

EXPERIMENTAL SYNTAX AND LINGUISTIC FIELDWORK

Maria Polinsky

1. The fieldworker's backpack and the experimenter's lab coat

Linguistic fieldwork is research conducted on a language that the linguist does not speak natively, through the collection of primary language data gathered in interaction with native-speaking consultants (Chelliah and de Reuse 2011: 7; Bochnak and Matthewson 2015: 2; Bowerman 2008). Experimental fieldwork is simply experimental work conducted in a natural setting (the location where a given language is spoken), rather than in the researcher's lab or online. Although nothing in this definition requires that the work be conducted on a language spoken in a remote or poorly accessible setting, or an endangered language, or an understudied language, one or all of these extra conditions are often implicit in our understanding of linguistic fieldwork—and these assumptions can get in the way of planning potential experimental work.

It is common to contrast linguistic fieldwork with lab-based experimental work on language, but the two are less different than they seem. Both lines of inquiry work to understand the mental representation of language by a native speaker; both are guided by testable hypotheses; both are designed to evaluate predictions based on theoretical considerations; both conduct those evaluations by constructing minimal contrasts; and both deal with variation within and across language users. Fieldwork essentially consists of tiny experiments that are fine-tuned in situ based on consultant feedback. In short, there is no irreconcilable difference between the fieldwork culture and the culture of laboratory linguistics—and experimental syntax, in particular.

However, differences between the two do exist. In the paragraphs below, I will address the main points of divergence: the baseline data used, the nature and role of the language consultant, and the degree of a researcher's involvement in the language community.

Experimental syntax typically relies on already-established data and uses established syntactic analyses as the springboard, whereas fieldwork relies on primary data that is collected and analyzed by the same group of researchers. Since experimental syntax relies on existing analyses, it has mainly dealt with well-described and thoroughly analyzed languages. But, it is helpful to remember that deep analyses of such languages also started with introspection and conversations between linguists as to whether structure X was possible, and if not, why. A fieldworker working on a lesser-studied language needs to collect the primary data first, and such collection involves both natural production and targeted elicitation; this collection is essentially an experiment, in the broad sense of that word. Reductively, then, the fundamental difference between lab work and fieldwork is that the fieldworker needs to do more preliminary work before s/he can start an experiment in the narrow sense.

While experimental research on languages is viewed as *modeling grammars*, fieldwork is often associated with *descriptive work* (within the latter, work on endangered languages often goes under the special rubric of *salvage work*). Yet descriptive work is just the first step in constructing the model of a new language available to a researcher in the field. This first step is simply taken for granted in experimental work; someone has

already done the original data collection, and that work can be relied on without much hesitation.

The better-described languages primarily represent educated, rich societies (particularly the English-speaking world), and it is from this population that experimental syntactic work has primarily drawn. In psychology, researchers have expressed concern that the oversampling of people from Western, educated, industrialized, rich, and democratic (WEIRD) societies—who represent as much as 80 percent of study participants, but only 12 percent of the world’s population—may be skewing our understanding of human behavior and culture (Henrich et al. 2010). Likewise, in experimental work, the emphasis has been on monolingual, young, available, and literate speakers (MYALSs), and that may be skewing our perception of native speakerhood. In some ways, the use of MYALSs is a matter of expediency. These days, it takes only an hour to collect online judgment data from English speakers, however narrowly defined, and to build an experimental paradigm based on those data. Fieldwork, on the other hand, often—though not always—involves speakers in more remote areas, who may be less literate or educated, are often bilingual, and may not be as comfortable with test-taking as MYALSs.

The two subfields differ both in the types of participants they recruit as well as in potential sample sizes. Experimental syntax is often based on large-scale comparison with many MYALSs (Sprouse and Almeida 2012, 2017), while a fieldworker settled in a small language community on the wane may be dealing with five remaining speakers of a given language. The reality is that, when working with an endangered language, the luxury of large numbers is simply not there. It may therefore be tempting to think that “the conclusions that can be drawn from [data from endangered languages] will be weaker and more speculative in nature than the conclusions based on quantitative data” (Gibson and Fedorenko 2013: 94) and that “obscure, little-studied languages [are an]... unsatisfactory data source” (Featherston 2007: 278) Yet, that’s not a sufficient reason to abandon an experimental approach to these languages. After all, studies of clinical populations, fMRI research, most studies in phonetics, and research on sign language also have tiny subject pools but are proudly counted among experimental approaches.

In experimental work, the language speaker is either a *participant* or a *subject*; in fieldwork, the speaker is called the *informant* or language *consultant*. The consultant’s role is much more active than that of the participant; the consultant is not just at the receiving end of the language experiment but contributes to the data and the flow of work. Both the fieldworker and the consultant are trying to get a glimpse of the consultant’s mind, so the consultant is simultaneously an object of testing and an active participant in that testing. The more trained a consultant becomes, the more eager s/he may be to offer opinions on the relative acceptability of similar examples or even the possible analysis of a particular structure—an analysis that may not involve linguistic lingo but can be insightful and informative. The fieldworker and consultant together perform iterative experiments, asking a similar question over and over again to the point where a semblance of statistical significance may arise (an issue I will revisit below). Both the fieldworker and the consultant engage in learning: not just about the consultant’s language, but about linguistics as well. In fact, some successful fieldworker–consultant partnerships have led to the training of native-speaker linguists, a model pioneered by Ken Hale in the USA (Hale 1972a, b, 1992) and Peter Skorik (Vaxtin and

Golovko 2005) and Alexander Kibrik (Kibrik 2005; Vaxtin and Golovko 2005) in Russia.¹ Unlike experimental study participants, fieldwork consultants may be older, taking the Y out of MYALS and potentially leading to additional challenges brought on by aging (on the effects of aging on language, see Kemper et al. 2001; Burke and Shafto 2008). Fieldwork consultants may also lack the educational savvy of MYALS encoded by the L in this abbreviation, as speakers of understudied or endangered languages often lack literacy in the traditional sense. This limitation, too, imposes additional requirements on researchers and their paradigm.

Related to these differences, fieldworkers and experimentalists differ in the degree to which they are involved in the community. Experimentalists and their MYALSs rarely forge long-lasting relationships. MYALSs (and other experimental participants) take part in an experiment, answer a few questions before and after, get compensated for their participation, and leave. They rarely come back, unless the return is for a follow-up experiment, and they are not connected to the researcher in a meaningful way. In linguistic fieldwork, on the other hand, it is crucial to build a strong relationship with the language consultant(s). Since this process takes time, trust, mutual understanding, patience, and strong motivation, fieldwork tends to be a self-selecting discipline; researchers who view the consultant as a mere machine for producing language do not last long, and typically switch to a different line of inquiry.

Given the preceding discussion, it is evident that in order to conduct experimental work in the field, one needs a team. A one-man orchestra will not do, as several strands of expertise are required. An *experienced fieldworker* with good ties to a community can provide the primary data, help establish contacts, and (in the optimal cases) train local native-speaker linguists in the ongoing and future work. A *syntactician* is needed to do what they do best: articulate a specific set of hypotheses and propose ways of collecting, norming, and analyzing data to test them. This role can, in theory, be fulfilled by just one person, though a team of syntacticians may often be involved—with some members focused on the theory and others on the experimental aspect of the study. Additionally, a *psycholinguist or neurolinguist* is needed to design and conduct the actual experiment, employing the appropriate and necessary methods given the research question at hand (I would like to underscore that I remain agnostic in terms of which methods need to be used—they could include judgments, reaction times, electrophysiology, etc.)

Every fieldwork situation is different, but if *native-speaker linguists* are available, they can provide an important link between the different strands of inquiry and become a major force in bringing experiments to the field. The team I described here is an idealization, reflecting both my own experiences as well as the main components that work to produce effective experimentation in the field.

No matter how many people are involved, the most successful fieldwork experimentation projects are those where all of the team members share some common assumptions and continue educating each other. In modern times, when it is possible for people to meet without being in the same room, regular meetings for an ongoing project are not only desirable, but doable. Projects are less successful when each person is

¹ Other names may be added to this list; there is no comprehensive accounting of all of the outside researchers who have overseen or encouraged the training of native-speaker linguists.

responsible for a narrow corner of the work and communication amongst the team is limited. One helpful outcome of team interaction in experimental fieldwork is that each type of researcher is forced to get out of their comfort zone. A theoretical syntactician may have to explain what parasitic gaps are and why they are relevant to teammates who do not have the concept at their fingertips; a fieldworker may have to clarify the role that different conjugations play in the language under discussion and why ‘boil’ cannot be expressed by just one word; a native-speaker linguist may be aware of gender differences in the use of proper names, and a person who knows eye-tracking may balk at the use of ambiguous referents in the visual stimuli. Although it appears obvious that teammates should be willing to educate each other, such interactions do not always happen. Like the work with native consultants, this type of communication requires trust, mutual respect, appreciation of the others’ expertise, and patience.

