To appear in Grant Goodall (Ed.), The Cambridge Handbook of Experimental Syntax, Chapter 13

# Acceptability Judgments at the Syntax-Semantics Interface

Jesse A. Harris

## 1 Introduction

This chapter reviews the theoretical and conceptual issues central to acceptability judgment tasks, and related paradigms, at the syntax-semantics interface, and provides a broad overview of core results obtained from research in this domain. Challenges faced by studies in experimental semantics are distinct from those in experimental syntax, which at times requires different linking hypotheses, research questions, or experimental paradigms. However, the current state of affairs suggests that acceptability and other offline judgments will continute to contribute highly informative and profitable tools for exploration of phenomena at the syntax-semantics interface. For comparison, we start with the role of syntactic acceptability judgments in establishing the shape of tacit grammatical knowledge.

Within the generative linguistics tradition, there is an explicit and long-standing endorsement of a *mental grammar*, i.e., a grammatical device that generates strings of the language, to which native speakers are assumed to have implicit access. An important methodological question is how we, as theorists, are to access this implicit knowledge to characterize key properties of the grammar

from native speakers of a language. In other words, *how can the empirical foundations of linguistic theory be best established*?

There are, naturally, a wide spectrum of answers to this question, although it is doubtful that any approach could avoid consulting native speaker judgments. While some approaches question the usefulness of introspection from the syntactician (Gibson & Fedorenko 2010; Gibson, Piantadosi & Fedorenko 2013), others have argued either that introspection is enough in principle, given the robust nature of syntactic judgments (Newmeyer 1983), or that theoretically-naive native speakers' judgments of core syntactic data reliably conform to the intuitions of syntacticians (Culicover & Jackendoff 2010; Sprouse & Almeida 2013, 2017, among others). In any event, interest in establishing methodological best practices is rapidly increasing, as is controlled experimental research in large-scale language judgment experiments, thanks, in part, to easier access to subject populations afforded by online crowdsourcing platforms. The increased interest has yielded an expanding array of tasks used in language research – e.g., magnitude estimation (Bard, Robertson & Sorace 1996; Featherston 2005) or thermometer tasks (Featherston 2008), among others. These developments have been reviewed in much greater depth elsewhere than will be mentioned here, for instance, in Part 1 of this handbook or in Schütze (2011).

The great majority of these studies concentrate on questions related to syntactic well-formedness (whether a sentence expression is syntactically generated by the grammar of the language). In this chapter, we address the role of acceptability judgments (and other experimental paradigms) in investigating the link between syntactic structures and linguistic meanings. Of course, this project presupposes a systematic, well-defined relation between syntactic and semantic representations. Though there are many conceivable relations that one might pursue, we will follow a fairly standard view of the syntax-semantics interface here, briefly reviewed in Section 2.1. Furthermore,

studies in syntax and semantics raise similar methodological and conceptual issues, including a need for explicit *linking hypotheses*: assumptions about how observed behavior or action is related to underlying, unobservable cognitive processes and mental states, introduced in Section 2.2. Methods for investigating the syntax-semantics interface are briefly discussed in Section 2.3, and case studies concentrating predominantly on the interpretation of quantified sentences are discussed in Section 3. The chapter concludes with some speculations on prospects and challenges faced in future studies.

# 2 Acceptability at the Syntax-Semantics Interface

## 2.1 The Relation between Syntax and Semantics

A fairly innocuous assumption is that a sentence of a language represents a pairing between a valid linguistic form (composed of sounds or signs) and a meaning. At its most abstract, then, a language $\mathcal{L}$ can be characterized as the rules governing how the set of forms $F$ are paired with a set of meanings $M$. Generative syntax traditionally seeks to characterize the syntactic devices that yield structures in $F$, i.e., the sentences of $\mathcal{L}$. The empirical questions around $F$ are relatively straightforward: what are the sentences that comprise $F$, and what theories of syntax make the correct predictions? Thus, common goals of a syntactic acceptability judgment experiment include (i) determining whether native speakers agree on the status of a sentence, and, to a lesser extent, (ii) providing evidence for a syntactic operation or constraint implicated in forming a structure. The many successes, and persistent challenges, of this enterprise have been thoroughly documented in this handbook.

In addition to a set of devices responsible for generating syntactic strings, we can posit a corresponding *semantic engine*, responsible for composing sentential meanings from the meanings of its subconstituents. Until recently, the empirical foundation of semantic theory and the syntax-semantics interface has received far less attention, though it is growing rapidly on many fronts (see Pylkkänen & McElree 2006; Bott et al. 2011 for review). While there are no inherent barriers to using judgment data in exploring the syntax-semantic interface, there are somewhat unique challenges faced by researchers in this area. The case studies reviewed below should reveal cause for cautious optimism, as such challenges are offset by the rich theoretical and empirical gains to be made in the study of the native speaker's knowledge of linguistic meaning.

What is knowledge of meaning? Intuitive answers are surprisingly illusive. A common starting place within semantic theory is to associate knowledge of meaning with knowledge of *truth conditions*: to understand the meaning of a sentence S is to understand the conditions under which S is true. For example, understanding the meaning of the sentence *Every cat has a black tail* (1a) amounts to understanding that the sentence is true just in case for any entity (within some contextually salient set of entities) that is a cat, that cat has a tail which is black (e.g., Larson 1995 for introduction).[1] Semantic competence requires that speakers be able to identify the situation or situations that make any given sentence true, and to correctly reject those that make it false. Semantic competence also requires that speakers understand how words and sentences are related to each other, as in entailment, synonymy, or antonymy. For example, the truth of the (1a) and (1b) together entail the truth of (1c).

[1] Although the role of truth-conditions in semantics is not without its critics (e.g. Dummett 2006; Schiffer 2015; Soames 1992), discussion of these criticisms here would take us too far afield for present purposes.

(1) a. Every cat has a black tail.

b. Mikey is a cat.

c. Therefore, Mikey has a black tail.

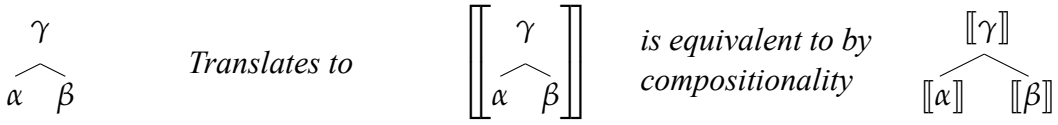
Experimentalists frequently capitalize on knowledge of truth conditions in exploring how comprehenders calculate the meaning of a sentence. In the truth-value judgment task, for example, subjects are asked to judge whether the sentence is true in a depiction of a situation (usually pictures or scenes acted out by an actor or puppet). Although the task was developed for studies of child language development (Abrams et al. 1978; Crain 1991; Gordon 1998), variants have been adapted for use with adult populations, typically to test how sentences with quantifiers are verified for truth or falsity (e.g., Lidz et al. 2011).

Another cornerstone of semantic theory is that meaning is *compositional*, as formulated by Partee (1995):

(2) **Principle of compositionality:** The meaning of an expression is a function of the meanings of its parts and of the way those parts are syntactically combined.

A natural consequence of compositionality is that subexpressions of sentences themselves have meanings, which compose with the meanings of other expressions in the sentence. This meaning is standardly represented as a form of (higher-order) predicate logic, a kind of *logical form* consisting of logical operators, functions, and their arguments. The meaning of an expression, e.g., $\alpha$, is conventionally represented by double square brackets $[\![\alpha]\!]$. For two expressions $\alpha, \beta$ whose mother node is $\gamma$, their meanings, $[\![\alpha]\!]$ and $[\![\beta]\!]$, may be combined together via a compositional operation $\circ$ to produce a composite meaning for $\gamma$: $([\![\alpha]\!] \circ [\![\beta]\!]) \Rightarrow [\![\gamma]\!]$.

(3) *A transparent syntax to meaning mapping:*




Compositionality comes in two basic variants (e.g., Janssen 1997; Pagin & Westerståhl 2010). On a strongly compositional view, every semantic element and operation corresponds to a unique element in the syntax. In other words, there is a one-to-one match between semantic representations and syntactic structures – every *syntactic* element $\gamma$ has a corresponding *semantic* element $[\![\gamma]\!]$, and *vice versa*. In case a semantic operator lacks an overt syntactic analogue, a silent, unpronounced operator is realized within the phrase structure tree, to preserve the one-to-one mapping between syntactic and semantic elements. On a weakly compositional view, in contrast, some semantic operations may have no syntactic analogue, so that there is a many-to-one relation between semantic and syntactic structures. On this view, meanings may be transformed through non-syntactic operations that change the semantic type or category of the element, in operations known as type-shifting (e.g., Partee & Rooth, 1983). Although the two views of compositionality cover much of the same empirical ground, they may do so with somewhat different mechanisms, and predict divergent mappings between syntactic and semantic representations.

What is the precise relation between a linguistic form and its meaning? Clearly, form and meaning are highly correlated: syntactic units typically form coherent units of interpretation. Most current models of the syntax-semantics interface assume forms and meanings are mediated by a syntactic level of representation, which feeds the compositional engine underlying interpretation. This level of *Logical Form* or LF is distinguished from a *logical form* in that LF is not yet semantically interpreted, and is still fundamentally syntactic nature (as discussed in May 1977, 1985).

