Techniques in Complex Semantic Fieldwork

M. Ryan Bochnak¹ and Lisa Matthewson²

¹Linguistics and English Language, University of Manchester, Manchester, UK, M13 9PL; email: ryan.bochnak@gmail.com
²Department of Linguistics, University of British Columbia, Vancouver BC, Canada, V6T 1Z4

Keywords

semantics, fieldwork, elicitation, contexts, storyboards, out-of-the-blue, de re/de dicto, verum

Abstract

The main goal of semantic fieldwork is to accurately capture the contribution of natural language expressions to truth conditions and to pragmatic felicity conditions, by interacting with native speakers of the language under investigation. Most semantic fieldwork tasks (including for example acceptability judgment tasks, elicited production tasks or translation tasks) require the researcher to present a discourse context to the consultant. The important question then becomes how to present that context to consultants, and how to best ensure that the consultant and the researcher have the same context in mind. We argue that phenomena which rely on controlling for interlocutor beliefs are particularly well suited for the storyboard elicitation methodology. This includes ‘out of the blue’ scenarios, which we treat as a special type of discourse context which must also be controlled for. We illustrate these claims by presenting novel storyboards targeting the de re/de dicto ambiguity and verum marking.
1. INTRODUCTION

Semantic fieldwork is the task of eliciting information about meaning in human language by interacting with native speakers. Researchers who engage in semantic fieldwork include for example formal semanticists, descriptive linguists, typologists, or writers of pedagogical materials. In this paper we begin by giving an overview of some of the core semantic fieldwork techniques; this discussion is aimed at a general audience. We then discuss one particular methodological issue: how semantic fieldworkers can best present their language consultants with contextual information for linguistic tasks that they perform.

We start with some basic assumptions about the goals and processes of semantic fieldwork. The overarching objective is to accurately capture the contribution of natural language expressions to truth conditions and to pragmatic felicity conditions. The central challenge is how one establishes these truth conditions and felicity conditions on the basis of responses that language consultants give to ‘tasks’ the fieldworker sets. There are many specific questions which need to be answered; some are listed in examples (1a)-(1e).¹

(1a) What types of tasks should consultants be asked to perform?
(1b) When should which type of task be used?

¹In the small but growing literature on semantic fieldwork methodology, contributions have been made to all of these questions. See for example Matthewson (2004), Krifka (2011) for general overviews, and the papers in Bochnak and Matthewson (2015) for treatments of specific issues and phenomena.
(1c) How does one best conduct each type of task?

(1d) What can a researcher conclude about truth and felicity conditions from various types of responses a consultant makes to each type of task?

(1e) In what form should stimuli (such as context descriptions) be given to the consultant?

Two core assumptions will underlie our discussion of these issues. First, we take it that truth and felicity conditions can only be calculated or judged by native speakers at the level of a complete utterance. One therefore cannot investigate the meanings of individual items directly, but only in terms of their effect on larger strings, and this is especially so for grammatical morphemes (e.g., determiners, tenses, evidentials). We furthermore take it as a given that truth and felicity conditions cannot be established based on sentences which are uttered, or understood, in isolation. Utterances are only true or false, felicitous or infelicitous, in context. This is why we focus our attention in the later part of this paper on the optimal ways to present discourse contexts to consultants.

The rest of the paper is structured as follows. We begin in section 2 with an overview of the current state of the art in semantic fieldwork techniques, based on the literature and on our own experience as semantic fieldworkers. We mainly concentrate on the question in example (1a), and give some pointers to the literature on the questions in examples (1b)-(1d). The rest of the paper focuses in on the issue in example (1e), i.e., the best methods for providing consultants with contextual information for linguistic tasks that they perform. There are a range of techniques in common use, including verbal context descriptions, single images, storyboards, and videos. All have advantages and disadvantages, many of which have been discussed (see the references in footnote 1). The question we pose here is one which to our knowledge has not been asked in a global way before: Are certain types of linguistic phenomena more suited to particular methods for presenting contexts?

After surveying a number of elicitation methods that have been used for specific phenomena, we advance the hypothesis that certain types of content can be probed easily enough by using verbal context descriptions or single images, while content that relates to the belief states of interlocutors is far more likely to require more complex presentation techniques such as storyboards. These include for example presupposition triggers, information-structuring elements, epistemic modals, attitude verbs and discourse particles. We also propose that establishing an “out-of-the-blue” context – ironically, perhaps – also usually requires a complex context presentation technique. This is because out-of-the-blue also crucially requires controlling for interlocutors’ belief states, e.g., lack of shared knowledge. Without explicitly

2Some complete utterances might consist of a single item, such as ‘Hello’ or ‘Wow’. These should be investigated in terms of their effect on the discourse.

3Throughout this paper, we use “verbal context descriptions” to refer to cases where contexts are provided to consultants either orally or in a written format.
controlling for lack of shared knowledge, the fieldworker risks consultants silently constructing their own contexts, which may have unintended negative impacts on the results.

2. SEMANTIC FIELDWORK TASKS

The job of a researcher investigating meaning in the field is to extract information about truth conditions and felicity conditions through interactions with native speakers of a language, potentially even one’s own. The challenge is that this information is neither consciously accessible to speakers, nor directly observable by the researcher (utterances do not come paired with an overt indication of their meaning!). As Bohnemeyer (2015) points out, the solution is that the fieldworker must create opportunities to observe correlations between utterances and things/events in the world. This may be done indirectly, by observing correlations between utterances and other utterances, in either the target language or a contact language.

In this section we give a brief overview of types of tasks that speakers are typically asked to perform in semantic fieldwork. The presentation is necessarily very much simplified and abridged; see Matthewson (2004), Krifka (2011), Bohnemeyer (2015), and Tonhauser and Matthewson (2017) for more detailed discussion.4

2.1. Three useful tasks

One important and widely-used elicitation method is an acceptability judgment task. In this task, the speaker is presented with some state of affairs paired with an utterance in the target language, and is asked to judge its acceptability. One case is given in example (2.1), with a verbal context description. In this example, the researcher is attempting to establish whether a verb which lacks any overt aspectual marking can report an event which is in progress at the utterance time. The context is set up so that an in-progress interpretation is appropriate, and the consultant judges the utterance containing the unmarked verb form as unacceptable.5

Atayal

(2) Context: You witness that Tali is pushing around a neighboring kid. You rush to inform Tali’s mother: “Come to check quickly...”

