

# Semantic elicitation in the field: on managing data and conducting sessions \*

Jérémy Pasquereau  
Université de Poitiers

**Abstract:** This paper is a report of personal research methodology. More specifically, it describes a few approaches I have developed and used in the course of investigating semantic topics via elicitation in the field. The first half deals with data for/from semantic elicitation: creating stimuli, handling variation in judgments, and keeping track of collected data. The second half focuses on working with consultants: finding them, training them, and conducting sessions.

**Keywords:** semantics, fieldwork, elicitation, logistics, implementation

## 1 Introduction

There exists a sizeable amount of resources providing stimuli to investigate different kinds of meanings using tasks of semantic elicitation which have become standard since Matthewson (2004). For instance, should one wish to study the meaning of a particular tense-aspect-mood (TAM) construction in the Tsugni language (Nakh-Daghestanian), one could start with a corpus study, then with the TAM questionnaire in Dahl (1985), then use storyboards (e.g. those available on the Totem Field Storyboards website or on personal webpages<sup>1</sup>) or the Progressive Aspect questionnaire (Bertinetto, Ebert, and Groot 2008) among many others. But for a linguist who may have never worked on collecting semantic data face-to-face with non-linguist speakers, it may not be obvious how to start using these methods. For one thing, given the logistics of data collection in the field (e.g. a lot of non-categorical data on different topics over several days with several speakers) keeping track of data (what has been collected or not) can be quite challenging. Secondly, as anyone who has ever taken a psycholinguistics class will know, the implementation of these methods (e.g. what instructions a speaker is given, how the question is asked) has significant repercussions on the data collected (e.g. Schütze (2016:chapter 5)).

I have carried out research on meaning in two types of environments. In one type of environment, I have easy access to amenities, colleagues, and a potentially large number of consultants through online experimental platforms (e.g. Ibexfarm (Drummond 2013)). In the other type of environment, access to amenities and colleagues is not as easy, I do not speak the language under study, and I cannot use online experimental platforms as my (potential) consultants would not be able to take part in them for lack of equipment. My experience of doing research on meaning has not been the same depending on the environment in which I have carried it out. In this paper I report on certain

---

\* I am very grateful to my various consultants over the years. For helpful discussion and feedback on previous versions of this paper, thanks to Patricia Cabredo Hofherr, David Gyórfi, Meg Harvey, Robert Henderson, Lisa Matthewson, and John Powell. For thought-provoking comments and questions, thanks to panellists and audiences at the Sinn und Bedeutung 25 Roundtable on Semantic Fieldwork Methodology, and to two anonymous reviewers for *Semantic Fieldwork Methods*, as well as to Matthew Baerman, Heidi Harley, Stephen Marlett, and Carolyn O'Meara.

Contact info: [jepasquer@gmail.com](mailto:jepasquer@gmail.com)

<sup>1</sup> For instance, Kilu von Prince's website include storyboards to elicit tense aspect data: <http://kiluvonprince.de/tag/aspect/>.

strategies I have developed to deal with these difficulties as they apply to semantic elicitation in this second type of environment.<sup>2</sup> These methods have been evolving since 2011 as part of my fieldwork on Karata (Nakh-Daghestanian, Russia) and especially Seri (Isolate, Mexico).

This paper is about semantic elicitation-based methods, and specifically about how I have implemented some of these methods when working with speakers who have not been trained in linguistics. There are of course other ways to obtain data that are also very informative to answer semantic questions (e.g. corpus data) but these fall outside of the remit of this paper. In section 2, I describe several strategies I have developed to deal with semantic elicitation data. In section 3, I discuss some attitudes and approaches I have found useful to enhance the success of semantic elicitation work with speakers. Section 4 concludes the paper.

## 2 Managing data

Following Tonhauser and Matthewson (2015), I take a piece of data in semantics to consist of at least four elements :

- a sentence
- a context in which the sentence is used
- a response to a task about that sentence used in that context
- information about the native speaker consultants

These are illustrated in (1)<sup>3</sup> from Pasquereau (2021:ex. 18). For this data point, the context was presented both orally and with the help of a graphical illustration<sup>4</sup>. Information about the native speaker consultants who provided the response is given in the form of an example identifier (sometimes in a footnote). In addition, for this example, consultants provided a comment which I wrote down; in that case it is preceded by the abbreviation SC for Speakers' Comment.

(1) **Me:** Context: Today there was a race between four groups of three girls. Each group of three girls is in a different running track. Figure 1.

**Main consultant:** Xicaquiziil cmajiic quih c-apxoj yopancojc  
 child.PL woman.PL DET SBJ.NMLZ-be\_three.PLUR RLYO.RUN.PL  
 'The girls ran in groups of three.'

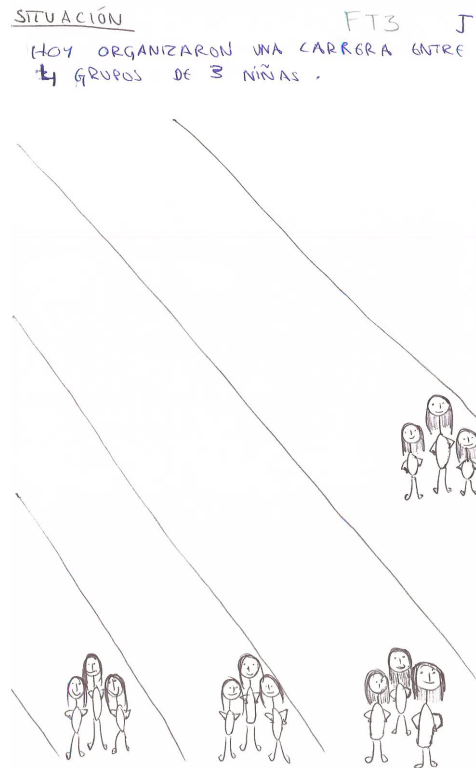
**Consultants:** TRUE SC: It's three here, three there, etc.<sup>5</sup>

<sup>2</sup> For more general aspects of fieldwork, see Bowern (2015); Chelliah and Willem (2010); Meakins, Green, and Turpin (2018); Newman and Ratliff (2001); Sakel and Everett (2012).

<sup>3</sup> The abbreviations used in this paper are: ABS = absolutive, AUG = augment, CAUS = causative, DEF = definite, DEM = demonstrative, DET = determiner, DIST = distal, DS = different subject, EMPH = emphatic, EQ = equative verb root, INDF = indefinite, INTR = intransitive, NMLZ = nominalizer, OBL = oblique, PASS = passive, PL = plural, PLUR = pluractional, POSS = possessive, RLT = realis t-form, RLYO = realis yo-form, SBJ = subject, SG = singular, UNSP = unspecified.

<sup>4</sup> Translation of the (Spanish) text on the picture "Situation: Today a race was organized between 4 groups of 3 girls."

<sup>5</sup> Example identifier: [EDSEI24OCT2018DRPM.GH.ATHF.LKPH]



**Figure 1:** Picture illustrating the context in (1)

The file from which each data point has been taken is indicated in the example identifier. Here's how to decode [EDSEI24OCT2018DRPM.GH.ATHF.LKPH, TVAL.CON] for example:

- ED: the initials of the village where the data was collected, here **El Desemboque**
- SEI: ethnologue code for Seri
- 24OCT2018: date of collection
- DRPM.GH.ATHF.LKPH: code for the speakers, here four speakers identified by codes
- TVAL: the type of elicitation task, here truth value
- CON: how the sentence was obtained, here it was originally constructed by me (and its acceptability checked with consultants before being used in a truth-value judgment task)

Speaker codes are matched to speaker biographical information in a word document called “SpeakersData” (see appendix A for the list of documents mentioned throughout the paper which I use on field trips). Often, in published work, I use the cover term “elicitation” rather than specify what type of elicitation task was used and how the sentence under judgment was obtained (see appendix B for more detail about data identifiers).

