5

Using questionnaires as a tool for comparative linguistic field research: Two case studies on Javanese

Jozina Vander Klok
University of Oslo

Thomas J. Conners
University of Maryland

In this paper, we discuss how written questionnaires for targeted constructions can be a beneficial tool for comparative linguistic field research through two case studies on Javanese (Austronesian; Indonesia). The first case study is based on a questionnaire designed to elicit how a language or a dialect expresses the semantic meaning of modality (Vander Klok 2014); we show how it can be implemented in three different ways for comparative linguistic field research. The second case study is based on a questionnaire which investigates the morphosyntax of polar questions across four Javanese dialects; we show how items can be designed to maximize direct comparison of features while still allowing for possible lexical, phonological, or morphosyntactic variation. Based on these two studies, we also address methodological challenges that arise in using questionnaires in comparative linguistic field research and offer best practices to overcome these challenges.

Key words: Fieldwork methodology; Questionnaires; Comparative linguistic studies; Javanese

1.   Introduction

The field linguist today has a wide variety of tools to draw on in assembling a toolkit; these tools include participatory observation, recordings of narratives or natural conversation, interviews, focus groups, elicitation, storyboards, questionnaires, and their variants.[1] Depending on the subfield of linguistics or type of research questions being investigated, the field linguist might also include various physiological experimental methods such as eye tracking, ERPs (event-related potentials) or ultrasound.[2] Given the relative affordability and portability of high-quality equipment to implement all of these tools, the field linguist is then faced with the selection of the most appropriate tool(s) for data collection at hand based on the research question, the research location and timing. Various factors will play a role in this decision: what the research project is targeting; how many speakers it is possible to work with; the time-frame of the project, among others.

The focus of this paper is to address how and when *written questionnaires for targeted constructions* can be the right tool, at the right place and the right time in a fieldwork setting. The term 'questionnaires' in this paper refers to a set of questions designed to elicit a constrained set of answers from multiple respondents. We use 'targeted constructions' in the sense of Burton and Matthewson (2015) whereby the written questionnaire is designed to target a specific linguistic phenomenon or a set of linguistic phenomena. This is further discussed in Section 2 in the context of situating the use of questionnaires for targeted constructions in comparative linguistic field research within a typology of written questionnaires.

We propose that questionnaires can be a particularly useful tool when the research engages with comparative linguistic fieldwork. We investigate this issue in Section 3 through two case studies on Javanese, an Austronesian language spoken mainly in Java, Indonesia, known for a high degree of dialectal variation (e.g., Hatley 1984). The first case study concerns a questionnaire designed to elicit how a language or a dialect expresses the semantic meaning of modality (Vander Klok 2014); we show how it can be implemented in three different ways. The second case study is based on a questionnaire which

---

[1] This list is not exhaustive, nor do we assume that these tools are necessarily discrete. For instance, written questionnaires can be used orally as elicitation. Further, some methods such as storyboards, combine the use of narration and elicitation (see Burton and Matthewson 2015 for details on storyboard methodology).

[2] An excellent introduction to experimental methods for linguists is Arunachalam (2013) as well as articles in Podesva & Sharma (2013). While they do not explicitly address using these methods in a fieldwork setting, one can see how they can be adapted. See also Krifka (2011) for an overview of experimental methods from a semantics perspective and Gick (2004) for ultrasound use for phonetic studies.

investigates the morphosyntax of polar questions in different Javanese dialects (Vander Klok, Ahsanah & Sayekti 2017).[3]

Conducting fieldwork with any tool is never without challenges. In Section 4, we address some challenges with conducting comparative linguistic fieldwork, including maximizing direct comparison of features while still allowing for lexical, phonological, or morphosyntactic variability across dialects; determining which language or dialect is best included at different stages of the questionnaire; and recruiting participants of said dialect or variety. In Section 5, we discuss ways in which field linguists can prepare for and overcome the above challenges for comparative linguistic field research, covering the design, metadata, language/variety selection, instructions, and implementation of questionnaires for such purposes. This section will be most useful for either new field linguists or those working specifically on comparative linguistic studies. However, since these suggestions for best practices are broadly applicable to most fieldwork settings, we also refer the reader to other current resources. Section 6 concludes.

