Using Matching Tasks in Semantic Fieldwork

Rafael Abramovitz and Maša Močnik
{rafabr,mocnik}@mit.edu
Draft – September 2020

Abstract

We discuss an elicitation technique we call a *matching task*, a type of acceptability judgment task with elements of a translation task. This method has been crucial to the results in our joint semantic fieldwork (Močnik and Abramovitz 2019), but has not been explicitly discussed in the literature. In the paper we compare it with more standard elicitation tasks that we have found unsatisfactory for various reasons. We conclude by enumerating a set of circumstances which resulted in matching tasks being necessary in our work, and under which this task may be useful to other linguistic fieldworkers.

1 Introduction

As formal semanticists have paid more attention to data from understudied languages (Bittner 1987, et.seq.), the question of how to collect data from naïve native speakers has played an increasingly significant role in work on semantics. Matthewson (2004) was the first to draw attention to this question, showing that the text-based methods of American structuralist linguistics were inadequate for answering many of the questions of interest to formal semantics. Further, she offered a then-novel (though now widely accepted) defense of the feasibility of doing semantic fieldwork through a contact language. Since then, work employing the techniques argued for by Matthewson has become standard, and further work explicitly discussing methodological points in semantic fieldwork has also been forthcoming.[1]

In much of the work in this tradition, judgments obtained in tasks involving translations have been deemphasized: while acknowledging their

Acknowledgments to be inserted.

Our transcription uses the IPA, except that we use *č* for the voiceless alveolo-palatal affricate. Our glossing schema follows the Leipzig Glossing Rules, except for: AP - antipassive, CF - counterfactual, CS - causative, E - epenthetic vowel, IRR - irrealis, RLS - realis, VB - verbalizer.

[1] For discussions of methodologies, see Vander Klok (2014, 2019), Deal (2015), Tonhauser and Matthewson (2015), Cable (2019), Vander Klok and Connors (2019), Bochnak and Matthewson (2020), among others.

usefulness in certain situations, Matthewson nonetheless notes that "translations should always be treated as a clue rather than a result...The only real evidence about truth conditions is truth value judgments in particular contexts" (Matthewson 2004, 389). What can linguists do, then, in situations where speakers are unable or unwilling to give truth value judgments in contexts? In this paper, we report on our experience working on the extremely polyfunctional attitude verb *ivək* in Koryak, a highly endangered Chukotko-Kamchatkan language of northeastern Russia, where we found it necessary to employ an elicitation methodology we call a "matching task." This task crucially implicates translation: a native speaker is asked to judge whether a sentence in the target language (Koryak) can be understood in the context as a given sentence in the contact language (Russian).

After giving some background information on Koryak and our prior work (§2), we introduce the matching task (§3), compare it to the other tasks that have been discussed in the semantic fieldwork literature and why they were insufficient to gathering the data we were interested in (§4), and speculate about the combination of circumstances that led to it being necessary for us to employ matching tasks (§5).

2 Background

2.1 Koryak

Koryak is a highly endangered Chukotko-Kamchatkan language with several hundred speakers that is spoken in and around the northern part of the Kamchatka Peninsula in the Russian Far East. It is a pluricentric language made up of at least two dialect continua with limited mutual intelligibility between one another: our work has exclusively been with speakers of the Chawchuven dialect, traditionally spoken by the nomadic reindeer herders of central and northern Kamchatka. Our consultants are female native speakers of Koryak aged 50-85 who were raised in herding communities where no Russian was spoken; most of our consultants' parents, while multilingual, spoke only indigenous languages of the area.[2] The implementation of Soviet colonial policies in the Russian Far East saw our consultants sent to Russian-medium boarding schools, and they are consequently fluent L2 speakers of and literate in Russian (our contact language). We suspect

[2] All of the consultants that we have worked with extensively are female as there are very few male native speakers of Koryak left. The two male native speakers who were capable of consultant work that we found were too busy hunting and gathering to meet regularly.

that some of them have a near-native command of that language.[3] Author Abramovitz has intermediate proficiency in Koryak and fluent L2 proficiency in Russian, and has been carrying out fieldwork on Koryak since 2014. The data for the joint semantic fieldwork was collected between August 2018 and the present, via electronic communications (video/audio and messages) as well as during three fieldwork trips by Author Abramovitz (August-September 2018, July-September 2019, and November 2019).

2.2 Močnik and Abramovitz (2019)

The phenomenon that led to the formulation of the matching task was our study of the variable-force variable-flavor attitude verb *ivək* in Koryak (Močnik and Abramovitz 2019). Example (1), which has a matrix *ivək* and the indicative mood in the embedded clause, and example (2), which has the counterfactual mood in the embedded clause, illustrate some of the construals of *ivək*. Example (1) also illustrates that *ivək* does not express just any attitude (e.g. it cannot mean 'know' or 'imagine'). The 'wish' construal is not available in (1), but it becomes available with the counterfactual mood on the embedded verb, as in (2). Similarly, a directive flavor ('order' and 'suggest'), not mentioned above, is available with an embedded imperative or an infinitive. There is no embedding or configuration that would yield 'know' and 'imagine' felicitous to the speakers in a matching task.

(1) meʎʎo ∅-k-iv-ə-ŋ-∅, (əno)
Melljo.ABS.SG 2/3.S/A.IND-PRS-*ivək*-E-PRS-3.S.IND that
∅-ku-muq-et-ə-ŋ-∅
2/3.S/A.IND-PRS-rain-VBLZ-E-PRS-3.S.IND
'Melljo {says, thinks, allows, hopes, fears, *knows, *imagines, *wishes} that it's raining.'

(2) meʎʎo ∅-k-iv-ə-ŋ-∅, (iwke)
Melljo.ABS.SG 2/3.S/A.IND-PRS-*ivək*-E-PRS-3.S.IND if.only
n-ə-ʔ-ə-muq-et-ə-n
2/3.S/A.CF-E-CF-E-rain-VBLZ-E-2/3.S/O.CF
'Melljo wishes it would rain.'

(Močnik and Abramovitz 2019)

[3] None of the consultants speak English. Some of our consultants have knowledge of Chukchi, a close relative of Koryak's, including one speaker who is a native bilingual, and at least one other speaker who was one as a child. Another consultant is a native bilingual of Koryak and Even, an unrelated language of the Tungusic family spoken in central Kamchatka, and at least one further consultant has passive knowledge of that language.