In sum, fieldwork and experimental syntax can and should be combined; they share a number of premises and they stand to enrich one another. But in order for this marriage of minds to be successful, it is important to plan carefully, and it is this planning that I will explore in the rest of this chapter. While my focus is on syntactic work, there are encouraging examples of fieldwork and experimentation converging outside of syntax, especially in the semantic realm (Arunachalam and Kothari 2011; Bohnmeyer 2015; Gil 2008; Bochnak and Matthewson 2015). The two fields can enhance each other by sharing approaches and tools.

2. Conceptual issues

Ultimately, since both fieldwork and experimental syntax work on *language*, many of the approaches used in the two disciplines are parallel—successful approaches to language in general are also often successful approaches to lesser-studied languages in an experimental setting. In this section, I will discuss some conceptual issues surrounding approaches to fieldwork and experimentation.

There are two schools of thought concerning fieldwork. According to American structuralists, fieldwork is all about the process of discovery—you approach the language as a complete unknown, ignoring any information that may already be available for fear of getting tainted with incorrect ideas. The beauty of this approach is that you learn by trial and error (a lot of it, too) and whatever you learn stays with you forever. The alternative approach is to gather as much information as possible prior to embarking on the project, in the hopes that your fieldwork will allow you to verify what you read (it is always good to question whether the other researcher got it right) and move ahead with interesting discoveries. Understanding the general theoretical landscape of a given phenomenon is important in both approaches; if you know how noun–adjective combinations are built, for example, you can ask better questions when encountering them in your language of study. The latter approach is particularly helpful for experimental work in general, and experimental work in syntax more specifically. Having a working knowledge of the existing work on a given language (if any such work is available) as well as syntactic theory allows for more effective exploration and falsification.

In the age of powerful statistics and great gadgets (many of which are improving faster than our theories!), experimental work is tempting. Experimental studies seduce us

- (6) a. [How many people]_i did Kim talk [to *t_i*] about Mary?
 b. ??/*[How many people]_i did Kim talk [about Mary] [to *t_i*]?

Intervention effects are notoriously varied and difficult, and they have raised many theoretical questions: Do intervention effects stem from semantics (as has been argued primarily for intervention effects in negative islands, cf. Kluender and Gieseman 2013)? Can they be reduced to economy conditions (Lechner 2013)? Do they have more to do with the nature of probing for syntactic features than with the intervener itself (Preminger 2014)? Should they be discounted as effects of linear order (Bruening 2014)? The theoretical hypotheses here are clear, and it would be relatively straightforward to set up an experiment differentiating at least the semantic and the syntactic accounts of intervention. (A theoretical syntactician or fieldwork linguist who is not well versed in principles of processing may be duly intrigued by this phenomenon.)

Knowledge of processing, however, turns the intervention phenomenon on its head. Processing operates incrementally, and one of its basic tenets is that, once a filler is identified in a non-argument position (the *wh*-position, in this case), it should be linked to a gap as soon as possible—a concept known as the Active Filler Hypothesis (Clifton and Frazier 1989). The opportunity to link a gap with an available filler supersedes the opportunity to identify a lexical phrase of category XP. Thus, with respect to the intervention example in (6), the Active Filler Hypothesis predicts that the parse below, where *how many people* is associated with the stranded preposition *about*, will be preferred to any other parses:

- (7) How many people_i did Kim talk about *t_i* ...

Once this processing constraint is taken into account, we can see that the introduction of *about Mary* into (6)b makes this sentence unacceptable or degraded for reasons orthogonal to intervention effects. Thus, any experiment on intervention that uses sentences like (6)a,b will be confounded by Active-Filler effects. Experimentation on this phenomenon is not warranted unless the full range of relevant effects are taken into consideration.

Although both the examples discussed in this section are from English, it is easy to imagine that similar issues may arise with respect to any language—indeed, these types of effects are likely to be even murkier when encountered in languages less well understood than English. So, these remarks are intended to be a cautionary note about the extent of preparation that is necessary before embarking on an experimental adventure in the field. If one does not have a clear hypothesis to test—or does not strike the right balance between language structure and the use of language in real time—an experiment will have to wait.

Let's assume that the question “Am I ready to conduct an experiment?” has been resolved, and an experiment is truly warranted. At this stage, it is useful to consider whether a lesser-known language that a fieldworker has access to is the best option for the experiment, or whether a more easily-accessible language could be examined first. For instance, a number of languages combine prenominal and postnominal modifiers; examples include Mosestén (Sakel 2002), Yimas (Foley 1991), Tongan (Churchward 1953), and French. If one sought to experimentally investigate differences in the

processing and structure of pre- and postnominal modifiers, it would make sense to start with French. Starting the experimental process with easy-to-access speakers in a forgiving laboratory setting allows for many of the kinks to be worked out before the methodologies and findings of the preliminary study are extrapolated to more exotic languages.

3. Is it worth the trouble?

Is it reasonable to run experiments in a fieldwork setting? My answer to this question is a cautious yes, and in this section, I will examine several examples where fieldwork experimentation is warranted. The list is by no means exhaustive, and hopefully with time it will grow. For the sake of exposition, I pass over the practical details of experimentation in the examples I discuss in this section; I will turn to the latter in section 4.

3.1. Phenomena over languages

As a general rule, when planning experimental work in a fieldwork setting, it is important to start with *phenomena* rather than *languages*. Say you wish to investigate the processing of *wh*-questions derived by A-bar movement in a novel language, language L. If your ultimate goal is to compare the processing of such *wh*-questions in L versus English (where we are confident that *wh*-question formation involves A-bar movement), then you would first need to ascertain that *wh*-questions in L are also formed via A-bar movement. Choosing a language where the *wh*-word is in the initial position is not informative enough; the *wh*-word may be in the initial position because it is a predicate of a pseudo-cleft with a silent copula, as shown in (8)b. This is expected in head-initial languages in particular (Potsdam and Polinsky 2011), so if the lesser-studied language in question tends towards head-initiality, special care must be taken to tell these two derivations apart.

- | | | | |
|-----|----|---|----------------|
| (8) | a. | Wh-word _i S t _i V X | A-bar movement |
| | b. | [PredP Wh-word] [DP [CP S V X]] | pseudo-cleft |

While finding exotic languages with initial *wh*-words is not difficult, rushing to experimentation before conducting a syntactic analysis of these candidate languages is premature. This is where the role of the fieldworker becomes indispensable; someone familiar with language L will know the details of its structure and be able to determine if it meets the desiderata for an experiment.

Assuming this general approach is adopted, what kinds of phenomena warrant experimental investigation in languages with remote access? The most obvious answer deals with phenomena that are not available in better-studied languages. In what follows, I will review two such phenomena that have already received attention at the intersection of fieldwork and experimental syntax: alignment and word order.

3.2. Examples of convergence

3.2.1. Alignment

English, German, Dutch, Italian, Spanish, Korean, and Japanese—all languages that have been studied extensively using the experimental-syntax paradigm—have been shown to share a number of processing constraints. Among these is the subject preference advantage (SPA): the observation that subject gaps (for example, in relative clauses) are easier to process than object gaps. Consider the familiar minimal pair below, where the relative clause in (9)a includes a gap in subject position, and the one in (9)b has a gap in object position. The latter clause is more difficult to process, as numerous studies have shown (see Kwon et al. 2010, 2013 for overviews):

- (9) a. the senator_i [that ____i attacked the reporter] admitted the error
 b. the senator_i [that the reporter attacked ____i] admitted the error

The SPA is quite robust in all of the languages mentioned above, but the reasons for this remain unclear. Problematically, the existing data come from nominative–accusative languages, in which subjects appear in the same case, regardless of transitivity, and the marked form is the object (accusative). This covariance of grammatical function (subject vs. object) and case (nominative vs. accusative) has prevented researchers from determining which of these two factors underlies the SPA. In addition to case, alignment may be expressed via agreement: all subjects, regardless of transitivity, can be cross-referenced on the verb, whereas objects do not determine verbal agreement.

A potential workaround to this problem is to investigate the SPA in languages with morphological ergativity. Ergative languages allow for the separation of case and grammatical function, since the subject position is associated with two cases: absolutive (intransitive subjects) and ergative (transitive subjects). Compare in Niuean (Polynesian):

- (10) a. Kua koli e ekekafo. *Niuean*
 PFV dance ABS doctor
 ‘The doctor danced.’
 b. Kua lagomatai he ekekafo e faiaoga.
 PFV help ERG doctor ABS teacher
 ‘The doctor helped the teacher.’

