must* be analyzed as GQs, not noun phrases which *can* be analyzed as GQs. Thus, we are interested in ‘essentially quantificational’ NPs, in the sense of Partee (1995). The reason for this is that any noun phrase *can* be analyzed as a GQ, including even proper names (Montague 1974). For example, the name *Henry* can be viewed as denoting a function of type $\langle \langle e, t \rangle, t \rangle$; the function returns the output ‘true’ for all sets of which Henry is a member, and false for all sets of which he is not a member.

How can we detect essentially quantificational noun phrases empirically? Two common diagnostic properties for GQs are proportional readings, and scopal ambiguities with respect to other scope-bearing operators. With regard to proportional readings, the English GQ *many A* has a reading where ‘*Many A B*’ requires a large proportion of As to B. This reading is facilitated by a partitive structure, as in (2). If 25 out of 30 students passed the test, then (2) is fine. It is decidedly more questionable if 25 out of 100 did.

- (2) Many of the students passed the test.

Many also has a cardinal reading—facilitated in (3) by a *there*-insertion structure—in which it suffices for there to be a large number of Ns. However, this reading does not require a quantificational analysis, as noted by Milsark (1974), Partee (1988), Szabolcsi (2010), among others.

- (3) *Context: There are 40,000 students at UBC. Yesterday there was a protest rally and 2,000 students turned up.*

There were many students at the rally yesterday.

The second diagnostic for GQs is their giving rise to scopal ambiguities, as illustrated in (4).

- (4) All the children built a raft.
 Surface scope reading: For each child *x*, there is a raft *y*, and *x* built *y*. (different rafts)
 Inverse scope reading: There is a raft *y*, and all the children built *y*. (only one raft)

As discussed by [Matthewson et al. \(2012\)](#), proportional readings and scopal ambiguity are both necessary conditions for an essentially quantificational noun phrase, and

together, they constitute sufficient evidence. That is, a noun phrase which displays proportional readings and participates in scopal interactions must be analyzed as a GQ.³ An initial hypothesis that a language possesses GQs is therefore falsifiable by determining that noun phrases in that language lack these properties. In the next subsection we attempt to find such evidence in St'át'imcets.

Let us turn now to evidence ~~evidence~~ that St'át'imcets lacks GQs. The discussion relies in large part on the insights of Davis (2010). St'át'imcets (a.k.a. Lillooet) is a highly endangered Northern Interior Salish language spoken in the southwest interior of British Columbia. The methodology of data collection for the material presented here was one-on-one intensive fieldwork. Standard elicitation methodologies were used, including primarily a Felicity Judgment Task, in which speakers judge the acceptability of sentences in specific discourse contexts (Matthewson 2004, 2011).

St'át'imcets possesses several elements which appear quantifier-like, including *tákem* 'all', *zízzeg* 'each', *sáq'ulh* 'half', *cw7it* 'many', *k'wík'wena7* 'few' and *t'qwaw's* 'both'.⁴ Matthewson (1998) argues that although these elements differ from English quantifiers in that they must appear inside a DP which also contains a determiner, the entire resulting phrase denotes a generalized quantifier. One piece of evidence for this claim is the presence of proportional readings. Consider (5), containing *cw7it* 'many':⁵

- (5) *úxwal'* [i=cw7ít=a plísmen]
 go.home [DET.PL=many=EXIS policeman]
 'Many of the policemen went home.' (Matthewson 1998: 304)

(5) is rejected in a context where 25 out of 100 policeman go home, but accepted if 25 out of 30 policeman go home. Based on a range of similar data, Matthewson (1998) argues that *cw7it* (as well as *k'wík'wena7* 'few') has *only* a proportional reading when it appears inside DP. This appears to be evidence that the resulting noun phrases are essentially quantificational.

However, Davis (2010) shows that phrases containing elements like *cw7it* do *not* satisfy the other diagnostic for GQ-hood: they do not participate in scopal interactions.⁶ He argues on the basis of this fact that St'át'imcets lacks generalized quantifiers.

³ On its own, scopal behaviour is not a sufficient condition, since elements other than GQs (e.g. negation, modals, intensional verbs) participate in scopal ambiguities. Matthewson (1998) assumed (following much prior literature) that proportional readings alone were a sufficient condition for generalized quantifierhood. We will see below that this is not correct.

⁴ St'át'imcets data are presented in the orthography designed by Jan van Eijk and used by community members; see van Eijk and Williams (1981). The symbol 7 represents a glottal stop.

⁵ Morpheme glosses not covered by the Leipzig Glossing Rules include DETC = deictic; DEON = deontic; EMPH = emphatic; EPIS = epistemic; EXIS = assertion of existence; RED = redirective applicative; STAT = stative; TEMP = temporal.

⁶ Matthewson (1999) showed that sentences containing one quantified phrase and one plain indefinite do not display scopal ambiguity. Davis takes the investigation a step further by looking at sentences containing *two* overtly quantified phrases.

246 *The measurement of semantic complexity*

Consider (6). Here, we have two apparently quantificational noun phrases, ‘all the children’ and ‘half the books’.

- (6) *Context: Four children are meant to read four books over the summer holidays.*
 [tákem i=sk’úk’wmi7t=a] paqwal’ikst-mín-itas [sáq’ulh i=púkw=a]
 [all DET.PL=child(PL)=EXIS] read-RED-3PL.ERG [half DET.PL=book=I
 ‘All the children read half the books.’

Davis argues that neither of the two potential scopal readings exists for (6). The scenario in (7) is one in which (6) should be accepted under a wide scope reading for the subject: For each child *x*, *x* read a (different) half of the books. However, consultants reject (6) in this context.