---


5In order to support the hypothesis that it is the absence of aspectual marking which leads to the unacceptability in example (2), Chen (2018) also tests minimally altered utterances in the same discourse context, which contain overt aspect markers.
Acceptability is a theory-neutral concept, and as such is easy for speakers to make a judgment on. Importantly, a negative judgment of ‘unacceptable’ does not provide any indication in and of itself about the reason why an utterance is unacceptable. If the linguist has already established that the utterance is syntactically well-formed, and its felicity conditions are satisfied (for example, that it does not contain any presupposition failures), and it is not pragmatically inappropriate for some other reason, then a judgment by the speaker of ‘acceptable’ can reasonably be assumed to give evidence that the sentence is true in this context, and a judgment of ‘unacceptable’ conversely would indicate that the sentence is false. In this way, we make progress towards establishing the truth conditions of the utterance.

One of the most interesting issues with an acceptability judgment task is how to present to the speaker the state of affairs/discourse context against which the utterance should be judged. In (2), the context was presented verbally, but non-verbal representations are often useful. This topic is addressed in detail in section 3.

A second useful task is an elicited production task, in which the speaker volunteers a target language utterance (or part of one) in response to a stimulus of some sort and possibly a prompt by the researcher. Examples of production tasks include the methodology used in the Topological Relations Picture Series (BowPed; Bowerman and Pederson 1992). BowPed offers a set of line drawings for the purpose of investigating topological relations, which are expressed in English and many other languages with prepositions. In each picture, a “figure” is highlighted in colour with respect to a “ground” (Talmy 2000), or pointed to with an arrow, and the researcher asks the consultant “Where is [the figure]?”, preferably in the target language; see Figure 1. The consultant is then supposed to describe the topological relationship using spatial language (e.g., “The coat is on the hook”, “The apple is in the bowl”).

---

6 Often, these pre-conditions are taken to be satisfied and speakers are asked to make judgments explicitly about truth, in which case we call it a truth value judgment task (e.g., Crain and Steedman 1985). Note that (2) is not a case where unacceptability necessarily signals falsity. Rather, there is a pragmatic mismatch between asking the addressee to come quickly, and a verb form which does not allow event-in-progress readings.

7 Some researchers report difficulties in using the BowPed materials in fieldwork. For instance, some consultants may not recognize the objects in the drawings, and some may find it easier to look at photographs of culturally appropriate objects rather than drawings (Cerqueglini 2018; Moro
descriptions. Once again, the question arises of how to best present the discourse context or stimulus for each research question; we take this up in section 3.

Thirdly, we have translation tasks. Here, a speaker volunteers a target language utterance that is in some sense “equivalent” to a given meta-language utterance (or vice versa). Although translations allow the fieldworker to collect positive data in the form of grammatical target language utterances, translations in and of themselves do not provide direct evidence for the semantics of morphemes and larger phrases (Matthewson 2004). See Deal (2015) for detailed discussion and illustration of what a fieldworker can safely conclude (or not) based on a translation task. In any case, translations, just like the other tasks, should be paired with discourse contexts. An example that is admirably explicit comes from Cable (2017) in example (3), where the context (“scenario” in Cable’s terminology) was presented verbally:

Tlingit

(3) Scenario: Your friend didn’t come to work today. Usually, if he’s not in, he has a very good reason. Someone asks you why your friend didn’t come in. You make a wild guess.

English sentence to translate: “Maybe he’s sick.”

Tlingit translation offered:

Gwál yanéekw.
DUB IMPFV.3S.sick
‘Maybe he’s sick.’

(Cable 2017, p. 628-629)

Ideally all the tasks, but especially judgment and production tasks, require discourse contexts. That’s why one of the most important questions to address is how to construct and

and Klamer 2018). As with any type of elicitation material, socio-cultural factors may require the researcher to adapt the materials in certain ways, or to train the consultants to get used to the task. For further discussion, see Bohnemeyer (2015).

---

Figure 1

BowPed pictures: e.g., “The cat is on the mat” and “The cloud is above the mountain”
present contexts to the consultant. We turn to this issue in section 3. Before that, however, we briefly address the concept of a questionnaire.

### 2.2. Collections of tasks: Questionnaires

It is not only important to use elicitation techniques which are robust and well-grounded for the language(s) and phenomena one is working on, it is also important to ensure that other researchers can apply them in the same or other languages, attempting to replicate results. One important step in facilitating such replicability is to make one’s elicitation techniques explicit (as Cable does in example (3)). A further step is to consciously design a set of elicitation materials intended for cross-linguistic use. Such collections of elicitation materials are called questionnaires, and they are frequently used by typologists, as well as increasingly by formal semanticists.⁸

As questionnaires are simply elicitation plans that can be adapted to be used for many languages, they may include within them any of the types of task we discussed above, as well as others, and they may include various techniques for presenting discourse contexts.

Something to bear in mind about questionnaires is that it can be very challenging to construct materials which are truly cross-linguistically valid. Language-specific properties, and/or socio-cultural issues, often mean that the stimuli need to be adjusted by individual researchers. A helpful discussion of the process of constructing and working with a questionnaire is given by Beck et al. (2009), who investigated comparison constructions in 14 languages (and whose work has inspired much subsequent cross-linguistic research). Beck et al. write that ‘To create questionnaires for the languages at hand we went through the following procedure steps ...: familiarising with the relevant structures and morphemes in the target language; constructing examples with the help of the primary informant(s); eliciting data from naïve secondary informants; analysing the elicited data, fixing the questionnaire and repeating the elicitation step if necessary’ (2009, p. 15). The process also involved the construction of appropriate discourse contexts for many sentences in individual languages (Beck et al. 2009, p. 16).

---

⁸Many repositories for questionnaires are available online. A non-exhaustive list includes the Max Planck Language and Cognition team’s Field Manuals and Stimulus Materials site (http://fieldmanuals.mpi.nl); Dahl’s (1985) tense and aspect questionnaire, which is available through the Max Planck Institute for Evolutionary Anthropology (https://www.eva.mpg.de/lingua/tools-at-lingboard/questionnaires.php); TulQuest (http://tulquest.huma-num.fr/en); and Skopeteas et al.’s (2006) Questionnaire for Information Structure (https://www.sfb632.uni-potsdam.de/quis/quis-materials.html).
3. PRESENTING CONTEXTS TO CONSULTANTS

As we pointed out above, all kinds of semantic fieldwork tasks require the researcher to present a discourse context to the consultant, since utterances are only true or false, or felicitous or infelicitous, with respect to particular contexts (with so-called ‘out-of-the-blue’ contexts as a special case; see below). An important desideratum is that the researcher and the consultant understand the same discourse context. That is, we need to minimize the risk that the consultant enriches the discourse context silently when judging or producing a response.

Establishing discourse contexts in a way which is maximally clear to the consultant and which avoids the risk of silent enrichment is not a trivial matter. In this section we review several methods for presenting contexts to speakers for judgment and production tasks, along with some of their most salient advantages and drawbacks.