If alignment is manifested in verbal agreement only, freestanding noun phrases may remain unmarked, whereas the form of the predicate varies depending on whether it agrees with the intransitive subject or direct object (absolutive agreement) or transitive subject (ergative agreement). Compare in Ch’ol (Mayan), where one set of affixes on the verb indexes the absolutive argument, and the other, the ergative:

- (11) a. Tyi y-il-ä-y=ety. *Ch’ol*
 PFV 3SG.ERG-see-TRANS.VERB-EPENTHESIS=2ABS
 ‘S/he saw you.’
 b. Tyi ts’äm-i-y=ety.

PFV bathe-INTRANS.VERB-EPENTHESIS=2ABS
 ‘You bathed.’ (Coon 2017: 101)

Ergative languages allow researchers to study the processing of case and/or agreement and grammatical function (that is, the syntactic position of an argument in clause structure) as independent phenomena in a way that accusative languages do not. In ergative languages, case marking does not co-vary with the subject/object distinction. If ergative languages are sensitive to differences between subjects and objects (regardless of case marking), this will provide strong and novel evidence that subjects constitute an independent concept in grammar.

Some ergative languages are consistently ergative (that is, their ergative alignment is found across all aspectual and tense forms), while others display “split ergativity”: their ergativity alignment is limited to certain aspectual or mood features (the perfective or irrealis, for instance) or to particular persons (non-pronominal expressions). See Coon and Preminger (2017) for an overview and discussion.

There is a clear need to test subject preference in consistently ergative languages, and this need has led to experimental fieldwork in Basque (Carreiras et al. 2010; Gutierrez-Mangado 2011; Laka et al. 2012), Avar (Polinsky et al. 2012; Polinsky 2016), Niuean (Longenbaugh and Polinsky 2016, 2017), Ch’ol, and Q’anjob’al (Clemens et al. 2015).³ For each language, the studies tested speaker preferences in the comprehension of subject and object gaps in relative clauses, following research on clauses like the English ones shown in (9) above. The languages listed were chosen for both conceptual and practical reasons. The conceptual reasons included: (i) the need to use consistently ergative languages where both ergative and absolutive arguments can extract without a gap (i.e., languages without syntactic ergativity); (ii) the need to compare and contrast ergative languages where the relative clause precedes the head noun (head-final languages) and follows the head noun (head-initial languages); and (iii) the need to compare languages where alignment is encoded on the nominal (via dependent-marking, i.e. case marking) versus the predicate (via head-marking, i.e. agreement) (see Nichols 1986 for the distinction). The comparisons are summarized in the table below.

	Dependent-marking (case marking)	Head-marking (agreement)
Head-initial	Niuean	Ch’ol, Q’anjob’al
Head-final	Avar, Basque	--

Table 1

Experimental paradigm for studying subject preference, morphologically ergative languages

Practical considerations for choosing these languages include: (i) the availability of pre-existing analytical work on the data in question, and (ii) at least in the Avar and Mayan

³ Q’anjob’al is partially syntactically ergative, but the study cited here examined the domain where only morphological ergativity holds.

studies, partnerships with native-speaker linguists. The native-speaker linguists played a crucial role in norming the stimuli for the experiments, finding participants, and offering explanations to community members about why the study was useful (I will return to this in section 4).

Studies of Basque and Avar were carried out using the self-paced reading paradigm, which has also been used in many MYALS-based studies of relative clauses. The results generally upheld the SPA, but there were some complications. First, because Basque and Avar speakers primarily use their languages orally, the average reading times for these speakers were about three times longer than in languages with a well-established reading tradition, such as German or Japanese. That led the researchers to look for alternative testing methods, including sentence-picture matching (SPM) (Bishop 2003). In this task, which can be used both offline and online, participants are presented with a series of pictures (usually two or four), listen to one sentence, and then have to decide which picture goes with the sentence. The results of this task were independently compared against the results from a self-paced reading task in a language where test participants were used to reading on a daily basis (Clemens et al. 2015). The results were comparable, which confirmed the utility of the SPM task.

The results of the SPM experiment (in Avar, Niuean, Ch'ol, and Q'anjob'al) upheld the SPA, and therefore offered novel evidence in favor of the privileged status of subjects regardless of alignment. The SPA was particularly apparent in head-marking languages, where it was essentially the main result. In dependent-marking languages, there was an additional cueing effect that followed from morphological informativity: as the marked case, the ergative in the relative clause served as the cue that an absolutive argument needed to be projected. This ergative cueing effect was observed in both prenominal (Basque, Avar) and postnominal (Niuean) relative clauses. As expected, when the relative clause contained only an absolutive argument, no cueing effects were found, since the absolutive can serve either as the subject of an intransitive or the object of a transitive clause. On the nominative–accusative side, cueing effects are observed in the presence of the accusative, which only reinforces the SPA.

These fieldwork experiments lead to new predictions: in an ergative language with prenominal relative clauses and head-marking (the missing cell in Table 1, above), the SPA should be particularly apparent. As far as I am aware, few languages fit this description. Among them is Abkhaz, a language spoken in the Northwest Caucasus (Hewitt 1979: 35–45), which may offer an excellent test case for connecting fieldwork and experimentation in the future.

Once the basic work on consistently ergative languages had been done (to establish basic patterns), it was reasonable to move to split ergative languages such as Hindi and Georgian. This work is currently being undertaken. Based on preliminary results obtained by Foley and Wagers (2017), Georgian also exhibits SPA effects.

The long-distance dependencies discussed in this section have long been at the center of attention in experimental syntax. However, alignment differences go well beyond the SPA, and future experimental work on ergative languages can include explorations into island constraints, licensing or prediction of case forms, agreement attraction, and other phenomena (see Longenbaugh and Polinsky 2017 for a discussion of several directions in experimental research on ergativity).

3.2.2. Word order

Most experimental syntactic research has been based on languages with the basic word orders SVO or SOV, which are by far the most common orders cross-linguistically. Many SVO and SOV languages, including the ones studied experimentally, allow *subject-before-object* (SO) orders and *object-before-subject* (OS) orders; for these languages, the main experimental result is that OS orders impose a greater processing load (Bader and Meng 1999; Kaiser and Trueswell 2004; Mazuka et al. 2002; Sekerina 1997; Kwon et al. 2006; Tamaoka et al. 2005). However, in the languages investigated, OS may be derived by scrambling.⁴ If we assume that scrambling is not base-generation, we can predict that the OS order should be syntactically more complex (as shown below). OS order is often less frequent, as in Korean and Japanese (Kwon et al. 2006). These considerations point to the SO order as the starting point.

- | | | | |
|------|----|---------------------|---|
| (12) | a. | $O_i S t_i V$ | SCRAMBLING, SOV LANGUAGE |
| | b. | $O_i S V t_i$ | SCRAMBLING, SVO LANGUAGE |
| | c. | $O_i V_k S t_k t_i$ | SCRAMBLING AND VERB MOVEMENT,
SVO LANGUAGE |

In the psycholinguistic literature, two general theoretical explanations for the SO preference have been spelled out (Koizumi et al. 2014). In one view, grammatical factors of individual languages (such as syntactic complexity) are the main driving force behind word order preferences; these preferences are therefore domain-specific. If this account is correct, SO is not a universally preferred order. In the alternative view, word order preferences follow from universal human cognitive features; if that is the case, SO word order should be preferred regardless of the basic word order of any individual language (Bornkessel-Schlesewsky and Schlewsky 2009; Tanaka et al. 2011). These views both correctly predict that SO word order is preferred in SO languages: SVO, SOV, and VSO. The deciding group are OS languages, of which VOS is the only reliable type attested cross-linguistically. To create a test environment, researchers must compare the OS order, which can be assumed to be basic, and the SO order. Within this comparison, the domain-specific approach predicts that OS order should be easier, and SO order should be associated with a higher processing burden. The universal approach predicts the opposite.

Koizumi et al. (2014) and Yasunaga et al. (2015) conducted several studies on word order processing preferences for the Mayan language Kaqchikel, in which the basic order is consistently VOS (regardless of dialect). These studies compared and contrasted SVO (SO) and VOS (OS) orders. These two orders differ along two dimensions: structural complexity (SVO is derived from VOS via scrambling, (13))⁵ and frequency

⁴ Although there are a number of analyses that uphold the base-generation approach to scrambling (Fanselow 2001; Neeleman and van de Koot 2008), it is not obvious that scrambling is syntactically more complex. Clear, incontrovertible evidence for scrambling as movement is surprisingly hard to come by, at least in Germanic.