The basic idea is that the syntax of an expression constrains how its subconstituents combine, by determining the order in which compositional operations may apply. Crucially, however, the location where an element is interpreted may diverge from where it is pronounced. Assuming that LF representations are generated by the same (or very similar) kinds of displacement operations as operate in syntax, the syntax yields a covertly expressed LF, which determines the way in which meanings may be composed: *Syntax* $\Rightarrow$ *Logical Form* $\Rightarrow$ *Semantic representations*.

Logical Form was motivated in large part by May's (1977; 1985) influential solution to a problem posed by the compositional interpretation of quantifiers in different syntactic positions. Semantically, a quantifier encodes a relation between two sets (Generalized Quantifier Theory: Barwise & Cooper 1981; Keenan & Stavi 1986), which are sometimes called the *restrictor* (the domain to which the quantifier applies) and the *nuclear scope* (a condition predicated of the domain). Furthermore, quantificational phrases like *every student* must bind a variable $x$ within its scope. It is straightfoward to give a compositional first-order logic representation for quantifiers in subject position; for example, *every* in (4b) translates into the logical quantifier $\forall$, where it takes scope over the variables it binds.

(4) a. Every [$_{\text{Restrictor}}$ student] ($_{\text{Nuclear scope}}$ laughed)

b. $\forall x : [student(x) \rightarrow laughed(x)]$

For every $x$ : if $x$ is a student, then $x$ laughed.

Problems arise when interpreting quantifiers in object position (5), as the denotation of the object (*every student*) cannot be directly composed with the verb without changing its meaning (Heim & Kratzer 1998, for discussion). May proposed the operation of *Quantifier Raising* (QR), in which a quantificational phrase like *every student* covertly adjoins above the sentence to [Spec, CP] where

it can bind the variable $x$ associated with the trace left by the movement operation (May 1977, 1985).

(5) *Example of quantifier raising*

a. *Syntax:* John greeted every student.

b. *Logical Form:* [$_{\text{Spec, CP}}$ every student ]$_1$ [$_{\text{TP}}$ John greeted $t_1$ ]

c. *Semantics:* $\forall x\,[student(x) \rightarrow greet(j, x)]$

For every $x$: if $x$ is a student, then John greeted $x$

Compelling evidence for QR initially came from the fact that covert quantifier movement appeared to parallel overt forms of movement, in particular *wh*-movement, as in *Who*$_1$ *did John greet* $t_1$, which was also argued to apply covertly in *wh-in-situ* languages like Chinese (Huang 1982). Though not identical, both forms of movement trigger crossover effects and show similar locality restrictions governing movement. Importantly for the discussion below, a QR account also provided an appealing treatment of quantifier scope ambiguity, in which one syntactic string permits multiple construals of quantificational scope, the surface and the inverse scope interpretations:

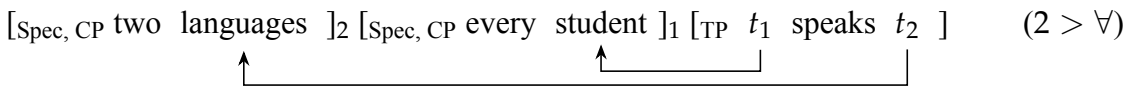
(6) Every student speaks two languages.

a. *Surface scope: The universal takes wide scope*

[$_{\text{Spec, CP}}$ every student ]$_1$ [$_{\text{Spec, CP}}$ two languages ]$_2$ [$_{\text{TP}}$ $t_1$ speaks $t_2$ ] $(\forall > 2)$

*Every student speaks two (possibly different) languages.*

b. *Inverse scope: The numeral takes wide scope*




*There are two languages that every student speaks.*

However, a host of empirical and theoretical challenges to QR were soon uncovered. Among the more major concerns were (a) that it is hard to reconcile differences between *wh*-movement and QR, (b) the fact that different quantifiers permit different scope-taking possibilities, and (c) the observation that indefinites are seemingly unaffected by constraints on movement. Nonetheless, many theorists assume that LF mediates the syntax-semantics interface, primarily to account for cases in which overt constituent order fails to map onto a compositional semantic derivation (notable exceptions include directly compositional approaches to scope and compositionality, e.g., Jacobson 1999 or Barker 2002, 2012).

The relationship between structure and meaning can take many forms, and the interface between them can be more or less direct. A good deal of experimental research has addressed properties of the syntax-semantics interface via studies on quantifier scope ambiguity, some of which is reviewed in Section 3.1.1 below. However, experimental research requires making key "linking hypotheses", i.e., theories that explain the relationship between processes responsible for language interpretation, on the one hand, and behavioral responses, on the other. These issues are discussed in the next section.

## 2.2 Linking Hypotheses

Although they are rarely formalized explicitly, linking hypotheses play a crucial role in establishing arguments for mental representations from observable repsonses or behaviors, such as performance on a questionnaire or the pattern of eye movements while reading. As a single action or behavior

may be associated with any number of causes, linking hypotheses are needed for any area within psychological sciences. To take a common sense example, a diner might compliment a chef's cooking because he truly enjoyed the meal, or because he did not wish to offend. Similarly, a participant's behavior on an acceptability judgment task might transparently reflect the state of her internal grammar, or it might instead reflect an application of prescriptive rules or how plausible she thinks the sentence is. Without direct access to mental states, the best we can do is infer representations indirectly on the basis of theory and behavior.

A case in point is the distinction between the twin concepts of grammaticality and acceptability. The *grammaticality* of a structure is a theory-internal construct, which is not directly accessible to native speakers (Chomsky 1969). Instead, language judgments ultimately require trading in the related notion of *acceptability*, which reflects the feelings or responses that native speakers have about sentences in their native language (see in particular Schütze 1996). Linking hypotheses assumed in most syntactic acceptability studies are relatively straightforward: performance on, for example, a sentence rating task is the product of the percept of acceptability, which itself is influenced not only by the subject's grammatical knowledge, but also by a host of unrelated factors (as reviewed in Sprouse 2018), such as those influencing the processing of the sentence (e.g., complexity, plausibility, frequency, etc.) or the task (e.g., comparison with other items in the experiment, saturation, habituation, etc.). By collecting judgments on various sentences from many subjects, a properly designed acceptability judgment study will expose the underlying factors of interest, typically the tacit grammatical knowledge shared by members of a linguistic community. Syntactic acceptability studies are thus usually designed to reduce the effect of non-grammatical factors on acceptability judgments.

Linking hypotheses in experimental semantics and pragmatics are currently in need of devel-

opment. Progress in this area demands a clear articulation of what a theory of interpretation entails, and researchers are faced with numerous, seemingly daunting questions not only about the *object*, but also about the *method*, of study. In terms of the object of study, experimentalists are confronted with a set of related questions: What are the mental representations that correspond to (sentence) meanings, and what are the mechanisms that yield those mental representations? How are those meanings formed compositionally? Are there distinct mental percepts for interpretation, similar to those posited for acceptability?

An equally thorny set of questions arise around the method of study. What are the methods that best identify the meaning or meanings that speakers access when performing the task? Are acceptability judgments an appropriate methodology for probing meaning? What is the semantic analogue to "How acceptable is this sentence?". It is certainly not "How meaningful is this sentence?" because meaning is not, at least *prima facie*, a gradable concept. Rather, experimentalists are usually more concerned with what meaning or meanings are (first) available to participants, and the extent to which those meanings are affected by structural considerations and contextual bias.

The primary domain of study within experimental semantics is the interpretation of ambiguous strings. How should a comprehender interpret a sentence string with multiple meanings? For example, does she compute all possible interpretations, or just the salient ones? Of the many possible answers, just two are mentioned here for illustration.

According to Grice's (1975) cooperative principle, typical speakers and comprehenders are engaged in a rational exchange of information. Comprehenders should assume that speakers intend to convey relevant information accurately, and may accordingly adopt the *Principle of Charity*, by which a sentence is disambiguated according to *the context of utterance*.

(7) **Principle of Charity:** Whenever possible, assume that the speaker speaks truthfully, and select whatever interpretation makes the sentence true.

(Davidson 1974; Gualmini, Hulsey, Hacquard & Fox 2008)

Alternatively, comprehenders may resolve ambiguity according to whatever interpretation is most accessible or salient, independent of the speaker's likely intentions. Of course, what determines accessibility itself requires a separate account, and may involve constraints on grammatical computations, discourse and performance factors, or some combination of the two.

(8) **Truth Dominance:** Whenever an ambiguous sentence S is true in a situation on its most accessible reading, we must judge sentence S to be true in that situation.

(Meyer & Sauerland 2009)

In some respects, linking hypotheses in experimental semantics parallel those raised in experimental syntax. However, the range of possible solutions diverge in important ways, especially as the notion of *acceptability* in syntax lacks an obvious analogue in current semantic theory. While syntactic acceptability judgment studies address whether (or to what extent) a sentence could have been generated by the grammar, there is less agreement on what mental representations are investigated in studies of sentence meaning. Possibilities include the existence of a semantic representation (a possible interpretation for a sentence), the existence of derivational mechanism (a syntactic or semantic operation) that yields a meaning, or a preference for a particular meaning in a context. The case studies reviewed below do not necessarily investigate the same issues, though all are valid approaches to the empirical study of interpretation.