3.1. Presenting contexts verbally

A classic strategy for conveying a discourse context to a consultant is for the researcher to describe the context verbally. This method was already exemplified above in examples (2) and (3). Matthewson (2004) advises that the fieldworker should explain the context first before giving the sentence for which a judgment is requested, rather than the other way around. Presenting the context first prevents the consultant from imagining their own context upon hearing the sentence – exactly what the fieldworker is trying to avoid – and also is more faithful to what happens in real life: we hear and interpret sentences relative to the contexts we already find ourselves in.

One important question for the fieldworker using the verbal presentation method is which language to use to present the context, the target language or the meta-language. A variety of factors go into the decision of which language to choose. Among these are practical considerations, such as the level of fluency of the researcher and consultant in the respective languages. Sociolinguistic factors may also come into play; see AnderBois and Henderson 2015.

A worry that has been raised in the literature is that elicitation using a meta-language will influence the results obtained (Harris and Voegelin 1953, and see for example Zhornik and Pokrovskaya 2018 for a recent claim to this effect). Matthewson (2004) argues against this view, claiming that the influence of the meta-language is negligible. She points out that even if speakers are influenced in their productions to assimilate to, for example, the word order of the meta-language, they will still reject sentences that are ungrammatical or infelicitous in context, and will never offer unacceptable sentences.

As AnderBois and Henderson (2015) argue, the linguistic phenomenon under investigation may influence the choice of language. Using the meta-language is favorable when it is impossible to avoid the linguistic phenomenon being investigated within the context de-
scription itself. Meanwhile, using the target language may be beneficial when a distinction being investigated cannot easily be paraphrased in the meta-language. See AnderBois and Henderson 2015 for concrete examples of each case.

As we have already stressed, the presentation of a context is necessary as the fieldworker and consultant need to share the discourse context against which a sentence is judged. There is always the risk that the consultant could still envision extra context beyond what the fieldworker verbally describes, and this needs to be controlled for as much as possible. But verbal context descriptions can become quite long and complicated when we want to control for specific factors. This challenge can be overcome by making the move to elicitation materials involving visual representations, to which we now turn.

3.2. Using individual pictures

Let’s say the researcher wants to investigate the semantics of comparison constructions like the following:

(4a) This ladder is taller than that one.
(4b) This ladder is tall compared to that one.

While the English sentences in examples (4a)-(4b) can both be used to make comparisons, they differ in which types of contexts they can be used in. In particular, sentences like example (4b) are odd in so-called crisp judgment contexts, where the two objects only differ in the relevant property very slightly (Kennedy 2007).

How could the researcher describe such a context to a consultant? One could say “Imagine that I have a tall ladder, and another that’s almost as tall, but slightly shorter. Could I say [target language sentence like those in examples (4a) or (4b)]?” A potential issue is that gradable predicates like tall - used in the verbal description - are vague and context-dependent, meaning that there is no precise cut-off point between being tall and being not-tall, and the classification depends on the sorts of things being compared (e.g., tall for a 10-year-old vs. tall for a skyscraper). The terms almost and slightly, also used in the context description, also have this property. So chances are high that the researcher and consultant may not have the same context in mind at this point. One could perhaps use measure phrases to get the point across (“Imagine that one ladder is 4 meters tall, and another ladder is 3.9 meters tall ...”). However, not all languages have such measure phrases, so this could be problematic depending on the language being used to describe the context, and even in languages that do, not all gradable predicates can be modified by measure phrases (cf. ??4 meters big).

Suppose instead that the researcher shows the consultant the picture in Figure 2. Now the researcher can just point to the two leftmost ladders in the picture and ask the consultant if a comparison construction in the target language can be used to compare them. (For English speakers, the sentence in example (4b) is typically judged unacceptable as a comparison of
Figure 2
Comparing heights of ladders (Bochnak and Bogal-Allbritten 2015); e.g., "This ladder is tall compared to that one."

The simplest type of visual presentation of a context is through the depiction of a situation in a single picture (photo or drawing). An important advantage of using pictures is that they can perform double duty: they can be used both for production tasks to elicit grammatical sentences describing the picture, and to probe for negative evidence, i.e., discovering what kinds of sentences cannot be used to describe the picture.

The single-picture method works quite well for investigating expressions that describe scenarios that can easily be represented in a static format. An early example is the BowPed picture series mentioned in section 2 (Bowerman and Pederson 1992). Importantly, in addition to using these pictures for an elicited production task, Bowerman and Pederson explicitly state that the materials should be used to probe for negative evidence, in order to determine the boundaries of use for spatial expressions, especially for those that may appear to be similar in meaning (e.g., in vs. inside).

In another example, Louie (2015) makes use of a picture prompt for eliciting judgments on distributivity, i.e., whether a property distributes over the entities that make up the plurality denoted by the subject. Louie used pictures like those in Figure 3 to test whether distributive readings were available for sentences in Blackfoot (Algonquian) which translate as “Three dogs chased two cats”. Arrows are used to represent the chasing relation between the plural subject (“three dogs”) and plural object (“two cats”). Context B in Figure 3 only supports a distributive reading, so if a speaker accepts a sentence in context A but not in context B, this suggests that a distributive reading is unavailable for that sentence.

Other researchers have successfully used the single-picture method for investigating a
A range of topics in meaning. Ameka et al. (1999) build on Bowerman and Pederson’s materials to develop a picture series targeting positional verbs (e.g., sit, stand, lie). Bruening (2008) developed a picture series to investigate the scopal properties of quantifiers (e.g., “Not every woman is riding a horse”). Bochnak and Bogal-Allbritten (2015) and Bowler (2016) argue that visual stimuli can be useful for exploring comparisons and superlatives of dimensional physical attributes (e.g., “This ladder is taller than that one”; see Figure 2). Lima (2014) likewise used picture stimuli for eliciting quantity comparisons (e.g., “This person has more rice than that person”).

Importantly, tasks involving pictures require more than simply presenting a picture and asking a speaker to give a sentence describing it. A picture is worth a thousand words, and there may be many sentences that truthfully describe the image but which do not contain the grammatical phenomenon that the researcher is targeting. For instance, the picture in Figure 2 can be truthfully described using the sentence “There are nine ladders”, but this is not a target sentence type for investigating comparisons or superlatives. Extra instructions are necessary in order to guide the types of responses requested from consultants. As mentioned above, Bowerman and Pederson (1992) instruct that the speaker should be shown a picture and asked the question, “Where is [the figure]?”, in order to guide the intended responses. Setting up extra context in addition to the content of the picture thus further constrains the types of responses a speaker might give, and allows the researcher to better target a particular grammatical phenomenon.