⁵ While the status of the left-hand subject as a topicalized constituent is relatively clear, the nature of this topicalization (scrambling vs base-generation) and the landing site of topicalization are subject to some debate. The authors do not commit to a particular landing site category; they denote the base position theoretically as a gap.

relying on extensive fieldwork, and modifying the experimental methodology in an ecologically sound way—in particular, by using auditory presentation. Further successful fieldwork experiments could be built on this model; for example, it would be valuable to directly compare VSO and VOS in Fijian, where they are equally possible (Dixon 1988; Aranovich 2013) or Tagalog, where VSO and VOS are both observed in Agent Voice (Kroeger 1993).

Another important word order consideration concerns *incrementality in production*. It is generally assumed that language users do not plan entire utterances before beginning to speak. Instead, as in parsing, planning unfolds step by step (Levelt 1989, Ferreira and Swets 2002). While incremental planning is itself uncontroversial, it is less clear whether structure or lexicon serves as the *starting point* in production planning. If structure is the starting point, the speaker will generate the syntactic skeleton of her utterance and then add lexical content in an incremental manner. If lexical encoding takes precedence, the reverse is true. English production seems to support the lexical model of encoding, but data from languages with more flexible word order support the structural model—or a combination of the two (Hwang and Kaiser 2015; Norcliffe et al. 2015b).

Until recently, all of the work in this subarea has focused on a small set of subject-initial languages, but lately verb-initial languages have been added to the data pool. These languages “offer an interesting test case for studying the effects of grammar on sentence formulation. In order to select a suitable sentence-initial verb, information about the relational structure of the event presumably must be planned early, possibly earlier than in subject-initial languages” (Norcliffe et al. 2015b: 1020). Recent experimental data from two such languages, Tselal (Norcliffe et al. 2015a) and Tagalog (Sauppe et al. 2013),⁸ suggest that the early position of the verb changes the order of encoding operations: relational information encoded in the transitive verb receives priority over either character associated with that verb. In both languages, the verbal morphology carries important information concerning the event and its participants, and this may give priority to grammatical structures. If these results are on the right track, they offer additional support for Hwang and Kaiser’s (2015) proposal that production is guided by both structure and lexical access, but that the relationship between these two components varies by language and can only be predicted based on a careful examination of each language’s grammatical system.

4. Practical issues

4.1. Starting point

Since any experimental work conducted in a fieldwork setting is experimental work on *language*, it needs to start with a specific hypothesis and a rationale for choosing one language over others—the previous section illustrated some of these rationales. Fishing-expedition experiments do not work well in a lab setting, and the situation in the field is no different.

⁸ In both studies, the experimental work relied on a detailed syntactic analysis of the language in question based on primary data.

Furthermore, when starting experimental fieldwork on a new language, it is advisable to begin by replicating experimental methodologies already used on more familiar languages. Say your ultimate aim is to use a visual world paradigm to explore the anaphoric use of classifiers in a lesser-studied language. You should begin with a simpler experiment to pave the way. For instance, you might replicate a study on Mandarin conducted by Huettig et al. (2010), in which speakers heard a noun and looked at pictures of objects that shared or did not share that classifier. (In the Mandarin study, the main finding was that classifier distinctions influence eye gaze, but only when classifiers are overtly present in the speech stream.) By replicating an existing study, the researcher can rely on an established experimental paradigm (which can be modified as needed) and minimize unknowns. Once the replication experiment has been done, a novel study is easier to conduct. (Note that most of the experiments described in section 3 replicated the experimental design of work conducted on better-studied languages.)

4.2. Participants

Fieldwork is not always about pitching a tent in a remote location, sharing exotic food with your consultants, and carrying coffers of recording equipment up a steep hill. As Claire Bowerman put it, one does not have to “be Indiana Jones in order to be a real linguist or fieldworker” (2008: 14). The main components of fieldwork are: (i) that the language has not been well described before, and therefore, (ii) data collection will need to be undertaken before completing an experiment. It is possible to find languages that fit this bill within the cities that house research universities. If there are enough speakers of a given language in a city, experiments can even be done in a familiar lab.⁹ The Endangered Language Alliance of New York City is an outstanding example of work on the linguistic diversity of a large urban area.

Still, bringing speakers of lesser-known languages into the researcher’s experimental setting is less common and less likely than traveling to those speakers. The trip can take a few hours or it can take several days. No matter how close or far the speakers are, you’ll need to plan carefully. Two main aspects of the interaction with participants deserve mention: *justification of the study to the participants* and *researcher involvement in the community*.

In experimental work in the lab, the former component barely plays a role. MYALSs rarely ask questions about the experiments they participate in. They are accustomed to tests and test-taking, they typically don’t expect explanation of the reasoning behind an experiment, and they normally show cooperative behavior in dealing with the experimenter. In a fieldwork setting, such cooperative behavior and magnanimous indifference are an exception, not the norm. When embarking on an experiment in the field, be prepared to be greeted with curiosity, suspicion, surprise, criticism for engaging in silly activities, or some other reaction that may be hard to predict—anything but immediate acceptance. Because participants are not likely to cooperate automatically, it is important to be able to explain why you are conducting your study and possible benefits to the community. The explanation can be presented at a

⁹ An important consideration in this situation is that speakers in such a setting will most likely be bilingual, something I will return to in section 4.3 below.

community meeting, built into the prompt of the experiment, or offered in an initial conversation with participants. It is often helpful to rely on the fieldworker, who may already have ties with the local community, and on native-speaker linguists, if they are available. If not, seek out community members who are in positions of authority. Once these people approve of your project, they will be able to serve as the link between your research team and the local participants. Teachers or priests often perform this role. Remember that, in justifying the study to the community, it is important to step outside of your technical frame of mind and couch the study in more general terms. Stating that the study will allow outsiders to understand the language of the community better is a valid justification. Small communities are often pleased by outside interest in their language. However, this is not always the case, and it is never a good idea to push an experiment upon a group that is not willing to accept it.

Quite a bit of experimental fieldwork has been done with Mayan languages; I have already mentioned Koizumi et al. (2014), Yasunaga et al. (2015), Clemens et al. (2015), and Norcliffe et al. (2015a), and there is also work on Yucatec Mayan (see Butler 2011 on number marking; Skopeteas and Verhoeven 2009 on information structure; Norcliffe and Jaeger 2016 on production). The confluence of research on Mayan is not accidental, and it owes its success to at least two practical considerations. First, there is an abundance of rich primary work on Mayan, pioneered by Judith Aissen, Nora England, Clifton Pye, Barbara Pfeiler, and Roberto Zavala (see Aissen et al. 2017 for a summary volume). Second, and equally important, is the strong pattern of indigenous activism in Mayan communities (Warren 1998; Fischer and Brown 2001). Local activists tend to be interested in collaborating with linguists (and other researchers) in promoting new work on their languages—as long as that work contributes to the recognition of the local communities and cultures. This type of community engagement is very important and not always available. A similar convergence of factors has favored experimental syntax research into the Austronesian language Chamorro (see Chung and Wagers, this volume). While this combination of first-rate primary research and community involvement does not mean that experimental work on Mayan and Chamorro is carried off without a hitch, it eases the path for researchers hoping to work with native speakers in the field.

Long-term involvement of the researcher in the language community is another crucial aspect of fieldwork, and another fundamental difference between experiments with MYALSs and experiments in the field. Fly-by-night studies do not work in the fieldwork settings, and there is an expectation that both sides should benefit from the experiment. The benefits to the community may be broad—like validating faith in the community and its language or promoting cultural awareness—or may be more direct. For example, researchers often succeed in bringing informants on board with the argument, “If someone wants to learn your language they will know what its most difficult aspects are.” In the work conducted on Mayan languages, popular presentations (on the value of promoting Mayan languages, Mayan diversity, or Mayan inscriptions) are frequent and greatly appreciated by the local communities. Community feedback and desire for involvement will vary across communities, but a sense of investment in the research should always be expected.

Being involved in a community means knowing and respecting the rules of cooperative behavior in that community. Nowhere is this knowledge more important than

when considering *compensation* for participation in an experiment. The field linguist and/or native consultant should make recommendations about culturally appropriate ways to approach this issue in a given community, and it is important to follow their advice (even if it may seem counterintuitive to an outsider). It is also imperative that all participants be compensated the same way; in tight-knit communities, where people like to talk, any sign of disparity or favoritism may sink the experiment.