## 2.3 Ways to Study Meanings: Acceptability and Beyond

Linguistic meaning depends on a complex interaction between linguistic form and context. Experimental studies tend to investigate meaning in one of two ways (Bott et al. 2011). First, one can manipulate one or more factors from context or form, in order to determine the effect on meaning. Conversely, one can manipulate the meaning to determine which contexts or linguistic forms are felicitous. Manipulating meaning requires that it be conveyed in another, non-linguistic fashion, so as not to confound form and meaning. As Bott et al. (2011) observed, the dependence of meaning on form is one of the reasons why experimental studies in meaning are often more complex than studies in syntax, which are designed to study form independently of its meaning.

Various methods are standardly employed in the empirical study of meaning, some of which are enumerated in 1–3 below. As with any experimental question, there is no single method or paradigm that is uniformly suited for all manipulations or is entirely free of confounds or alternative interpretations. However, tasks used in experimental syntax do not always translate to experimental semantics well, and often adjustments must be made. For example, studies that use fixed point scales (e.g., the standard Likert rating) usually ask participants to judge not acceptability, but *naturalness* or *appropriateness* within a context. Another possibility is the extent to which a paraphrase that disambiguates a sentence is acceptable, with or without a context. Other offline tasks include asking subjects to paraphrase the meaning of a sentence, to select a paraphrase, or to complete a sentence whose completion will disambiguate the interpretation. Some of the major kinds of tasks are summarized below, though the list is certainly not exhaustive:

1. **Target sentence in context**

    i. *Continuation acceptable in context:* Judge whether a target is an *acceptable* continuation

of the context or not (e.g., Kurtzman & MacDonald 1993).

ii. *Verification in context:* Judge whether the target sentence is *true* in a context or not (e.g., Ionin 2010; Brasoveanu & Dotlačil 2012).

iii. *Ratings in context:* Rate *how acceptable* a target sentence follows from a context (e.g., Harris, Clifton & Frazier 2013).

iv. *Forced choice interpretation:* Select which paraphrase best corresponds to the interpretation (e.g., Frazier, Clifton, Rayner, Deevy, Koh & Bader 2005; Harris & Potts 2009; Amaral 2010).

2. **Picture matching:** Select which picture, scene, video, etc. best corresponds to the situation described by the target sentence (e.g., Radó & Bott 2012; Lidz et al. 2011).
3. **Diagrams:** Complete or draw a diagram that best conveys the (first, most natural, etc.) interpretation of the target sentence (e.g., Gillen 1991; Bott & Radó 2007; Gyuris & Jackson 2018).

These tasks primarily address which meaning *M* is appropriate given a form from *F* (and some context of utterance), and therefore the mapping between a form and a meaning. They are less well suited, however, to investigating the semantic engine that drives the computation of meanings *within the set M itself*. With some exceptions, the question of *how* a meaning is computed is better addressed with online measures, which can capture delays in performance that are associated with recruiting additional linguistic operations or other cognitive resources needed to generate a meaning. The relation between offline and online measures is discussed briefly in connection with complement coercion in 3.3 below.

# 3 Case Studies

I now review studies that highlight major themes in the syntax-semantics interface: the interpretation of quantified sentences, evidence for incremental interpretation, and resolving semantic mismatch with complement coercion.

## 3.1 Interpreting Quantification

### 3.1.1 Quantifier Scope Ambiguity

One of the better studied areas in the syntax-semantics interface is the interpretation of expressions with multiple quantifiers (e.g., Dayal 2013, for review). For example, a sentence like *Every kid climbed a tree* (9) is ambiguous with respect to the relative order of its quantified phrases: the universal DP *every kid* and the existential DP *a tree*. On one construal, the universal DP takes scope over existential (9a); on another, it is the existential that takes wide scope over the universal (9b).

(9) Every kid climbed a tree.

a. $\forall x\,[kid(x) \rightarrow\ \exists y[tree(y) \wedge climb(x)(y)]]$ $(\forall > \exists)$

For any x: if x is a kid, then there is some tree y such that x climbed y.

b. $\exists y\,[tree(y) \wedge \forall x[kid(x) \rightarrow\ climb(x)(y)]]$ $(\exists > \forall)$

For some y: y is a tree, and for all x, if x is a kid, then x climbed y.

For simplicity, suppose that the contextually salient domain of kids consist of *Abbey*, *Ben*, and *Carol*, and the domain of trees consists of three individual trees (🌲, 🌲, 🌳). A *wide-scope universal construal* (9a) is true in a situation where all the contextually salient kids climbed some tree or

other, and so a universal wide scope interpretation is validated if Abbey, Ben and Carol all climbed one of the trees in the domain. However, the truth conditions of (9a) are satisfied whether one tree or multiple trees were climbed. The tree can be the same tree for all three kids, or possibly distinct trees, as shown in Figure 13.1a. In contrast, a *wide-scope existential construal* (9b) is true in a situation where just one tree was climbed by all the contextually salient kids – in our example, just in case Abbey, Ben, and Carol all climbed the same tree, as shown in Figure 13.1b. Therefore, the wide-scope universal construal (9a) entails the wide-scope existential (9b) construal, but not *vice versa*.


(a) Sample situation supporting only the *wide scope universal* reading: All the kids climbed a different tree.

(b) Sample situation supporting both a *wide scope universal* and a *wide scope existential* reading: All the kids climbed the same tree.

Figure 13.1: Pictorial representations of scope interpretations

Linguists have long been interested in scope ambiguity for many reasons. As discussed above, a prominent strand of research addresses the relation between the grammar and the interpretive system, particularly whether interpretations are constrained by syntactic operations at LF. At minimum, this requires understanding which readings are grammatically permissible. A correspondence between the availability of an interpretation and the availability of its associated syntactic representation would lend strong support for the idea that similar (or even identical) operations are involved in producing overt syntactic structures and covert LF structures that are interpreted by the semantics.

Although the ambiguity of doubly quantified sentences like (9) is widely recognized, judgments about which interpretations are available for related sentences have varied widely in the literature. A sentence containing *n* number of quantifiers has, in principle, *n*! logically possible scope construals. Yet, the sentence below (10) contains three quantifiers, and has been argued to have fewer than the predicted six scope configurations: five by Hobbs & Shieber's (1987) count, and four by Park's (1995).

(10) Every representative of a company saw most samples.

What is the correct answer? How many interpretations *does* such a sentence have? One might be tempted to simply poll linguistically-naïve subjects to settle the issue. But several of the possible readings are difficult to articulate or distinguish without training in logic or semantics. If experts cannot agree, where does that leave our participants? In addition, simply because an interpretation is not accessible to a particular participant does not mean that it is not a valid interpretation, as a particular meaning may be avoided because it is hard to compute, is implausible without proper context, or is simply not necessary for a shallow or incomplete parse of the sentence that may be sufficient for the task (e.g., Bott et al., 2011). What, then, can experiments tell the theorist when semantic intuitions are so variable?

Ioup (1975) and others have addressed this issue head on, by constructing accounts not of which readings are *available* to the grammar, but of which readings are *most accessible*, and which gradient discourse factors (topicality, specificity, discourse-linking, etc.) contribute to the prominence of a scope construal. In a series of speeded naturalness experiments, Kurtzman & MacDonald (1993) explored which scope construals are preferred in doubly quantified sentences. They adopted a constraint-based approach in which multiple interpretations are initially available, but the com-

prehender must ultimately choose the one that best satisfies the most relevant constraints. Several different constraint types were considered; some were based on syntactic or string order preferences, whereas others referenced the semantic or information structure properties of the language. Examples of the former include the *external subject preference* (assign wide scope to the subject) and the *linear order preference* (preserve surface order of quantificational elements), each of which predict that the subject in canonical SVO word order would favor wide scope of the subject. However, the preferences make distinct predictions for non-canonical word orders, like the passive: a principle favoring wide scope for the external subject would yield different interpretations than one favoring wide scope of the first quantifier. In addition, discourse-based preferences, like the *topic principle* (roughly, what the sentence is about), can be defined to favor a wide scope construal for the topic. It should be clear that no single one of these principles can, nor was intended to, capture all interpretations – for example, in sentences with three quantifiers, the topic principle would have nothing to say about the relative order of any non-topic constituents. Instead, the interpretation that satisfies the most important relevant principles or constraints is selected.

In Kurtzman & MacDonald's (1993) study, participants rated two-sentence discourses in which a doubly quantified sentence was followed by a continuation that was assumed to disambiguate towards a particular scope construal. The singular continuation (11a) disambiguated the previous sentence (11) to the wide scope existential interpretation, in that it strongly implies that one tree was climbed by all the kids. In contrast, the plural version (11b) supports the wide scope universal interpretation, in that multiple trees are made salient. Both orders of the quantifiers were tested (12).