As we have already hinted, the suitability of using single pictures for delivering the context for an utterance depends on the extent to which a scenario can be represented in a static format. In many of the case studies listed above, the target utterances describe stative, non-dynamic eventualities (e.g., the position or location of an object, a physical property of an
object, or an amount of a substance). However, more than a single picture might be needed to effectively depict certain dynamic eventualities. To address this issue, Bruening (2008) designed some of his materials as two-picture “before” and “after” sequences. For instance, as a target to elicit “A woman picked up every tool”, the “before” picture features four women standing with four tools lying on the floor, and in the “after” picture each woman is holding a tool.

Bruening’s materials also contain a few sequences of between four and six pictures, which are intended to convey a short story. The picture sequences are each accompanied by a short script intended to guide the production of particular sentence types. These sequences are intended to target the possible scope relationships between indefinites and negation (e.g., “The man didn’t catch a fish”). The method of using sequences of pictures to convey more elaborate discourse contexts has been further developed as the storyboard methodology, to which we turn next.

3.3. Storyboards

A storyboard is, simply put, a series of pictures which tell a story. In this section we overview several ways in which storyboards can be used in semantic fieldwork, and we illustrate the method with reference to one particular storyboard (from Vander Klok, to appear). We also outline some recent new developments in the use of storyboards, and conclude by briefly addressing some potential challenges with storyboard-based fieldwork.

One way in which series of pictures are frequently used is to elicit a connected narrative from a language consultant. A significant advantage of the method is that it has the potential to elicit fluent, connected speech, with no danger of translation interference. An early example of this technique is found in Berman and Slobin 1994. In Berman and Slobin’s Frog Stories project, children who were acquiring a number of different languages were invited to tell a story based on the pictures in a published storybook (Mercer Mayer’s *Frog, Where Are You?*). The resulting narratives were analysed and compared for features including tense and aspect marking, and devices which contribute to narrative structure.

If one’s goal as a fieldworker is to test specific hypotheses or generalizations, then a more directed approach to designing one’s series of pictures is called for. Burton and Matthewson (2015) dub the hypothesis-testing version of the storyboard technique ‘targeted construction storyboards’. Targeted construction storyboards differ from Berman and Slobin’s methodology in two core respects. First, they are designed by the researcher to elicit specific linguistic elements, or to create a context which supports particular truth and/or felicity conditions. There is thus at least one separate storyboard for each research question. Second, targeted construction storyboards are crucially intended to be used in tandem with follow-up elicitation tasks such as acceptability judgment tasks. The follow-up elicitation enables the collection of negative data, and it allows the researcher to check forms which did not spontaneously arise.
when the story was first told.

An illustrative example of the targeted construction storyboard technique comes from Vander Klok (to appear). Vander Klok’s storyboard, ‘Bill vs. the weather’, was designed to elicit epistemic modal claims with past temporal perspective and future orientation. These are claims that at some past time \( t \), it was epistemically possible for some agent (roughly: the agent believed it was possible) that an event would happen at some time after \( t \).

Epistemic modals with past temporal perspective are a good candidate for storyboard elicitation for a couple of reasons. The question of whether these readings exist has been controversial in the literature, with disagreement about the empirical situation in various languages. Furthermore, the readings are challenging to elicit using verbal context descriptions, because the background required is complex and involves keeping track of interlocutors’ changing beliefs over time (see also Burton and Matthewson 2015 for discussion).

In ‘Bill vs. the weather’, Bill is a forgetful person who often arrives at work having forgotten his umbrella when it was raining, or other necessary clothing items for the weather. One sunshiny day, he shows up with a backpack full of weather-protection items. The day after that, his colleague asks him why he had brought each item the day before, and the response is that ‘It might have rained/snowed/been windy.’ These responses involve epistemic modals with past temporal perspective and future temporal orientation: at some past time \( t \) (yesterday morning), it was epistemically possible for Bill that there would be bad weather later that day. At the utterance time the following day, it is no longer epistemically possible for Bill that there was bad weather yesterday (since he knows that it in fact turned out to be sunny). Sample pictures from the storyboard are given in Figure 4 (Bill arriving at work with his backpack; Bill’s colleague asking him why he brought an umbrella; Bill replying that it might have rained later).

When this storyboard was used with speakers of Paciran Javanese (Austronesian; Indonesia), Vander Klok was able to establish that modals which are independently known to have exclusively epistemic interpretations are acceptable in the target contexts. This was established in follow-up elicitation, since speakers did not volunteer the epistemic modals when
originally telling the story (instead using other strategies including evidential markers, for example).

The popularity of storyboards as a targeted elicitation method has increased substantially in recent years, and many researchers are sharing their materials, supporting the possibility of replicating parallel results across languages.\(^9\) As more semantic fieldworkers are creating targeted construction storyboards, the methodology for their use has also developed in interesting ways. As well as being used to elicit spontaneous narratives, storyboards can also be treated simply as visual presentations of discourse context which lead directly into standard elicitation tasks (such as acceptability judgment tasks). See for example Sadlier-Brown (2018), and discussion in section 4 below. And while in their original conception, storyboards were intended to avoid translation, and were also intended to be usable with consultants who are not literate in the relevant language(s), recent work is exploring the use of written prompts in conjunction with, or within, the images; see for example Nouri Hosseini et al. (2018), and section 4.1 below.

As with any method, storyboards can involve challenges, and may not be useful in every fieldwork context. Most challenging, perhaps, are cultural contexts which do not support the concept of a series of pictures as underlying a story. In such cases (or even with new consultants in any cultural context), the consultants may describe the pictures individually, rather than creating a connected narrative. See for example Moro and Klamer (2018) on their work with Alorese (Austronesian, Indonesia). Even among cultures in which picture-narratives are an accessible concept, not all are familiar with the Western convention of speech or thought bubbles (see Figure 4 for the use of speech bubbles in the Bill vs. the weather storyboard). Lesser cultural complications include the fact that text-supplemented storyboards can obviously not be used where literacy is absent, and that individual storyboards may often need to be redrawn to adapt to the local situation (using appropriate ethnicity, clothing, objects, flora and fauna, etc.). That said, with patience and creativity, storyboards can often be used even in the most challenging fieldwork situations, involving monolingual speakers without a