The *number of participants* in an experiment is a serious question that comes up in all types of experiment planning, not just experimental fieldwork. The *appropriate* number largely depends on the methodology, the task at hand, and the goals set by a researcher. On the other hand, the *feasible* number may be constrained by the context, and the potential pool of participants may not be as large as it is when working with MYALS in a research center. Unfortunately, the difficulty of recruiting participants often becomes a deterrent for researchers hoping to work on a language in a fieldwork setting. Here, I would like to offer two considerations.

First, as noted in section 1, there are experimental fields in linguistics where the number of participants is very low, but nobody contests the validity of the results. Such fields include aphasiology, brain imaging (which involves expensive fMRI techniques), some sign language research, and most research in phonetics. In these fields, researchers have learned to work around their small sample sizes by modifying their statistical analyses. A common solution is random sampling with replacement (Groves et al. 2009), but researchers may also choose to simply adjust their tests for sample size, making sure to use nonparametric tests when they do not have enough data to be assured of a normally distributed dataset. A standard workaround is to obtain a large number of data points per participant. This strategy can easily be carried out in fieldwork settings by testing each structure under examination with multiple lexicalizations and on multiple days. The experiment will take longer, but the results will be as useful and usable as those of a phonetic-recording experiment done in the lab with three speakers over the course of one day.

Second, not all “fieldwork languages” have a limited number of speakers. For languages with robust speaker populations, it is desirable to recruit *more* participants than are typically recruited in laboratory settings, since there is a greater likelihood that fieldwork participants will not be used to test-taking and may not completely follow the protocol. Unforeseen factors such as lack of vision correction or bad dentistry may cause noise within your data. It may be hard for the researcher to anticipate such situations, but it is possible to conduct an experiment in a society where enhancers like glasses and dentures are a luxury—and it may be culturally inappropriate to turn away someone with no teeth who wants to participate in a production study, or someone who can barely see but wants to take part in an SPM task. To account for such noise in the data, a good rule of thumb is to increase the number of participants by about 20 percent, as compared to numbers in a lab setting. On the flip side, in some communities, people who are asked to participate will bring friends and family members along, and you may wind up with more participants than you actually need. Again, if it is culturally inappropriate to turn these participants away, you may have to accommodate them. Some of these participants’ data will have to be discarded—a small price to pay for the collection of good data.

In a lab, MYALSs show up *on schedule*, and the idea of arriving unannounced to an experiment is equally strange to researcher and participant. But things are different in the

field. “Don’t expect people without clocks and watches to be concerned about very specific times of the day. It’s pointless arranging a meeting for 10:30 when no one has a clock. It’s much better to be flexible in your work hours. Be aware too that in some cultures an agreement to meet isn’t like making an appointment... and doesn’t necessarily obligate the person to turn up” (Bower 2008: 135). For many participants in the field, the experiment is a social event. As mentioned above, some may bring their friends or neighbors to watch or participate. These social considerations also mean that the “quiet testing room” we tend to take for granted in a lab setting may be completely alien to your fieldwork participants—and you may find yourself, quite literally, with noise in your data. No matter what, always be prepared for chaos and commotion.

Finally, those *consent forms* and *questionnaires* that our MYALSs fill in and sign without thinking twice may cause significant consternation among people who are not used to taking tests and signing documents on a daily basis. Again, these forms should be designed in consultation with people who have worked in the given cultural setting. In most fieldwork situations, including in the context of experiments, oral consent is preferable to written consent. Most ethics committees in Western universities are amenable to this option. Make sure the consent form is prepared in the language you are targeting in your experiment. If a fluent speaker is involved in your research team, that speaker can explain the consent form to each participant; if not, the consent form and experimental instructions can be recorded in advance. Questionnaires should include standard biographical information as well as information about the participant’s knowledge of different languages and literacy—an issue I address in the next section.

4.3. Participants and literacy

In the work on subject preference advantage in Mayan (Clemens et al. 2015), we found a sharp contrast between bilingual (Spanish–Ch’ol, Spanish–Q’anjob’al) and monolingual Mayan participants. The trends in the data were the same for both groups, but the SPA and other effects were stronger in the bilingual cohort as compared to the monolingual speakers. The monolingual speakers were significantly less accurate than the bilingual speakers even on the syntactically unambiguous clauses. In addition, the standard error and level of noise in the data was greater for the monolinguals than the bilinguals in each of our analyses.

We interpreted these results as an indication that the monolinguals, who had no experience with literacy, faced greater challenges in the SPM task because they lacked general skills that belong to the playbook of “cooperative research behaviors”: following instructions with less context than one receives in the “real world,” interacting with technology, interpreting abstract or hypothetical questions, and imagining unlikely situations. These behaviors develop in general educational settings, regardless of language, and may improve as people engage with literacy on a daily basis.

Additional support for our hypothesis came from the monolingual participants’ performance on longer sentences. We found that in the monolingual cohort alone, our experimental results became less accurate as the auditory stimuli became longer. Although this lower performance on longer stimuli was observed in both bilinguals and monolinguals, it was again greater among the monolinguals. A longer stimulus imposes a greater memory load, and there is independent evidence that educational experience

correlates with working memory capacity (Gathercole et al. 2004). This variance would have been negligible in a population more skilled at test-taking, but it played a negative role in our pool, particularly with Mayan-speaking monolinguals.

These observations concerning test-taking skills and the ability to engage in metalinguistic deliberations are not unique to exotic languages. Existing work on gradience in English judgments shows that such gradience is relativized to the participants' educational levels, again offering support for the correlation between general literacy and cooperative research behavior. Subjects with higher levels of formal education produce cleaner data in an elicitation or experimental setting (cf. Dąbrowska 1997, 2012; Street and Dąbrowska 2010). All told, experimenters in the field need to be prepared for difficulties in test-taking that are not directly related to the participants' competence.

Lastly, it is worth touching on a perennial question raised by our bilingual/monolingual discrepancy in the Mayan study: all things being equal, which group of speakers should researchers rely on, monolingual or bilingual? Which participants should we test? Although there are opposing views on this topic (see, e.g., Vaux and Cooper 1999; Bower 2008), I suggest that, in an ideal world, it is a good idea to test both groups, while remembering to keep track of their literacy, education, and multilingual experience, as in Clemens et al. (2015) in Mayan.

4.4. Materials

Planning an experiment on a well-studied language, with MYALSs as participants, takes a long time, and is usually more labor-intensive than running the experiment itself. With a lesser-known language, that preparation time increases even more. Plan to double your preparation time before a fieldwork experiment; you may find that there are more unknowns, more people who need to be involved, and more confounds to be discovered along the way.

What makes preparation for fieldwork experiments so complex? For one thing, most standard experiments rely on existing dictionary and corpus data to determine the frequency of items or constructions, establish plausibility conditions, and choose between alternative stimuli. But many lesser-known languages lack dictionaries, large annotated and tagged corpora, or even a decent collection of texts. As a result, this stage in preparation for a fieldwork experiment may become an experiment in its own right: the fieldworker will need to collect a corpus, analyze it, and use that new data to construct experimental stimuli. At the very least, a fieldworker could collect narratives in the given language from several consultants—using typical prompts such as the Frog story (Mayer 1969), the Pear story (Chafe 1980), Totem Field Storyboards (which are designed to elicit a body of tokens of specifically targeted constructions in fieldwork),¹⁰ or traditional folklore stories—or pre-test some participants on sentence completion tasks in order to build up a pre-normed collection of materials.¹¹ For example, for an experiment on

¹⁰ <http://totemfieldstoryboards.org/>

¹¹ The advantage of Frog and Pear stories is in that these stories have already been used to collect data from a wide variety of speakers and languages (cf. Berman and Slobin

adjunct islands, the fieldworker may get speakers to provide continuations of sentences like “My neighbor was happy when...” and “The plants bloomed later because...” Collecting data like this gives the researcher the necessary set of adjuncts from which to build a set of experimental stimuli on *wh*-questions.

Another bottleneck in the creation of materials concerns the norming of the experimental data. With larger, better-known languages, this task is often done online (and quite efficiently), but that option may not be available for lesser-known languages. Again, an extra field trip may be needed to norm the stimuli. In preparation for the work on Niuean reported in Longenbaugh and Polinsky (2016, 2017), we took two trips to Auckland to work with Niuean speakers on creating stimuli, norming them, and running a pilot; only after that, on our third trip, were we able to conduct the experiment.