(7)

A reads	B reads	C reads	D reads
books 1,2	books 2,3	books 3,4	books 1,4

(8) represents a context in which (6) should be accepted on an inverse scope reading: For a particular half of the books (1 and 2), each child read that half. However, consultants also reject (6) in the context in (8).

(8)

A reads	B reads	C reads	D reads
books 1,2,3	books 1,2,4	books 1,2,3,4	books 1,2

It turns out that (6) is accepted in all and only situations in which all the children read at least one book, and a *total* of two out of the four titles are read.⁷ These contexts correspond to a *cumulative* reading. As discussed by Scha (1981), cumulative readings are scopeless readings, in which a predicate relates two sets which each have a certain total size. In our example, what is required is that all the four children participated in reading, and a total of half the four books were read (but exactly who read what is left vague). Scha (1981) shows that this reading cannot be captured using an ordinary GQ analysis. See Davis (2010) for data involving other combinations of quantifiers in St’át’imcets, showing in each case that the cumulative reading is the only available reading.

Davis’s findings are quite significant: They mean that a St’át’imcets sentence containing two quantifier-like elements does not have the same semantics as its translation ‘equivalent’ in English. St’át’imcets really lacks the GQ readings of all such sentences. Notice that the absence is far from immediately obvious. Any sentence

⁷ Support for the claim that it is the total number of books which matters comes from one consultant’s comment on context (7): “No—they read all the books, so you couldn’t say they read half the books” (Davis 2010).

with just one quantifier will seem similar to its English translation. For example, the St’át’imcets version of ‘All the children ate fish’ will, just like the English version, require every contextually salient child to have eaten fish. One can only detect the absence of GQ readings by means of the targeted fieldwork done by Davis.

Aside from the absence of GQs, St’át’imcets displays another absence in the area of quantification: There are no true generics. As mentioned earlier, Matthewson (1998) shows that quantifier-like elements in Salish always co-occur with determiners. Importantly, these determiners explicitly limit the domain of quantification by placing a deictic restriction on the individuals being quantified over. For example, (9a) contains the ‘absent’ plural determiner *nelh...a*, because the people being talked about are deceased at the time of speech, while (9b) contains the ‘present’ determiner *i...a* and picks out a particular group of present people.

- (9) a. tqílh=wi7 tákem pináni7 nelh=ucwalmícw=a láku7
 almost=EMPH all TEMP.DEIC DIST.DET.PL=person=EXIS DEIC
 lil’wat7úl=a snek’wnúk’wa7-s
 Mount.Currie=EXIS relatives-3POSS
 ‘He was related to almost everyone in Mount Currie.’ (Matthewson 2005: 393)
- b. tqílh=wi7 tákem i=ucwalmícw=a láku7 lil’wat7úl=a
 almost=EMPH ALL DET.PL=person=EXIS DEIC Mt.Currie=EXIS
 snek’wnúk’wa7-s
 relatives-3POSS
 ‘He is related to almost everyone in Mount Currie.’

The explicit domain restriction which is obligatory with St’át’imcets quantification is incompatible with generic quantification, which by its very nature makes a claim about an entire class of individuals. Matthewson (1998) shows that attempts to elicit generic statements in St’át’imcets result in ordinary universal constructions, as in (10). She points out that the quantified phrase in such statements is identical in structure to quantified phrases involving a specific set of individuals, as in (11).

- (10) *tákem* i=twéw’w’et=a ama-mín-itas k=wa píx-em’
all DET.PL=boys=EXIS good-APPL-3PL.SBJ DET=IPFV hunt-INTR
 ‘All boys love hunting.’ (Matthewson 1998: 333)
- (11) *tákem* i=cácl’ep=a twéw’w’et nás=tu7 píx-em’
all DET.PL=Fountain=EXIS boys go=then hunt-INTR
 ‘All the boys from Fountain went hunting.’ (Matthewson 1998: 333)

It appears that although St’át’imcets speakers can translate English generics into their language, the sentences they produce are not truly general. Support for this is given by (12). A speaker commented about this sentence that “there’s a bunch of men there; it doesn’t pertain to all the men in the world.” “However, when asked how she *would*

refer to all the men in the world, [12] was the only way it could be done” (Matthewson 1998: 333). See also Gillon (2006) for similar evidence that the related language Skwxwú7mesh lacks generics.

- (12) léxlex s=Henry [lhél=ki=tákem=a sqáyqeycw]
 intelligent NMLZ=Henry [from=DET.PL=all=EXIS men]
 ‘Henry is the most intelligent of all the men.’ (Matthewson 1998: 333)

The discussion in this section has shown that at least one language exists which lacks essentially quantificational noun phrases. We also saw that some languages lack generic readings. The question now is what the implications of these findings are for complexity. In the next section I lay out some criteria for measuring semantic complexity, and in section 12.5 I apply the criteria to the St’át’imcets quantificational system.

12.4 What is semantic complexity?

As mentioned above, there has been relatively little formal work on semantic complexity, and there is certainly no generally accepted definition. Semantic complexity is mainly discussed in the acquisition, psycholinguistic, or processing literature, and in these contexts it is usually not explicitly defined. Here I will lay out some possible ways in which we could measure semantic complexity.

Consider a fairly standard picture of what is included in the semantic competence of a native speaker of a language. The speaker must know at least (a) the lexical entries for all the morphemes in their vocabulary, (b) the mechanisms which are used to compose meanings together, in a way which interfaces with syntax, and (c) how to form pragmatic inferences (such as presuppositions and implicatures) and interpret context-dependent meaning (indexicality, vagueness, etc.). A speaker who knows all this will know, for any utterance in his or her language, its truth conditions, along with any pragmatic inferences it gives rise to in particular discourse contexts.