---

\(^9\)Online repositories of storyboards include the Totem Field Storyboards website (http://totemfieldstoryboards.org); the MelaTAMP storyboards, designed for testing tense-aspect-mood categories in Melanesian languages (https://wikis.hu-berlin.de/melatamp/Elicitation_materials (see von Prince et al. (2018) for cross-linguistic results using these storyboards); the Tübingen Elicitation Materials for Linguistic Fieldwork site (https://fieldworkhub.wordpress.com/storyboards/); and the Eh-Lab storyboards for speech acts and particles (https://syntaxofspeechacts.linguistics.ubc.ca/). Skopeteas et al.’s (2006) questionnaire also contains several storyboard tasks (https://www.sfb632.uni-potsdam.de/quis/quis-materials.html), as does the Max Planck Field Manuals and Stimulus Materials site (http://fieldmanuals.mpi.nl; see e.g. Eisenbeiss et al. 1999). A related though slightly different methodology is Katie Sardinha’s Story Builder (http://www.story-builder.ca), which consists of a set of flashcards with individual images that can be chained together to form novel stories interactively with consultants. Storyboards have also been used in language acquisition research, e.g., Gagarina et al. (2012).
tradition of picture-aided storytelling. See for example Salles (2018), who successfully used storyboards in her work with the Pirahê (Mura; Brazil).

### 3.4. Videos

Videos are also used in fieldwork for both production and judgment tasks. An early example is the Pear Story video (Chafe 1980). This six-minute film contains sound effects but no words or dialogue. It features a farmer harvesting pears, and a boy passing by who steals some of the pears. A language consultant is supposed to watch the film, and then tell the story in their own words. Similar to the Frog Stories described in the previous section, this task is not designed for eliciting any particular linguistic phenomenon; the intent is to elicit a connected narrative from the consultant. However, the Pear Story task has been used to investigate aspects of cross-linguistic comparison, for instance comparing the use of classifiers across Chinese dialects (Erbaugh 2001).

The use of targeted video stimuli has recently been advocated for by Bar-el (2007, 2015) for investigating aspectual distinctions. For instance, there is cross-linguistic variation in whether accomplishment predicates lexically entail a culmination or merely implicate culmination (Bar-el et al. 2005; Travis 2010). In English, asserting that a past perfective accomplishment did not reach its culmination is contradictory as in example (5b), indicating that the culmination inference is an entailment. In Stát’ímcets (Lillooet Salish), accomplishments also give rise to a culmination inference, indicated by the speaker’s comment in example (6a), but the continuation in example (6b) denying the culmination does not lead to contradiction, suggesting that this inference is an implicature (see Martin 2019 for an overview of this phenomenon cross-linguistically).

(5a) Mrs Ramsay wrote a letter.

(5b) #Mrs Ramsay wrote a letter, but she didn’t finish writing it.

(6a) Ts’áqw-an’=lhkan ta=n-kiks=a
    eat-TR=1SG.SU DET=1SG.POSS-cake=EXIS
    ‘I ate my cake.’
    Speaker’s comment: “Sounds like you ate all of it.”

(6b) Ts’áqw-an’=lhkan ta=n-kiks=a
    lhkúnsa ku=sq’ít t’u7
    eat-TR=1SG.SU DET=1SG.POSS-cake=EXIS now DET=day but

(Stát’ímcets)

(Smith 1997)
qelh-cál=lhkan ku=k’wík’wena7 t’u natcw
save-ACT=1SG.SU DET=few until tomorrow
‘I ate my cake today, but I saved a little for tomorrow.’

(adapted from Bar-el et al. 2005)

This distinction is quite subtle, and is unlikely to be discovered on the basis of translations of English sentences containing accomplishments. Such a task would be awkward anyway, since the English translation equivalent of the St’át’imcets sentence in example (6b) is contradictory. Bar-el argues that this difficulty can be overcome by designing video clips that depict accomplishments that do or do not culminate (e.g., an agent crossing a bridge the whole way, or turning back part way across). The videos can be used both for elicited production (the speaker describes what happens in the video) and for judgment tasks (asking the speaker if a sentence can felicitously describe what happens in the video).

Compared to the Pear Story video, Bar-el’s videos are more similar to storyboards, in that they are targeted towards eliciting a particular linguistic phenomenon. At the same time, her videos are in a sense more like individual pictures than storyboards because there is no over-arching story line – Bar-el’s short video clips are only intended to elicit the production of or a judgment about a single sentence. But in principle, videos can also be targeted at particular phenomena and have an over-arching story line to make them more storyboard-like.¹⁰

Videos do have some drawbacks compared to storyboards though. First, an advantage of using storyboards is that it allows speakers to go at their own pace and allow for easy back-tracking, whereas a video would need to be paused or cued to the right scene for follow-up questions. Second, videos require more extensive production and specialized equipment to create, whereas storyboards just require a pen and paper. This situation is changing however, thanks to new technologies such as mobile phone apps that make it easier to create video-based materials. Finally, for storyboards, it’s easy to draw whatever you want (e.g., ghosts, talking animals), whereas filmed videos need to be “realistic”, unless one has the capabilities for creating animations or other special effects.¹¹


¹¹Drawings including “unrealistic” characters were used to good effect by Koch (2008), for example. In his investigation of focus in Neʔkepmxcin (Thompson River Salish), Koch had computer characters ask questions which the consultants answered.
4. CONTROLLING FOR BELIEF STATES

Now that we have reviewed several techniques for presenting contexts to speakers for production and judgment tasks, we turn to the question of whether certain methods are better suited for investigating particular linguistic phenomena.

We propose that phenomena which involve reference to interlocutors’ epistemic states are particularly well served by the storyboard technique. This is because a belief state cannot easily be represented in a single picture, and can quickly become too complex to describe verbally. Verbal descriptions of belief states can be especially tricky if there is incomplete control of the meta-language (on the part of the fieldworker or consultant). In a storyboard, a character’s belief state can be naturally conveyed by what they say, how they interact with others, and potentially through the use of thought bubbles (which do not have to contain language; see Figure 4). Videos in principle have all these advantages as well, except that thought bubbles are not necessarily available.12

The phenomena we have in mind that have to do with interlocutors’ epistemic states include, but are not limited to, presupposition triggers, focus-sensitive elements, discourse particles, and epistemic modals. Furthermore, we claim that “out-of-the-blue” contexts belong to this class of phenomena, since a lack of shared knowledge between interlocutors needs to be controlled for. Note that the class of phenomena that we intend to cover with our proposal is partially similar to, but does not completely overlap with, distinctions such as semantics vs. pragmatics, or at-issue vs. not-at-issue content. Indeed, certain phenomena fall squarely under the rubric of truth-conditional semantics, but still have to do with interlocutor belief states (e.g., the past epistemics discussed above). Note also that we are not claiming that storyboards shouldn’t be used for phenomena that don’t involve interlocutor beliefs, though it could be the case that simpler context presentation methods may be sufficient for such phenomena.