As we progressed in our work on the Niuean project, we made certain that the syntactic analysis of Niuean relative clauses was clear to us. This experience underscores another critical requirement in combining fieldwork and experimentation: you cannot analyze the language and run an experiment at the same time. The analysis has to come first, no matter how long it takes. Recall Yasunaga et al.’s (2015) work on Kaqchikel, discussed in section 3.2.2; the researchers’ analysis of Kaqchikel SVO presupposed the topicalization of the subject argument.¹²

Once the materials for the anticipated experiment are assembled, it is desirable to conduct a pilot experiment with two or three language consultants and to collect their opinions and recommendations on the stimuli and time course of the experiment. Sometimes, even though your materials may be perfectly grammatical and well-formed, the pilot participants identify finer points that would not be apparent to any of the outsiders, including the fieldwork linguist. For example, in a subsistence culture, it may be important to specify what kind of fruit, vegetable, or animal is mentioned in the stimuli: just calling it a *mango* or a *goat* may not be enough. It may be culturally appropriate to use proper names in some societies but not in others, and it may be necessary to identify if the transfer of an object from one person to another is permanent or temporary. Sometimes, just an accidental resemblance between a person in the set of visual stimuli and a member of the local community may become a source of confusion, discomfort, or amusement.

Participants in the pilot study cannot be the same consultants that helped with the initial stimuli construction, and they cannot participate in the full experiment later on, so it is important to be judicious in choosing the participants for your pilot (of course, a lot will depend on which speakers, and how many, are available).

1994; Chafe 1980). The resulting data are highly comparable as they are based on the same plot. Yet these stories are culture-specific, and they may not work in a new fieldwork setting without significant modification.

¹² It is fine to have two (or more) alternative hypotheses regarding the structure in question; sometimes primary data may not distinguish the two well enough, and the experiment can come to the rescue. Regardless, the consequences of each analysis must be spelled out.

While it is important to tailor your stimuli to your particular experiment, it is also useful to remember the many existing materials. Stimuli from well-known languages can be combined with fieldwork stimuli (for example, Benjamin Bruening’s Scope Fieldwork Project¹³ or the MPI Language and Cognition field materials¹⁴) and then subjected to further selection.

4.5. Methods

There is no need to invent new methodologies in a fieldwork setting; what works well in a lab setting should also work in the field. The two main desiderata are (a) a reliance on an auditory rather than visual presentation (because of the likelihood that lesser-known languages will exist primarily in a spoken medium, with low literacy), and (b) an expectation that the testing environment may be less “clean” and more disrupted than in the lab (in terms of both the environment—too hot, too cold, too dark—and the risk of non-participants standing around watching the experiment or walking into the testing room talking; for more on this issue, see section 4.2). There may be no single place where the experiment can be conducted; instead, the research team will have to move their equipment from one participant’s home to another. Even an experienced fieldworker may not anticipate the numerous practical issues that come up in testing, and this may create a need for multiple field trips—adding more time to the project.

It is ideal to start testing with simple behavioral data; such tests are known as “paper-and-pencil tasks” in studies of MYALSs, but there should be no paper or pencil in the fieldwork version (see den Dikken et al. 2007 for a discussion). Instead, the data can be recorded then analyzed. If this stage is productive and there is justification for doing something more elaborate, it is reasonable to follow up with an eye-tracking experiment and then a neuro-imaging experiment. Eye-trackers, and even ERP machines, are becoming ever more portable, but before jumping on a plane or boat with the latest EyeLink or Brain Products amplifier, it is worth asking the questions raised in section 2: Is the experiment warranted? What can language L deliver that cannot be obtained by studying a different language? Can a simpler methodology be used to answer the same questions?

4.6. Language endangerment and experimental work

Many lesser-studied languages are also endangered languages. In these cases, often the only speakers left are so-called *semi-speakers*, also known as *passive (recessive) bilinguals* or lower-proficiency *heritage speakers*. The special status of these speakers in fieldwork was first raised in a seminal paper by Hans-Jürgen Sasse. Sasse observed that differentiating native grammars “from the ... situation of language decay is essential for the evaluation of data elicited from last generation speakers in a language death

¹³ <http://udel.edu/~bruening/scopeproject/scopeproject.html>

¹⁴ <http://fieldmanuals.mpi.nl/volumes/2001/>

situation... How reliable is the speech of the last speakers [of a given community] and how much does it reveal of the original structure?" (Sasse 1992: 76). Semi-speakers present unique challenges in the fieldwork context:

If you are working on a highly endangered language, conversation data might be very difficult to obtain. People might not speak the language on a daily basis, or they might feel uncomfortable about speaking spontaneously while being recorded. (Bowerman 2008:122)

Beyond the sociocultural context that accompanies endangered languages, there may also be issues with production and comprehension. In terms of production, endangered language data is likely to include a higher-than-typical occurrence of (i) long pauses, due to lexical-access problems; (ii) disfluencies or retractions; (iii) multiple redundancies and repetitions; and (iv) short segments, with few, if any, embedded structures. It is also reasonable to expect variation in production across and within speakers due to the speakers' uncertainty about some forms—a side effect of a small or fragmented speech community with reduced communication in the target language. Heritage morphology is typically rife with overmarking and overregularization, with few, if any, null pronominals (see Benmamoun et al. 2013 for details of heritage language production).

In comprehension, the *yes-bias*, which is the tendency to over-accept questionable data while being reluctant to reject ungrammatical or infelicitous language material, appears to be one of the strongest telltale signs of heritage language status (Orfitelli and Polinsky 2017; Polinsky 2018). In manifestations of the *yes-bias*, heritage speakers generally give higher ratings to well-formed and felicitous segments, but are loath to reject ill-formed or infelicitous structures because of their own uncertainty. When such speakers do reject a particular linguistic form, this can be taken as a sign of solid, strong judgment. Yet when the much-coveted star on a linguistic example does not materialize, its absence may not provide clear information; the example may actually be correct, or the speaker may be so uncertain that she cannot commit to a decision.

If the language you wish to study has few remaining speakers and those speakers show signs of being semi-speakers or heritage speakers, should you still go ahead with your experiment? The answer depends on the hypothesis and questions underlying your research. Investigations of such speakers can provide useful new data on heritage language structures. If you do choose to approach the lesser-studied language as a heritage language, it may be valuable to compare its behavior to other heritage languages for which a baseline dialect is available—for example, Heritage Spanish or Heritage Korean. Existing experimental work on heritage languages in the field includes behavioral studies on Heritage Inuttitut (Sherkina-Lieber 2011, 2015; Sherkina-Lieber et al. 2011).

Conclusions

This chapter has surveyed the main conceptual and practical aspects of experimental work in the field involving lesser-known languages. Connections between experimental work and fieldwork are bound to grow, given the increased interest in documenting lesser-studied languages, the progress in analytical tools and diagnostics in theoretical

syntax, and the rapid development of experimental approaches to sentence structure that are becoming more integrated with syntactic theory.

I have argued that there are not any insurmountable differences between experimental syntax and fieldwork per se. Both fields work with massive amounts of data, both rely on hypothesis testing, and both constantly refine and update their techniques. It is probably more apparent with experimental work than fieldwork, but great strides in both disciplines have been made since the 1980s in terms of both sophistication and depth of analysis. Although the two disciplines may use different tools and different vocabularies, they address the same fundamental issues. Furthermore, while lab-based experimental syntax may seem the more glamorous practice these days (after all, it is a relatively new field, and novelty always attracts attention), applying the experimental approach to lesser-known languages in a fieldwork setting is an equally challenging—and highly rewarding—pursuit.

Perhaps the most crucial takeaway from this discussion is the ideal composition of an experimental fieldwork team: I have argued, and I firmly believe, that the most effective way of combining experimentation with fieldwork is by building research teams where different members (the fieldworker, the experimentalist) bring different skillsets and areas of expertise, but share common goals, such as a willingness to rely on each other's strengths. Engagement of native speakers as members of the team, or major stakeholders in the ongoing project, is also an important ingredient of successful experimental work in the field.

Acknowledgments

This work was completed during my visit at the Hungarian Academy of Sciences under the Distinguished Guest Scientist Program in 2017. I am grateful to Ryan Bochnak, Marcel den Dikken, Grant Goodall, Jon Sprouse, and two anonymous reviewers for their helpful comments on the earlier draft of this chapter.

References

- Aissen, N. C. England, and R. Zavala Maldonado (eds.) *Mayan languages*. London: Routledge.
- Aranovich, R. 2013. Transitivity and polysynthesis in Fijian. *Language* 89: 465–500.
- Arunachalam, S., and A. Kothari. 2011. An experimental study of Hindi and English perfective interpretation. *Journal of South Asian Linguistics* 4: 27–42.
- Bader, M., and M. Meng. 1999. Subject-object ambiguities in German embedded clauses: An across-the-board comparison. *Journal of Psycholinguistic Research* 28: 121–143.
- Benmamoun, E., S. Montrul, and M. Polinsky. 2013. Heritage languages and their speakers: Opportunities and challenges for Linguistics. *Theoretical Linguistics* 39: 129–181.
- Bishop, D. V. M. 2003. *Test for reception of grammar, Version 2 (TROG-2)*. London: The Psychological Corporation.
- Bochnak, R., and L. Matthewson. 2015. Introduction. In R. Bochnak and L. Matthewson (eds.) *Methodologies in semantic fieldwork*, 1–12. Oxford: Oxford University Press.