This picture points to a few different levels at which we could measure complexity. We could calculate the complexity of the formal representation of lexical entries or of the truth conditions of entire utterances. Another possibility would be to calculate the complexity of everything that goes into creating truth-conditional meaning (including composition mechanisms). We could also measure the paradigm complexity of particular areas of the grammar, by checking whether more or fewer semantic distinctions are expressed in the same semantic domain. We could include pragmatic inferences as part of complexity, and finally we could consider overall expressive complexity (whether some languages lack certain meanings altogether).

Let us first consider how to measure the complexity of formal representations of truth conditional content. There are two obvious ways in which this could be done: in terms of the length of the description (cf. Miestamo 2008, among many others), or in

terms of semantic strength.⁸ An example of the first approach comes from Gennari and Poeppel (2003), who assume that verbs denoting externally caused events (such as *break*) are more complex than verbs denoting internally caused events (such as *grow*). The reasoning is that “externally caused verbs, but not internally caused ones, conceptually require two participants” (Gennari and Poeppel 2003: B28) and hence their lexical entries are longer.

One problem with the length criterion is that there are often many different ways to write logically equivalent semantic formulas. For example, the two Russellian lexical entries for the English definite article in (13) are different lengths, but are equivalent.

- (13) a. $[[\textit{the}]]$ = $\lambda f_{\langle \textit{et} \rangle} . \lambda g_{\langle \textit{et} \rangle} . \exists x [\forall y [f(y) \leftrightarrow x = y] \ \& \ g(x)]$
 b. $[[\textit{the}]]$ = $\lambda f_{\langle \textit{et} \rangle} . \lambda g_{\langle \textit{et} \rangle} . \exists x [f(x) \ \& \ \forall y [f(y) \rightarrow x = y] \ \& \ g(x)]$

Another example is given in (14). These two possible denotations for *most*, which differ slightly in length but which are truth-conditionally equivalent, are discussed by Hackl (2009).

- (14) a. $[[\textit{most}]](A)(B) = 1$ iff $|A \cap B| > |A - B|$
 b. $[[\textit{most}]](A)(B) = 1$ iff $|A \cap B| > \frac{1}{2} |A|$

If the correct truth conditions can be obtained from two different formal analyses, we cannot measure the complexity of an element based on its formal properties until we have found a way to choose between the equivalent formal analyses.⁹

What about the ‘semantic strength’ criterion? Gennari and Poeppel (2003) also use this one. They assert that complex elements are those which have ‘more entailed properties’ than simpler ones. Thus, eventive verbs are more complex than stative verbs, because ‘eventive verbs entail simpler conceptual units such as CAUSE, BECOME or CHANGE and resulting STATE, corresponding to the event’s internal dynamics they denote, while stative verbs lack any such causal entailments’ (Gennari and Poeppel 2003: B28). However, under an analysis in which eventive verbs contain operators like CAUSE or BECOME, the lexical entries for eventive verbs are also longer than those for stative verbs. In this case, Gennari and Poeppel’s second criterion (number of entailments) can be reduced to the first (length).

The length criterion and the entailment criterion do not always give the same results, however. Consider numerals. The lexical entries for all numerals are the same length, yet (15a) entails (15b) and (15c), (15b) only entails (15c), and (15c) entails none of the others. It seems unlikely that we want to conclude that *four* is more complex

⁸ An element A is stronger than an element B if a sentence containing A asymmetrically entails a parallel sentence containing B.

⁹ Hackl (2009) uses processing evidence to distinguish between the two potential analyses in (14). He argues that only (14a) is correct for *most*, while (14b) represents the meaning of *more than half*.

than *three*. This suggests that semantic strength does not (alone) determine complexity.

- (15) a. There are **four** cabbage trees in the park.
 b. There are **three** cabbage trees in the park.
 c. There are **two** cabbage trees in the park.

Another case in which length and semantic strength do not match up involves the English past tense compared with the St’át’imcets non-future tense (as analyzed by Matthewson 2006a). As seen in (16–17), the formulas contain the same number of symbols, and thus are equally complex according to the first criterion. However, a semantic strength criterion would classify the English past as more complex, since a sentence containing a dedicated past tense entails one containing a general non-future tense, but not vice-versa.

- (16) $[[\text{PAST}_i]]^{\text{st}}$ is only defined if $g(i) < t$. If defined, $[[\text{PAST}_i]]^{\text{st}} = g(i)$.
 (adapted from Kratzer 1998)

- (17) $[[\text{NON-FUT}_i]]^{\text{st}}$ is only defined if $g(i) \leq t$. If defined, $[[\text{NON-FUT}_i]]^{\text{st}} = g(i)$.

Although semantic strength per se is probably not a good complexity measure, semantic strength often correlates with another measure, namely, paradigm complexity. Paradigm complexity embodies the idea that for the same semantic space S , a language which encodes sub-divisions within S is more complex than a language which does not (McWhorter 2001; Kusters 2003; Miestamo 2006b, 2008; Lindström 2008, among others). By this criterion, the English tense system is more complex than the St’át’imcets one, because there are more tenses in English than there are in St’át’imcets. Notice that a system which encodes more distinctions within a given domain will also tend to have elements with more specific (hence semantically stronger) lexical entries.

So far we have dealt only with truth conditional content, and have ignored complexity of pragmatic inferences.¹⁰ Some pragmatic phenomena, such as presuppositions and conventional implicatures, straightforwardly add formal complexity. For example, the Fregean analysis of the English definite article in (18) contains the underlined presuppositional material. This is the information which must be common knowledge in the discourse situation at the time of utterance.

- (18) $[[\textit{the}]]$ = $\lambda f_{\langle e, t \rangle}$ and there is exactly one x such that $f(x) = 1$. the unique y such that $f(y) = 1$ (Heim and Kratzer 1998: 75)

¹⁰ I set complexity of composition mechanisms completely aside, as hardly any research has been done into whether languages can vary in their composition rules (see Chung and Ladusaw 2004 for one exception, and Gil 2005b for a more radical view of how languages can vary in composition strategies).