Before elicitation can take place, sentence-context pairs need to be prepared by the linguist (section 2.1). In elicitation, speakers will be asked to consider the linguistic expression in the provided context and give a judgment as to its acceptability, truth, or appropriateness (section 2.2). Sometimes these judgments do not concord from one speaker to another and even from time to time for the same speaker (section 2.3). Such variation in speakers' judgments should be kept track of as it can turn out to be quite meaningful (section 2.4). The following sections present how I handle these situations in my fieldwork practice.

## 2.1 Creating stimuli

Some stimuli are ready when I arrive in my field site, and I can start work on them immediately (e.g. data points collected in previous field trips that I want to double-check). Indeed, since the previous field trip, I have been thinking about what questions I want to investigate and what data I want to collect or double-check. Every time I think of a question, hypothesis, or idea for collecting data, I write it in a document named 'pensieve'<sup>6</sup> in the folder dedicated to this specific field trip. I have several pensieves divided roughly by linguistic domain, e.g. morphology, semantics, syntax. I have found the use of these documents very helpful since they are where I put all my ideas, and all the suggestions and questions I get from colleagues. The contents of the pensieve documents can be anything from the following list:

- previously collected sentences, sentences extracted from texts, or published examples—either because I want to try to replicate them or use them as bases for constructing other data points to test a hypothesis
- sentences and context ideas in the contact language
- hypotheses and generalizations that need to be tested or checked

I then draw from these documents to create questionnaires. As questions, hypotheses, or data that I want to double-check move from one of the pensieves to a questionnaire, they are highlighted in yellow in the pensieve. If some of the pensieve becomes irrelevant (e.g. because newly collected data makes questions in the pensieve no longer applicable), these are highlighted in red. Ideally, by the end of a field trip, the whole pensieve will have been highlighted; if not, what has not been highlighted is copied over into the next pensieve (see section 2.4).

Other stimuli are not ready when I arrive as I need to work with my main consultant to obtain part or all of them. First, I choose a question or hypothesis and possible data point ideas from the pensieve files, and move them into a new file which I name according to the conventions described above. On the basis of this question or hypothesis, I create a few data points per hypothesis, e.g. context/sentence pairs for truth-conditional judgments, about five data points for which I expect one type of answer and five for which I expect another type of answer. This is so that both positive and negative predictions of the hypothesis can be tested. Then, in preparation sessions with my main consultant, we go over these context/sentence pairs to translate sentences into the target language, if needed, and detect any unintended oddness. Sometimes, preparation sessions also involve drawing if I intend to use visual stimuli. Once all the stimuli have been checked, I put them in a new file named

---

<sup>6</sup> The term *pensieve* was coined in the Harry Potter book series to refer to a magical object into which wizards can drop and store memories so they can later review them.

“Questionnaire” designed to be used in one session (although it often lasts more than one). A given questionnaire will contain several data sets, each pertaining to generally three or four hypotheses or questions (and each questionnaire uses a single elicitation task). I then use an online list randomizer to randomize these data points. Both my having materials covering a few different hypotheses within one questionnaire/session and my presenting them in a randomized order are my efforts to lower the probability that speakers will catch on to the phenomena under study, which could lead them to develop response strategies.

## 2.2 Constraining judgment responses

Different tasks give rise to different judgments; for instance, the truth-value judgment task gives rise to truth-value judgments while an acceptability judgment task gives rise to acceptability judgments. In my practice, I aim to clearly segregate different tasks and their corresponding judgments, and explain to my consultants how the tasks and their corresponding judgments differ. For instance, an unacceptable sentence is a sentence for which there is no context in which it can be used appropriately (i.e. truthfully, felicitously, and in a pragmatically normal way), whereas a false (acceptable) sentence in a given context will be true in another. In addition, for each task, I ask consultants to choose their judgment from a finite set of possibilities I provide for each session.

My providing a finite set of responses (e.g. for truth-value judgments: true, lie, not a lie but could be more specific) is meant to diminish the amount of interpretation I have to do between the speakers’ responses and what it means for the research hypothesis under investigation. It does not limit speakers in further explaining what they mean by their response: in fact I ask speakers to follow up on their response with explanations or comments. For instance, for truth-value judgments, if the response is “true”, I ask consultants to repeat the target language sentence; if the response is “lie” I ask them to explain what the lie is given the situation; if the response is “not a lie but could be more specific” I ask them what it is that could be made more specific.

## 2.3 Dealing with variation

On the one hand, most hypotheses in linguistic research on meaning make categorical predictions, but on the other hand the data we collect typically do not always wholly conform to a categorical classification within or between speakers and items. What should the linguist do faced with inter- or perhaps even intra-speaker variation? As linguists, faithfully reporting the data is non-negotiable but reporting all variation at face value is not desirable either. This is because some of the observed variation could be an effect of the task, or the way the linguist (mis)presented the materials, or the way it was (mis)perceived by the consultant—in other words, far from making the description of the data more precise, reporting variation of this type would in fact make the description more wrong. I will call this type of variation *non-meaningful variation*. I list a few of the many possible causes of observed non-meaningful variation I have encountered in my fieldwork experience in (2).

- (2) a. influence of other variety/grammar (e.g. influence of prescriptivism)
- b. the speaker does not clearly understand what the linguist says or wants (e.g. the task has not been understood by the consultant, the linguist’s pronunciation is not understood)
- c. lack of imagination at the time of elicitation (e.g. when doing acceptability judgments in syntax, a sentence will be judged unacceptable if the speaker cannot imagine any context)

in which s/he would use the sentence; this may be because the sentence is ungrammatical, but it could also be due to the speaker’s ‘lack of imagination’)

- d. the speaker has a theory of how things work in their language which leads them to adjust their judgments so they stay consistent with their theory
- e. when using context-sentence pairs (for acceptability or truth-conditional judgments) contexts may be sub-optimal (e.g. the context might be too unspecific thus leaving room for speakers to interpret it differently, which has an effect on their judgment)

When faced with variation in a data set, the challenge is precisely to tease apart what is linguistically non-meaningful from variation that actually reflects something about the language and how its speakers use it. The primary tool I use to “weed out” non-meaningful variation is replication both across and within speakers. Within-speaker variation can be detected by replicating data over several sessions, sometimes over several field trips. For instance, in my work on pluractional verbs in Seri (Cabredo Hofherr, Pasquereau, and O’Meara 2018; Pasquereau and Cabredo Hofherr 2020; Pasquereau and Cabredo Hofherr 2021), one question was whether they were acceptable at all with a singular object or whether they required a plural object. To test this, on the basis of examples previously offered by speakers, I constructed examples containing a pluractional verb form and a singular object like (3) and submitted them to the judgment of my consultants (some consultants judged it more than once at different points in time).

- (3) Maria quih hapaspoj                    iiqui                    icaaca    z  
Maria DET SBJ.NMLZ.PASS.write 3POSS.toward 3POSS.[OBL.NMLZ].UNSP.SBJ.send INDF  
iyaasipl.  
3>3.RLYO.write.PLUR  
‘Maria wrote a letter.’

Importantly, I gave the sentence to be judged first without context; this is because I was still at a stage where I was not sure what the sentence would mean if it was acceptable and I was hoping that speakers would themselves come up with a context (see Speaker Comment in (4)), which would then help me formulate hypotheses. The first judgment I got was positive—the sentence was judged acceptable (4a)—however when I tried it on a second consultant, they were adamant that the sentence was unacceptable (4b). No matter how much I tried to provide contexts that had been offered to me as possible verifying scenarios and how much I insisted<sup>7</sup>, the speaker justified their response with a hypothesis about the grammar of their own language (reasons 2d and e). In subsequent sessions with other consultants (4c-e), some found it unacceptable, but another found it acceptable and even offered another version of the sentence with the same construction I was testing.