- Bohnmeyer, J. 2015. A practical epistemology for semantic elicitation in the field and elsewhere. In R. Bochnak and L. Matthewson (eds.) *Methodologies in semantic fieldwork*, 13-46. Oxford: Oxford University Press.
- Bolinger, D. 1987. The remarkable double IS. *English Today* 9: 39-40.
- Bornkessel-Schlesewsky, I., and M. Schlewsky. 2009. *Processing syntax and morphology: A neurocognitive perspective*. Oxford: Oxford University Press.
- Bowern, C. 2008. *Linguistic fieldwork. A practical guide*. New York: Palgrave Macmillan.
- Bruening, B. 2014. Defects of Defective Intervention. *Linguistic Inquiry* 45, 707—719.
- Burke, D. M., and M. A. Shafto. 2004. Aging and language production. *Current directions in psychological science* 13: 21–24.
- Butler, L. K. 2011. The morphosyntax and processing of number marking in Yucatec Maya. Ph.D. Dissertation, U of Arizona.
- Carreiras, M., J. A. Duñabeitia, M. Vergara, I. de la Cruz-Pavía, and I. Laka. 2010. Subject relative clauses are not universally easier to process: Evidence from Basque. *Cognition* 115: 79-92.
- Chafe, W. (ed.). 1980. *The pear stories: Cognitive, cultural and linguistic aspects of narrative production*. Norwood: Ablex.
- Chelliah, S. L. and W. J. de Reuse. 2011. Definition and goals of descriptive linguistic fieldwork. *Handbook of Descriptive Linguistic Fieldwork*, 7-31. Dordrecht: Springer.
- Churchward, C.M. 1953. *Tongan grammar*. Tonga: Vava'u Press.
- Clemens, L., J. Coon, P. Mateo Pedro, A. M. Morgan, M. Polinsky, G. Tandet, and M. Wagers. 2015. Ergativity and the complexity of extraction: A view from Mayan. *Natural Language and Linguistic Theory* 33: 417-467.
- Clifton, C., and L. Frazier. 1989. Comprehending sentences with long distance dependencies. In G. N. Carlson and M. K. Tannenhaus (eds.) *Linguistic structure in language processing*, 272-317. Dordrecht: Kluwer.
- Coon, J. 2017. Little- v^0 agreement and templatic morphology in Ch'ol. *Syntax* 20: 101-137.
- Coon, J., and O. Preminger. 2017. Split ergativity is not about ergativity. In J. Coon, D. Massam, and L. Travis (eds.) *The Oxford handbook of ergativity*, 226-252. Oxford: Oxford University Press.
- Coppock, E. and L. Staum. 2004. Origin of the English Double-is construction. Unpublished Manuscript, Stanford University.
- Coppock, E., L. Staum, J. Brenier, and L. Michaelis. 2006. ISIS: It's not a Disfluency, but How do We Know That? Paper presented at the 32nd Annual Meeting of the Berkeley Linguistic Society, Berkeley, CA.
- Curzan, A. 2012. Revisiting the reduplicative copula with corpus-based evidence. In T. Nevalainen and E. C. Traugott (eds.) *The Oxford Handbook of the History of English*, 211-221. Oxford: Oxford University Press.
- Dąbrowska, E. 1997. The LAD goes to school: A cautionary tale for nativists. *Linguistics* 35: 735-766.
- Dąbrowska, E. 2012. Different speakers, different grammars: Individual differences in native language attainment. *Linguistic Approaches to Bilingualism* 2: 219-253.

- den Dikken, M., J. Bernstein, C. Tortora, and R. Zanuttini. 2007. Data and grammar: Means and individuals. *Theoretical Linguistics* 33: 335-352.
- Dixon, R.M.W. 1988. *A grammar of Bouma Fijian*. Chicago: University of Chicago Press.
- Fanselow, G. 2001. Features, θ -roles, and free constituent order. *Linguistic Inquiry* 32: 405-437.
- Federmeier, K. D., Segal, J. B., Lombrozo, T., and M. Kutas. 2000. Brain responses to nouns, verbs and class-ambiguous words in context. *Brain* 123: 2552-2566.
- Featherston, S. 2007. Data in generative grammar: The stick and the carrot. *Theoretical Linguistics* 33: 269-318.
- Ferreira, F. and B. Swets. 2002. How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language* 46: 57-84.
- Fillmore, C. 1992. "Corpus linguistics" or "Computer-aided armchair linguistics". In J. Svartvik (ed.) *Directions in Corpus Linguistics. Proceedings of Nobel Symposium 82, Stockholm, 4-8 August 1991*, 35-60. Berlin: de Gruyter.
- Fischer, E.F., and R. M. Brown (eds.). 2001. *Maya cultural activism in Guatemala*. Austin: University of Texas Press.
- Foley, W. A. 1991. *The Yimas language of Papua New Guinea*. Stanford: Stanford University Press.
- Foley, S., and M. Wagers. 2017. The subject gap preference in a split-ergative language: Reading time evidence from Georgian. NELS 41.
- Gathercole, S. E., S. J. Pickering, C. Knight, and Z. Stegmann. 2004. Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology* 18: 1-16.
- Gibson, E., and E. Fedorenko. 2013. The need for quantitative methods in syntax and semantics research. *Language and Cognitive Processes* 28: 88-124.
- Gil, D. 2008. How complex are isolating languages? In M. Miestamo, K. Sinnemäki, F. Karlsson (eds.) *Language complexity: Typology, contact, change*, 109-132. Amsterdam: John Benjamins.
- Groves, R.M., F.J. Fowler, M. Couper, J. Lepkowski, E. Singer, and R. Tourangeau. 2009. *Survey methodology*. Hoboken: Wiley.
- Gutierrez-Mangado, M. J. 2011. Children's comprehension of relative clauses in an ergative language: the case of Basque. *Language Acquisition: A Journal of Developmental Linguistics* 18: 176-201.
- Hale, K. 1972a. A new perspective on American Indian linguistics, In A. Ortiz (ed.) *New perspectives on the pueblos*, 87-110. Albuquerque: University of New Mexico Press.
- Hale, K. 1972b. Some questions about anthropological linguistics: The role of native knowledge. In D. Hymes (ed.) *Reinventing anthropology*, 382-397. New York: Pantheon.
- Hale, K. 1992. Language endangerment and the human value of linguistic diversity. *Language* 68: 35-42.
- Henrich, J., Heine, S. J., and A. Norenzayan. 2010. The weirdest people in the world? *Behavioral and Brain Sciences* 33: 61-135.
- Hewitt, B. G. 1979. *Abkhaz*. Amsterdam North Holland.