In addition to expressing different flavors (assertive, doxastic, bouletic, etc.), sentences with *ivək* can also differ in the strength (force) of the claim. For example, *ivək* can convey a low credence ('allow for the possibility') in the truth of the embedded proposition, but it can also convey a higher one ('think'). As will be discussed in §4.1 with example (5), we used the Russian attitude verb *dopuskat'* ('allow for the possibility') to probe for this weaker doxastic meaning.[4] To foreshadow that data in (5), here is a different example of a weaker reading of *ivək*:

(3) Context: Two people went out hunting and haven't come back. Hewngyto said that it's possible that they got lost, but he also said that it's possible that they hadn't.

ʔewŋəto ∅-iv-i əno
Hewngyto.ABS.SG 2/3.S/A.IND-*ivək*-AOR COMP
taɣəjɲiŋ-ə-lʔ-ə-t ∅-təmŋew-ɣəʔe.
hunt-E-S/O.PTCP-E-ABS.DU 2/3.S/A.IND-get.lost-3DU.S.AOR
ənno ʔopta ∅-iv-i əno əčč-i
3SG.ABS also 2/3.S/A.IND-*ivək*-AOR COMP 3NSG-ABS.DU
jatan ∅-ko-pel-aɲ-ŋ-e
only 2/3.S/A.IND-PRS-remain-VBLZ-PRS-3DU

'Hewngyto suggested that the hunters had gotten lost. He also suggested that they are just late.'

In Močnik and Abramovitz (2019), we modeled this flexibility of flavor and strength in a way that is parallel to the analyses in the modal verb domain. We followed Rullmann et al. (2008) in positing a universal quantifier over possible worlds with a domain restriction that derives the weaker readings. The distinction between the assertive ('say' and 'suggest') and the doxastic ('think' and 'allow for the possibility') flavor was modelled with a free variable, like a modal base. Unlike the doxastic-assertive distinction, the bouletic flavor ('hope', 'fear', and 'wish') was shown to arise because of further material in the embedded clause, which is overtly signaled by the counterfactual mood with 'wish' and which we proposed is covert with 'hope' and 'fear'. This material combines with the doxastic interpretation of *ivək*, yielding the bouletic reading of the sentence and giving

[4] We followed Močnik (2019a,b), who discusses the Slovenian existential belief verb *dopuščati* ('allow for the possibility'), which the Russian one is a cognate of. See also footnote 10.

the illusion that *ivək* is itself a bouletic verb. The conclusion of our investigation was that *ivək* is polyfunctional in flavor (doxastic and assertive flavors) and polyfunctional in force (encoding a strong force with a weakening mechanism).

3 Matching task

We start by describing the matching task, before proceeding to a comparison with other tasks (§4) and a more general discussion of its usefulness in fieldwork (§5). The matching task presents the consultant with a situation c, described either in the target or in the contact language, and two sentences: a sentence p in the target language and a sentence p', typically in the contact language.[5] We refer to p' as the matching sentence. The goal is to elicit a judgment of p in c under the intended interpretation given by the matching sentence p'. Below, we sketch the basic idea behind the matching task (§3.1) and discuss it from a theoretical perspective (§3.2).

3.1 Basic idea

We build up to the results of the matching task in two steps. The speaker first performs an acceptability judgment task for p in c. This makes the speaker familiar with some of the material (namely, p and c). It also provides us with an opportunity to obtain volunteered information and commentary about p (e.g. a volunteered translation).[6] For example, we might eventually learn that a particular reading is salient, or that the sentence as we proposed it is syntactically ill-formed. The second step brings us to the actual matching task: we check whether p can be interpreted as p' in c. The way this task is executed depends heavily on the properties of the contact language (in particular, its lexical inventory) and possibly on other cultural factors. For our work with Koryak speakers, we used the Russian expression 'thought' (*mysl'*) and asked whether "p can express the same thought as p' in c." We found this to be a good way of checking whether p can be understood as p' in c since it yielded consistent results and it also produced

[5] The contact language is typically distinct from the target language, but this need not be so (as in monolingual fieldwork). Furthermore, p' could in principle be given in a third language, distinct from the contact and the target language. Perhaps we should consider the following common practice to be an instance of the matching task: an English-speaking semanticist asks their non-native-English-speaking colleague whether a particular sentence in their native language can mean "this" (pointing to a formula on the board).

[6] Such comments are also relevant data points in semantic fieldwork (Matthewson 2004).

negative answers. By contrast, if we mentioned 'meaning' (*značenije*) to try to ask whether "p can express the same meaning as p' in c," we instead triggered a word-for-word translation.

Here is a toy example from French to illustrate the matching task. Suppose that we are interested in knowing whether the French pronoun *elle* imposes gender restrictions on its referent. Let our contact language be English. In order to probe for how *elle* behaves, we could set up a situation with two potential referents, one male and one female, and use the matching task to see whether *elle* can refer to the female one:[7]

(4) a. Ann and Bob are fighting on the playground. (c)
b. *Elle est fâchée*. (p)
c. Ann is angry. (p'_1)
d. Bob is angry. (p'_2)

We ask the consultant to imagine a situation where Ann and Bob are fighting on the playground (c). We ask whether *elle est fâchée* is acceptable in this situation. Suppose they respond with 'yes' and offer no further comments. We then proceed to ask whether *elle est fâchée* can express the same thought as the matching sentence 'Ann is angry' in this situation. Presumably the answer is affirmative. If we had asked whether *elle est fâchée* could express the same thought as 'Bob is angry' here, the answer would have been negative.

3.2 The role of the matching sentence

As outlined above, the consultant in a matching task is ultimately asked to verify whether the sentence p (in the target language) can, in a situation c, express the "same thought" as p' (in the contact language) in c. What are we actually asking the consultant to do?

Here is a plausible way to understand what the matching task really does: the contact language sentence p' serves to enrich the context by specifying (some of) what is left open by the target sentence p. Take example (4): the matching sentence 'Ann is angry' contains reference to a particular female individual. By asking whether *elle est fâchée* can express "the

[7] It is obvious in this case that there are simpler ways of checking what *elle* means because English pronouns make gender distinctions. This is a lucky fact of English – if the contact language had been Koryak, which does not have a gender distinction in the pronouns, we would not be able to simply ask whether *elle* can mean 'she.'

same thought" as 'Ann is angry', the consultant is in effect asked to evaluate whether *elle* can refer to the entity designated by 'Ann'. In this case, the presence of the matching sentence allows us to set the assignment function to one where the pronoun refers to Ann. While we could have chosen a different method to achieve this, such as pointing to the entity Ann in a visually-depicted situation, it is sometimes challenging (if not impossible) to use other means to control for underspecification, as we discuss in §4.