- (4) a. Acceptable [EDSEI27NOV2017DRPM, elicitation] SC: She didn’t finish it, came back to write it.  
b. Unacceptable [EDSEI16ABR2018GH, elicitation] SC: No it’s not good, it’s *Maria quih*

<sup>7</sup> The specific example/judgment in (4b) is drawn from an informal one-on-one session that was carried out in a car with this particular consultant. While driving, I repeatedly submitted this sentence to this consultant and their family, each time providing a context that emphasized a particular aspect, but this consistently elicited the same response which I report above.

*hapaspoj iiqui icaaca pac iyaasipl, iyaasipl* it's only for several letters because *iyaasipl* is plural, if it's one letter you have to use the singular *iyaaspoj*

- c. Unacceptable [EDSEI18ABR2018AIMR, elicitation]
- d. Acceptable [EDSEI18ABR2018AMMO, elicitation] SC: Maria is writing a letter, you could also say *Maria quih hapaspoj iiqui icaaca zo caasipl iha*. 'Maria was writing a letter.'
- e. Unacceptable [EDSEI18ABR2018GBMR, elicitation] SC: But you could say *Maria quih hapaspoj iiqui icaaca pac iyaasipl*. 'Maria wrote letters.'

At this point I was not sure why some people easily accepted the sentence while others categorically rejected it. One evening, as I was sitting informally with several Seri people, including the speakers who had given the judgments in (4b, c), I decided to casually submit the sentence again explaining that I was interested in whether there was inter-speaker variation. I thought I'd let speakers argue it out and I'd see if there was genuine variation or if the variation I had thus far observed was just the result of my not specifying the context enough, and people accommodating different contexts for the sentence. The speakers thought long and hard about it, and one of them came up with the context in (5a); at that point, after some discussion among them, one of the speakers who I had already worked on that sentence with let out a long "Aaah" and confirmed that indeed the sentence could be acceptable but only under a very specific set of circumstances. Over the remainder of the field trip, I continued checking whether this sentence (and others similar to it) was judged acceptable (5b-e). This allowed me to ascertain that the sentence was indeed acceptable and to make a hypothesis as to what context the pluractional form of the verb *caaspoj* 'write' required, which I then tested using truth value judgments (not shown here).

- (5) a. Acceptable [EDSEI18ABR2018GH.ATHF, elicitation] ATHF: It's good if it's a very long letter and she writes it in several stages. GH: Aahh yes, that's true, but it's very specific, if the letter is really long and she does not finish it in one go, but she stops and comes back to it several times, yes it's possible. You can specify by saying *Maria quih hapaspoj iiqui icaaca zo cötiinim iyaasipl*. 'Maria is writing a letter over various times. (lit. coming and going)'
- b. Acceptable [EDSEI19ABR2018GH, elicitation] SC: you could specify by saying *Maria quih hapaspoj iiqui icaaca zo cötiinim iyaasipl*. [ELAB] Maria escribe una larga carta.
- c. Acceptable [EDSEI19ABR2018ATHF, elicitation] SC: It's good, Maria is writing a letter.
- d. Acceptable [EDSEI23ABR2018LKPH, elicitation] SC: Maria is writing a letter.
- e. Acceptable [EDSEI24ABR2018MOEA, elicitation] SC: Maria is writing a letter. But it's better to say *hapaspoj iiqui icaaca pac* with *iyaasipl*

For a long time, I would not insist or question the judgment a speaker had given me. This stemmed from the fact that speakers have the knowledge and the linguist is a student learning from the speaker-teacher. However I have now come to think of it differently. First, speakers (of any language) are not infallible, and for all the reasons in (2) may give a judgment at a particular time, or even more than once, that they will change once some things appear in a different light. Thus,

chatting and, to some extent, challenging the speaker, respectfully and in good humour, can be a way to clear up any one of the reasons of non-meaningful variation in (2).

Ways of insisting or challenging speakers that I have used include offering another sentence or context for their consideration as an alternative to one they favor or organizing a meeting between speakers who gave me differing judgments and submitting to them the data point(s) on which they disagree.

Second, I have found that variation can be due to properties of the context that the linguist isn't aware of and has not controlled for. There again, discussions of judgments and contexts with speakers can help understand possible sources of inter-speaker and intra-speaker variation.

Of course, this way of working is dependent on what the communicative conventions are in the particular community one works with<sup>8</sup>. In my first field trip to the Seri villages, I did not do much "challenging"; in fact I was still getting used to the language, so I essentially listened to what people said, e.g. I would work with speakers and try to transcribe whatever they told me as well as I could. But, in the process of working with Seri people, perhaps as we were getting to know each other a little better on a personal level, I noticed that they like teasing each other and me and do not fear challenging each other's responses in collective sessions. This made me more comfortable "challenging" them myself in later field trips, and since I saw that they seemed to respond well, I judged this way-of-doing appropriate.

Ideally, one would manage to exclude reasons like those in (2) as the sources of observed variation, and one could therefore be confident that whatever variation is found reflects something about the language itself and how it is used. But, any kind of data collection will always contain some randomness that is not explainable (e.g. for some reason, in one instance, a speaker didn't like a construction, and in another they liked it). For this reason, the best way to ensure reliability of the data is to replicate it within and across speakers.

## 2.4 Keeping track of findings

When working with consultants, I take notes directly on my computer, one file per work session. These files are saved in a "Transcriptions" folder and named according to the conventions described in section 2. Transcribed texts are saved in the "Texts" folder. I keep track of these files and the correspondence between audio and written files in a document called "SessionsData". I record all work sessions (preparation with the main consultant, data collection sessions): recordings are named according to my conventions and saved every evening to my computer in the "Recordings" folder. Questionnaires are kept in "Questionnaires" folder. When an elicitation session is based on a questionnaire, I copy/paste the questionnaire into a new Word file saved in the "Transcriptions" folder and named accordingly. I then record my consultants' comments, responses, and suggestions in that same file.

At the end of the day, I do two things. First, if I've worked on a previously prepared questionnaire on that day, I make a copy of this questionnaire in my "Questionnaires" folder and append the name of the copy/pasted file with the suffix "results" (e.g. QuestionnaireFT4results). I then proceed to transfer all the judgments I have written down in the transcription file(s) onto the QuestionnaireResults file along with a data identifier for each response. This is particularly useful as it allows me to see the collected data in a non-randomized hypothesis-based order and in one place, especially when I have used the same questionnaire for different sessions and possibly with different groups

---

<sup>8</sup> Thanks to an anonymous reviewer for drawing my attention to this point.



of consultants. Secondly I make a copy of all my files every evening on an external hard drive.

To keep track of the questions answered and hypotheses tested as I create questionnaires and collect data based on those questionnaires, I update a constantly evolving Latex file called “report”. This document is roughly organized by topic (e.g. comparatives, causatives). For each topic, I present the problem I am trying to account for and the hypotheses I entertain and, as I find or collect relevant data, I write them down in the document to support or argue against a particular hypothesis.

While in the field, I have access to the internet, and I correspond with a few collaborators via email on projects for which I want to collect data. I often receive emails with suggestions of things to test, and likewise I’ll keep collaborators abreast of any new discovery or difficulty. I save these email exchanges in a special subfolder named “EmailExchangeFieldwork”.

After I have come back from the field trip, I create a folder for the next field trip with new pensieve files to which I start adding as soon as I have ideas for the next field trip. Since I’m also in touch with my consultants through a messaging application, I also create a file “PostFT” where I copy/paste the contents of our interactions on the app. In the weeks following my return, I go over the “report” file to write thematic reports (i.e. the report contains thematic sub-reports for each topic).