- Huetting, F., Chen, J., Bowerman, M., and A. Majid. 2010. Do language-specific categories shape conceptual processing? Mandarin classifier distinctions influence eye gaze behavior, but only during linguistic processing. *Journal of Cognition and Culture* 10: 39–58.
- Hwang, H., and E. Kaiser. 2015. Accessibility effects on production vary cross-linguistically: Evidence from English and Korean. *Journal of Memory and Language* 84: 190–204
- Kaiser, E., and J.C. Trueswell. 2004. The role of discourse context in the processing of a flexible word-order language. *Cognition* 94: 113–147.
- Kemper, S., M. Thompson, and J. Marquis. 2001. Longitudinal change in language production: Effects of aging and dementia on grammatical complexity and propositional content. *Psychology and Aging* 16: 600-614.
- Kibrik, A.E. 2005. Opyt OTiPLa (filfak MGU) v izučenii maloopisannyx jazykov. In A. E. Kibrik (ed.) *Malye jazyki i tradicii: Suščestvovanie na grani*. Vypusk 1. *Lingvističeskie problemy soxranenija i dokumentacii malyx jazykov*, 53-72. Moscow: Novoe izdatel'stvo.
- Kluender, R., and S. Gieselmann. 2013. What's negative about negative islands? A re-evaluation of extraction from weak island contexts. In J. Sprouse and N. Hornstein (eds.), *Experimental syntax and island effects*. Cambridge: Cambridge University Press.
- Koizumi, M., Yasugi, Y., Tamaoka, K., Kiyama, S., Kim, J., Ajsivinac Sian, J. E., and García Matzar, P. O. 2014. On the (non-)universality of the preference for subject-object word order in sentence comprehension: A sentence processing study in Kaqchikel Maya. *Language* 90: 722–736.
- Kroeger, P. 1993. *Phrase structure and grammatical relations in Tagalog*. Stanford: CSLI.
- Kwon, N., M. Polinsky, and R. Kluender. 2006. Subject preference in Korean. In D. Baumer, D. Montero and M. Scanlon (eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics*, 1-14. Somerville, MA: Cascadilla Press.
- Kwon, N., Y. Lee, P. C. Gordon, R. Kluender, and M. Polinsky. 2010. Cognitive and linguistic factors affecting subject/object asymmetry: An eye-tracking study of pre-nominal relative clauses in Korean. *Language* 86: 546-582.
- Kwon, N., Kluender, R., Kutas, M., and Polinsky, M. 2013. Subject/Object processing asymmetries in Korean relative clauses: Evidence from ERP data. *Language* 89: 537-585.
- Laka, I., M. Santesteban, K. Erdocia, and A. Zawiszewski. 2012. The Basque language in the minds of native and non-native bilinguals. In P. Salaburu and X. Alberdi (eds.) *The challenge of a bilingual society in the Basque Country*. (Current Research Series No. 9.), 157-172. Reno: University of Nevada.
- Lechner, W. 2013. Diagnosing covert movement: The Duke of York and reconstruction. In L. Cheng and N. Corver (eds.) *Diagnosing syntax*, 158-189. Oxford: Oxford University Press.
- Levelt, W. J. M. 1989. *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Longenbaugh, N. and M. Polinsky. 2016. The processing of long-distance dependencies in Niuean. In H. Hsieh (ed.) *AFLA 22: The Proceedings of the 22nd Meeting of*

- the Austronesian Formal Linguistics Association*, 98-120. Canberra: The Australian National University.
- Longenbaugh, N. and M. Polinsky. 2017. Experimental approaches to ergative languages. In J. Coon, D. Massam, and L. Travis (eds.) *The Oxford handbook of ergativity*, 709-736. Oxford: Oxford University Press.
- Massam, D. 1999. *Thing is* constructions: the thing is, is what's the right analysis? *English Language and Linguistics* 3: 335-352.
- Mayer, M. 1969. *Frog, where are you?* New York: Dial Books.
- Mazuka, R., Itoh, K., and Kondo, T. 2002. Costs of scrambling in Japanese sentence processing. In M. Nakayama (ed.) *Sentence processing in East Asian languages*, 131-166. Stanford, CA: CSLI.
- McConvell, P. 1988. To be or double be? Current changes in the English copula. *Australian Journal of Linguistics* 8: 287-305.
- Neeleman, A., and H. van de Koot. 2008. Dutch scrambling and the nature of discourse templates. *The Journal of Comparative Germanic Linguistics* 11: 137-189.
- Nichols, J. 1986. Head-marking and dependent-marking grammar. *Language* 62: 56-119.
- Norcliffe, E., and F. Jaeger. 2016. Predicting head-marking variability in Yucatec Maya relative clause production. *Language and Cognition* 8: 167-205.
- Norcliffe, E., Konopka, A. E., Brown, P., and Levinson, S. C. 2015a. Word order affects the time course of sentence formulation in Tzeltal. *Language, Cognition and Neuroscience* 30: 1187-1208.
- Norcliffe, E., F. Jaeger, and A. Harris. 2015b. Cross-linguistic psycholinguistics and its critical role in theory development: Early beginnings and recent advances. *Language, Cognition and Neuroscience* 30:1009-1032.
- O'Neill, T. 2015. The domain of finiteness: Anchoring without Tense in copular amalgam sentences. Ph.D. Dissertation, CUNY Graduate Center.
- Orfitelli R., and Polinsky M. 2017. When performance masquerades as comprehension: Grammaticality judgments in non-native speakers. In M. Kopotev, O. Lyashevskaya, and A. Mustajoki (eds). *Quantitative approaches to the Russian language*, 197-214. London: Routledge.
- Phillips, C. 2010. Should we impeach armchair linguists? In S. Iwasaki, H. Hoji, P. Clancy, and S.-O. Sohn (eds.) *Japanese-Korean Linguistics* 17, 49-64. Stanford, CA: CSLI Publications.
- Polinsky, M. 2016. *Deconstructing ergativity*. Oxford: Oxford University Press.
- Polinsky, M. 2018. *Heritage languages and their speakers*. Cambridge: Cambridge University Press.
- Polinsky, M., C. Gomez Gallo, P. Graff and E. Kravtchenko. 2012. Subject preference and ergativity. *Lingua* 122: 267-277.
- Potsdam, E., and M. Polinsky. 2011. Questions and word order in Polynesian. In C. Moyse-Faurie and J. Sabel (eds.). *Topics in Oceanic morphosyntax*, 83-109. Berlin: Mouton de Gruyter.
- Preminger, O. 2014. *Agreement and its failures*. Cambridge, MA: MIT Press
- Sakel, J. 2002. *A grammar of Mosestén*. Berlin: De Gruyter.
- Sasse, H.-J. 1992. Language decay and contact-induced change: Similarities and differences. In M. Brenzinger (ed.) *Language death: Factual and Theoretical Explorations with Special Reference to East Africa*, 59-79. Berlin: Mouton de

Gruyter.

- Sauppe, S., Norcliffe, E., Konopka, A. E., Van Valin Jr., R. D., and Levinson, S. C. 2013. Dependencies first: Eye tracking evidence from sentence production in Tagalog. In M. Knauff, M. Pauen, N. Sebanz, and I. Wachsmuth (eds.) *Proceedings of the 35th Annual Meeting of the Cognitive Science Society (CogSci 2013)*, 1265-1270. Austin, TX: Cognitive Science Society.
- Sekerina, I.A. 1997. The syntax and processing of scrambled constructions in Russian. Ph. D. dissertation. City University of New York.
- Sherkina-Lieber, M. 2011. Knowledge of Labrador Inuttit functional morphology by receptive bilinguals. Ph. D. dissertation, University of Toronto.
- Sherkina-Lieber, M. 2015. Tense, aspect, and agreement in heritage Labrador Inuttit. Do receptive bilinguals understand functional morphology? *Linguistic Approaches to Bilingualism* 5, 30-61.
- Sherkina-Lieber, M., A. T. Pérez-Leroux, and A. Johns. 2011. Grammar without speech production: The case of Labrador Inuttit heritage receptive bilinguals. *Bilingualism: Language and Cognition* 14: 301–317.
- Skopeteas, S. and E. Verhoeven. 2009. Postverbal argument order in Yucatec Maya. *STUF - Sprachtypologie und Universalienforschung*, 58: 347-373.
- Sprouse, J. and D. Almeida. 2012. Assessing the reliability of textbook data in syntax: Adger's Core Syntax. *Journal of Linguistics* 48: 609-652.
- Sprouse, J. and D. Almeida. 2017. Design sensitivity and statistical power in acceptability judgment experiments. *Glossa* 2(1), 14.
<http://doi.org/10.5334/gjgl.236>.
- Street, J. and E. Dąbrowska. 2010. More individual differences in language attainment: How much do adult native speakers of English know about passives and quantifiers?. *Lingua* 120: 2080-2094.
- Tamaoka, K., Sakai, H., Kawahara, J., Miyaoka, Y., Lim, H., and M. Koizumi. 2005. Priority information used for the processing of Japanese sentences: Thematic roles, case particles or grammatical functions? *Journal of Psycholinguistic Research* 34: 281–332.
- Tanaka, M. N., Branigan, H. P., McLean, J. F., and Pickering, M. J. 2011. Conceptual influences on word order and voice in sentence production: Evidence from Japanese. *Journal of Memory and Language* 65: 318–330.
- Vaux, B., and J. Cooper. 1999. *Introduction to linguistic field methods*. Munich: Lincom Europa.
- Vaxtin, N.B, and E. V. Golovko. 2005. Isčezajuščie jazyki i zadači lingvistov-severovedov. In A. E. Kibrik (ed.) *Malye jazyki i tradicii: Suščestvovanie na grani*. Vypusk 1. *Lingvističeskie problemy soxranenija i dokumentacii malyx jazykov*, 40-52. Moscow: Novoe izdatel'stvo.
- Warren, K. 1998. *Indigenous movements and their critics: Pan-Maya activism in Guatemala*. Princeton: Princeton University Press.
- Yasunaga, D., M. Yano, Y. Yasugi, and M. Koizumi. 2015. Is the subject-before-object preference universal? An event-related potential study in the Kaqchikel Mayan language. *Language, Cognition and Neuroscience* 30: 1209-1229.