The matching sentences we have been using in our fieldwork contain linguistic material that serves to disambiguate or resolve some aspect of the target sentence. We have found this to be a good use of the matching task since the consultants were able to pick up on the intended difference between the target and the contact language sentence, and disregard other potential differences (e.g. between *est fâchée* and 'is angry').[8]

In sum, the matching task has not been intended to test the equivalence of meaning between two sentences, which can be a problematic starting point (Deal 2015). Instead, the matching sentence provides contextual support, which makes the matching task at its core a kind of acceptability judgment task.

4 Comparison with other tasks

We consider three types of elicitation methods in turn: acceptability judgment tasks, elicited production tasks, and translation tasks (Bochnak and Matthewson 2020). In this section we will look at a number of cases where the matching task has given us an advantage over the standard elicitation methods, making it a useful supplement to the existing tools in fieldwork.

4.1 Acceptability judgment tasks

In an acceptability judgment task, the consultant is asked whether a sentence in the target language is acceptable in a given situation (presented

[8] One might object that the presence of the matching sentence influences the speaker to accept Koryak-Russian pairings more readily, and that the variety of interpretations of *ivək* that we have obtained follows from a methodological error. This makes the prediction that consultants should not consistently reject certain readings when they are presented in the matching task, which is incorrect. Furthermore, as Roger Schwarzschild (p.c.) suggests, one could carry out an experiment to test whether the presence of the matching sentence has an effect on speakers' judgments by using matching tasks on a well-studied language and with consultants who additionally speak a foreign language fluently, and comparing the data gathered via the matching task with the data gathered via acceptability judgment tasks reported in the literature.

verbally or visually). In this section we illustrate how this approach has been insufficient for our semantic fieldwork on Koryak. Consider the following example, initially presented as an acceptability judgment task. In this task, the consultant was read the story in (5).[9]

(5) Hewngyto is walking down the street. Melljo sees him and asks: 'Menno ɣənin ŋevətqet? Metke kotavareŋjaŋəŋ jajak?' (Where is your wife? Is she making jam at home?) He replies:

qoo. t-ə-k-iv-ə-ŋ-∅ əno
dunno 1SG.S/A-E-PRS-*ivək*-E-PRS-1SG.S that
∅-ko-ta-vareŋja-ŋ-ə-ŋ-∅ jaja-k
2/3.S/A.IND-PRS-make-jam-make-E-PRS-3.S.IND house-LOC

'I don't know. I allow for the possibility that she's making jam at home.'
He continues walking. Qechghylqot sees him and asks: 'Menno ɣənin ŋevətqet? Metke keluŋ umkək?' (Where is your wife? Is she picking berries in the forest?) Hewngyto replies:

qoo. t-ə-k-iv-ə-ŋ-∅ əno
dunno 1SG.S/A-E-PRS-*ivək*-E-PRS-1SG.S that
∅-k-elu-ŋ-∅ umk-ə-k.
2/3.S/A.IND-PRS-pick.berries-E-PRS-3.S.IND forest-E-LOC

'I don't know. I allow for the possibility that she's in the forest picking berries.' (Močnik and Abramovitz 2019)

The consultant was asked to judge the felicity of this story, and at first rejected it, presumably having interpreted *ivək* with its strong quantificational force, which makes Hewngyto's replies contradictory. When she was asked explicitly whether the sentence could express the same thought as the one with *dopuskat'* ('allow for the possibility'), i.e. the matching task, she readily confirmed this and stated that the story was acceptable on this reading of Hewngyto's replies. Hence, the matching task revealed a different acceptability judgment from the simple acceptability task.

[9] The following parts of the situation were read in Russian: *Hewngyto is walking down the street. Melljo sees him and asks*, *He replies*, *He continues walking*, *Qechghylqot sees him and asks*, and *Hewngyto replies*. The parts in the parentheses in (5) were not read, they merely provide the translation from Koryak for the reader.

It is convenient that Russian lexicalizes a weak doxastic verb *dopuskat'* ('allow for the possibility') in addition to the stronger *dumat'* ('think').[10] Using *dopuskat'* in a matching task enabled us to isolate the weaker construal of *ivək*. While our consultants were sometimes able to access the weaker interpretation in an acceptability judgment (especially if they had been exposed to *dopuskat'* earlier in the session), this access was not sufficiently stable, as shown above. We were therefore hesitant to rely only on acceptability judgments in further tasks. Given the many interpretations of sentences with *ivək*, it was very important that we be sure of which reading was being accessed by the consultant in a given context.

A more general issue we have had with acceptability judgments is that our consultants often seem to ignore the discourse contexts we provide, and instead give judgments as to whether the sentence is an acceptable sentence of the language.[11] One type of interaction with our consultants that has caused us to come to this conclusion is that, upon being asked for an acceptability judgment, instead of giving a 'yes' or 'no' answer, consultants will often simply provide a translation into Russian, and that translation may be incompatible with the context we presented them with. Now, this could be the result of the consultant not paying attention to the context, or it could be an artifact of their non-native knowledge of Russian (we suspect both play a role), but either way it is clear that their translation cannot simply be taken at face value. Trying to insist that speakers not respond to queries on the acceptability of sentences in context by translating the sentences into Russian is uncomfortable for the elicitor, and, doubtless, the speaker too, as they would often respond to this by remaining silent and ignoring the elicitor's question. Contextual acceptability judgments therefore proved not useful in eliciting judgments on attitudes in Koryak, both for reasons specific to the verb *ivək*, and because of our consultants' reactions to acceptability judgment queries themselves.

[10] That being said, we do not think it would be impossible to carry out this work in the absence of a weak doxastic verb in the contact language. Since the word *dopuskat'* as a doxastic attitude is somewhat literary in Russian (it is more commonly used to express the deontic meaning 'allow'), we wanted to be sure that our consultants were correctly understanding the reading of *dopuskat'* that we had in mind. We would therefore sometimes check with our consultants that the Koryak sentences that had been accepted with *dopuskat'* were also acceptable on the meaning of 'think that maybe' (R. *dumat' čto vozmožno / dumat' čto možet byt'*), which they were. The combination of a strong doxastic attitude with an embedded weak epistemic modal could therefore stand in for a weak doxastic attitude in a language that does not lexicalize one.