### 3 Working with speakers

In the Seri village where I have done the most semantic fieldwork, the majority of adults are native speakers of Seri and use it every day (approximately 900 speakers (Marlett 2006); see O’Meara, Perales, and Pasquereau (2022) for an evaluation of the vitality of Seri). The set of speakers I work with is mostly constituted of the people that I was introduced to on my first field trip<sup>9</sup> and those that I then met through them. I don’t work on everything with every speaker. For instance, although the document referencing all the Seri speakers I have worked with contains details of 23 speakers, I have done semantic fieldwork with only nine of them. This selection largely reflects my opinion of what consultants are best suited for. Among these speakers, the person who has been my main consultant since the beginning has turned into more of a research assistant, who I in fact hire full time now when I am in the field. In what follows I refer to them as *the main consultant* even when their role is not that of a consultant.

#### 3.1 Finding speakers

In general, I subscribe to the view expressed in Dimmendaal (2001:61) that a consultant should be a co-investigator or colleague. This is the ideal situation, and this is the situation I have with my main consultant in my Seri field site. I have found that not all my consultants are quite as interested in semantic work as my main consultant but they are still good consultants with whom I work on other things (e.g. morphology, documentation). But when one starts working with a new speaker, it is not obvious right away what kind of consultant they will turn out to be. And so, I have developed a kind of routine for how I acquaint new consultants with semantic fieldwork.

When I start working with a speaker I’ve never worked with before, I usually schedule a one-on-one work session (with or without my main consultant) where we have an informal conversation for us to introduce ourselves and informally explain what I’m interested in doing. We then go over

---

<sup>9</sup> I am grateful to Carolyn O’Meara and Stephen Marlett for introducing me to people they had been working with in Desemboque de los Seris and Punta Chueca.

the informed consent form and start working on different elicitation tasks over what is left of the initial session and the next. The way I start doing elicitation with a new speaker follows roughly the following progression: translation of simple words from the contact language to the target language, translation of simple sentences in context, acceptability judgments of simple sentences (one clause) and only then do I introduce truth-value judgments.<sup>10</sup>

These initial sessions allow me to show the speaker the kind of work that we will be doing following a sequence that I have designed to be progressive, in the sense that the first tasks require fewer explanations and are generally less taxing than the ones at the end. These initial sessions are also a way to test whether a speaker is more comfortable with or interested in some tasks than others. For instance, I can see whether the speaker responds to my questions easily or whether they do not really talk, and whether the speaker has strong ideas about what is the ‘right way’ to talk as opposed to ‘wrong ways’. Having strong ideas about one’s language is not necessarily an indicator of a good or bad consultant for elicitation: if the speaker unbendingly rejects, on the basis of (perceived) prescriptism, sentences that I know to be acceptable in the language, I’m not encouraged to work with them on elicitation tasks; if however the speaker has strong opinions about which sentences are true or false in one or another context, I take this to be an encouraging trait. After one or two initial one-on-one sessions, I usually invite the speaker to take part in a collective session.

Of course, it is problematic to have to make such decisions after such a short time, and so I often revise my working decisions. In particular, I never exclude a speaker who wants to work with me; rather I will invite the speaker to work on tasks which involve a different type of work (than elicitation) or for which I need more consultants. For instance, some speakers I have worked with are really shy, however I found out that they had been doing some transcription and translation work (from Spanish into Seri) for someone else. Since I had the project of creating a book presenting a few Seri games, I asked them whether they would be interested in writing down the rules of the game with me: after I gave them a rough template, we would meet once or twice every week to monitor the work.

### **3.2 Training speakers**

As implied by the title of this section, my (current) view is that speakers should be trained for elicitation tasks to the extent that it is feasible, for two reasons: first, it’s important for the investigator to know that they are doing the intended task as otherwise the collected data would be of questionable value; second, it’s important for the speakers themselves to understand what the investigator expects, as lack of comprehension could impact them and the working relationship negatively.

I train speakers when I first start working with them and when we use a particular task for the first time at the beginning of every field trip. During any one field trip, I repeat instructions and training more or less often depending on how long it has been since the last time we used the particular task we are about to use again. At a minimum, I start every session with a short introduction of what we are about to do, and give them an explanation of the specific task we will be using. As a rule I do not mix different elicitation tasks within a session to avoid confusion.

In this section I focus on how I train speakers for the truth-value judgment task. The subset of speakers I do semantic elicitation with also do other types of elicitation with me at other times (e.g. morphology, syntax). I therefore find it important to mark the difference in tasks explicitly. In particular, I have found that semantic elicitation tasks require the most attention. Since I mostly

---

<sup>10</sup> This is also the progression I followed when I first started working on Seri to get used to the language.

conduct collective elicitation (which can get noisy and full of banter) I also like to flag that this task will be more taxing than others we sometimes use. I typically start with instructions as in (6).

- (6) Today what we are going to do is a little bit complicated. For this reason, we'll need to really concentrate, and of course, please ask questions if you're not sure about something. We [I and my main consultant] are going to give you two elements: a situation in Spanish and a sentence in *cmiique iitom*. What we want to know is whether the sentence can be used to describe the situation to someone without lying.

At this point I give them (easy) illustrative examples as in (7).<sup>11</sup> For each example, I start by reading the context (if there is one) in the contact language I use while I pass around the corresponding illustration (if there is one), then my main consultant reads out the sentence in their native language, and at this point consultants are asked to provide a judgment.

- (7) **Me:** Context: It's July in Desemboque and I'm really hot. I tell my friend:  
**Main consultant:** Ihpxahaapl.  
 I SG.INTR.EMPH.AUG.COLD  
 'I am cold.'  
**Me:** What do you think? Would I be lying to my friend if "It's July in Desemboque and I'm really hot" and I tell him:  
**Main consultant:** Ihpxahaapl. ['I'm cold.']

In general, I repeat the context and my main consultant the sentence at least once. If we see that speakers do not respond after a while, my main consultant repeats both the context (that we prepared together previously) and the sentence in the target language. This however has not happened with illustrative examples of the type in (7) as they are meant to be obvious: usually speakers, once they have understood what the task is about, will say that the sentence is a lie in the situation given that the situation says the opposite of what the sentence says.

At this point, since they have responded "lie", I follow up by asking them to explain why the sentence is a lie in the given context. I do this every time a speaker gives me the judgment "lie". This is for several reasons: (i) it can provide an indication as to whether the sentence really is false or whether the speaker finds it infelicitous or unacceptable, and (ii) it can provide clues as to what the meaning of the sentence is. An example of such a comment is provided in (8).

- (8) **Me:** Context: Every Sunday, he would make one tube, no more.  
**Main consultant:** Icatoomec quih thaa ma x, Gabriel quih hehe z  
 Sunday DET RLT:EQ DS UT Gabriel DET WOOD INDF.SG  
 iyahoiilc.  
 3>3:RLYO:CAUS:be\_tubular.PLUR  
 'On Sundays, he made tubes (out of one tree).'
- Consultants :** LIE SC: because there it says *iyahoiilc* and the sentence gives to understand that each Sunday he made several tubes, here it should be *iyahoiij*. [Questionnaire2FT5results, ED-SEI16NOV2019DRPM.GH.GHF.LKPH.ATHF2, CON]

<sup>11</sup> I make sure to check these examples beforehand with several people to make sure there are no issues with them and to make sure that the judgments are easy and obvious to most speakers.

The second illustrative example is one where the sentence is true in the context (9). This again is designed to be straightforward. In general, when speakers give the judgment “true”, I follow up by asking them to at least repeat the sentence (and sometimes to repeat the context back to me).

- (9) **Me:** Context: It’s January in Desemboque, it’s very windy and I’m really cold. I tell my friend:  
**Main consultant:** Ihpxahaapl. [‘I’m cold.’]

These two straightforward examples have illustrated what a context or situation is, what kind of sentences will be given, what types of responses they can give, and what the task is that they will be asked to do. At this point, I point out that the previous examples show that a particular sentence can be true or a lie in a particular situation, but there is a third possibility: namely that in a given situation or context, a sentence is not a lie but lacks details and could be more specific. This is then further illustrated by examples, one for each judgment value (10). Of course, I do not use the phenomenon under investigation in the illustrative examples; rather I try to choose a distinct but similar phenomenon (with the same level of difficulty and the same types of contrast if possible).