2. The place of written questionnaires for targeted constructions within a typology of questionnaires

Questionnaires represent an important tool utilized not only by field linguists when collecting primary linguistic data from a range of speakers (Bowern 2015:92), but also by researchers across a variety of linguistic subdisciplines (cf. Dollinger 2015:12). The aim of this section is to underline the inherent advantages and disadvantages of questionnaires (Section 2.1) and then contextualize how *written questionnaires for targeted constructions* fit into a typology of written questionnaires (Section 2.2).

A brief note is first necessary on the use of the terms 'survey' vs. 'questionnaire'. While some researchers, such as Dollinger (2015), use *questionnaire* exclusively, others such as Shilling (2013) use the term *survey* exclusively. We suggest that it is useful to make a distinction between written *surveys* and *questionnaires*. Overall, our view is that the main goal of written questionnaires is to gather linguistic data, whereas the primary goal of *surveys* is to collect information on language demographics, use, attitudes, and language backgrounds. This distinction is in line with how the term 'survey' is used by SIL (e.g., Nahhas 2007), the ILI (Indigenous Language Institute; Linn 2004), or the FPCC in Canada (First Peoples' Cultural Council; Franks & Gessner 2013). These organizations

---

[3] These case studies focus on the use of questionnaires as a tool for comparative field research; as such the results are not discussed in detail. See Vander Klok (2013b) for results of the modal questionnaire in Paciran Javanese, and the results of the questionnaire of polar questions are partially presented in Vander Klok (2017).

have developed surveys primarily on language attitudes and use for the purposes of language planning and obtaining funding. Of course, some questionnaires that focus on collecting data on linguistic phenomena might also necessarily gather information on the participants and/or their language use or attitudes; thus, what we consider to be typical 'survey' data is not excluded from questionnaires.[4]

2.1 Inherent advantages and disadvantages of questionnaires

Questionnaires have a number of inherent advantages, including an easy setup for comparative studies, the possibility of gathering data from multiple participants/respondents, and options for different types of implementation. Concerning the latter, a *written questionnaire for targeted constructions* can, for instance, be implemented as an acceptability judgment task (i.e., yes-no task, Likert scale rating task, Magnitude Estimation task, Thermometer task; see e.g., Schütze & Sprouse 2013), a fill-in-the-blank task, a translation task, or a correction task. We will discuss some of these implementations in Section 3.

One inherent disadvantage of written questionnaires is that they are written, therefore only reflecting one 'mode' of language use, and limiting the collection of data for phonetic or phonological studies. This can be solved by recording the sessions and by using the questionnaire as the basis for an oral elicitation task.[5] Other possible limitations discussed in Schilling (2013:102-103)—such as ordering effects; the fact that participants could confound grammaticality with 'correctness'; the fact that the questionnaire could create an artificial setting where respondents do not respond in the way they truly use language—are not viewed as inherent disadvantages, but rather methodological challenges. We discuss some of these in Section 4.

What we aim to show in this paper is that questionnaires can be productively used to collect data in the field for comparative linguistic research. Overall, whether the goal is typological comparison, dialectology, theoretical inquiry, or descriptive documentation, questionnaires are useful field tools, as long as care is taken in their development, implementation, and interpretation.

---

[4] We do not go into detail regarding surveys on individual language speaker metadata, but we do outline the relevant metadata we consider to be required for comparative research from a cross-dialectal perspective in Section 5.5.

[5] In sociolinguistics, questionnaires were traditionally conducted as face-to-face interviews, since for instance the *Atlas linguistique de la France by* Gilliéron & Edmont (1901-1910); see Dollinger (2015: Chapter 2) for a historical overview.

2.2   How do written questionnaires for targeted constructions fit into a typology of questionnaires?

Concerning a typology of questionnaires, Dollinger (2015:12) identifies three criteria whereby questionnaires can be classified: by type of reporting, type of information sought, and subject area, as summarized in (1). For Dollinger's research focus, the classification by subject area is based on questionnaires directly relevant to social dialectology and the use of written questionnaires.