On the other hand, conversational implicatures do not add formal complexity, since they are not represented in the truth conditions.¹¹ Yet intuitively, an utterance for which an implicature is generated should count as more complex than an utterance without one. Some researchers are attempting to detect the added complexity of conversational implicatures using processing evidence (Bott and Noveck 2004; Breheny *et al.* 2006, among others). The idea is that conversational implicatures must be calculated, and this user complexity should have an effect on processing times.

However, research into the processing of pragmatic inferences is in its infancy, as pointed out by Sedivy (2007: 481). Sedivy writes (2007: 493) that “We do not yet have a clear picture of the hearer’s processing costs of computing implicatures; while some studies show a measurable processing cost, others suggest that Gricean inferencing can be used without detectable cost.”

Another type of pragmatic ‘work’ which may add to complexity involves the inferences required by semantically underspecified or general elements. For example, we saw above that the St’át’imcets non-future tense is semantically weaker and paradigmatically simpler than the English past tense. However, precisely because the St’át’imcets non-future tense is more general, it may impose more demands on the hearer (who does not have the temporal reference narrowed down for them in the semantics). And indeed, semantically general elements can be shown to have a processing cost (Breedin *et al.* 1998).¹² Again, though, research is just beginning, and is heavily concentrated on English.

When thinking about the role of processing evidence, it is important to remember that merely measuring processing times (without having any formal analysis of the relevant constructions) would not lead to interesting results. This is particularly so since our different complexity measures can make different predictions for the same elements. A long (complex) description may correspond to a semantically strong (complex) truth-conditional contribution, which however requires little (simple) pragmatic inferencing work. There are interesting questions about what we predict for the processing of—or what processing might tell us about—cases where the formal criteria give conflicting results.

The final aspect to semantic complexity I will discuss has been called ‘expressive complexity’ by Yanovich (2012). A system A is expressively more complex than a system B if there are meanings which A can express and B cannot. Expressive complexity is different from paradigm complexity, which compares systems in which the relevant elements cover the *same* semantic space, but make more or

¹¹ At least in the (neo-)Gricean tradition. Some recent work argues that scalar implicatures are calculated compositionally rather than post-compositionally; see, for example, Chierchia *et al.* (2012).

¹² Jackendoff and Wittenberg (this volume, chapter 4) make a similar point when they write that “simpler grammars [...] put more responsibility for comprehension on pragmatics and discourse context.”

fewer distinctions within that space. For example, the St’át’imcets non-future tense covers the same semantic space as the English past plus present tenses. But what if a language had no way to invoke past reference times? That would be a lack of expressive complexity when compared to either English or St’át’imcets. See Yanovich (2012) for other examples of potential ways in which languages could vary in expressive complexity.

In this section I have outlined several possible ways to measure semantic complexity, including the length of the formal description, paradigm complexity, the extent of pragmatic inferencing required, and expressive complexity. Armed with these, we now turn to the question of whether English and St’át’imcets differ in the complexity of their quantificational systems.

12.5 Is quantification in St’át’imcets less complex than in English?

Let’s start by comparing the formal analyses of two particular quantifier denotations in the two languages. (19) gives a typical denotation for the English quantifier *half*, which creates a generalized quantifier. According to this denotation, *half* takes two sets, one provided by the common noun phrase and one provided by the verb phrase. It requires the cardinality of the intersection of the NP and VP sets to be half the cardinality of the NP set.

$$(19) \quad [[\textit{half}]] = \lambda P_{\langle \text{et} \rangle} . \lambda Q_{\langle \text{et} \rangle} . |P \cap Q| = \frac{1}{2} |P|$$

As shown above, the St’át’imcets equivalents of quantifiers like *half* do not create GQs. Davis’s (2010) denotation for *sáq’ulh* ‘half’ is given in (20).

$$(20) \quad [[\textit{sáq’ulh}_i]]^g = \lambda P_{\langle \text{et} \rangle} : |(g(i))(P)| = \frac{1}{2} |P| . (g(i))(P)$$

Sáq’ulh introduces a choice function, whose value is contextually given via the assignment function *g*. The choice function applies to the set denoted by its sister (a plural DP), and picks out a subset of that set; the entire phrase refers to the subset chosen by the choice function. The condition before the period in (20) requires the cardinality of the subset chosen to be half of that of the superset; if this fails to hold, the result is undefined.

In spite of being non-GQ-creating, the denotation for St’át’imcets *sáq’ulh* is not obviously simpler than that for English *half*. The denotation for *sáq’ulh* is not shorter, and it contains a presupposition lacking from the denotation for *half*. However, St’át’imcets *is* simpler if we adopt a metric of *type-complexity*. GQ-creating quantifiers are of type $\langle \langle e, t \rangle, \langle \langle e, t \rangle, t \rangle \rangle$; they take two set arguments. The St’át’imcets quantifiers in Davis’s analysis are of the simpler type $\langle \langle e, t \rangle, \langle e, t \rangle \rangle$, as they take one set argument, and return a smaller set. The metric of type-complexity was not discussed so far, but see for example Landman (2006) on issues of type-complexity and the drive for simpler types in certain areas of the grammar.

In terms of paradigm complexity, St’át’imcets does have a simpler quantificational system than English, because it possesses a smaller set of quantifiers than English does. As discussed by Matthewson (1998), St’át’imcets lacks DP-internal quantifiers corresponding to *most*, *some* or *no*. It also lacks many of the morphologically complex quantifiers that English is so good at creating, like *exactly five*, *fewer than five*, or *more than half*.