In the following sections we will argue for our hypothesis on the basis of two small pilot studies involving de re/de dicto contrasts and out-of-the-blue contexts.

4.1. A pilot study: the de re vs. de dicto contrast

The de re vs. de dicto contrast is surely one of the most challenging to elicit in a fieldwork context. After all, most linguistics students experience some level of puzzlement or confusion when trying to access the judgments. Moreover, the de re vs. de dicto contrast is all about interlocutors’ beliefs, and as such is predicted by our hypothesis to be a prime candidate for storyboard elicitation.

12A character’s private beliefs might be presented through "voice-overs" in a video, but this would reduce the opportunities for cross-linguistic parallelism in the stimuli.
The contrast is illustrated with respect to the sentence in example (7), with descriptions below closely following Maier (2009:432-433).

(7) John believes the president of PepsiCo is rich.

(Maier 2009, p. 432)

Under a de dicto reading, John might believe that the president of PepsiCo is rich because, e.g., he generally believes presidents of large corporations are rich, but he might not know the identity of this person in the actual world. This means that we cannot substitute the president of PepsiCo with Indra Nooyi in the belief report, even if Indra Nooyi happens to be the current president of PepsiCo. Under a de re reading, John believes of a certain individual, who happens to be the actual president of PepsiCo, that she is rich. On such a construal, the report is felicitous if John met Indra Nooyi at a party and they talked about her yacht, so he now believes “Wow, this woman is rich.” On this reading, we can substitute the definite description with a co-referential noun phrase, even though John still might have no idea that the individual is the president of PepsiCo, or that her name is Indra Nooyi.

A simple thought experiment should reveal that conveying to a consultant what is wanted for a de re/de dicto task using only verbal descriptions is likely to be quite challenging, to say the least. In order to investigate whether storyboards could successfully elicit the contrast, we created two new storyboards: ‘Willem-Alexander’ (de dicto) and ‘Miranda Smith’ (de re).

The premise for these two storyboards is that Mary is conducting a survey about who people think is rich. Mary learns about what her interviewee believes and knows through his speech, and afterward she records her survey results using different descriptors to refer to the protagonists. The fieldwork consultant is then asked to judge whether what Mary writes in her report is true. Sample pictures from these storyboards are given in Figures 5 and 6 below, and the full versions are available at [URL to be filled in]. One feature of these storyboards is that they are text-heavy, unlike prototypical storyboards which have little or no text (e.g., ‘Bill vs the weather’, discussed in section 3.3). However, our de dicto and de re storyboards still involve simpler language in the prompts than pure verbal descriptions would (as the matrix layer ‘X believes that ...’ is conveyed visually).

The results of a small pilot study (conducted with five non-linguist native speakers of English) were suggestively positive. As predicted, all speakers judged that Mary’s report was true when she repeated the original description, or used a description which was clearly available to her interviewee. For the test sentences where Mary substituted a name which was unknown to her interviewee, responses were mixed, but trended in the right direction. Three participants liked the de re substitution, while only one participant fully accepted the de dicto substitution.

We suspect that for such a subtle distinction as the de re/de dicto contrast, clear results will require a larger number of participants, randomized presentation order, and filler conditions, following standard psycholinguistic methods (see e.g., Abbenhuis et al. 2014; Arunachalam
We suspect that for such a subtle distinction as the de re/de dicto contrast, clear results will require a larger number of participants, randomized presentation order, and filler conditions, following standard psycholinguistic methods (see e.g., Abbuhl et al. 2014; Arunachalam 2013). For fieldwork with few consultants, the task may need to be repeated at different intervals to check for consistency of the judgments over time. However, presenting the contexts in the form of storyboards appears very likely to be successful, whereas verbal context descriptions are likely to each be a paragraph long and risk becoming extremely complex.

Amy Rose Deal has recently conducted a detailed investigation into the availability of de re and de dicto readings of attitude reports in Nez Perce (Sahaptian) using mainly verbal context descriptions with her consultants (Deal 2018). Her results appear convincing, but she has reported to us (personal communication, December 21, 2018) that the data collection was challenging, and required testing many contexts (many of which are indeed a paragraph in length) in order for the empirical generalizations to emerge. Also noteworthy is the fact that in
4.2 Out-of-the-blue contexts

So-called out-of-the-blue (OOTB) contexts are used by fieldworkers to address a range of research questions. As discussed by Tonhauser and Matthewson (2017), OOTB contexts are sometimes used to establish what the context-independent meaning of an expression is (for example, what the default temporal interpretation of an unmarked form is (Smith et al., 2007), or what the context-independent contribution of scalar items is (van Tiel et al., 2016)). Another use of OOTB contexts is when attempting to prove that a particular element places constraints on the context (such as requiring a particular piece of information to be already in the interlocutors' common ground). Such elements are predicted to be infelicitous in an OOTB context. And all-new sentences in an information structure context are also typically elicited with OOTB utterances.

How does one elicit out-of-the-blue judgments? It might seem that a good method would simply be to present the consultant with utterances without any discourse context. However, this is not the case. Tonhauser and Matthewson (2017) argue that 'It is problematic to draw conclusions from utterances presented in isolation, since the researcher has no control or knowledge of what discourse context the consultant may be imagining while performing these tasks'; see also Crain and Steedman 1985. The point here is that asking a consultant to perform a task involving a contextless utterance essentially leaves the consultant the leeway to similar to our de dicto storyboard, in her Nez Perce research (Amy Rose Deal, p.c.).

Nez Perce, there are dedicated syntactic constructions for de re attitude reports, which might make the judgments a bit easier for consultants (or at least easier for consultants to detect differences between varieties of attitude reports). In any case, while using verbal context descriptions for testing de re/de dicto attitudes can be successful, we nevertheless believe that creating storyboards will help to make the judgment tasks easier for the consultant and fieldworker alike.\textsuperscript{13}

\textsuperscript{13}Interestingly, although she did not report this explicitly in her paper, Deal also used short storyboards, similar to our de dicto storyboard, in her Nez Perce research (Amy Rose Deal, p.c.).
4.2. Out-of-the-blue contexts

So-called out-of-the-blue (OOTB) contexts are used by semantic fieldworkers to address a range of research questions. As discussed by Tonhauser and Matthewson (2017), OOTB contexts are sometimes used to establish what the context-independent meaning of an expression is (for example, what the default temporal interpretation of an unmarked form is (Smith et al. 2007), or what the context-independent contribution of scalar items is (van Tiel et al. 2016)). Another use of OOTB contexts is when attempting to prove that a particular element places constraints on the context (such as requiring a particular piece of information to be already in the interlocutors’ common ground). Such elements are predicted to be infelicitous in an OOTB context. And all-new sentences in an information structure context are also typically elicited with OOTB utterances.