(1)   Typology of written questionnaires based on Dollinger (2015:12)

Classification 1. Type of reporting
i.   Community-reporting

ii.   Self-reporting

Classification 2. Type of information sought
i.   Language attitudes and perceptions

ii.   Linguistic behaviour

Classification 3. Type of subject area
i.   Questions concerning language attitudes and perceptions
ii.   Questions concerning regional language variation and social language variation
iii.   Questions using acceptability judgments of grammaticality

According to this typology, although Dollinger presents each classification as distinct, it is clear that Classifications 2 and 3 overlap considerably. For instance, questionnaires concerning regional language variation and social language variation could easily involve both types of information sought in (1). Additionally, the third type within Classification 3, 'Questions using acceptability judgments of grammaticality', does not fit well within the classification by 'subject area', as acceptability judgments are an implementation type rather than a subject area.[6] Note further that under the proposed distinction between *questionnaires* and *surveys*, the subject area relating to questions concerning language attitudes and perceptions could be considered a survey.

We return to *written questionnaires for targeted constructions* as defined in this paper to see how these questionnaires fit into Dollinger's typology. *Written questionnaires for targeted constructions*, as stated above in Section 1, are a set of questions designed to elicit

---

[6] Dollinger's (2015:12) description of 'Questions using acceptability judgments of grammaticality' underlines the implementation and not the 'subject area' of questionnaires. He writes: "...originally a mainstay in generative linguistics on a binary scale, WQs [written questionnaires] have come to be used on gradient scales outside the generative domain since Bard et al.'s (1996) Magnitude Estimation Method."

a set of answers from multiple respondents based on a particular linguistic research question or hypothesis. According to Dollinger's typology in (1), using a questionnaire in this manner would potentially satisfy either sub-type of reporting under Classification 1 and would be classified under 'linguistic behavior' in terms of the type of information sought in Classification 2. However, this type of questionnaire does not appear under the 'Classification 3, type of subject area' since it is not required to be implemented using acceptability judgments, nor is it appropriate to leave this classification open. Concerning the implementation, we purposely leave this choice open, as it one of the inherent advantages of using a questionnaire.

Given the above issues, we propose a revised typology of questionnaires in (2). A new addition is 'type of implementation' as a separate classification, in line with this being an inherent advantage of written questionnaires. Further, given the considerable overlap between 'information sought' and 'subject type', we collapse these two into one criterion as 'type of information sought', now under Classification 2. We also allow for the type of self-reporting to be mediated (e.g. by a research assistant), or non-mediated (the respondent directly fills out the questionnaire without any assistant present, such as online). Finally, we acknowledge that a written questionnaire could be identified with one or more of the subtypes within each of these classifications. For example, the same questionnaire could inquire about language attitudes as well as linguistic phenomena or it could be implemented using different types of implementation (as long as the content is the same); see our first case study in Section 3.3 for an example.

(2)   Revised typology of written questionnaires

Classification 1. Type(s) of reporting
i.   Community-reporting
ii.   Self-reporting (Mediated vs. Non-mediated)

Classification 2. Type(s) of information sought
i.   Language attitudes and perceptions
ii.   Social language variation
iii.   Linguistic phenomena, including descriptive documentation

Classification 3. Type(s) of implementation (non-exhaustive list)
i.   Acceptability judgments of grammaticality or felicity
ii.   Fill-in-the-blank
iii.   Translation
iv.   Correction, etc.

Lastly, while our focus is on questionnaires, we take it that no single tool in the toolkit can satisfy every research situation and every researcher's needs. For instance, in response to Featherston's (2007) position paper that experimental methods such as acceptability

judgment tasks are necessary beyond data collected from introspection, Bornkessel-Schlesewsky & Schlesewsky (2007:331) offer additional experimental methods and write:

> *"From an empirical perspective, there cannot be 'one perfect method' for the investigation of linguistic knowledge. It is important to recognize the limitations of individual methods and to capitalize upon the insights that can be gained from their combination."*

Bowern (2015:85) also argues in her comprehensive fieldwork manual for the merits of elicitation beyond text collections:

> *"...some aspects of a language are only discoverable through elicitation—they will appear in texts so seldom that it will be almost impossible to get enough information about them".*

A final example from Dollinger (2015:53), who advocates for the use and validity of written questionnaires in social dialectology, underlines the practicality of having multiple methods at our disposal. He argues that,

> *"...a combination of methodologies would possibly lead to the most reliable results as the advantages and disadvantages of each method would become apparent and, ideally, balanced by another method".*

In line with these perspectives, our main goal is to show how and when questionnaires can be useful in conducting fieldwork.