[11] We did not manage to resolve this issue with visual stimuli. There are also some independent difficulties with pictorial representations, see §4.3.

One final note on (5) is in order. It is very striking that the existence of the weak reading does not in and of itself render the dialogue in (5) acceptable (that is, non-contradictory). Usually, when we suspect that a sentence may have two readings, A and B, and wish to isolate reading B, we propose a context that excludes reading A, assuming that this will force the speaker to consider reading B. In other words, we give reading B 'a fair "pragmatic chance" of being recognized' (Matthewson 2004, p. 406). If the speaker rejects the sentence in that context, we take this as evidence against the existence of reading B. What our experience with (5) and others like it suggests is that this may not always be the case. In (5), the consultant declared the sentence incoherent instead of accessing the coherent reading. This happens not only in highly endangered languages. Consider, for example, the garden path sentence in (6).

(6) ? The horse raced past the barn fell. (Bever 1970, p. 316)

The grammatical interpretation of (6) contains a reduced relative clause with a transitive *race* (the horse that was raced past the barn fell). This parse is very hard to access, though; most naturally, *raced* is read as a matrix intransitive verb, which clashes with the actual matrix clause verb *fell*. Even though (6) is ungrammatical with the intransitive *race*, the existence of a grammatical parse does not make the sentence easily acceptable. Bever contrasts (6) with the completely natural (7). Unlike *race* in (6), *send* in (7) cannot be understood intransitively.

(7) The horse sent past the barn fell. (Bever 1970, p. 316)

The parallel that we see between (5) and (6) is that the sentence allows for two possible parses and the "good" parse (grammatical and felicitous) is not prominent enough to render the sentence acceptable. In the Koryak example (5), at least, the context was not sufficient to guarantee that the acceptability judgment targeted the desired parse. This is a crucial point. What we were able to achieve with the matching task is to bring attention to the less prominent parse and to elicit an acceptability judgment on it.

One might object that the difficulty with (5) is due to the complexity of the verbally-presented context, which perhaps requires the speaker to remember too much information at once. Could the issue have been avoided if we had presented our situations pictorially? Bochnak and Matthewson (2020) suggest that it might have been, arguing that with content pertaining to the interlocutors' belief states it is insufficient to present descriptions of the context verbally, and that more involved stimuli, such as storyboards,

are likely to be necessary. However, it is difficult to imagine a simple way to express the distinction between strong ('think') and weak ('allow for the possibility') doxastic attitudes pictorially. This is especially so because, as discussed in greater detail in §4.3, our consultants had issues with thought bubbles, which we tried to use to represent the non-assertive uses of *ivək*. They instead interpreted what we had intended to be thought as speech.

4.2 Translation tasks

Translation tasks in semantic fieldwork do not enjoy the same good reputation as acceptability judgment tasks (Matthewson 2004, Tonhauser and Matthewson 2015, Deal 2015). A noted difficulty arises when the contact and the target language lack a one-to-one mapping between linguistic forms.[12] For example, as mentioned in footnote 7, Koryak 3rd person pronouns (unlike Russian ones) do not carry gender marking, which makes a translation of a sentence containing such a pronoun from Russian to Koryak incomplete and a translation from Koryak to Russian overly specific. In this section we discuss two further issues that we encountered and explain how the matching task remedied them.

4.2.1 Contact-to-target translation

The contact-to-target translation task consists of asking the consultant to produce a target language translation, given some contextual information and a contact language sentence. This task is not particularly useful when the fieldworker wishes to conduct research on a particular lexical item in the target language, and the target language contains other ways of expressing approximately the same meaning. The fieldworker has little control over which item is used by the consultant. For example, when we asked a southern Koryak consultant how to say 'I hope it will rain', she offered the sentence in (8).

(8) t-ə-ko-məčwən-at-ə-ŋ-∅
1SG.S/A-E-PRS-hope-VBLZ-E-PRS-1SG.S
∅-je-muq-et-ə-ŋ-∅
2/3.S/A.IND-FUT-rain-VBLZ-E-FUT-3.S.IND
'I hope it will rain.'

[12] We set aside the situation of monolingual fieldwork (Sarvasy 2016), where we might ask the speaker to try to paraphrase our target sentence.

This sentence uses the lexical verb *məčwənatək*, which is a verb of hoping in the southern dialect that is not recognized by our northern Koryak consultants.[13] While we eventually found out via matching tasks that southern Koryak speakers accept *ivək* on the 'hope' reading, the verb *məčwənatək* was usually produced (by the speakers who had it) in contact-to-target language translations. Similarly, the existence of *četkejuŋkə* or *čečkejuŋkə* ('think (about?)') and *ɣajmatək* ('wish') hindered the investigation of *ivək*. Since we did not have the luxury to "fire" consultants (see §5.2), we had to gather data even from the speakers who, for example, produced *məčwənatək* in Russian to Koryak translations. The matching task allowed us to circumvent this issue since it fixed the lexical item that we wished to obtain judgments for.

Despite these disadvantages, we echo Matthewson (2004) in stressing that the translation task can provide important clues. We discovered some of the readings of *ivək* because we asked the consultants to translate Russian sentences with attitude verbs into Koryak. Additionally, speakers' providing Koryak sentences with *ivək* in response to a Russian-to-Koryak translation query provided independent confirmation of the existence of the various interpretations of that verb. Nevertheless, given the issues we have encountered, the matching task gave us a more reliable tool to use with the consultants.

4.2.2 Target-to-contact translation

Consider now the case where a translation into the contact language is requested. The main issue that we have encountered with this task is that the consultants sometimes imported features of the target language into the contact language. To illustrate, consider (9).[14]

(9) Situation: People are talking about whether God exists. They ask Qechghylqot, who says:

t-ə-k-iv-ə-ŋ-∅	əno	amu	aŋaŋ
1SG.S/A-E-PRS-*ivək*-E-PRS-1SG.S	COMP	might?	God.ABS.SG
∅-ko-tva-ŋ-∅,	to		
2/3.S/A.IND-PRS-be-PRS-3.S.IND	and		

[13] In addition to the 'hope' meaning, southern consultants have also offered that *məčwənatək* can mean 'be sure' (Russian *uveren*) and 'doubt' (Russian *ne verit'/somnevat's'a*). We have not had a chance to explore this in more detail.