How does one elicit out-of-the-blue judgments? It might seem that a good method would simply be to present the consultant with utterances without any discourse context. However, this is not the case. Tonhauser and Matthewson (2017) argue that “It is problematic to draw conclusions from utterances presented in isolation, since the researcher has no control or knowledge of what discourse context the consultant may be imagining while performing these tasks”; see also Crain and Steedman 1985. The point here is that asking a consultant to perform a task involving a contextless utterance essentially leaves the consultant the leeway to imagine their own context. This therefore does not satisfy the desideratum that the fieldworker and the speaker have the same context in mind.

Here we make two claims about eliciting OOTB contexts. First, interlocutor ignorance must be explicitly established. (This claim is independent of the method of presenting the context.) And second, storyboards are the best, or at least a very reliable, way to explicitly establish ignorance.

We’ll start by looking at some attempts in the literature to establish interlocutor ignorance. One of the most rigorous examples of setting up discourse contexts to control for interlocutor knowledge comes from Tonhauser et al. (2013), in their investigation of what they call “strong contextual felicity.” Elements which are subject to a strong contextual felicity constraint impose a restriction on the discourse context. Presupposition triggers, for example, are felicitous in contexts which entail or imply the presupposition, and infelicitous in contexts which don’t.

In (8) and (9), the status of the Paraguayan Guaraní element avei ‘too’ is being tested, specifically whether its presence gives rise to the implication that somebody besides the bus driver is eating empanadas. In both cases, the context is controlled in that the utterance is set up to be out-of-the-blue/discourse-initial. In example (8), the context does not entail or imply either that somebody else is eating empanadas besides the driver, or that nobody else is. The utterance containing avei is rejected. However in example (9), the context also sets up shared knowledge of the presence of empanadas: somebody is eating empanadas besides
(8) Context: Malena is eating her lunch, a hamburger, on the bus going into town. A woman who she doesn’t know sits down next to her and says:

#Ñande-chofeur o-karu empanáda avei.
A1PL.INCL-driver A3-eat empanada too
‘Our bus driver is eating empanadas, too.’

(Tonhauser et al. 2013)

(9) Context: Malena is eating her lunch, empanadas, on the bus going into town. A woman she doesn’t know sits down next to her and says:

a. Ñande-chofeur o-karu empanáda avei.
A1PL.INCL-driver A3-eat empanada too
‘Our bus driver is eating empanadas, too.’

(Tonhauser et al. 2013)

The acceptability contrast between examples (8) and (9) suggests that avei imposes a strong contextual felicity constraint. This minimal pair illustrates that there is not just one monolithic ‘out-of-the-blue’ context: there are varieties of OOTB/discourse initial contexts that need to be controlled for. Tonhauser et al.’s strong contextual felicity diagnostic is successful in explicitly setting up the absence of prior discourse, and an additional advantage is that it has the collection of minimal pairs built in (since contexts which do entail or imply the relevant piece of meaning are contrasted with contexts which do not).

One issue we would like to draw attention to is that Tonhauser et al. avoid explicit statements about interlocutor (lack of) knowledge. For instance in (8), it is simply not mentioned whether or not somebody other than the bus driver is eating empanadas. While we agree that it may be preferable to avoid explicit statements to this effect, we also think there is a case for always explicitly controlling - perhaps indirectly - for the lack of addressee knowledge, in order to avoid consultants potentially adding extra information in their minds. That is, we suggest that ignorance cannot always safely be assumed on the basis of a piece of information simply not being mentioned in the context description.14

To illustrate the fact that verbal context descriptions of OOTB contexts can be subject to consultant context-enrichment, in spite of the researcher’s best efforts to clarify, we offer

---

14Consider in this light example (i), designed by Tonhauser et al. (2013) to test whether appositives impose a strong contextual felicity constraint.

(i) Context: The children in a history class have to give presentations about famous people. Malena has to talk about the pope. She starts with:
an example from the fieldwork of one of this paper’s authors. The phenomenon being tested in this case was verum. Verum is, roughly speaking, marking of emphasis on the truth of the proposition expressed; it is encoded by stress on an auxiliary or finite verb in English (and German; Höhle 1992 and much subsequent work). An example is given in (10), where all caps represents stress accent. Verum is infelicitous out-of-the-blue, as shown in example (11).

(10) Speaker A: Good thing it wasn’t a dark and stormy night when you were camping yesterday.
    Speaker B: It WAS a dark and stormy night.

    It was / #WAS a dark and stormy night.

Now consider what happened when our second author was testing whether the Gitksan (Tsimshianic) particle $k’ap$ marks verum, and therefore is infelicitous in out-of-the-blue contexts. Example (12) shows the utterance to be tested; the prediction if $k’ap$ is a verum marker is that the utterance will be infelicitous in this context.

(12) Context to test: Hector walks into the room and we have some surprising news for him:

\[ K’ap \text{ yukw}=hl \text{ dim nek}=s \text{ Aidan! } \]
\[ \text{VERUM IPFV}=\text{CN PROSP marry}=\text{PN Aidan} \]
\[ ‘Aidan IS getting married!’ \]

(Gitksan sentence to be tested)

When the $k’ap$-version of this utterance was first presented to the consultant, with the discourse-initial context as described, he accepted it, commenting “Yeah. It means ‘No, he really is.’ You’re not kidding.” This goes against the predictions of the hypothesis that $k’ap$ marks verum. However, when the researcher actually uttered the sentence in a situation where somebody had just entered the room, the same consultant laughed and emphatically rejected the utterance. This shows that even a verbally-described scenario which the researcher believes establishes a discourse-initial context may not be viewed as such by a consultant.

Papa Bendécto 16, o-nasé-væ-kue Alemánia-pe, oi-ko Róma-pe.
Pope Benedict 16 A3-born-RC-NOM.TERM Germany-in A3-live Rome-in

‘Pope Benedict the 16th, who was born in Germany, lives in Rome. (Paraguayan Guaraní; Tonhauser et al. 2013)

We believe it is possible that if appositives did have a strong contextual felicity constraint, some consultants might enrich the context, accommodating the information that the class had previously been taught where the pope was born. Such consultants would accept the utterance, obscuring the fact that a strong contextual felicity constraint was present.
We have just seen that establishing interlocutor ignorance can be challenging to describe verbally in a foolproof way. And ‘acting out’ a scenario on the spot, although it can be useful (as with the Gitksan verum case just discussed), is not always possible. In order to investigate the idea that establishing interlocutor ignorance is well-suited to the storyboard technique, we designed a set of storyboards to test verum marking in OOTB contexts, and conducted a small pilot study on their effectiveness.