3. Two case studies on the use of questionnaires for comparative linguistic field research

We discuss in this section two case studies on Javanese (Austronesian) which use questionnaires as the methodology for comparative linguistic field research. The first case study uses a questionnaire on modality (Vander Klok 2013b) and discusses three ways we have implemented this questionnaire for comparative research across two Javanese varieties. The second case study is based on a questionnaire that investigates the morphosyntax of polar questions in four different Javanese varieties (Vander Klok, Ahsanah & Sayekti 2017). These written questionnaires target specific linguistic phenomena: the first targets how the semantics of modality is lexically expressed for any language, and the second targets how polar questions are morphosyntactically well-formed in Javanese.

Before discussing the case studies themselves in Sections 3.3 and 3.4, we first provide a brief background of the relevant details on Javanese in Section 3.1. We further specify why questionnaires are an advantageous method to use with this language for comparative linguistic research in Section 3.2.

3.1   Javanese (Austronesian)

Javanese is an Austronesian language spoken primarily on the island of Java, Indonesia by some 70 million people.[7] Javanese has a high degree of variation across dialects in all areas of the grammar (e.g. Hatley 1984; Wedhawati et al. 2006; Hoogervorst 2010). Even though it is the Austronesian language with one of the highest number of native speakers, it remains underdocumented and understudied, particularly with respect to its dialectal variation (Conners & Vander Klok 2016). It is mainly spoken in the provinces of Central Java and East Java and can be divided into three broad dialectal groups: West, Central, and East Javanese (e.g. Hatley 1984). Figure 1 shows the languages spoken on the island of Java and neighboring islands: Sundanese in West Java; Madurese in Madura and East Java; and Balinese in Bali. Indonesian, the national language, is spoken throughout Java and especially in Jakarta, the capital city. Figure 1 also shows Osing and Tenggerese in East Java, Banyumasan in Central Java, and Banten in West Java, which are considered by many to be languages distinct but closely related to Javanese. The dialect spoken in the royal court centers of Yogyakarta and Surakarta/Solo can be referred to as the prestige variety or Standard Javanese given its (historical) influence as well as the fact that it is sanctioned and used across Java in the educational system.[8]

Javanese also has an extensive speech level system – a system of potentially asymmetrical exchange where selection of linguistic features (lexicon, morphology, morphosyntax) is dependent on the relative social status of interlocutors. There are three 'basic' levels: *ngoko* 'low Javanese', *madya* 'mid Javanese', and *krama* 'high Javanese' (Poedjosoedarmo 1968; Errington 1988), although *krama* is currently endangered due to a changing socio-cultural environment and influence from Indonesian (see, e.g., Oetomo 1990; Errington 1998; Conners 2010; Zentz 2015).[9]

---

[7] According to the 2010 census report, there are 68,044,660 Javanese speakers from age 5 years and above, from a population of 95,217,022 self-identified Javanese people (Kewarganegaraan, Suku Bangsa, Agama dan Bahasa Sehari-hari Penduduk Indonesia – Hasil Sensus Penduduk 2010. Badan Pusat Statistik. 2011. ISBN 978-979-064-417-5). Note that these numbers are based on self-reporting and have been going down (see Abtahian et al. 2016).

[8] Pertaining to terms concerning language, some use 'dialect' to refer to a variety of a language that is characterized by different grammatical features, while 'variety' is associated with a group according to some external factor, perhaps geographical or social (e.g., Wardhaugh 2015). For instance, in reference to the external factor of prestige, it may be more appropriate to say "prestige variety" instead of "prestige dialect". In the Indonesian context, however, these terms can often be used interchangeably since a different geographical area or social setting in Java typically results in different grammatical features. In this case, 'variety' or 'dialect' can refer to a variant of a language that is (somewhat) mutually intelligible with a variant of the same language.