What about expressive complexity? We have seen that there are readings—the GQ scopal readings—that St’át’imcets quantifiers simply cannot express. Instead, only cumulative readings are possible. To be sure, there are ways of getting across the scopal readings, but they involve extensive paraphrasing. For example, a St’át’imcets speaker who wants to express wide-scope distributive universal quantification (each student read a potentially different book) can list the names of each student.¹³ This need for time-consuming paraphrase suggests that St’át’imcets is less expressively complex here.

However, we mustn’t forget to ask the question the other way around: Does English lack readings the St’át’imcets quantificational constructions have? The answer is no and yes. Obviously, English allows cumulative readings. Scha’s famous examples show that, as in (21).

- (21) Exactly 600 Dutch companies use 5000 American computers.

However, at least in my judgment, cumulative readings are not easily available in English in many cases where St’át’imcets allows them. Take for example (22). In the context given, the St’át’imcets sentence is fully acceptable, but the English sentence sounds odd.

- (22) *Context: Alvin and Betty were supposed to read four books each over the summer holidays. Alvin read books 1 and 2; Betty read books 3 and 4.*

[t’qaww’s i=sk’wemk’úk’wmi7t=a] paqwal’ikst-mín-itas [tákem i=púkw=a] 
 [both DET.PL=children=EXIS] read-RED-3PL.ERG [all DET.PL=book=EXIS]
 ‘Both the children read all the books.’

The infelicity of the English sentence in a context which supports the cumulative reading suggests that English sentences lack readings which St’át’imcets ones have, and therefore is expressively less complex in its own way. Just like St’át’imcets does with the scopal readings, English has to paraphrase, saying something like ‘The two children together read four books in total.’

What about the absence of true generics in St’át’imcets? That represents a lack of expressive complexity vis-à-vis English. As discussed above, St’át’imcets speakers have sentences which they can use in generic contexts, but the constructions they are



This is similar to a strategy used in Pirahã, according to Everett (2005: 624).

forced by the grammar to use do not express fully general statements. This is similar to the situation described by Deal (2011) for Nez Perce modals. Nez Perce possesses a single non-epistemic modal *o'qa*. Deal argues that *o'qa* is a possibility modal, which can be used in necessity contexts, even though the truth conditions are not that of a necessity claim. In downward entailing contexts, *o'qa* always expresses negated possibility (*You can't go*), and Nez Perce speakers cannot easily express negated necessity (*You don't have to go*), having to resort to paraphrases. Deal explicitly argues that Nez Perce and English differ in effability (i.e. in expressive complexity). The same appears to be true of St'át'imcets with respect to English generics.¹⁴

In this section we have seen that although St'át'imcets lacks generalized quantifiers, we cannot automatically conclude that its quantificational system is less complex than that of languages which possess them. St'át'imcets quantifiers are formally less complex than English ones only if we take type complexity into account. St'át'imcets quantifier words themselves are not expressively less complex than English; the only lack of expressive complexity derives from a syntactic restriction which forces all quantifiers to co-occur with domain-restricting determiners. And St'át'imcets is paradigmatically less complex only because it happens to have a smaller inventory of quantifier words than English does. This last point has nothing to do with lacking GQs. A language could easily have only cumulative quantifiers and nevertheless possess a very large quantifier inventory.

12.6 Overall language comparison between St'át'imcets and English

In this section I compare St'át'imcets and English in a wider range of semantic sub-domains. I briefly introduce the ways in which the languages differ, and discuss whether the differences reflect complexity differences. The discussion in each subsection will necessarily be brief and simplified, but the process should give us a feel for whether complexity is roughly equal between these two unrelated languages.

12.6.1 Determiners

St'át'imcets determiners have more paradigm complexity than English ones, in that they encode deixis, number and referentiality. The St'át'imcets determiner system is given in (23), taken from Matthewson (1998) and adapted from van Eijk (1997).

¹⁴ Although St'át'imcets and Pirahã appear to share an absence of generics, I do not draw the same conclusions as Everett (2005) does about culture or cognition. The absence of generics in St'át'imcets does not derive from, or entail, any lack of ability to generalize, or to use intensionality or displacement. It derives from a syntactic restriction (the fact that all quantifier elements must appear within a full DP containing a determiner), and from the determiner system, which requires that all determiners encode deictic information.

(23)

	REFERENTIAL			 N-REFERENTIAL
	PRESENT	ABSENT	REMOTE	
SINGULAR	<i>ti...a</i>	<i>ni...a</i>	<i>ku...a</i>	<i>ku</i>
PLURAL	<i>i...a</i>	<i>nelh...a</i>	<i>kwelh...a</i>	

St’át’imcets does lack any equivalent of the English definite article, as all St’át’imcets determiners are indefinite (Matthewson 1998, 1999, 2009). However, this does not lead to a lack of expressive complexity in the language. Matthewson 1999 shows that in a language with only indefinite determiners, the determiners cover the same total semantic space as a language with both definite and indefinite determiners.

In terms of formal complexity, there seems to be basically a tie between the two languages, although we need the caveat that in both languages, there are a range of potential analyses to be considered. Some possible denotations for English Ds are given in (24), and two alternative denotations for the St’át’imcets singular present determiner are given in (25). At least the English definite and ~~the~~ St’át’imcets *ti...a* are in the same formal complexity ballpark: Both impose conditions which restrict the contexts of use, and both apply to a (possibly intensional) one-place predicate and output an individual.