(11) Every kid climbed a tree.

a. The tree was full of apples. (*Bias to inverse scope / existential wide scope*)

b. The trees were full of apples. (*Bias to surface scope / universal wide scope*)

(12) A kid climbed every tree.

a. The kid was full of energy. (*Bias to surface scope / existential wide scope*)

b. The kids were full of energy. (*Bias to inverse scope / universal wide scope*)

A surface preference was observed for both quantifier types (11b, 12a), but was stronger for the existential than the universal (which was interpreted in terms of a single reference principle). However, when the quantified sentence was presented in the passive voice (e.g., *A tree was climbed by every kid*), there was no consistent preference for one interpretation over another. Kurtzman & MacDonald proposed that the null effect resulted from a conflict between constraints favoring wide scope of the first DP (constraints favoring scope setting by linear order, surface subject, or topic preferences) and those favoring wide scope of the second DP (constraints favoring wide scope of the external argument or the thematically most salient constituent).

There were, as many others have noted, problems with the study (e.g., Tunstall 1998; Bott & Radó 2009, among others). Most crucially, the continuations do not fully disambiguate the scope construals. As discussed in connection with Figure 13.1, the truth conditions of a wide scope universal construal entail the truth conditions of a narrow scope reading, and a situation where all the kids climb a single tree supports either scope configuration, and so the singular continuation in (11a) does not fully exclude the wide scope universal construal. Other studies have confronted this issue by employing alternative methods for disambiguation (Gillen 1991; Bott & Radó 2007; Gyuris & Jackson 2018). For instance, in some of Gillen's (1991) studies, participants were asked

to draw or choose between diagrams representing their preferred interpretation of doubly quantified sentences, not unlike Figure 13.1 above. In others, participants rated how well the diagrams cohered with their interpretation of the sentence. Her results paralleled judgments observed in the literature, as well as those observed in corpus studies. Bott & Radó (2007) explicitly compared language judgments in response to (a) disambiguating linguistic contexts (question-answer pairs) and (b) two kinds of diagrams (depicting abstract vs. concrete scenarios). Linguistic contexts and abstract diagrams elicited scope judgments similar to those observed in previous studies. In contrast, diagrams depicting more concrete scenarios displayed an across-the-board universal wide scope preference, which they attributed to the interference of extralinguistic verification strategies, which favor the observed interpretation.

Diagram methods have several advantages over disambiguation by rating sentence completions or by asking for a linguistic paraphrase, both of which require the comprehender to calculate and retain one or more meaning, while simultaneously judging the sentence for contextual felicity. Diagram methods avoid ambiguity left open by context sentences and paraphrases, and may also offload some of the cognitive difficulty that may be associated with calculating and remembering interpretations. This last point is especially important, as difficulty imagining a situation that would support a particular scope construal might make the reading less accessible, but not necessarily unavailable as a possible meaning for the sentence. Thus, a result favoring situations that are easier to imagine might not reveal much about what interpretations are available for the sentence, although it would be all too easy to conflate the two. This again highlights how the challenges that studies in the syntax-semantics interface must contend with are somewhat distinct from syntactic acceptability judgment studies.

Another issue is that not all quantifiers pattern alike with respect to scope-taking possibilities, in

contrast to May's (1977) early QR approach, in which QR can apply to quantifiers in any sequence. For instance, Beghelli & Stowell (1997) organize quantifiers into distinct semantic classes, and propose that quantifier types are located in different structural positions at LF (along the lines of the series of functional projections advocated by Cinque 1999 and others). The structural position of the functional projection determines the scope construal possibilities: A quantificational element Q1 will scope above another quantificational element Q2 if (a) the functional position of Q1 structurally dominates the functional position of Q2, and (b) Q1 can move to its functional position (e.g., is not contained with an island, etc.) at LF. This account was motivated by several intriguing empirical gaps left open by an unconstrained QR approach. For instance, Beghelli & Stowell (1997) observed that group-denoting quantifiers (those that denote groups of individuals, like *some* and numerals) must take wide scope over sentence negation in (13b), even though both scope construals seem to be possible in (13a). They account for the asymmetrical scope relations by positing that group-denoting subjects must move to a position above the negation at LF and cannot reconstruct into their $\theta$-marked position (where the thematic relation is assigned).

(13) a. No student read two books. (*No > two; Two > no* )

b. Two students read no books. (*Two > no; * No > two* )

The strong predictions of configurational accounts are still in need of exploration. However, a few studies concentrating on distributivity suggest that configurational restrictions alone are not sufficient to explain preferential scope construals (Bott & Radó 2009; Brasoveanu & Dotlačil 2012).

### 3.1.2 Indefinites

It is not clear that all determiners behave uniformly as quantifiers. Indefinites like *a professor* in (15) pose a particularly complex problem for the theory of scope construal; they can be interpreted in positions where quantifiers cannot, and are said to take "exceptional wide scope" (Fodor & Sag 1982; Ruys 1992, among others). Positions allowing a wide scope construal of indefinites crucially include syntactic islands, from which movement is normally prohibited (Ross 1967), such as the relative clause island below (14), or the antecedent of a conditional (see Ionin 2010 or Schwarz 2011 for review).

(14) *Relative clause islands*

a. John read every article [$_{RC}$ that Mary assigned ].

b. * Who$_1$ did John read every article [$_{RC}$ that $t_1$ assigned ] ?

(15) *Exceptional wide scope of indefinite:*

John read every article [$_{RC}$ that a professor assigned ].

i. *Wide scope for indefinite:* There is a (particular) professor, such that John read every one of the articles that professor assigned. (*A professor > every article*)

ii. *Narrow scope for indefinite:* For every article that was assigned by John's professor, John read it. (*Every article > a professor*)

The ambiguity of the indefinite *a professor* in the relative clause contrasts with sentences in which a universally quantified phrase *every professor* can only be interpreted within the relative clause (16.ii).

(16) *No wide scope construal of universal in relative clause:*

John read an article [$_{RC}$ that every professor assigned ].

i. * *No wide scope for universal:* Every professor assigned some (possibly distinct) article, which John read. (* *Every professor > an article*)

ii. *Narrow scope for universal:* There was an article that every professor assigned, and John read that article. (*An article > every professor*)

Many approaches to the exceptional nature of indefinites have been pursued, starting with Fodor & Sag's (1982) *referential account*, in which an indefinite is directly interpreted as an individual the speaker has in mind, not as a quantifier. However, problems with interpreting the indefinite as a particular individual arise when the speaker does not (or cannot) have a particular referent when uttering the sentence (Ludlow & Neale 1991). Several current approaches now adopt a variant of the *choice function* analysis (Reinhart 1997; Winter 1997; Kratzer 1998; cf. Geurts 2000). Informally, a function $f_{CF}$ is a choice function if it selects some non-empty member of a set and applies it to its domain. For example, a choice function $f_{CF}$ with a domain of professors $f_{CF}(\text{professor})$ would select some non-empty member from the set of professors, e.g., Dr. Jones. The choice function effectively mimics the existential by picking out at least one individual that makes the sentence true, without necessarily contributing quantificational force to the expression. A variant of the choice function account employs a mixed approach, where an indefinite can be interpreted as a quantifier in some environments and a choice function in others (e.g, Winter 1997; Schwarz 2001; Ionin 2010). These approaches are motivated, in part, by the fact that when indefinites are modified by a *certain*, they permit very different scope construals (Hintikka 1986).

In a series of questionnaire studies, Ionin (2010) tested the predictions for scopal interpretations

of *single* (only quantificational or only choice function) versus *mixed* (both quantificational and choice function) mechanism accounts of indefinites. She compared bare *a* indefinites with *a-certain* indefinites in contexts biasing towards wide, intermediate, and narrow scope construals. Target sentences, e.g., *Every coach thinks that a (certain) player is missing*, were judged as appropriate or not given their preceding contexts, biased towards widest scope, intermediate, or narrow scope construal of the indefinite.

Whereas *a* indefinites showed high acceptance rates in all contexts in simple matrix clauses, *a-certain* indefinites were more often rejected, showing that they are incompatible with narrow scope interpretations. However, *a-certain* indefinites were accepted more often when they were to be interpreted outside the scope of a relative clause or the antecedent of conditional island. The results lend initial support to accounts that posit different mechanisms for deriving bare *a* and *a-certain* indefinites (e.g., Schwarz, 2001), suggesting that at least some kinds of indefinites may be interpreted as non-quantificational. Given the great deal of cross-linguistic variation between indefinites, particularly for those that appear to lexically encode an element of speaker-ignorance (e.g., Kratzer & Shimoyama 2002; Alonso-Ovalle & Menéndez-Benito 2013, among many others), the issue of just how indefinites get their scope continues to be an active area of research.

### 3.1.3 Quantification and Anaphora

Sentences with a single quantifier can also be ambiguous. In the second sentence of (17), the quantificational phrase *three ships* and its bare quantifier variant (*three*) can be interpreted anaphorically with respect to a previously mentioned group of ships (*the five ships*). This is the *subset* or *presuppositional* interpretation. An alternate, and intuitively less preferred, interpretation is the *new set* or *existential* reading, in which the quantifier refers to another three ships, in addition to the five

mentioned in the preceding sentence.

(17) Five ships appeared on the horizon. Three (ships) sank.