[14] By contrast to (5), the weaker force of *ivək* in (9) was accessible to the consultant. As mentioned in §4.1, the weaker reading is sometimes accessible outside of the matching task.

t-ə-k-iv-ə-ŋ-∅ əno amu aŋaŋ
1SG.S/A-E-PRS-*ivək*-E-PRS-1SG.S COMP might? God.ABS.SG
ujŋe a-tva-ka
NEG.RLS NEG-be-NEG
∅-k-it-ə-ŋ-∅
2/3.S/A.IND-PRS-INTR.AUX-E-PRS-3.S.IND

speaker translation: *Ja dumaju, čto est' bog, i ja dumaju, čto jego net* ('I think that God exists and I think that he does not exist')

our interpretation: 'I allow for the possibility that God exists and I allow for the possibility that he does not exist.'

Qechghylqot's utterance was accepted in this situation, and the consultant commented that a person doubting the existence of God might say this.[15] Nevertheless, notice that the volunteered sentence in the contact language is contradictory. This cannot be attributed to a cultural difference, though, such as Koryaks not objecting to contradictions. The sentence in (10), which was elicited immediately before the sentence in (9), was judged by the consultant to be unnatural, and, upon being asked by the elicitor whether it was a contradiction, the consultant agreed.

(10) # ujŋe e-muq-et-ke
NEG.RLS NEG-rain-VBLZ-NEG
∅-k-it-ə-ŋ-∅, ʔam
2/3.S/A.IND-PRS-INTR.AUX-E-PRS-3.S.IND but
ʔewŋəto-na-k liɣi
Hewngyto-OBL.SG-ERG know
∅-ku-lŋ-ə-ŋ-nin, əno
2/3.S/A.IND-PRS-consider-E-PRS-3SG.A>3.O COMP
∅-ku-muq-et-ə-ŋ-∅
2/3.S/A.IND-PRS-rain-VBLZ-E-PRS-3.S.IND
intended: 'It's not raining, but Hewngyto knows that it's raining.'

When we subsequently presented the sentence in (9) to the consultant, she explicitly drew a contrast between it and (10), saying that (9) sounds

[15] The consultant also said that the sentence sounded 'a bit Russian' (a comment she tends to make regarding sentences that sound somewhat unnatural, regardless of whether they have features of Russian identifiable to us or not) because it repeated a lot of words, and suggested a more natural-sounding variant with pro-drop and auxiliary-drop in the second conjunct.

fine in comparison to (10) (Rus. *normal'no zvučit po sravneniju s predyduščim* 'sounds fine in comparison to the previous one').

Here is another example. Certain epistemic modals appear in Koryak sentences that are accepted as non-contradictory by the consultants, but are translated into Russian in a way that is contradictory to native Russian speakers. The sentence in (11) was volunteered by a Koryak speaker as an example of a felicitous sentence with *amu*, an epistemic modal that we think may have variable force, but when asked to explain what the sentence meant, the speaker provided the contradictory translation in (11).[16]

(11) amu ∅-ku-muq-et-ə-ŋ-∅
might? 2/3.S/A.IND-PRS-rain-VBLZ-E-PRS-3.S.IND
peterburg-a-k to amu ujŋe
St.Petersburg-E-LOC and might? NEG.RLS
e-muq-et-ke ŋanko
NEG-rain-VBLZ-NEG there
speaker's translation: *Naverno dožd' idët v Peterburge, i naverno dožd' ne idët tam* ('It's probably raining in St. Petersburg and it's probably not raining there')
our interpretation: 'It might be raining in St. Petersburg, and it might not be raining there.'

A potentially related phenomenon is noted by Rullmann et al. (2008, fn. 32). The St'at'imcets sentence in (12) is reported as having been "easily" accepted by the consultant, who offered the contradictory English translation below.

(12) t'ak k'a tu7 k Elvis, t'u7 cw7aoz k'a t'u7 kw
go INFER then DET Elvis but NEG INFER just DET
s-t'ak-s
NOM-go-3POSS
'Maybe Elvis left, but maybe he didn't.' (2008, ex. (64))
'Elvis left, but I'm sure he didn't.' (volunteered)

[16] It is possible that the Koryak consultant simply has an incorrect Russian lexical entry, i.e. that their *naverno* ('probably' in colloquial L1 Russian) only has the 'might' reading, rather than it having an epistemic strength that ranges from 'might' to 'probably'. This type of alternative explanation (that the L2 item is weak rather than it having a variable force) seems less likely for (9), since the Koryak speakers can otherwise use the Russian *dumat'* correctly as 'think' in speech.

More generally, Koryak-to-Russian translation tasks involving *ivək* almost exclusively led to the 'say' translation, and occasionally led to the 'think' translation, but almost never led to any of the other translations.[17] Given that other elicitation tasks (matching, Russian-to-Koryak translation) have established the existence of additional interpretations of this verb, we did not find such translation tasks to be adequate to investigating the meaning of *ivək*.[18]

4.3 Elicited production task

One technique that has recently gained attention among linguists doing semantic fieldwork on understudied languages is the elicited production task. In this task, the consultant is asked to talk about a stimulus, e.g. a picture or a coherent series of pictures (a storyboard). This provides a more natural way of setting up the situation, in particular for those speakers whose traditional culture includes writing or pictures.[19]

In our fieldwork we have attempted to use storyboards to elicit sentences with attitude verbs, as well as to set up situations for acceptability judgment tasks that involve attitude verbs. As argued by Bochnak and Matthewson (2020), storyboards are useful for precisely this type of semantic content (see §5.1 below). In this section we speculate as to why this has failed in our case.

[17] Another example of this is that expressions where *ivək* embeds a counterfactual complement, which correspond to English 'wish' (2), were often translated as 'say if only' or 'think if only' (Rus. *dumajet/govorit, xotja by*) which is a locution that is not recognized by native Russian speakers. By contrast, when asked to translate 'wish' (Rus. *želat'*) into Koryak, the speakers had no trouble using *ivək*.

[18] The contradictory translations are interesting in their own right, and we might wonder how exactly the consultants understand the Russian lexical items since their translations are non-contradictory to the them, but not to L1 Russian speakers. In the case of *ivək*, we suspect that our consultants simply think that the Russian verbs 'say' (R. *govorit'*), and possibly also 'think' (R. *dumat'*), have variable flavor and variable force. Our impression is that the consultants have, by contrast, a native-like understanding of the attitude verbs that correspond to the other flavors of *ivək*. In part, we suspect this because some of the meanings are also independently lexicalized in Koryak (see the discussion in §4.2.1), and in part because we have tested some of the consultants by asking to describe what, e.g., the Russian *nadejat's'a* ('hope') mean and they correctly described a hoping attitude. However, their understanding of the force of Russian attitude verbs appears to be sketchier, and we have therefore relied on contradiction tests to determine the existence of weak readings.