- (24) a. $[[a]] = \lambda f_{\langle et \rangle} . \lambda g_{\langle et \rangle} . \exists x [f(x) = 1 \ \& \ g(x) = 1]$
 b. $[[the]]^g = \lambda f_{\langle \langle s, e \rangle, \langle s, t \rangle \rangle} . \lambda s : \exists ! x f(\lambda s.x)(s) = 1 . \text{ix } f((\lambda s'.x)(s)) = 1$ (Elbourne 2005: 51)

- (25) a. $[[ti \dots a_k]]^{g,c} = \lambda f_{\langle et \rangle} : (g(k))(f)$ is proximal to the speaker in *c* . $(g(k))(f)$
 (adapted from Matthewson 2001)
 b. $[[ti \dots a]]^{g,c} = \lambda f_{\langle \langle s, e \rangle, \langle s, t \rangle \rangle} . \lambda s : \exists ! x f((\lambda s.x)(s_o)) = 1$ where s_o is proximal to the speaker in *c* . $\text{ix } f(x)(s_o) = 1$ (Matthewson 2009, adapted from Elbourne 2005)

The results so far, including the results from our case study on quantifiers, are listed in (26). The language listed in each cell is the one which is more complex in the relevant area.

(26)

	FORMAL COMPLEXITY	PARADIGM COMPLEXITY	EXPRESSIVE COMPLEXITY
quantifiers	tie	English	English
determiners	tie	St’át’imcets	tie

256 *The measurement of semantic complexity*

12.6.2 *Tense*

As we saw in section 12.4, St’át’imcets does not distinguish present from past tense, but has a single non-future tense morpheme, which happens to be phonologically covert. The non-future tense is shown in (27–28), and the denotation is repeated in (29). The tense picks out the contextually salient reference time, which must be non-future.¹⁵

(27) táyt-kan
 hungry-1SG.SBJ
 ‘I was hungry / I am hungry / * I will be hungry.’

(28) k’ác-an’-lhkan
 dry-TR-1 SG.SBJ
 ‘I dried it / I am drying it / * I will dry it.’

(29) [[NON-FUT_i]]^{g,t} is only defined if $g(i) \leq t$. If defined, [[NON-FUT_i]]^{g,t} = $g(i)$.

The St’át’imcets tense system is paradigmatically less complex than that of English, but equally formally complex and equally expressively complex. The results so far are given in (30).

(30)

	FORMAL COMPLEXITY	PARADIGM COMPLEXITY	EXPRESSIVE COMPLEXITY
quantifiers	tie	English	English
determiners	tie	St’át’imcets	tie
tense	tie	English	tie



12.6.3 *Modals*

The question here is whether modals in our two languages encode modality type (the difference between epistemic interpretations, deontic interpretations, and so on), and whether they encode modal force (the strength of the quantification).

English modals lexically encode modal force, but often leave modality type up to context (Kratzer 1981, among many others). This is shown in (31). The modal *must* can receive either epistemic or deontic interpretations—modality type is free—but is

¹⁵ Future time-reference is achieved by the combination of NON-FUT with a future modal *kelh*, which Matthewson (2006b) argues is an overt counterpart of English WOLL, the element which combines with tense to give *will* and *would* (Abusch 1985). I therefore take the two languages to be paradigmatically equally complex with regard to future time reference.

always interpreted as a necessity modal, with strong modal force. *May* is similarly unrestricted with respect to modality type, but is unambiguously a possibility modal, with weak modal force.

- (31) a. She *must* be in her office. EPISTEMIC / DEONTIC ONLY NECESSITY
 b. She *may* be in her office. EPISTEMIC / DEONTIC ONLY POSSIBILITY

The situation in St’át’imcets is exactly reversed: The modals lexically encode modality type, but leave modal force up to context (Matthewson *et al.* 2007; Rullmann *et al.* 2008, Davis *et al.* 2009). Examples are given in (32–33), contrasting the unambiguously epistemic *k’a* with the unambiguously deontic or irrealis *ka*. Both of these modals (like all modals in the language) are felicitous in either necessity or possibility contexts; they do not form strong/weak pairs, like English modal auxiliaries do.

- (32) wá7=k’a s-t’al l=ti=tsítcw-s=a s=Philomena
 be=EPIS STAT-stop in=DET house-3SG.POSS=EXIS NMLZ=Philomena
 ‘Philomena must/might be in her house.’ ONLY EPISTEMIC NECESSITY / POSSIBILITY

- (33) lán=lhkacw=ka áts’x-en ti=kwtámts-sw=a
 already=2SG.SBJ=DEON see-TR DET=husband-2SG.POSS=EXIS
 ‘You must/can/may see your husband now.’ ONLY DEONTIC NECESSITY / POSSIBILITY

In terms of paradigm complexity, the languages are roughly tied. While English does have a few more modal auxiliaries than St’át’imcets has modal clitics, this is mostly due to remnants of past-tense inflection on English modals (as in *shall* vs. *should*, *will* vs. *would*, *can* vs. *could*, *may* vs. *might*). As outlined above, St’át’imcets lacks a distinction between past and present tense, a fact which holds equally for its modals. But in terms of modality type and modal force, the languages are equal in complexity, as indicated in the simplified tables in (34–35). Each language chooses one distinction to encode and one distinction to leave up to context. These tables also show that the languages have equivalent expressive complexity, in that the same total semantic space is covered.