[9] Many current Javanese scholars use the term 'speech level' or 'speech styles' (e.g. Errington 1998) given that there are grammatical differences between the levels (e.g. Wolff & Poedjosoedarmo 1982). As such
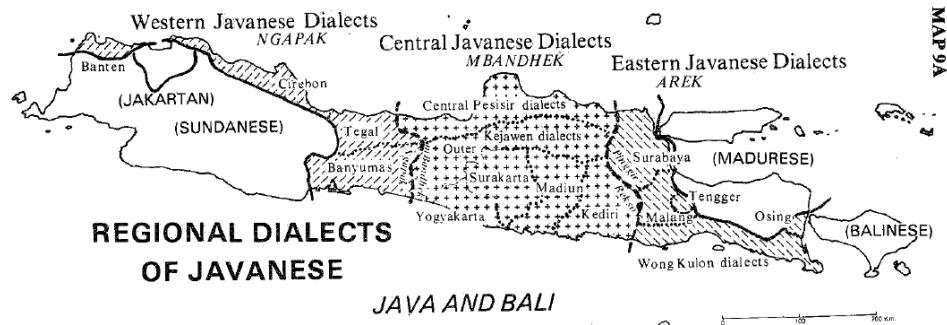
Figure 1. Map of languages and some Javanese dialects spoken on Java, Indonesia. (Hatley 1984: 24, Map 9A)

Among Javanese speakers, there is now widespread bilingualism with Indonesian, which has become—especially in city settings—the language of administration and education. As this is not a case of stable bilingualism (see Abtahian et al. 2016), any linguistic research on Javanese must also consider the role of Indonesian and the language profile of the speakers.

Research on Javanese dialects must therefore take into account a number of linguistic and social layers: Indonesian as the national language; Standard Javanese as the prestige variety; geographical distance from the courtly centers; and an intricate speech level system.

3.2 Why are questionnaires useful in the context of field research on Javanese?

Beyond the inherent advantages of questionnaires as discussed in Section 2, in the case of conducting fieldwork on Javanese specifically, using questionnaires is an advantageous tool for at least two reasons. First, it is fairly easy to run numerous participants given the high number of speakers available and the general willingness of participants. Furthermore, from a comparative linguistic research perspective, questionnaires make it possible to collect comparable sets of answers, or at least always involve the same questions (cf. Bowern 2015:92), a necessary condition when studying linguistic phenomena across dialects or varieties of a language.

---

the speech level system is not simply a socially stratified system, and the terms 'basi-, meso-, or acrolectal' as suggested by a reviewer are not appropriate.

3.3   First case study: Questionnaire on modality

The first case study concerns a questionnaire on modality (Vander Klok 2013b, 2014), which was conducted on two East Javanese varieties, one spoken in the village of Paciran, Lamongan Regency, and the other spoken in the city of Malang, Malang Regency. Before discussing how the questionnaire was used as a tool for comparative semantic research across Javanese dialects, we first briefly situate our study within the Javanese language and the linguistic expression of modality.

Javanese is not a heavily inflected language overall (Conners, *In press*). In the verbal paradigm, verbs are marked for voice or focus through prefixation, and there are two inflecting applicative suffixes with a range of functions (Robson 2014:38-54;74-78). Verbs are not grammatically marked for tense, aspect or modality; instead, there is a relatively rich inventory of auxiliaries and adverbs that optionally mark aspect and modality (Robson 2014:54).[10] For instance, the verb *mateni* in (3) does not indicate tense, aspect or mood; it includes information about argument structure including the actor voice prefix *m-* (assimilated in place of articulation to the root *pate* 'die') and the applicative suffix *–(n)i*. Instead, the marker *wes* 'already' gives information about when this event has taken place relative to a contextually salient reference time.[11]

(3) *Pak Suwanan wes   mate-ni   lampu.*   PACIRAN JAVANESE
     Mr. Suwanan already AV.die-APPL light

     'Mr. Suwanan has turned off the light.' (Vander Klok 2012:57)

As for modality, few careful semantic studies have been conducted on Javanese. Research on modality shows that modals in natural language lexically differ in expressing three dimensions: *modal force*, expressing a possibility or necessity modal claim; *modal flavor*, such as epistemic, based on the speaker or agent's knowledge or deontic, based on a body of rules or regulations; and *modal strength*, expressing a modal claim weaker or stronger than possibility or necessity, such as 'weak necessity' (*should* or *ought to* in English).[12]