Frazier's Minimal Lowering account explains the preference for subset interpretations by a general injunction against unmotivated movement at LF (Frazier 1999, 2000; Frazier et al. 2005). For illustration, assume that subjects are generated within the Specifier of VP / $v$P and move to subject position [Spec, TP ] for independent reasons (e.g., McCloskey 1997, among others), and need not adjoin above the sentence via Quantifier Raising (May 1985) or to a dedicated quantifier position (Beghelli & Stowell 1997) at LF. Sentential subjects (*a dog* in (18)) could then be interpreted in two positions at LF: either at a surface subject position (18a), or a 'lowered' position within the VP, where they are reconstructed for interpretation (18b).

(18) A dog was in the garden.

a. $[_{\text{Spec, TP}}$ A dog$_i$ $[_{\text{TP}}$ was $[_{\text{VP}}$ $t_i$ $t_v$ $[_{\text{PP}}$ in the garden ]]]] (*Subject position*)

b. $[_{\text{Spec, TP}}$ DP$_i$ $[_{\text{TP}}$ was $[_{\text{VP}}$ a dog$_i$ $t_v$ $[_{\text{PP}}$ in the garden ]]]] (*Lowered into VP*)

The position of the subject was hypothesized to have consequences for interpretation. According to Diesing's (1992) *Mapping Hypothesis*, the interpretation of a subject DP (including quantifiers) depends on its position at LF. DPs interpreted outside the VP at LF (18a) are understood as presupposed or quantificational (yielding the subset reading). In contrast, those interpreted inside the VP at LF (18b) are interpreted as existential, and contribute a new (non-presupposed) referent into the discourse. Minimal Lowering proposes that DPs are preferentially interpreted at LF in their surface positions, in order to avoid potentially unnecessary operations (19). Subset interpretations are thus preferred because the interpretive system avoids generating LF structures that would yield the new set interpretation, i.e., a structure with a lowered subject.

(19) **Minimal Lowering:** 'Lower' only when necessary, e.g., interpret a DP in its surface position if possible. (Frazier 2000)

When Frazier et al. (2005) tested how subjects interpreted sentences with ambiguous quantifiers (17) using simple Yes/No questions, e.g., "Were the three ships that sank among the five ships that appeared on the horizon?", subset interpretations were selected about 60% of the time. Similar results were obtained with questions probing cardinality, e.g., "How many ships appeared on the horizon? Five or Two". Eye tracking results revealed that the subset interpretation was also preferred during real-time processing, as readers slowed down when they were forced to adopt a new set interpretation later in the text. The general findings appear to hold across multiple languages (German and Korean, Frazier et al. 2005; Dutch, Wijnen & Kaan 2006 and Kaan, Dallas & Barkley 2007).

Minimal Lowering makes crucial reference to constraints governing the preferred syntactic configuration of elements at LF, which in turn yield distinct discourse interpretations for quantifiers. As a consequence, it is a relatively indirect explanation of the subset preference. More direct alternatives have also been proposed, which attribute the subset preference to discourse biases that avoid introducing new referents into the discourse (e.g., Gordon & Hendrick, 1998). Paterson, Filik & Liversedge (2008) found reading penalties for quantifiers that introduce a new discourse referent, regardless of whether the quantifier was structurally ambiguous (Kaan et al. 2007; see also Paterson, Filik & Moxey 2009 for review). Although the two approaches are not mutually exclusive, they do presuppose potentially different linking hypotheses, differentiated by whether LF structures or discourse representations are most relevant.

### 3.1.4 Quantificational Domain Resolution

Harris, Clifton & Frazier (2013) addressed yet another aspect of quantification. In addition to interacting with other linguistic elements, quantifiers rely on context to determine the domain of objects in their range that is relevant to its evaluation (e.g., Bach, Jelinek, Kratzer & Partee, 1995). In (20), for example, quantifiers like *every* are typically restricted by an overt noun phrase *kid*, which limits what kinds of elements the quantifier ranges over.

(20) Every kid laughed.

Adverbs of quantification typically lack an explicit domain restriction, leaving open the kinds of elements they range over (Berman 1987; Lewis 1975; Hinterwimmer 2008). For example, *mostly* in (21) is ambiguous between a reading in which the adverb ranges over subsets of a contextually salient group of students (21a), and one in which it ranges over times or events of walking (21b).

(21) Students mostly walk to school.

a. On occasions when they go to school, most often they walk. (*Quantification over times*)

b. Most students walk to school. (*Quantification over individuals*)

Harris et al. adopted Majewski's (2014) *No Extra Times* Principle, which states that sentences are preferentially interpreted as describing a single event when possible (22), as a general conceptual economy principle. Assuming that sentences referencing multiple occasions also require multiple mental representations of those situations, comprehenders should avoid postulating complex representations unless necessary.[2]

[2]Evidence for such a a single event preference has been observed in areas as diverse as collective/distributive ambiguities (Clifton & Frazier 2013), coordination (Clifton & Frazier 2012), reciprocals (Majewski 2014), and serial verb constructions in Russian (Harris & Korotkova 2019). Space constraints prohibit a detailed discussion here.

(22) **No Extra Times:** A sentence describes a single occasion (unless there is evidence to the contrary). (Majewski 2014)

Pairs of sentences were constructed with one of three quantified sentences as the first sentence: an ambiguous adverb of quantification (23a), a quantifier with the domain explicitly quantifying over individuals (23b), or a quantifier with the domain explicitly quantifying over times (23c). The second sentence disambiguated to an interpretation consistent with multiple individuals (24.i) or multiple times (24.ii). For example, an item with an ambiguous quantificational adverb disambiguated towards a multiple times interpretation would be *The students were mostly here. Some of the time they weren't however*.

(23) *Quantified sentence*

a. The students were mostly here. (*Ambiguous*)

b. Most of the students were here. (*Quantification over individuals*)

c. Most of the time the students were here. (*Quantification over times*)

(24) *Disambiguation sentence*

i. Some of them weren't however. (*Disambiguation to individuals*)

ii. Some of the time they weren't however. (*Disambiguation to times*)

Ambiguous sentences (23a) patterned with unambiguous quantification over individuals sentences (23b), as both elicited a ratings penalty when followed by a sentence disambiguating to times (24.ii) compared to sentences disambiguating to individuals interpretations (24.i). Sentences that quantified over times showed the opposite pattern: a penalty for disambiguation to individuals. The results suggest that readers were not simply leaving the domain of quantification vague

or underspecified, at least by the end of the sentence, nor were they picking an interpretation at random.

Harris et al. (2013) provide additional evidence that readers selected a domain for the quantifier that respected the No Extra Times principle was observed in an eye-tracking study. Reading was disrupted when subjects were presented with a sentence for which a single event interpretation was implausible (e.g., *The inspector was mostly in the capital*), which suggest that the inspector was located in the capital on many occasions, not that parts of the inspector were mainly distributed in the capital, compared to sentences that permit or are biased towards a multiple times interpretation (e.g., *The army / attack was mostly in the capital*). In all, these findings support the idea that a domain of quantification is selected incrementally during sentence processing.

## 3.2 Incremental Interpretation

Most research in online (real-time) sentence processing assumes that sentences are interpreted incrementally, word by word (Just, Carpenter & Woolley 1982). In general, offline methods lack the proper sensitivity to capture the more temporarily subtle factors in structure building, and such questions are usually explored with online measures, such as self-paced reading, eye tracking, electrophysiological measures, and so on (reviewed in Part 4 of this handbook). However, particularly robust incremental processes have been shown to appear in end-of-sentence judgments – for example, when a structure violates a very strong preference for parsing *wh*-movement (Fanselow & Frisch 2006; Hofmeister et al. 2007; Sprouse 2008). Far less research has addressed the extent to which incremental semantic processes can be detected with offline methods.

One example involves the licensing of Negative Polarity Items (NPIs), which appear only in

particular semantic contexts (Fauconnier 1975; Ladusaw 1979).[3] In a downward entailing (DE) environment like negation, one can infer the truth of a sentence *The teacher didn't drink red wine* from another *The teacher didn't drink wine*, when the situation described by the first sentence is entailed by the situation described in the second. The NPI construal of *ever* is licensed in DE environments, such as negation (25a) or relative clauses (25b), but not postive counterparts (25c).

(25) a. The teacher didn't **ever** drink wine in class.

b. A teacher who **ever** drinks wine in class will be fired.

c. * The teacher **ever** drank wine in class.

Clifton & Frazier (2010) used the licensing properties of *ever* to investigate whether the language processing system identifies DE environments incrementally (*local computation*) or only after the entire sentence has been processed (*global computation*). If computed locally, *ever* should appear to be unlicensed when the syntactic analysis is not locally compatible with a DE environment, e.g., when it is temporarily tempted to parse the sentence incorrectly as a non-DE environment. In contrast, global computation would predict that a temporarily unlicensed *ever* would fail to affect sentence processing, as the local ambiguity has been resolved by the end of the sentence.