(34)

English	NECESSITY	POSSIBILITY
DEONTIC	<i>must</i>	<i>may</i>
EPISTEMIC	<i>must</i>	<i>may</i>

(35)

St’át’imcets	NECESSITY	POSSIBILITY
DEONTIC	<i>ka</i>	<i>ka</i>
EPISTEMIC	<i>k’a</i>	<i>k’a</i>

In order to measure formal complexity, we as usual need analyses of both systems. Versions of these are given in (36–39). The first argument of the English modals is a conversational background function from worlds to sets of worlds; the conversational background is given by the context and determines the modality type.¹⁶

$$(36) \quad [[\textit{might}]]^{c,w} = \lambda g_{\langle s, st \rangle}. \lambda p. \exists w' \in g(w) [p(w') = 1]$$

$$(37) \quad [[\textit{must}]]^{c,w} = \lambda g_{\langle s, st \rangle}. \lambda p. \forall w' \in g(w) [p(w') = 1]$$

In St'át'imcets, the conversational background is lexically restricted to being either epistemic or deontic. All modals are necessity modals (i.e. they introduce universal quantification over worlds), but variable modal force is achieved by a choice function which narrows down the set of worlds being quantified over. The lexical entry for *k'a* is given in (38). *Ka* is exactly the same except the word 'deontic' replaces 'epistemic'. See Rullmann *et al.* (2008) for details.

$$(38) \quad [[\textit{k'a}]]^{c,w} \text{ is only defined if } c \text{ provides an epistemic conversational background } g \text{ and a modal choice function } f \text{ such that } f(g(w)) \subseteq g(w).$$

$$\text{If defined, } [[\textit{k'a}]]^{c,w} = \lambda f_{\langle st, st \rangle}. \lambda g_{\langle s, st \rangle}. \lambda p. \forall w' \in f(g(w)) [p(w') = 1]$$

(adapted from Rullmann *et al.* 2008)

The St'át'imcets denotations are formally more complex than the English ones. The St'át'imcets modals take more arguments (as a modal choice function is required), and the St'át'imcets modals have presuppositions not present in English. An interesting thing to think about, however, is whether the complexity of the St'át'imcets denotations has partly to do with the way predicate logic is constructed. Predicate logic gives us single symbols for both universal and existential quantification—thus making the English denotations simple—but does not provide single symbols for either epistemic or deontic modality. This adds to the point made in section 12.4 regarding the difficulty inherent in measuring formal complexity by a length criterion.

The results so far are summarized in (39).

(39)

	FORMAL COMPLEXITY	PARADIGM COMPLEXITY	EXPRESSIVE COMPLEXITY
quantifiers	tie	English	English
determiners	tie	St'át'imcets	tie
tense	tie	English	tie
modals	St'át'imcets	tie	tie

¹⁶ In the Kratzerian system (Kratzer 1981), there are actually two conversational backgrounds, a modal base and an ordering source. I am ignoring this here for the sake of simplicity.

12.6.4 *Evidentials*

Evidentials encode the type of evidence the speaker has for the claim expressed. St’át’imcets possesses a set of these, as shown in (40); see Matthewson *et al.* (2007), Matthewson (2012) for analysis.¹⁷

- (40) *k’a* inferential (any kind of indirect evidence)
- an’* perceived evidence (sensory witness of results)
- lákʷ7a* sensory non-visual
- ku7* reportative

Although English epistemic modals may all convey the evidential notion of ‘indirect evidence’ (von Stechow and Gillies 2010; Matthewson, in press), they do not encode sub-distinctions between different types of evidence. St’át’imcets is therefore both formally and paradigmatically more complex in this area. Expressively, there is no difference; English modals are compatible with all types of indirect evidence, and therefore the same total semantic space is covered by English epistemic modals and St’át’imcets evidentials.

(41)

	FORMAL COMPLEXITY	PARADIGM COMPLEXITY	EXPRESSIVE COMPLEXITY
quantifiers	tie	English	English
determiners	tie	St’át’imcets	tie
Tense	tie	English	tie
Modals	St’át’imcets	tie	tie
Evidentials	St’át’imcets	St’át’imcets	tie

12.6.5 *Presuppositions*

A presupposition is a proposition which is ‘taken for granted’ at the time the sentence containing it is uttered (Stalnaker 1974). Presuppositions thus impose contextual felicity constraints. If a presupposition is not common knowledge at the time of utterance, the hearer may express surprise by means of responses like ‘Hey, wait a minute! I didn’t know that ...’ (von Stechow 2004). In English, certain items lexically encode presuppositions; examples include the definite article, and words like *again*, *too*, and *stop*.

¹⁷ St’át’imcets evidentials encode not only evidence source, but also epistemic modality (Matthewson *et al.* 2007). Hence, the modal *k’a* discussed in the previous section is included in (40).

St’át’imcets has been argued to lack presuppositions which impose contextual felicity constraints in this way (Matthewson 2006a, 2009). Evidence for this claim includes the failure in St’át’imcets of the ‘Hey, wait a minute!’ test; there are no elements in St’át’imcets which induce surprise responses when some aspect of their meaning fails to be taken for granted by the hearer. St’át’imcets does have versions of words like *again*, *too*, and *stop*, and the relevant parts of their meanings share with English presuppositions the property of being not-at-issue (being outside the main asserted content of the utterances; Potts 2005, Tonhauser *et al.* 2013, among others). In other words, St’át’imcets possesses only a subset of the kinds of not-at-issue meaning available in natural language, namely, the non-presuppositional kind. These non-presuppositional not-at-issue meanings correspond to Potts’s (2005) definition of conventional implicatures.

Which is more formally complex: a presupposition, or some other kind of not-at-issue content? The answer is not clear, especially since the field has not reached agreement on how non-presuppositional, not-at-issue content should be modeled. Potts places conventional implicatures in a separate dimension of meaning, which would seem to add formal complexity, but the jury is still out on many issues surrounding multi-dimensionality.

The question of paradigm complexity does not really arise here, since the languages do not differ in the number of elements, but only in whether the elements impose contextual felicity constraints. What about expressive complexity? This is actually an interesting case. On the one hand, St’át’imcets seems to lack expressive complexity, because it lacks an entire class of meanings. On the other hand, suppose we take any particular translation pair (such as English *again* and St’át’imcets *múta7* ‘again’). In all cases where it *is* common knowledge that an event has happened before, both presuppositional *again* and non-presuppositional *múta7* are felicitous, and no difference between them should be detectable. In cases where the relevant information is *not* already known to the hearer, St’át’imcets *múta7* will be straightforwardly felicitous, and English *again* should either require presupposition accommodation, or will require paraphrases which lose the not-at-issue status of the information (e.g. ‘This has happened before and...’). In the last case, it is English which lacks expressive power compared to St’át’imcets.