- (10) **Me:** Now we are going to give you a few examples of situation/sentence pairs in which, in the given situation, the sentence is true, a lie, or it lacks details but it’s not a lie.

*I acted out the following contexts in addition to reading them out.*

- a. **Me:** Context: Juan, Gabrielito and Alina are small children, and they like it when someone spins them around. Gabriel spun Juan around once, then he spun Gabrielito around once, and at last he spun Alina around once.

**Main consultant:** Gabriel quih xicacaziil quih iyofizotim.

Gabriel DET child.PL DET 3>3.RLYO.lift.PLUR

‘Gabriel lifted the children.’

**Expected judgment:** TRUE or NOT A LIE BUT COULD BE MORE SPECIFIC

- b. **Me:** Context: Juan, Gabrielito and Alina are small children, and they like it when someone spins them around. Gabriel spun Juan around once, then he spun Gabrielito around once, and at last he spun Alina around once.<sup>12</sup>

**Main consultant:** Gabriel quih xicacaziil quih hant iyojeaatim.

Gabriel DET child.PL DET land 3>3.RLYO.drop.PLUR

‘Gabriel let the children fall.’

**Expected judgment:**LIE

- c. **Me:** Context: Juan, Gabrielito and Alina are small children, and they like it when someone spins them around. Gabriel spun Juan around once, then he spun Gabrielito around once, and at last he spun Alina around once.

**Main consultant:** Gabriel quih xicacaziil quih iyocaailim.

Gabriel DET child.PL DET 3>3.RLYO.spin.PLUR

‘Gabriel spun the children around.’

**Expected judgment:** TRUE

<sup>12</sup> Although the written rendition of the context here does not specify whether the children did or did not fall, the actual acted out description of the context did not involve any dropping of the (imaginary) children. (Thanks to Lisa Matthewson (p.c.) for helping clarify this point.)

Once we have worked through these examples, we are ready to begin. I usually end the explanations with a summary of the different types of judgments they can give me (11).

- (11) **Me:** We'll continue the session with this kind of work. Don't forget, pay attention to both the situation and the sentence and tell us if:
- the sentence is true in the situation;
  - the sentence is a lie given the situation, or;
  - the sentence is not a lie but it could be more specific.

Semantic analyses make predictions in terms of truth and falsity (and infelicity and pragmatic anomaly—but I focus on truth/falsity here), and this is why a piece of data in semantics has, as one of its components, a truth value judgment, with response values “true” or “false” (Matthewson 2004). Practices of how such judgment values are collected seem to differ. In one type of practice, consultants do not necessarily utter the words “true” or “false/lie”; rather, linguists record their consultant's response, whatever it may be (e.g. verbal, visual, a particular sentence like “you could say that but I would not”), and then, on the basis of their experience with this consultant, further interpret this response as corresponding to the true or false judgment value. In another type of practice (an implementation of which has been illustrated here), consultants directly utter the words “true” or “false/lie” as judgments. Because the second type of practice has been possible to implement in my particular fieldwork situations, I have favored it. The advantage of letting speakers themselves use the judgment values “true”, “lie”, or “not a lie but could be more specific” is that I do not need to interpret whatever else they might say instead as corresponding to “true” or “false”. This thus reduces the amount of interpretation I have to do of their responses, thus reducing the probability that their responses are mapped to the wrong judgment value because of e.g. misinterpretation or bias towards a particular judgment value on my part.

Speakers sometimes comment that I seem to be asking them the same questions several times, which can be frustrating or at least unsettling to them as they interpret it as my not paying attention or not trusting them. Indeed many situation/sentence pairs over time are similar—this is because I collect minimal pairs—and sometimes they are indeed identical from one time to the next; this is because I try to replicate data over time. I try to reassure consultants that I do pay attention but that (i) I am interested in subtle differences in their language and although sometimes sentences might seem identical to a previous one, they are not, they just differ in one word, and that (ii) it is very important that I double-check some sentences more than once to make sure I do not make mistaken claims about Seri.

Are these explanations enough for speakers to understand such a complex task? With repetitions and more explanations along the way, I think so. With some speakers that are now used to doing this kind of work, I don't need to give a whole set of instructions, and the judgments can be collected more informally. Of course, every session is an occasion to practice more and give more explanations: both from me to the speakers in terms of the kinds of things that I'm looking for, and from them to me in terms of the things that I may have overlooked in creating elicitation stimuli.

### 3.3 Conducting elicitation sessions

The practice of fieldwork can generate a tension between the rigor suggested by the word “work” and the flexibility required by the circumstances of the field. This is perhaps especially true of elicitation, as the linguist follows a preestablished line of research, and speakers are asked to perform specific

tasks on specific linguistic expressions. Two ways I have found to help ease that tension are fostering a pleasant work environment and conducting collective work sessions.

### 3.3.1 Fostering a productive atmosphere

In both places where I have done fieldwork, I usually meet consultants at the place I'm staying in during the trip. When consultants arrive, I offer them coffee or tea and place some snacks on the table. Throughout the session, we take small breaks for small talk and refills. I find doing this helps relax my consultants and makes the atmosphere more casual. This makes the sessions more fruitful and improves the quality of our work. In fact, some of the epiphanies of my research have been triggered by comments speakers have made in relaxed informal conversations, where they felt they could talk at leisure. Even though speakers may often give details that do not appear directly relevant, I always show interest in whatever they say and write it down because I want to foster this type of easy-going conversation in my work sessions.

For instance, in one particular collective session dedicated to getting paradigms of possessed and non-possessed nouns, a speaker's comment led to a breakthrough in my understanding of the meaning of verb forms. At the end of the session, as consultants were leaving, one lingered on and offered another carrier sentence for the noun *haaho* 'path' which we had ended the session on. The speaker seemed to have good intuitions about this sentence and, since I was also interested in understanding the meaning of verb forms in Seri for another project, I informally checked whether he'd also accept a version of the sentence he had offered with a different form of the main verb—by hypothesis a pluractional form—the verb *cajoene* 'raise dust, be dusty'. He commented that it could be used if the dust was in several separated areas (12).

(12) Notes taken from [EDSEI10MAY2019DRPM.AMMO.GH]

**Consultant:** Haaho tintica yajoene [ELAB, GH]

path DEM RLYO.raise\_dust

'That path is dusty.'

**Me:** Is the following sentence acceptable? Haaho tintica yajoene-tim [CON]

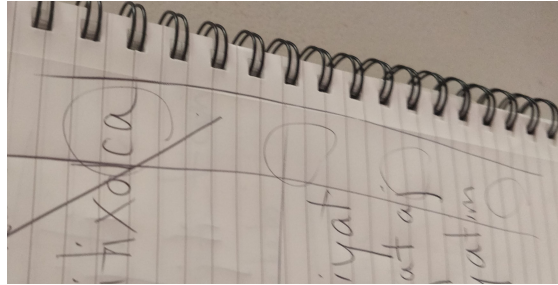
path DEM RLYO.raise\_dust-PLUR

'That path is dusty.'

**Consultant** It's acceptable yes it exists, it means that the path has dust in various parts.

*The consultant at this point took the notebook I use, and drew a rough illustration of a verifying scenario they had in mind for this sentence (figure 2), after which they made the following comment.*

**Consultant:** *yajoenetim* is good if the dust is separated.



**Figure 2:** Picture drawn by consultant during an elicitation session<sup>13</sup>

This comment gave me the idea for a hypothesis on the meaning of pluractional stative positional verbs, e.g. *caii* ‘stand’ and *coom* ‘lie’, which had eluded me until then, namely that such pluractional forms require of the plural referent of their subject that they be individuatable (13).