[19] Traditional Koryak culture does not, admittedly, make use of either writing or drawing. However, all of our consultants are literate in Russian and have at least a middle-school level of education, such that the written medium would be very unlikely to be cause them problems.

Consider the storyboard in Figure 1, with a glossed translation of the text in (13). This storyboard sets up a situation where the teacher thinks one thing (that his students are doing very poorly in school), but tells the principal the opposite (that they are doing very well) since he is afraid that the principal might get mad at him. Given that the verb *ivək* can mean both 'think' and 'say' (§2.2), we have used this type of situation to test whether certain conjunctive structures with *ivək* can be interpreted as being about both thinking and saying eventualities. See Močnik and Abramovitz (2019) for details.

(13) a. qok ečwej! ɣəm-nine-w jejɣučewŋ-ə-lʔ-u
alas 1SG-POSS-3PL study-E-S/O.PTCP-ABS.PL
tətteʌ qewwa-ŋ
very bad-ADV
∅-ko-jajɣočawŋ-ə-la-ŋ-∅!
2/3.S/A.IND-PRS-study-E-PL-PRS-3.S.IND
direktor-ə-na-k
principal-E-OBL.SG-ERG
∅-j-ena-kətʔajŋa-ŋ-∅!
2/3.S/A.IND-FUT-1SG.O-scold-FUT-3.S.IND
'Alas! My students are doing very badly! The principal will scold me!'

b. meŋiɲet ∅-ko-jajɣočawŋ-ə-la-ŋ-∅
how.well 2/3.S/A.IND-PRS-study-E-PL-PRS-3.S.IND
ɣə-nine-w jejɣučewŋ-ə-lʔ-u?
2SG-POSS-ABS.PL study-E-S/O.PTCP-ABS.PL
'How are your students doing?'

c. əčč-u metʔa-ŋ
3NSG-ABS.PL beautiful-ADV
∅-ko-jajɣočawŋ-ə-la-ŋ-∅!
2/3.S/A.IND-PRS-study-PL-PRS-3.S.IND
'They are doing very well!'

d. e-mel-ke!
ADV-good-ADV
'Good!'

e. t-ə-tinm-et-ə-k direktor-ə-na-ŋ!
1SG.S/A-E-lie-VBLZ-E-1SG.S principal-E-OBL.SG-DAT
'I lied to the principal!'

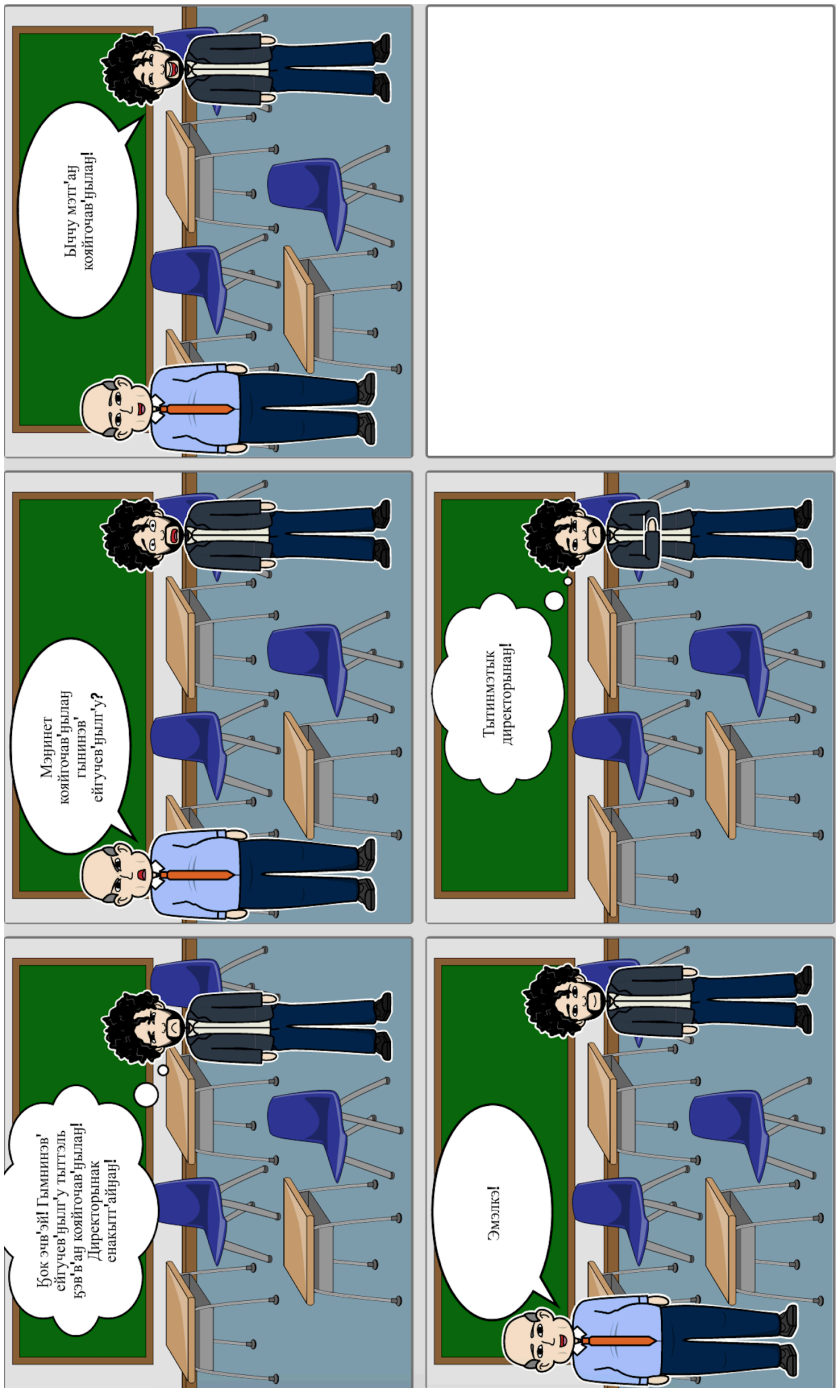

Figure 1: Lying teacher storyboard

Using this storyboard with our consultants has proven to be very difficult, as the consultants that we tested this with did not seem to be able to follow the narrative arc of the story. Specifically, when they were looking at the pictures, it did not seem that they had grasped the distinction between what the characters were saying and what they were thinking. Subsequently, we realized that it was likely that our consultants did not understand what the difference was between thought bubbles and speech bubbles.[20] Due to the pervasiveness of comic books, Western people (and, nowadays, likely most literate non-Western people) are familiar with the conventional interpretation of a cloudy bubble with circles (the person is thinking) and an elliptical bubble with straight lines (the person is speaking). However, comics did not become popular in the Soviet Union until well into our consultants' adult years, and the examples of earlier Soviet comics that we have been able to find do not employ thought bubbles at all. We suspect that a lack of familiarity with the relevant distinction is what confused our consultants. Given that the difference between thought and speech is crucial for testing the meaning of an attitude verb that can be mean both 'think' and 'say,' storyboards failed to give us usable results in acceptability tasks and elicited production tasks.[21]