In the absence of any clear winner in expressive complexity here, I’m going to call it a tie.

(42)

	FORMAL COMPLEXITY	PARADIGM COMPLEXITY	EXPRESSIVE COMPLEXITY
quantifiers	tie	English	English
determiners	tie	St’át’imcets	tie

tense	tie	English	tie
modals	St'át'imcets	tie	tie
evidentials	St'át'imcets	St'át'imcets	tie
presuppositions	?	English	tie

12.6.6 *Argument structure and control*

In English, the valency of predicates is not overtly encoded; both (43a) and (43b) contain the same verb.

- (43) a. Roland wrote. INTRANSITIVE
 b. Roland wrote a story. TRANSITIVE

Not so in St'át'imcets, which obligatorily marks (in)transitivity on every predicate, as shown in (44).

- (44) a. mets-*cál* kw=s=Roland
 write-*INTR* DET=NMLZ=Roland
 'Roland wrote.' INTRANSITIVE
- b. mets-*en-ás* kw=s=Roland ti=sqwélqwel'=a
 write-*TR-3ERG* DET=NMLZ=Roland DET=story=EXIS
 'Roland wrote a story.' TRANSITIVE

The St'át'imcets (in)transitivizers also encode (roughly) whether the subject argument is an agent in full control of the event (van Eijk 1997, Davis 2006). This is shown in (45), with the transitivizers glossed more fully to indicate their control status.

- (45) a. zík-*in'*-as ta=sráp=a i=sám7=a
 fall-*CONTROL.TR-ERG* DET=tree=EXIS DET.PL=white.person=EXIS
 'The white people cut the tree down (deliberately).' (Davis 2006: ch. 39)
- b. zikt-*s-ás* ta=sráp=a ts7a ku=xwélmen
 fall-*NON.CONTROL.TR-ERG* DET=tree=EXIS DEM DET=saw
 'This saw cut the tree down.' (Davis 2006: ch. 39)
- c. zík-*am* i=sám7=a
 fall-*CONTROL.INTR* DET.PL=white.person=EXIS
 'The white people did some tree-cutting.'

St'át'imcets is both formally and paradigmatically more complex in this area than English. Expressively, there is no difference; both languages have transitive and intransitive verbs, and can use them in cases where the subject is in control, and when s/he is not. The final summary of results is in (46).

(46)

	FORMAL COMPLEXITY	PARADIGM COMPLEXITY	EXPRESSIVE COMPLEXITY
quantifiers	tie	English	English
determiners	tie	St'át'imcets	tie
tense	tie	English	tie
modals	St'át'imcets	tie	tie
evidentials	St'át'imcets	St'át'imcets	tie
presuppositions	?	English	tie
argument structure	St'át'imcets	St'át'imcets	tie
control	St'át'imcets	St'át'imcets	tie

Of course, I have not presented a complete language comparison, since I have looked only at a small sub-part of each language, and have set pragmatic inferencing complexity aside due to the embryonic state of our knowledge in this area. From what we have looked at here, St'át'imcets is winning the complexity race on points.

It is perhaps more interesting, however, to think about the way complexity is distributed across the different ways of measuring. In the area of formal complexity, St'át'imcets is the clear winner, but the only clear expressibility difference involves English being more complex. With respect to paradigms, each language seems to concentrate its complexity in a different subset of areas. While English has a more complex quantifier system, St'át'imcets has a more complex determiner paradigm, and so on. From a functional standpoint, this distribution of complexity is probably not a coincidence. That is, we might expect that languages gravitate towards a roughly equal overall number of paradigmatically complex areas. This leads to the impression that languages 'care about' different semantic distinctions. For example, some languages 'care about' evidence source and have rich evidential paradigms; these are usually disjoint from the languages which 'care about' modal quantificational force and have paradigmatically rich modal systems which encode strength distinctions.

12.7 Conclusions

The first lesson from our case study on English and St'át'imcets was that determining whether languages differ in their semantics is a task which requires targeted, hypothesis-driven fieldwork. We then saw that even once we have established how the

languages in question vary, we should still not leap to conclusions about complexity. The fact that a language lacks some set of meanings which another language possesses (such as generalized quantifiers) does not automatically mean the former language is more simple. I have argued that in semantics, complexity must take into account not only formal complexity and paradigm complexity, but also pragmatic inferences, and the range of meanings a language can achieve (expressive complexity). We saw that while St'át'imcets noun phrases do not participate in scopal ambiguities, they freely allow cumulative readings in situations where English noun phrases do not, and are thus in one sense expressively more complex than their English counterparts. However, St'át'imcets does lack real generic readings, an absence which represents an expressive complexity difference with English.

The fact that languages vary in expressive complexity is a non-trivial result. It means that Fromkin *et al.* were wrong to say (2010: 34) that 'All languages are equally complex and equally capable of expressing any idea,' and it means that Katz's (1976: 37) Strong Effability Hypothesis, that "every proposition is the sense of some sentence in each natural language", is false. Expressive complexity is intuitively the most radical way in which languages can differ, and at the same time it is the type of complexity the human language faculty is least likely to be able to constrain. That is, Universal Grammar does not plausibly contain a list of meanings which all languages must be able to express. The weak expressive equivalence which does exist probably derives from general cognition and our status as social and cultural beings. Universal Grammar places formal and substantive constraints on meaning, but not on semantic complexity per se. So some languages could be, overall, semantically more complex than others.