(13) Notes taken from [EDSEI10MAY2019DRPM.AMMO.GH]

**Me:** Imagine that you’re on top of a mountain and you see a forest all around, it’s all green (figure 3a). In this situation, could you say the following without lying:

**Main consultant:** Hehet pac, taa yaayolca. [ELAB]  
 tree.PL INDF.PL DEM.DIST.PL RLYO.stand.PLUR  
 ‘There are trees.’

**Consultants:** No, that’s a lie. It means that the trees really stand out, but if the trees do not stand out, you can’t say *yaayolca* because I can’t focus on any individual trees. You could say that if some trees somehow stand out, maybe they’re taller. Here it’s only

Hehet pac taa yaii. [ELAB]  
 tree.PL INDF.PL DEM.DIST.PL RLYO.stand  
 ‘There are trees.’

**Me:** Okay, now imagine you’re on top of a mountain and you see a forest through which three rivers flow in different directions (figure 3b). Would you be lying if you said:

**Main consultant:** Hehet pac taa yaayolca. [ELAB]

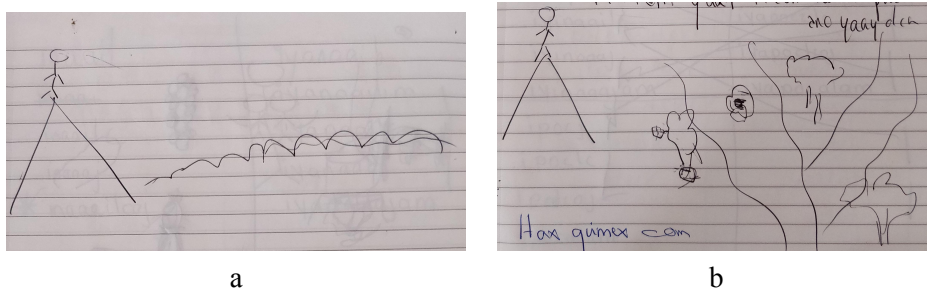
**Consultants:** Now that’s true. You could even say

Hax quimex com hehet pac iti yaayolca. [ELAB]  
 ABS.water SBJ.NMLZ.flow DEF.SG tree.PL INDF.PL 3[POSS].in RLYO.stand.PLUR  
 ‘In between rivers, there are trees.’ (lit. in the river, trees stand)

Because these judgments were obtained informally from just one consultant at the end of a session in which another task had been used, I further worked on these contrasts by integrating the suggested context-sentence pairs (and other similar ones I later created with the help of my main consultant) in a questionnaire we worked on in later sessions with more consultants. Though most elicitation sessions are questionnaire-led, comments and seemingly off-topic exchanges like the one above are fortunately very frequent.

I usually aim for a session to last about one hour but this can vary widely depending on the work we are doing, the state of fatigue my consultants and I are in, and whether I conduct collective or one-on-one sessions.

<sup>13</sup> The picture was drawn over previous writings; still, a path (two parallel lines) as well as several spatially separated dusty areas (circles) can be distinguished.



**Figure 3:** Pictures drawn by me during an elicitation session

### 3.3.2 Collective vs individual sessions

I aim to have every data point judged by at least six speakers (more if there is variation and/or if the contrast under investigation is very subtle)<sup>14</sup>. I have done both one-on-one and many-on-one elicitation sessions, though in recent years, I have tended to do more collective elicitation sessions. Both types of sessions have pros and cons (summarized in Table 1), but I have found that, at least for my fieldwork with the Seris, the pros of collective elicitation outweigh its cons, and that the latter can be controlled for to some extent.

	PROS	CONS
one-on-one	-no biasing from other speakers	-random variation due to e.g. lack of imagination, internal theory of the language
many-on-one	-speakers can help/challenge each other -more relaxed/pleasant working environment	-biasing from other speakers (social dynamics of groups) can flatten meaningful variation -hard to keep track of individual judgments/comments

**Table 1:** Pros and cons of individual and collective sessions

The obvious advantage of one-on-one sessions is that the speaker does not get influenced by other speakers, which is essential if, for instance, one is interested in potential dialectal differences. On the other hand, the absence of other speakers means that the speaker's beliefs do not get challenged at all, which can become problematic if said speaker, at the time of the elicitation, just happens to e.g. misinterpret the context provided by the linguist in a truth-value judgment task.

Collective elicitation sessions provide an instant fix to this shortcoming of one-on-one sessions: not only can speakers in a collective session challenge each other, they can also help each other in cases of momentary memory lapses or misunderstandings. Having several speakers at once also helps create a more pleasant atmosphere (provided the speakers get along of course), as speakers can converse among themselves.

On the other hand, social dynamics unknown to the outsider fieldworker can negatively impact a collective elicitation session. Having several speakers and judgments at the same time can have the

<sup>14</sup> Why six speakers? It's a rule of thumb inspired by the Small N Acceptability Paradigm (aka SNAP) judgments method proposed in (Mahowald, Hartman, Graff, and Gibson 2016) with 5 speakers.



unfortunate effect of inhibiting and even eclipsing an individual speaker's responses, especially in cases of disagreement. This in turn can have the effect of artificially homogenizing the grammar of the language under study by inhibiting individual speakers' potential variation. For instance, I once invited a speaker to join the group of speakers I usually work with. My usual consultants became a lot quieter in that session, and when they would speak it was to say they agreed with the consultant most recently arrived; this was made all the more obvious as I inserted here and there stimuli which I had checked several times before. In this particular case, the change in social dynamics was particularly obvious but it could be very subtle.

The second potential disadvantage of collective sessions, not unrelated to the first, is that the more speakers there are in one session, the more difficult it becomes to keep track of speakers' individual judgments and comments. This could be because some speakers are louder than others, or because the linguist can't keep up with speakers who might be talking at the same time. Both this disadvantage and the former can result in meaningful variation in the grammars of speakers going unnoticed.

Still, I have tried to counterbalance these potential drawbacks of collective sessions. As mentioned earlier, I hire my main consultant full time as my assistant when I'm in the field. In addition to the many ways in which this makes my work there more productive, they can warn me of social dynamics which would undermine collective sessions; they also provide insights on conversations going on between consultants, which may reveal helpful information about the task we are working on.

When I begin a work session (whether individual or collective in fact), I always give background on what the session is going to be about, that there may very well not be a single answer, and that it is likely consultants will have different opinions. In order to further encourage people to give their own judgment (regardless of what other consultants say) and to keep track of each speaker's judgments, I have started using sets of answer cards that each speaker can use and which help me not to lose track of who said what (see appendix C). So far, I have trialed answer cards for:

- truth-value judgments, with values "true", "not a lie but could be more specific", "lie"
- lexical recognition judgments, with values "yes, it's a word", "maybe, it could be a word, I understand what it means", "no, it's not a word"

I regularly work with eight consultants in addition to my main consultant. When I work in collective sessions, I limit the number of participants to four maximum (in addition to my main consultant). The composition of groups therefore varies depending on who is available when. I find it useful to keep track of which task/questionnaire I have done with which consultant(s). To do this, I have a document called "Schedule" which I use to plan, to the extent possible, when I will do which task with which consultant, and another document "Work.Informants" to keep track of what I have already done with each consultant.

#### **4 Conclusion**

Elicitation requires collecting speakers' intuitions about their language using prompts designed by linguists. But these intuitions can be affected by a host of factors that are not linguistically relevant. It is therefore necessary to control for the manner in which the prompts are presented and the intuitions are collected, beyond ensuring the adequacy of the particular prompts/methods used.

This paper has presented the ways I have implemented semantic elicitation-based methods in my fieldwork so far. The first section gave an overview of how I manage data (create stimuli, deal with variation in judgments, and keep track of findings), and the second section focused on working with speakers in elicitation (how to find them, train them, and organize an elicitation session).

Different implementation methods might be more appropriate or successful for different fieldwork environments. It is also likely that how a linguist implements elicitation methods influences what results they collect. Providing details about the logistics of how the examples used in a paper were collected would therefore allow readers to evaluate their confidence in the data.