Clifton & Frazier tested these predictions in an acceptability ratings task, using reduced clause garden path sentences (26), in which a verb (*arrested*) in a reduced relative clause structure is preferentially interpreted as the main clause verb (Bever 1970; Frazier 1979, among many others). Sentences with reduced relative clauses are known to disrupt online processing, and may require reinterpretation of the structure (e.g., Ferreira & Clifton 1986). Assuming that readers take *arrested*

[3]The work on NPIs is expansive and suggests that the licensing of NPIs is more complicated than presented here (e.g., Giannakidou, 1998). Proposals include syntactic, semantic, and pragmatic conditions on licensing, or that NPIs may be (spuriously) licensed in other environments (Saddy, Drenhaus & Frisch 2004; Xiang, Grove & Giannakidou 2016).

as the main verb in (26a) on first encounter, *ever* should be initially interpreted as appearing in a main clause environment and would therefore appear to be unlicensed, along the lines of (25c).

(26) a. A man (**ever**) arrested in this country won't break the law here again.

b. A man who was (**ever**) arrested in this country won't break the law here again.

The crucial finding was an interaction in acceptability between the two conditions, so that the acceptability penalty for *ever* was increased when it appeared inside a reduced relative clause – i.e., just when *ever* would have been temporarily unlicensed by the structure, compared to a full relative clause control (26b). In a second experiment, the penalty for *ever* was eliminated when a negative quantifier created a licit downward entailing environment. The results are highly compatible with recent findings that pragmatic implicatures are computed incrementally during online sentence processing. They also offer evidence that the online computation of semantic context may, in some cases, be detectable with offline acceptability measures. It should be noted, however, that the correct licensing conditions in this experiment also depended on the proper identification of the correct structural analysis of the reduced relative clause. As syntactic and semantic properties travelled together in this design, it is hard to know whether the semantic anomaly would have been detected in offline measures without a corresponding syntactic misanalysis. Our final case study is one in which a semantic operation arguably has no direct correlate in the syntax, as a more transparent illustration of semantic composition.

## 3.3 Semantic Mismatch

Thus far, this chapter has concentrated on the mapping between form and meaning $\langle F, M \rangle$, and the ways in which comprehenders prefer to resolve ambiguous structures, without having dedicated

much discussion to the generation of semantic representations within *M* itself. One particularly interesting question in this area is the resolution of semantic mismatch, in which two elements combine even though they cannot be (transparently) combined compositionally. These mismatches between form and meaning pose an apparent challenge to the strongest forms of compositionality (Pylkkänen & McElree 2006). One such example is *complement coercion*, in which the selectional requirements of an event-selecting verb (*start* or *begin*) clashes with its entity-denoting object (*the book*).

(27) a. The author read the book.

b. The author started the book.

c. The author started (writing / reading / copying) the book.

Though uncommon, a syntactic account might assert that a covert verb form is introduced into the structure, so that any cost associated with interpreting the selectional mismatch would be attributed to recovering this missing, essentially elliptical, structure (see Pylkkänen & McElree 2006 for arguments against this brand of approach). On a semantic coercion account, event-selecting verbs "coerce" entity-denoting complements into an event by enriching their semantic representation. Such accounts propose that the enrichment occurs only at the level of interpretation, i.e., is a non-syntactic operation, in which the ontological type of the noun is shifted into a salient event. The precise mechanisms vary between approaches, and include type coercion (Pustejovsky 1991, 1995) and underspecification (Egg 2005) of the event, as well as richer aspectual dimensions encoded within the lexicon (Piñango & Deo 2015).

Penalties for complement coercion structures have been observed in many online processing studies employing a diverse range of methods, including self-paced reading (e.g., Traxler, Picker-

ing & McElree 2002; McElree et al. 2001), eye tracking (e.g., McElree, Frisson & Pickering 2006; Pickering et al. 2004; Pickering, McElree & Traxler 2005), and neuroimaging techniques (e.g., Baggio et al. 2010; Kuperberg et al. 2010; Pylkkänen & McElree 2007; Pylkkänen et al. 2009; Husband, Kelly & Zhu 2011). Online studies tend to use plausibility ratings to control differences between items, in order to remove the possibility that disruptions in online measures could reflect overall differences in acceptability. The materials in these studies are therefore designed specifically to avoid offline differences in acceptability. In addition, comprehenders may lack the motivation to commit to a specific interpretation in offline tasks, if they can recover enough of the message to form a basic judgment about its felicity. Nonetheless, some studies have observed both offline and online processing costs for coerced complements: for example, sentences with complement coercion have been found to elicit offline naturalness ratings penalties, as well as disruptions during reading (Frisson & McElree 2008; Kuperberg et al. 2010). These findings indicate that readers are sensitive to at least some kinds of local selectional mismatches that are resolved with non-syntactic mechanisms, though a fuller range of non-syntactic operations must be explored to determine precisely how sensitive subjects are to semantic processes. This project is just one of many in need of further exploration.

# 4 Future Directions

This chapter introduced some of the core themes behind language and acceptability judgments at the syntax-semantics interface. We noted that studies that focus on meaning must confront a different set of methodological and theoretical issues than studies that address form. As sentence meaning depends on its structure, experimental research in semantics inherit a great deal of the concerns

that arise in experimental syntax. However, as the percept of acceptability is likewise influenced by meaning as well as form, studies of syntactic phenomena must also ultimately contend with semantic factors, as well.

This review has concentrated primarily on quantification, in which a rich array of theoretically important issues can be investigated. In particular, we highlighted studies that addressed quantifier scope ambiguity, the interpretation of quantificational anaphora, and the selection of a quantificational domain. Of course, an introductory chapter of this size cannot do justice to the field, and a great many issues were omitted by necessity.

Although the empirical coverage of the syntax-semantics interface is rapidly growing, the theory of semantic judgments is still in a relative infancy. Currently, there is very little theory detailing which classes of semantic phenomena will be observable in different kinds of measures, even at the coarse-grained level of comparing offline and online measures. Although different measures are likely to reflect distinct cognitive processes, some degree of overlap between methodologies is expected. As a starting place, a theory of semantic judgments should reference the relevant percept that gives rise to judgments of infelicity. As with native-speaker judgments about syntactic acceptability, semantic judgments might be heavily influenced by the detection of an error signal and/or processes recruited to resolve semantic conflict. The perception of an error may originate from many potential sources, including violating preferred interpretations of the discourse, goodness of fit with previous context, or some syntactic constraint needed to produce a particular meaning. In general, it is possible that such errors would be detected, thereby requiring resolution, when they have been calculated and the comprender has committed to the interpretation to some extent. Even if the interpretation is discarded or suppressed by the time of the language judgment, the resources needed to resolve the conflict may contribute to the percept of semantic well-formedness used in

making the judgment.

In all, the empirical base of semantic theory is well-positioned to develop the necessary set of robust results, methodologies, and linking hypotheses to make major advances in the future. Although there are a great many remaining challenges, the current state of affairs promises to lead to ever more informative findings and predictive theories.

## Acknowledgments

Many thanks to Grant Goodall and two anonymous reviewers for comments on this paper. Thanks also to Jessica Rett for comments on an early draft. Any mistakes or misrepresentations are my own.

## References

Abrams, K., Chiarello, C., Cress, K., Green, S. & Ellelt, N. (1978). The relation between mother-to-child speech and word-order comprehension strategies in children. In R. N. Campbell & P. T. Smith, eds., *Recent Advances in the Psychology of Language*, Plenum Press New York, NY, vol. 4, pp. 337–47.

Alonso-Ovalle, L. & Menéndez-Benito, P. (2013). Two views on epistemic indefinites. *Language and Linguistics Compass*, 7(2), 105–22.

Amaral, P. (2010). Entailment, assertion, and textual coherence: The case of almost and barely. *Linguistics*, 48(3), 525–45.

Bach, E., Jelinek, E., Kratzer, A. & Partee, B. H. (1995). *Quantification in Natural Languages*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Baggio, G., Choma, T., Van Lambalgen, M. & Hagoort, P. (2010). Coercion and compositionality. *Journal of Cognitive Neuroscience*, 22(9), 2131–40.

Bard, E. G., Robertson, D. & Sorace, A. (1996). Magnitude estimation of linguistic acceptability. *Language*, 72(1), 32–68.

Barker, C. (2002). Continuations and the nature of quantification. *Natural Language Semantics*, 10(3), 211–42.

Barker, C. (2012). Quantificational binding does not require c-command. *Linguistic Inquiry*, 43(4), 614–33.

Barwise, J. & Cooper, R. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4(2), 159–219.

Beghelli, F. & Stowell, T. (1997). Distributivity and negation: The syntax of *each* and *every*. In A. Szabolcsi, ed., *Ways of scope taking*, Springer, pp. 71–107.

Berman, S. (1987). Situation-based semantics for adverbs of quantification. In J. Blevins & A. Vainikka, eds., *University of Massachusetts Occasional Papers in Linguistics*, GLSA Publishing, vol. 12, pp. 45–68.

Bever, T. G. (1970). The cognitive basis for linguistic structures. *Cognition and the Development of Language*, 279(362), 1–61.

Bott, O., Featherston, S., Radó, J. & Stolterfoht, B. (2011). The application of experimental methods in semantics. In C. Maienborn, K. von Heusinger & P. Portner, eds., *Semantics: An International Handbook of Natural Language Meaning*, Berlin: Walter de Gruyter, vol. 33.1, pp. 305–21.

Bott, O. & Radó, J. (2007). Quantifying quantifier scope: a cross-methodological comparison. In S. Featherston & W. Sternefeld, eds., *Roots. Linguistics in Search of Its Evidential Base*, Berlin: Walter de Gruyter, pp. 53–74.