5 User's guide: When might you use this task?

In this section, we step back from examining the matching task *per se* and consider what factors led to its use being necessary in our work. Based on this, we suggest under what circumstances other linguists working in semantics may find this task useful. We distinguish between (i) phenomenon-specific factors, (ii) sociolinguistic factors, and (iii) language-internal factors.

5.1 Phenomenon-specific factors

Some difficulties come from the nature of the empirical domain under investigation: in our case, semantic fieldwork and fieldwork on mental states.

[20] The possibility that they were simply unable to see the difference between thought and speech bubbles also exists, though we consider that unlikely as they were able to read the Koryak text.

[21] We only developed this hypothesis regarding thought and speech bubbles after leaving the fieldwork site, and therefore never thought to train the consultants on the relevant distinction. Perhaps such training of the consultants would yield reliable results with tasks involving visually-depicted situations.

Such challenges are conceptually independent of the choice of target and contact language, and instead revolve around the inherent challenge of doing fieldwork on phenomena like attitude ascriptions.

Semantic and pragmatic phenomena are widely acknowledged to be among the most challenging fieldwork domains. As noted by Matthewson (2004), a consultant might reject a target item in an acceptability judgment task because it is false, infelicitous, pragmatically odd for independent reasons, or even simply ungrammatical. Conversely, a positive response requires the target item to pass with respect to these various dimensions: it needs to be morpho-phonologically and syntactically well-formed, it needs to be true in the situation, and it needs to be pragmatically felicitous. As discussed in §4.1, the relevant reading needs to be sufficiently prominent to the consultant as well. In sum, judgments of semantic and pragmatic phenomena build on other aspects of the grammar, which is reflected in the widely-shared experience of linguistic fieldworkers that not all of the native speakers who are reliable consultants for fieldwork targeting morpho-syntactic questions are reliable consultants for fieldwork targeting semantic questions.[22]

Similarly, certain areas of semantic/pragmatic fieldwork can be more difficult to investigate than others. Bochnak and Matthewson note:

> 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. (2020, pp. 262–263)

While the matching task may not be the method to probe at each phenomenon in semantics or pragmatics (see §5.3), we have found it to be able to take away some of the inherent difficulties of doing semantic fieldwork on phenomena like attitude verbs.

[22] Dmitry Privoznov (p.c.) reports the same observation from his fieldwork with speakers of Uralic, Turkic, and Mongolic languages in Central Siberia, as does Luiz Fernando Ferreira (p.c.) from his work with Karitiana (Tupian) speakers in the Brazilian Amazon.

5.2 Sociolinguistic factors

One significant sociolinguistic factor in our work is the poor state of preservation of Koryak, which has given us a rather small pool of consultants to draw from, none of whom use Koryak as their primary language anymore. The total number of fluent native speakers of Koryak is unknown, though we suspect it is somewhere in the vicinity of 600, most of whom are at least 65 years old. In Palana, which is the largest village in Koryakia and the place where most of our fieldwork was conducted, there are only fifteen or so fluent native speakers of the Chawchuven dialect of Koryak, of which only six were able to work with us regularly. Additionally, Koryak is rarely used anymore even by fluent native speakers who speak it better than they speak Russian: all of our consultants are the only remaining members of their household to speak Koryak, and some of them have health issues that make it difficult or impossible for them to leave the house to interact with other speakers. Even when two native speakers interact in Palana, it is far from certain that they will speak Koryak to each other.[23] The infrequency with which they use the language is plausibly a factor in their difficulty with giving semantic judgments.

The challenge of having such a small number of possible consultants means that it is not possible to simply exclude those speakers who have trouble with the more standard fieldwork methods, such as providing contextual felicity judgments (see §4.3). Luiz Fernando Ferreira (p.c.), for example, began working with a variety of consultants that he had worked with on morphosyntactic topics on a questionnaire that required the speakers to judge the felicity of Karitiana modals in various contexts. Consultants whose judgments on the modal expressions were sensitive to the contexts he provided on the initial portion of the questionnaire were the only ones that he went through the full questionnaire with; consultants whose judgments were not sensitive to the contexts were not asked the full questionnaire, and their data was not taken into account in developing the analysis.

Finally, we think that we have encountered difficulties with visual materials due to the speakers being unfamiliar with or uncomfortable with the visual stimuli (see also Bochnak and Matthewson 2020, p. 164, fn. 7). In particular, we suspect that this was at issue in our consultants' inability to

[23] In the first author's experience, native speakers of Koryak usually speak either entirely in Russian or code-switch between Koryak and Russian when speaking to each other in public, though in private Koryak (without code-switching) is heard much more frequently. The effect that the presence of a linguist had on the language used in the conversations, i.e. the observer's paradox (Labov 1972), is unknown.

make sense of the storyboard that relied on a distinction between thought bubbles and speech bubbles (see §4.3).

Any (and possibly all) of these sociolinguistic issues may have contributed to our consultants' difficulty in giving judgments in more standard elicitation methods.

5.3 Language-inherent difficulties

Languages differ in their lexical and functional inventories, which can pose a problem for some elicitation methodologies (§4.2). The matching task was built for this type of asymmetry. In particular, the task is useful when the target language contains linguistic material (see below) with several construals while the contact language has explicit ways of bringing out the individual construals.

For instance, the Koryak *ivək* is used in the place of English *say*, *suggest*, *think*, *allow for the possibility*, *hope*, *fear*, and *wish*, among others. A language like Russian, which also makes these fine-grained distinctions, can be used (and has been used) in a matching task to probe for a particular reading of *ivək*, as well as to fix the reading of *ivək* while its other properties are being tested.