## References

- Bertinetto, Pier Marco, Karen H. Ebert, and Casper Groot. 2008. The progressive in Europe. In *Tense and aspect in the languages of Europe*, 517–558. Berlin: De Gruyter Mouton.
- Bowern, Claire. 2015. *Linguistic fieldwork: A practical guide*. New York: Palgrave Macmillan.
- Cabredo Hofherr, Patricia, Jérémy Pasquereau, and Carolyn O’Meara. 2018. Event plurality in Seri. In *Proceedings of the Tenth Conference on the Semantics of Under-represented Languages in the Americas 10*, ed. Kimberly Johnson and Alexander Göbel, 1–16. Amherst: University of Massachusetts, Graduate Linguistic Student Association.
- Chelliah, Shobhana L, and Jules Willem. 2010. *Handbook of descriptive linguistic fieldwork*. Heidelberg: Springer.
- Dahl, Östen. 1985. *Tense and aspect systems*. London and New York: Basil Blackwell.
- Dimmendaal, Gerrit J. 2001. Places and people: field sites and informants. In *Linguistic Fieldwork*, ed. Paul Newman and Martha Ratliff, 55–75. Cambridge: Cambridge University Press.
- Drummond, Alex. 2013. Ibex farm. URL <http://spellout.net/ibexfarm>.
- Mahowald, Kyle, Jeremy Hartman, Peter Graff, and Edward Gibson. 2016. Snap judgments: A small n acceptability paradigm (snap) for linguistic acceptability judgments. *Language* 92:619–635.
- Marlett, Stephen A. 2006. La situación sociolingüística de la lengua seri en 2006. In *Situaciones sociolingüísticas de lenguas amerindias*, ed. Stephen A. Marlett. Lima: SIL International and Universidad Ricardo Palma. URL [http://lengamer.org/publicaciones/trabajos/seri\\_socio.pdf](http://lengamer.org/publicaciones/trabajos/seri_socio.pdf).
- Matthewson, Lisa. 2004. On the methodology of semantic fieldwork. *International Journal of American Linguistics* 70:369–415.
- Meakins, Felicity, Jennifer Green, and Myfany Turpin. 2018. *Understanding linguistic fieldwork*. Understanding Language series. London and New York: Routledge.
- Newman, Paul, and Martha Ratliff. 2001. *Linguistic fieldwork*. ERIC.
- O’Meara, Carolyn, Debora Perales, and Jérémy Pasquereau. 2022. Evaluación preliminar de la vitalidad de la lengua cmiiique iitom (seri). *Verbum et Lingua* 19:59–74. Guadalajara: Universidad de Guadalajara.
- Pasquereau, Jérémy. 2021. A unique operator for verbal pluractionality and numeral distributivity. In *Proceedings of Sinn und Bedeutung 25*, ed. Patrick Grosz, Luisa Martí, Hazel Pearson, Yasutada Sudo, and Sarah Zobel, 702–719. London: University College London and Queen Mary

University of London.

Pasquereau, Jérémy, and Patricia Cabredo Hofherr. 2020. Using semantics to probe paradigm structure: the case of multiple event marking in seri. *Lingue e linguaggio* 19:35–60.

Pasquereau, Jérémy, and Patricia Cabredo Hofherr. 2021. Two types of pluractionality in Seri. In *Proceedings of the Eleventh Conference on the Semantics of Under-represented Languages in the Americas II*, ed. Kimberly Johnson and Alexander Göbel, 1–16. Amherst: University of Massachusetts, Graduate Linguistic Student Association.

Sakel, Jeanette, and Daniel L Everett. 2012. *Linguistic fieldwork: A student guide*. Cambridge UK: Cambridge University Press.

Schütze, Carson T. 2016. *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Berlin: Language Science Press.

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

## **A Documents needed for a field trip**

- Administrative documents
  - Expenses spreadsheet
  - Forms
    - ◇ Consent forms
    - ◇ Introduction letter
- Pensieves
- Questionnaires
- Data
  - Recordings
  - Transcriptions
  - Texts
  - SessionsData (file that keeps track of the context of this folder by date)
  - PostFT4 (document containing data collected remotely after the field trip, here field trip 4, has ended)
- EmailExchangeFieldwork (document containing copies of emails received and sent during fieldwork)
- Report (of results)
- Schedule (work schedule in the field)
- Work.Informants
- SpeakersData

## **B Data tagging**

For each data point I collect, I indicate whether the data point occurred spontaneously (without any injunction from me) or whether it is the result of elicitation, in which case I sometimes specify what kind of elicitation task triggered it (Table 2). In addition to translation, I distinguish the following elicitation tasks by the type of data given by the speaker: judgments (e.g. acceptability, truth-value), elaboration, contextualization (verbal or drawing), description.

I use the word ‘elaboration’ to refer to a task where a sentence in the target language is produced as a result of the linguist’s presenting the consultant with a target language expression. Elaborations can be sentence examples that speakers are asked to come up with on the basis of one word or one expression. In this case, the sentence is not a judgment, nor a translation or a spontaneous example since speakers were asked to make it up on the basis of a prompt given by the linguist. I also call elaborations examples proposed by speakers as correct or better versions of a sentence that they were asked to judge. Contextualizations are tasks where on the basis of a target-language sentence, the

consultant is asked to come up with a context in which this sentence can be used truthfully. The context can be provided verbally or visually via drawing. Finally, I call descriptions target language sentences produced to describe a non-verbal stimulus presented by the linguist (e.g. storyboards).

Stimulus	Data (dependent variable)	Task
not from linguist	sentence in target language	spontaneous
<b>in contact language:</b> sentence	sentence in target language	translation
<b>in target language:</b> sentence (+ context)	judgment sentence in contact language sentence in target language drawing of the situation described by the sentence context in which sentence can be used truthfully	judgment translation elaboration? contextualization (drawing) contextualization (verbal)
expression (sub-sentential)	sentence in target language	elaboration
<b>non-verbal:</b> film, picture	sentence in target language	description

**Table 2:** Sources of collected data as a function of the stimulus presented to the speaker and the nature of the data obtained

Since judgments involve presenting a target language sentence (possibly with a context) and asking the consultant to assess e.g. its acceptability, it can be useful to specify how the presented target language sentence has been obtained: whether the sentence is from spontaneous speech, translation, elaboration, description, published data and/or whether it's been constructed by the linguist. This gives the possible combinations and tags in Table 3.

A constructed example is any sentence in the target language that has been created in part or fully by a non-native speaker. It can be based on a sentence produced by a native speaker (either a spontaneous one, or a sentence obtained as a result of translation, elaboration or description) that is further modified by a non-native speaker.

Context	task		sub-task	abbreviation	
Spontaneous				SPON	
Elicitation	translation	from [LANG]		TRANS[LANG]	
	elaboration			ELAB	
	description			DESC	
	drawing			DRAW	
	contextualization			CTXT	
	judgment	of produced sentence	of spontaneous speech	acceptability judgment	ACC.SPON
				truth-value judgment	TVAL.SPON
				felicity judgment	FEL.SPON
			of translation from [LANG]	acceptability judgment	ACC.TRANS[LANG]
				truth-value judgment	TVAL.TRANS[LANG]
				felicity judgment	FEL.TRANS[LANG]
			of elaboration	acceptability judgment	ACC.ELAB
				truth-value judgment	TVAL.ELAB
				felicity judgment	FEL.ELAB
			of description	acceptability judgment	ACC.DESC
truth-value judgment				TVAL.DESC	
felicity judgment				FEL.DESC	
	of constructed sentence		acceptability judgment	ACC.CON	
			truth-value judgment	TVAL.CON	
			felicity judgment	FEL.CON	
	of published data		acceptability judgment	ACC.PUB	
			truth-value judgment	TVAL.PUB	
			felicity judgment	FEL.PUB	