Bott, O. & Radó, J. (2009). How to provide exactly one interpretation for every sentence, or what eye movements reveal about quantifier scope. In S. Featherston & S. Winkler, eds., *Fruits of Empirical Linguistics. Volume 1: Processes*, De Gruyter Mouton, Studies in Generative Grammar 101, pp. 25–46.

Brasoveanu, A. & Dotlačil, J. (2012). Licensing sentence-internal readings in English. In M. Aloni, V. Kimmelman, F. Roelofsen, G. W. Sassoon, K. Schulz & M. Westera, eds., *Logic, Language and Meaning*, Berlin: Springer-Verlag, pp. 122–32.

Chomsky, N. (1969). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Cinque, G. (1999). *Adverbs and functional heads: A cross-linguistic perspective*. Oxford, UK: Oxford University Press.

Clifton, Jr., C. & Frazier, L. (2010). When are downward-entailing contexts identified? The case of the domain widener *ever*. *Linguistic Inquiry*, 41(4), 681–9.

Clifton, Jr., C. & Frazier, L. (2012). Interpreting conjoined noun phrases and conjoined clauses: Collective versus distributive preferences. *The Quarterly Journal of Experimental Psychology*, 65(9), 1760–76.

Clifton, Jr., C. & Frazier, L. (2013). Partition if you nust: Evidence for a No Extra Times principle. *Discourse processes*, 50(8), 616–30.

Crain, S. (1991). Language acquisition in the absence of experience. *Behavioral and Brain Sciences*, 14(4), 597–612.

Culicover, P. W. & Jackendoff, R. (2010). Quantitative methods alone are not enough: Response to Gibson and Fedorenko. *Trends in Cognitive Sciences*, 14(6), 234–5.

Davidson, D. (1974). On the very idea of a conceptual scheme. *Proceedings and Addresses of the American Philosophical Association*, 47, 5–20. Published in Donald Davidson (1984). *Inquiries into Truth and Interpretation: Philosophical Essays*. Oxford, UK: Clarendon Press Oxford.

Dayal, V. (2013). The syntax of scope and quantification. In M. den Dikken, ed., *The Cambridge Handbook of Generative Syntax*, Cambridge, UK: Cambridge University Press, pp. 827–59.

Diesing, M. (1992). *Indefinites*. Linguistic Inquiry Monographs. Cambridge, MA: MIT Press.

Dummett, M. (2006). *Thought and Reality*. Oxford University Press.

Egg, M. (2005). *Flexible Semantics for Reinterpretation Phenomena*. Palo Alto, CA: CSLI Publications Stanford.

Fanselow, G. & Frisch, S. (2006). Effects of processing difficulty on judgments of acceptability. In G. Fanselow, C. Féry, M. Schlesewsky & R. Vogel, eds., *Gradience in Grammar: Generative Perspectives*, Oxford, UK: Oxford University Press, pp. 291–316.

Fauconnier, G. (1975). Pragmatic scales and logical structure. *Linguistic Inquiry*, 6(3), 353–75.

Featherston, S. (2005). Magnitude estimation and what it can do for your syntax: Some wh-constraints in German. *Lingua*, 115(11), 1525–50.

Featherston, S. (2008). Thermometer judgments as linguistic evidence. In C. M. Riehl & A. Rothe, eds., *Was ist linguistische Evidenz*, Aachen, Germany: Shaker Verlag, pp. 69–89.

Ferreira, F. & Clifton, Jr., C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, 25(3), 348–68.

Fodor, J. D. & Sag, I. A. (1982). Referential and quantificational indefinites. *Linguistics and Philosophy*, 5(3), 355–98.

Frazier, L. (1979). *On comprehending sentences: Syntactic parsing strategies*. Ph.D. thesis, University of Connecticut, Storrs, CT.

Frazier, L. (1999). *On sentence interpretation*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Frazier, L. (2000). On interpretation: Minimal 'lowering'. In M. W. Crocker, M. Pickering & C. Clifton Jr, eds., *Architectures and Mechanisms for Language Processing.*, Cambridge, UK: Cambridge University Press, pp. 303–23.

Frazier, L., Clifton, Jr., C., Rayner, K., Deevy, P., Koh, S. & Bader, M. (2005). Interface problems: Structural constraints on interpretation? *Journal of Psycholinguistic Research*, 34(3), 201–31.

Frisson, S. & McElree, B. (2008). Complement coercion is not modulated by competition: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(1), 1–11.

Geurts, B. (2000). Indefinites and choice functions. *Linguistic Inquiry*, 31(4), 731–738.

Giannakidou, A. (1998). *Polarity sensitivity as (non) veridical dependency*. Amsterdam-Philadelphia: John Benjamins Publishing.

Gibson, E. & Fedorenko, E. (2010). Weak quantitative standards in linguistics research. *Trends in Cognitive Sciences*, 14(6), 233–234.

Gibson, E., Piantadosi, S. T. & Fedorenko, E. (2013). Quantitative methods in syntax/semantics research: A response to Sprouse and Almeida (2013). *Language and Cognitive Processes*, 28(3), 229–240.

Gillen, K. (1991). *The Comprehension of Doubly Quantified Sentences.* Ph.D. thesis, Durham University, Durham, UK.

Gleitman, L. R., Osherson, D. N. & Liberman, M. (1995). *An Invitation to Cognitive Science: Language*, vol. 1. Cambridge, MA: MIT Press.

Gordon, P. (1998). The truth-value judgment task. In D. McDaniel, H. Smith Cairns & C. McKee, eds., *Methods for Assessing Children's Syntax*, Cambridge, MA: MIT Press, pp. 211–232.

Gordon, P. C. & Hendrick, R. (1998). The representation and processing of coreference in discourse. *Cognitive Science*, 22(4), 389–424.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan, eds., *Syntax and Semantics*, Academic Press, vol. 3.

Gualmini, A., Hulsey, S., Hacquard, V. & Fox, D. (2008). The question–answer requirement for scope assignment. *Natural Language Semantics*, 16(3), 205–237.

Gyuris, B. & Jackson, S. R. (2018). Scope marking and prosody in Hungarian. *Glossa: A Journal of General Linguistics*, 3(1), 1–32.

Harris, J. A., Clifton, Jr., C. & Frazier, L. (2013). Processing and domain selection: Quantificational variability effects. *Language and Cognitive Processes*, 28(10), 1519–44.

Harris, J. A. & Korotkova, N. (2019). Preference for single events guides perception in Russian: A phoneme restoration study. In E. Ronai, L. Stigliano & Y. Sun, eds., *Proceedings of the 54th Annual Chicago Linguistic Society*. Chicago, IL, pp. 149–63.

Harris, J. A. & Potts, C. (2009). Perspective-shifting with appositives and expressives. *Linguistics and Philosophy*, 32(6), 523–52.

Heim, I. & Kratzer, A. (1998). *Semantics in Generative Grammar*. Oxford, UK: Blackwell Oxford.

Hinterwimmer, S. (2008). *Q-Adverbs as Selective Binders: The Quantificational Variability of Free Relatives and Definite DPs*. Berlin/New York: Walter de Gruyter.

Hintikka, J. (1986). The semantics of *a certain*. *Linguistic Inquiry*, 17(2), 331–336.

Hobbs, J. R. & Shieber, S. M. (1987). An algorithm for generating quantifier scopings. *Computational Linguistics*, 13(1-2), 47–63.

Hofmeister, P., Jaeger, T. F., Sag, I. A., Arnon, I. & Snider, N. (2007). Locality and accessibility in wh-questions. In W. Sternefeld & S. Featherston, eds., *Roots: Linguistics in search of its evidential base*, Berlin: Mouton de Gruyter, pp. 185–206.

Huang, C. T. J. (1982). Move WH in a language without *wh*-movement. *The Linguistic Review*, 1(4), 369–416.

Husband, E. M., Kelly, L. A. & Zhu, D. C. (2011). Using complement coercion to understand the neural basis of semantic composition: Evidence from an fMRI study. *Journal of Cognitive Neuroscience*, 23(11), 3254–66.

Ionin, T. (2010). The scope of indefinites: An experimental investigation. *Natural Language Semantics*, 18(3), 295–350.

Ioup, G. (1975). *The Treatment of Quantifier Scope in a Transformational Grammar*. Ph.D. thesis, City University of New York, New York.

Jacobson, P. (1999). Towards a variable-free semantics. *Linguistics and Philosophy*, 22(2), 117–185.

Janssen, T. M. V. (1997). Compositionality. In J. van Bentham & A. ter Meulen, eds., *Handbook of Logic and Language*, Amsterdam, The Netherlands: Elsevier Press, pp. 417–74.

Just, M. A., Carpenter, P. A. & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, 111(2), 228–38.

Kaan, E., Dallas, A. C. & Barkley, C. M. (2007). Processing bare quantifiers in discourse. *Brain Research*, 1146, 199–209.

Keenan, E. L. & Stavi, J. (1986). A semantic characterization of natural language determiners. *Linguistics and Philosophy*, 9(3), 253–326.