The matching task has also proved useful in doing research that did not specifically target phenomena relevant to questions in formal semantics, but where issues involving default interpretation of homophonous affixes were at play. For example, the Koryak exponents of the unwitnessed past tense and the resultative participle are homophonous, as in (14)–(15); both are expressed with the circumfix *ɣe- -lin* (*ɣa- -lin* or *ɣa- -len* by vowel harmony) with a 3rd person singular subject. As in many languages, the resultative participle can only have the internal argument as a pivot, making it an important test for unaccusativity in Koryak (Williams 1981 et seq.).

(14) ɣ-ajal-len-∅ utt-ə-ut
UW.PST-fall-3.UW.PST-SG tree-E-ABS.SG
'The/A tree fell.' (R. *derevo upalo*)

(15) ɣ-ajal-len-∅ utt-ə-ut
RES.PTCP-fall-3.RES.PTCP-SG tree-E-ABS.SG
'a fallen tree' (R. *upavšeje derevo*)

Without a matching task, most speakers translate and interpret utterances in elicitation with this morphology as past tense finite clauses (14), even when the syntactic environment privileges the reading in (15), such

as when the phrase *ɣajallen uttəut* is an argument of a finite verb (in which case this tended to be interpreted as two sentences joined paratactically). Using the matching task, which brought to the forefront the less prominent participial reading of this circumfix, made investigating unaccusativity in Koryak substantially less time-consuming, and, seemingly, easier for the consultants, as it was immediately clear to them what they were being asked about.

6 Conclusion

In this paper, we have discussed an elicitation technique we call a matching task, which is a type of acceptability judgment task that contains some features of a translation task. It consists of a context/situation, a target language sentence, and a contact language sentence. The consultant is asked to judge whether the target language sentence is acceptable in the situation with the intended meaning given by the contact-language sentence. We have been using this elicitation method in our fieldwork to investigate the underspecification and polysemy of Koryak modals and attitudes.

To our knowledge, this is the first explicit examination of the matching task in semantic fieldwork. One of the reasons we found this discussion to be necessary is that there does not appear to be agreement in the field as to the validity of this methodology. For some, this task seems to be standard practice in investigating meaning (Maria Polinsky and Seth Cable, p.c.). Polinsky (p.c.), for example, has pointed out that matching tasks have been employed by linguists from Moscow State University since the 1960s, when Aleksandr Kibrik first began to lead linguistic field trips to the North Caucasus. However, the matching task does not appear in any surveys of tools in formal semantic fieldwork (e.g. Bochnak and Matthewson 2020), and Lisa Matthewson (p.c.), for example, has informed us that she was not previously aware of such a task being used in semantic fieldwork, nor had she considered using it in her own work.

One goal of this paper has been to show that the standard methodologies of semantic fieldwork (acceptability tasks, translation tasks, and elicited production tasks) are insufficient for certain speakers or fieldwork environments. Another goal has been to describe and defend the matching task, a method for investigating meaning that we have found to produce consistent results across speakers. We have presented our work on the Koryak attitude verb *ivək* as a proof of concept of a way to elicit semantic data from speakers who are not able to provide it using standard methods. Ide-

ally, this technique will also allow other linguists to expand their pool of reliable consultants for semantic fieldwork, and thereby extend the scope of research in the semantics of understudied and endangered languages.

References

Thomas G. Bever. The cognitive basis for linguistic structures. In John R. Hayes, editor, *Cognition and the Development of Language*, pages 279–362. John Wiley & Sons, Inc., New York, 1970.

Maria Bittner. On the semantics of the Greenlandic antipassive and related constructions. *International Journal of American Linguistics*, 53 (2):194–231, 1987.

M. Ryan Bochnak and Lisa Matthewson. Techniques in Complex Semantic Fieldwork. *Annual Review of Linguistics*, 6(1):261–283, 2020.

Seth Cable. Describing future eventualities in Tlingit: The storyboards Hawaii Trip and Imagining the Future. *Semantic Fieldwork Methods*, 1 (2):1–35, 2019.

Amy Rose Deal. Reasoning about equivalence in semantic fieldwork. In Ryan Bochnak and Lisa Matthewson, editors, *Methodologies in Semantic Fieldwork*, pages 157–174. Oxford University Press, New York, 2015.

William Labov. *Sociolinguistic patterns*. Number 4. University of Pennsylvania Press, 1972.

Lisa Matthewson. On the methodology of semantic fieldwork. *International journal of American linguistics*, 70(4):369–415, 2004.

Maša Močnik. Where force matters: Embedding epistemic modals and attitudes. In *Formal Approaches to Slavic Linguistics (FASL) 27*, Ann Arbor, 2019a. Michigan Slavic Publications.

Maša Močnik. Where force matters: Embedding epistemic modals (and attitudes). In *Proceedings of Sinn und Bedeutung 23*, 2019b.

Maša Močnik and Rafael Abramovitz. A variable-force variable-flavor attitude verb in Koryak. In Julian Schlöder, Dean McHugh, and Floris Roelofsen, editors, *Proceedings of the 22nd Amsterdam Colloquium*, pages 494–503. ILLC, 2019.

Hotze Rullmann, Lisa Matthewson, and Henry Davis. Modals as distributive indefinites. *Natural Language Semantics*, 16:317–357, 2008.

Hannah Sarvasy. Monolingual fieldwork in and beyond the classroom: the Logooli experience at UCLA. In Ksenia Ershova, Joshua Falk, Jeffrey Geiger, Zachary Hebert, Robert Lewis, Patrick Munoz, Jacob Phillips, and Betsy Pillion, editors, *Proceedings of CLS 51*. Chicago Linguistic Society, 2016.

Judith Tonhauser and Lisa Matthewson. Empirical evidence in research on meaning. *Ms., The Ohio State University and University of British Columbia*, 2015.

Jozina Vander Klok. On the use of questionnaires in semantic fieldwork: A case study in modality. In *Proceedings of language documentation and linguistic theory 4*, pages 1–11, 2014.

Jozina Vander Klok. Exploring modality and temporality interactions through the storyboard Bill vs. the weather. *Semantic Fieldwork Methods*, 1(1):1–29, 2019.

Jozina Vander Klok and Thomas J Connors. Using questionnaires as a tool for comparative linguistic field research: Two case studies on Javanese. *Language Documentation & Conservation*, pages 62–96, 2019.

Edwin Williams. Argument structure and morphology. *The linguistic review*, 1(1):81–114, 1981.