**Table 3:** Proposed abbreviations

C Answer cards

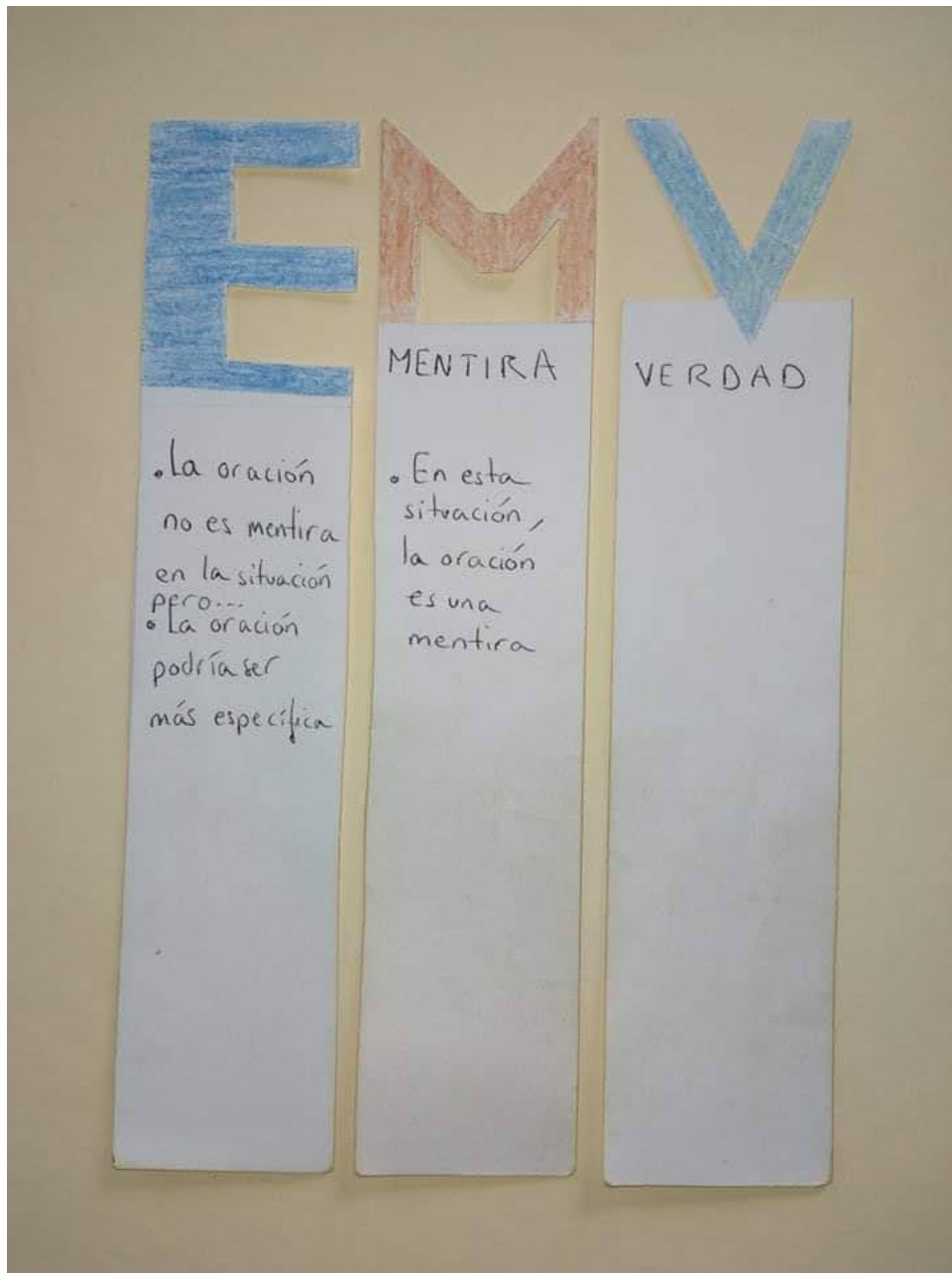
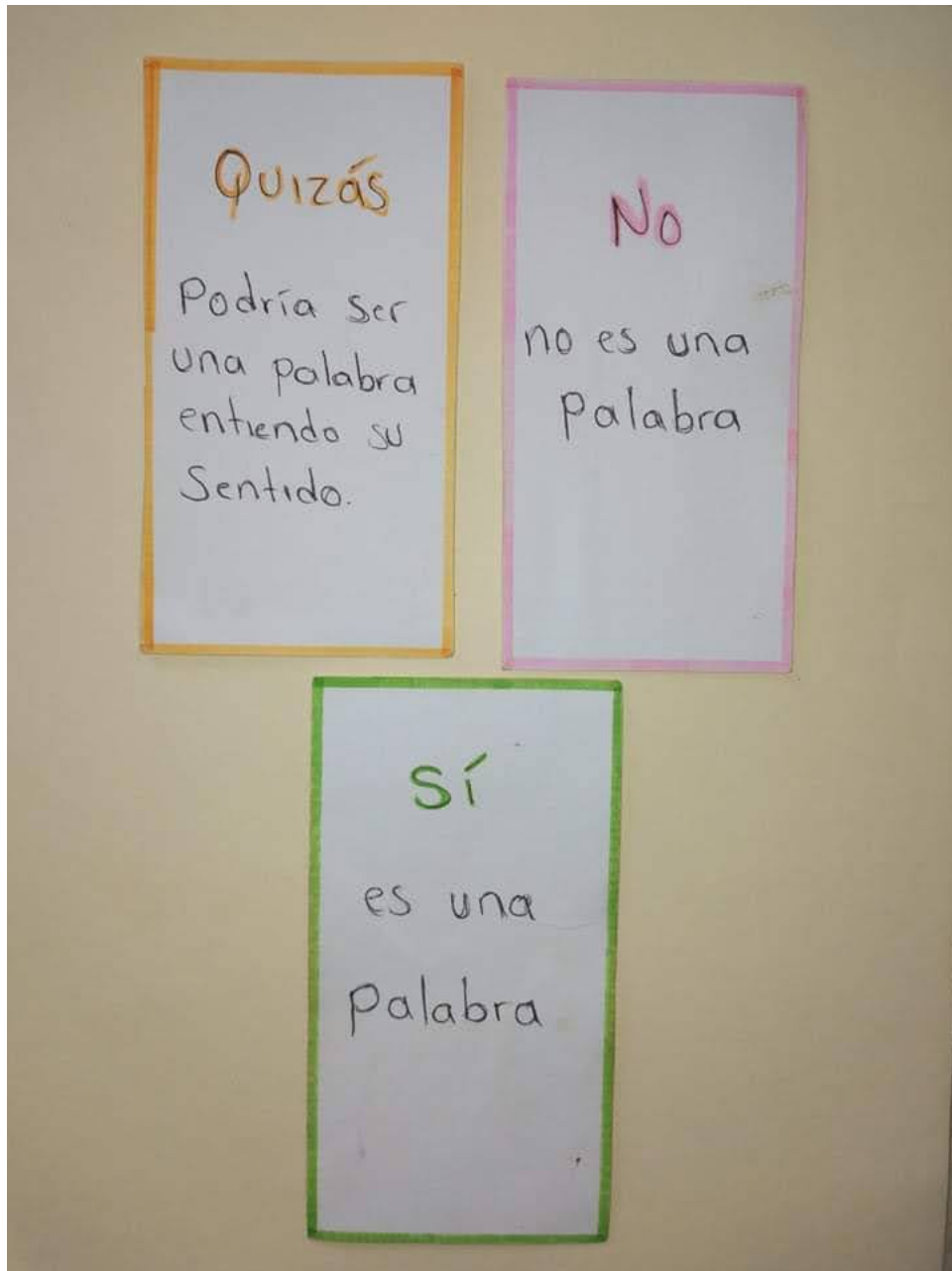


Figure 4: Set of answer cards for truth-value judgments



**Figure 5:** Set of answer cards for lexical recognition judgments