Kratzer, A. (1998). Scope or pseudoscope? Are there wide-scope indefinites? In S. Rothstein, ed.,

*Events and Grammar*, Dodrecht, The Netherlands: Kluwer Academic Press, pp. 163–96.

Kratzer, A. & Shimoyama, J. (2002). Indeterminate pronouns: The view from Japanese. In Y. Otsu, ed., *Paper presented at the 3rd Tokyo Conference on Psycholinguistics*. Tokyo: Hituzi Syobo, pp. 1–25.

Kuperberg, G. R., Choi, A., Cohn, N., Paczynski, M. & Jackendoff, R. (2010). Electrophysiological correlates of complement coercion. *Journal of Cognitive Neuroscience*, 22(12), 2685–701.

Kurtzman, H. S. & MacDonald, M. C. (1993). Resolution of quantifier scope ambiguities. *Cognition*, 48(3), 243–79.

Ladusaw, W. A. (1979). *Polarity Sensitivity as Inherent Scope Relations*. Ph.D. thesis, University of Texas, Austin, Austin, TX.

Larson, R. (1995). *Semantics*, pp. 361–380. Vol. 1 of Gleitman, Osherson & Liberman (1995).

Lewis, D. (1975). Adverbs of quantification. In E. Keenan, ed., *Formal Semantics of Natural Language*, Cambridge: Cambridge University Press, pp. 3–15.

Lidz, J., Pietroski, P., Halberda, J. & Hunter, T. (2011). Interface transparency and the psychosemantics of most. *Natural Language Semantics*, 19(3), 227–56.

Ludlow, P. & Neale, S. (1991). Indefinite descriptions: In defense of Russell. *Linguistics and Philosophy*, 14(2), 171–202.

Majewski, H. (2014). *Comprehending Each Other: Weak Reciprocity and Processing*. Ph.D. thesis, University of Massachusetts, Amherst, Amherst, MA.

May, R. (1977). *The Grammar of Quantification*. Ph.D. thesis, MIT. Reprinted by the Indiana University Linguistics Club, 1982.

May, R. (1985). *Logical Form: Its Structure and Derivation*. Cambridge, MA: MIT press.

McCloskey, J. (1997). Subjecthood and subject positions. In L. M. V. Haegeman, ed., *Elements of Grammar*, Dordrecht, The Netherlands: Kluwer Academic Publishers, Kluwer International Handbooks of Linguistics, pp. 197–235.

McElree, B., Frisson, S. & Pickering, M. J. (2006). Deferred interpretations: Why starting Dickens is taxing but reading Dickens isn't. *Cognitive Science*, 30(1), 181–92.

McElree, B., Traxler, M. J., Pickering, M. J., Jackendoff, R. S. & Seely, R. E. (2001). Coercion in on-line semantic processing. *Cognition*, 78, B17–B25.

Meyer, M.-C. & Sauerland, U. (2009). A pragmatic constraint on ambiguity detection. *Natural Language & Linguistic Theory*, 27(1), 139–50.

Newmeyer, F. J. (1983). *Grammatical Theory: Its Limits and its Possibilities*. Chicago, IL: University of Chicago Press.

Pagin, P. & Westerståhl, D. (2010). Compositionality I: Definitions and variants. *Philosophy Compass*, 5(3), 250–264.

Park, J. C. (1995). Quantifier scope and constituency. In *Proceedings of the 33rd Annual Meeting on Association for Computational Linguistics*. Association for Computational Linguistics, pp. 205–12.

Partee, B. (1995). Lexical semantics and compositionality. In Gleitman et al. (1995), pp. 311–360.

Partee, B. & Rooth, M. (1983). Generalized conjunction and type ambiguity. In P. Portner & B. Partee, eds., *Formal Semantics: The Essential Readings*, Blackwell Publishers, chap. 14, pp. 334–56. Originally published in R. Bäuerle, C. Schwarze, A. & von Stechow, eds., (1983).

*Meaning, Use and the Interpretation of Language*, Walter de Gruyter, pp. 361–93.

Paterson, K. B., Filik, R. & Liversedge, S. P. (2008). Competition during the processing of quantifier scope ambiguities: Evidence from eye movements during reading. *The Quarterly Journal of Experimental Psychology*, 61(3), 459–73.

Paterson, K. B., Filik, R. & Moxey, L. M. (2009). Quantifiers and discourse processing. *Language and Linguistics Compass*, 3(6), 1390–402.

Pickering, M. J., Frisson, S., McElree, B. & Traxler, M. J. (2004). Eye movements and semantic composition. In M. Carreiras & C. Clifton, Jr., eds., *The on-line study of sentence comprehension: Eyetracking, ERPs, and beyond*, Psychology Press New York, NY, pp. 33–50.

Pickering, M. J., McElree, B. & Traxler, M. J. (2005). The difficulty of coercion: A response to de Almeida. *Brain and Language*, 93(1), 1–9.

Piñango, M. M. & Deo, A. (2015). Reanalyzing the complement coercion effect through a generalized lexical semantics for aspectual verbs. *Journal of Semantics*, 33(2), 359–408.

Pustejovsky, J. (1991). The generative lexicon. *Computational Linguistics*, 17(4), 409–41.

Pustejovsky, J. (1995). *The Generative Lexicon*. Cambridge, MA: MIT Press.

Pylkkänen, L., Martin, A. E., McElree, B. & Smart, A. (2009). The anterior midline field: Coercion or decision making? *Brain and Language*, 108(3), 184–90.

Pylkkänen, L. & McElree, B. (2006). The syntax-semantics interface: On-line composition of sentence meaning. In M. Traxler & M. A. Gernsbacher, eds., *Handbook of Psycholinguistics*, Elsevier, pp. 539–579. 2nd edn.

Pylkkänen, L. & McElree, B. (2007). An MEG study of silent meaning. *Journal of Cognitive Neuroscience*, 19(11), 1905–21.

Radó, J. & Bott, O. (2012). Underspecified representations of scope ambiguity? In M. Aloni, V. Kimmelman, F. Roelofsen, G. W. Sassoon, K. Schulz & M. Westera, eds., *Logic, Language and Meaning*, Berlin: Springer-Verlag, pp. 180–9.

Reinhart, T. (1997). Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy*, 20(4), 335–97.

Ross, J. R. (1967). *Constraints on Variables in Syntax*. Ph.D. thesis, MIT, Cambridge, MA. Later published as *Infinite Syntax* (1986). Ablex Publishing Corporation. New Jersey.

Ruys, E. G. (1992). *The Scope of Indefinites*. Ph.D. thesis, Utrecht University.

Saddy, D., Drenhaus, H. & Frisch, S. (2004). Processing polarity items: Contrastive licensing costs. *Brain and Language*, 90(1-3), 495–502.

Schiffer, S. (2015). Meaning and formal semantics in generative grammar. *Erkenntnis*, 80(1), 61–87.

Schütze, C. T. (1996). *The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology*. Chicago, IL: University of Chicago Press.

Schütze, C. T. (2011). Linguistic evidence and grammatical theory. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(2), 206–21.

Schwarz, B. (2001). Two kinds of long-distance indefinites. In R. van Rooy & M. Stokhof, eds., *Proceedings of the 13th Amsterdam Colloquium*. Amsterdam, The Netherlands, pp. 192–7.

Schwarz, B. (2011). Long distance indefinites and choice functions. *Language and Linguistics Compass*, 5(12), 880–97.

Soames, S. (1992). Truth, meaning, and understanding. *Philosophical Studies*, 65(1-2), 17–35.

Sprouse, J. (2008). The differential sensitivity of acceptability judgments to processing effects. *Linguistic Inquiry*, 39(4), 686–94.

Sprouse, J. (2018). Acceptability judgments and grammaticality, prospects and challenges. In N. Hornstein, C. Yang & P. Patel-Grosz, eds., *Syntactic Structures after 60 Years: The Impact of the Chomskyan Revolution in Linguistics*, Mouton de Gruyter, pp. 195–224.

Sprouse, J. & Almeida, D. (2013). The empirical status of data in syntax: A reply to Gibson and Fedorenko. *Language and Cognitive Processes*, 28(3), 222–8.

Sprouse, J. & Almeida, D. (2017). Setting the empirical record straight: Acceptability judgments appear to be reliable, robust, and replicable. *Behavioral and Brain Sciences*, 40, e311.

Traxler, M. J., Pickering, M. J. & McElree, B. (2002). Coercion in sentence processing: Evidence from eye-movements and self-paced reading. *Journal of Memory and Language*, 47(4), 530–47.

Tunstall, S. (1998). *The Interpretation of Quantifiers: Semantics and Processing*. Ph.D. thesis, University of Massachusetts Amherst, Amherst, MA.

Wijnen, F. & Kaan, E. (2006). Dynamics of semantic processing: The interpretation of bare quantifiers. *Language and Cognitive Processes*, 21(6), 684–720.

Winter, Y. (1997). Choice functions and the scopal semantics of indefinites. *Linguistics and Philosophy*, 20(4), 399–467.

Xiang, M., Grove, J. & Giannakidou, A. (2016). Semantic and pragmatic processes in the comprehension of negation: An event related potential study of negative polarity sensitivity. *Journal of Neurolinguistics*, 38, 71–88.

# List of Figures