The four storyboards in this set involve two in which ignorance is explicitly controlled for (in that the characters confirm that they have no information about the relevant issue), and two where ignorance is only implied (via the characters simply not discussing the issue). In each pair of storyboards, there is one whose test sentence involves verum marking, and a minimally different utterance without verum. Our predictions about the results are as follows. The non-verum utterances should be judged as acceptable (since there is no constraint on the context imposed by these utterances), a verum-marked utterance should be firmly rejected in the case where ignorance is explicitly established, and there may be variable judgments in the case where verum is present and ignorance is not sufficiently controlled for (due to the fact that some but not all consultants may enrich the context). Sample pictures are provided in Figure 7. In this storyboard, ignorance about Charlie’s recent life is explicitly mentioned, and therefore the use of verum in the final panel is predicted to be infelicitous. The full storyboards can be found at [URL to be filled in].

In a pilot study on six native speakers of English and one native speaker of Gitksan, our predictions were upheld. Participants accepted the non-verum utterances, unanimously rejected the verum utterance in the case where ignorance was explicitly established (as in Figure 7), and were split down the middle for the case where verum followed only an implicit absence of prior discourse.

Moreover, follow-up interviews supported our hypothesis about the reason for the ‘acceptable’ judgments in the case where verum was used in an apparently OOTB context. We asked participants who had accepted the verum utterance in this storyboard if they wished to expand on their answers; their responses clearly indicated that they had mentally enriched the context to include relevant prior conversation between the interlocutors. For example, one participant explained that she was ‘assuming they have discussed it’; another stated that he made ‘the assumption that the background included contact with, and conversation about, Charlie.’ Importantly, this contextual enrichment was not possible in the storyboard where ignorance was explicitly controlled for. This was also pointed out by one participant, who stated that in the more strictly controlled storyboard, the verum utterance ‘could only have been OK if there had been a particular previous conversation and we know there hadn’t been.’

This case study reinforces our proposal that when testing out-of-the-blue contexts, storyboards are a very useful method. It also supports our claim that it is better to explicitly
control for interlocutor ignorance rather than leave this issue vague or unexpressed. As we have seen, our consultants’ comments in the verum pilot test clearly reveal that when it is possible for extra contextual background to be accommodated, a certain proportion of consultants will do so. For this reason, the storyboards which explicitly controlled for ignorance were more robust and successful at establishing that verum requires a particular prior discourse context.\footnote{The possibility for contextual enrichment by the consultant, as well as the importance of providing a forum for consultants to provide comments, is further illustrated in example (i). Neither consultant outright rejected this OOTB verum-marked utterance, but both gave comments indicating that prior}

Figure 7
Sample pictures from verum storyboard (‘Charlie DID get married’)
This case study also highlights the fact that creating a storyboard may not by itself immediately lead to one’s desired results. It must be a “good” storyboard, that controls for the context in the appropriate way. The development of a “good” storyboard often requires some trial and error, and it can be helpful to try out the storyboard on a few native speakers of one’s own language, so that some of the bugs can be worked out before using it in the field. The helpfulness of the consultants’ comments furthermore reveals an advantage of face-to-face elicitation rather than elicitation “at a distance”, e.g., via an online survey. Working face-to-face facilitates a feedback loop where consultants’ comments can signal potential problems to the fieldworker; this opportunity is missing from a large-scale experiment with no opportunity for feedback.

5. CONCLUSION

In this paper, we set out first to overview the state of the art in techniques for semantic fieldwork elicitation, and second to address the question of which methods are best suited for presenting contextual information to consultants. We focused on the presentation of contextual information for two reasons: (i) utterances are only true or false, or felicitous or infelicitous, relative to discourse contexts; and (ii) making sure the fieldworker and the speaker have the same discourse context in mind is one of the trickiest things to get right.

We suggested that certain methods for establishing discourse contexts are better suited than others for investigating certain types of semantic phenomena. Specifically, we hypothesized that setting up the discourse contexts for content that relates to interlocutors’ belief states is easier with more complex context presentation techniques such as storyboards. We do not claim that it is impossible to convey speaker beliefs or ignorance through other means. However, we do believe that storyboards, when done right, make the task of controlling for these factors much easier.

As a final note, we want to emphasize that there is no one-size-fits-all approach to semantic fieldwork methodology. We hope to have provided some useful techniques, some general conversation is crucial to its acceptability.

(i) Context: Adam and Betty are eating dinner quietly. Nobody has said anything yet. Betty suddenly says:

#K'ap siipxw=t Charlie/Tsaalii.
(#VERUM) sick=DM Charlie
‘Charlie is/#IS sick.’

Consultant 1’s comment: “Maybe he doubts it and had asked ‘Really?’ and she says ‘Yes, he’s really sick.’”

Consultant 2’s comment: “It probably came up in the conversation earlier.”

\[\]
operating guidelines, and some food for further thought. But each semantic fieldworker must construct their own methodological toolkit, appropriate for their own fieldwork situation. As fieldwork situations have a tendency to be ever-changing, researchers will find themselves constantly needing to revise and refine their methods, and to improvise.

ACKNOWLEDGMENTS

We are very grateful to St’át’imcets consultants Carl Alexander, the late Beverley Frank, the late Gertrude Ned, the late Laura Thevarge, and the late Rose Agnes Whitley, and to Gitksan consultants Vincent Gogag, Jeannie Harris, Hector Hill, and Barbara Sennott.

For helpful comments on an earlier draft of this paper, we are grateful to Eva Csipak, Henry Davis, Amy Rose Deal, Judith Tonhauser, Jozina Vander Klok, and an anonymous reviewer. We are also grateful to the audience at the Fieldwork: Methods and Theory Conference, University of Gothenburg, and the UBC Tense and Aspect in the Pacific Lab.

This research was funded in part by the Social Sciences and Humanities Research Council of Canada (grant #435-2016-0381), the Jacobs Research Fund and the Alexander von Humboldt Stiftung.

LITERATURE CITED


Beck, S., S. Krasikova, D. Fleischer, R. Gergel, S. Hofstetter, C. Savelberg, J. Vanderelst, and


Crain, S. and M. Steedman (1985). On not being led up the garden path: The use of context by the psychological parser. In D. Dowty, L. Karttunen, and A. M. Zwicky (Eds.), *Natural Language Parsing: Psychological, Computational and Theoretical Perspectives*, pp. 320–354. Cambridge:
Cambridge University Press.


Vander Klok, J. (to appear). Exploring the interaction of modality and temporality through the storyboard *Bill vs. the weather*. *Semantic Field Methods*.