To our knowledge, Ekowardono et al. (1999) is the most complete study of modals in Standard Javanese. Beyond this study, no formal semantics study exploring both the force and flavor of modals had been conducted prior to our fieldwork. This is likely due to the

---

[10] The following glosses are used in this paper: AV 'actor voice', APPL 'applicative', CIRC 'circumstantial modality', DEF 'definite', DEM 'demonstrative', DEON 'deontic modality', EPIS 'epistemic modality', FUT 'future', KE 'accidental or adversative passive', NEC 'necessity', NMLZ 'nominalizer', POS 'possibility', PRT 'particle', PST 'past tense', ROOT 'root modality'.

[11] See Vander Klok and Matthewson (2015) for an account of *wes* as 'already' and not the perfect aspect.

[12] For introductions to these dimensions of modality, see, e.g., Palmer (1986), Portner (2009), Hacquard (2011).

difficulty in eliciting nuanced, contextualized semantic judgments. Instead of simply collecting an inventory of occurring lexical terms, we were interested in exploring the lexicalization of the semantic space encoded by those forms, particularly given the importance of modals in the language. Since documentation had already suggested that modality lexically varies considerably across Javanese dialects (e.g. Ekowardono et al. 1999; Robson 2014; Cole et al. 2008), we were also interested in whether other dialects also showed similar variation with Standard Javanese.[13]

In order to facilitate data collection on modality in Javanese, Vander Klok (2013b, 2014) developed a questionnaire designed to elicit how modality is semantically expressed in natural language.[14] The questionnaire on modality has a total of 41 items (33 target and 8 fillers), which are contextualized for semantic (felicity) judgments. This follows the methodology advocated in Matthewson (2004) for semantic fieldwork, which argues that felicity or truth value judgments must be contextualized and that translations cannot be interpreted as linguistic evidence for semantic meaning.

The modal questionnaire was implemented in two different ways for comparative linguistic field research: as an acceptability rating task and a semi-forced choice task in Paciran by the first author; and as a semi-forced choice task in Malang by the second author. We also used elicitation first in a pilot study in Paciran and then as a supplement to the semi-forced choice task in Malang.[15] We outline how these types of implementation were conducted in the following three subsections.

3.3.1 Questionnaire on modality: Elicitation task

The elicitation implementation was used at two different times. In the first case, Vander Klok worked with one native speaker in one instance and a group of three native speakers in a second instance in Paciran in the initial stages of research on the modal questionnaire

---

[13] It is beyond the scope of this paper to provide a full inventory of modals in Paciran and Malang Javanese. The most frequent modals in both varieties include: *mesthi* 'EPIS.NEC', *kudu* 'ROOT.NEC', *paleng/paling* 'EPIS.POS', *oleh* 'DEON.POS', and *iso* 'CIRC.POS' (See Vander Klok 2013a for a detailed study on the possibility and necessity modals in Paciran Javanese).

[14] An English version is freely available online for cross-linguistic use, hosted on the Max Planck Institute for Evolutionary Anthropology (MPI EVA) website 'Typological tools for field linguistics' as well as on TULQuest. For MPI EVA, it is found under the sub-heading 'Modality' at http://www.eva.mpg.de/lingua/tools-at-lingboard/questionnaires.php; and at TULQuest: http://tulquest.huma-num.fr/fr/node/70.

[15] A reviewer points out that questionnaires implemented as elicitation sets seem to be no longer questionnaires since elicitation is implemented orally with a researcher and tends to be individual or with a small group, while questionnaires are written and can be either based on self-reporting or working with a researcher, and tends to be with a larger number of respondents. We agree that these are different implementations and can be called elicitation vs. questionnaires, but since the content is the same, we include the elicitation task under the discussion of how the modality questionnaire was conducted.

in 2011. The individual and group elicitation sessions were used as pilot studies in order to fine-tune the discourse contexts, to discard any items that were too confusing, and to ensure that all items were comprehensible. These pilot studies were crucial to make the questionnaire as clear as possible since it was relatively long (participants took between 20 and 45 minutes) and to only use the Paciran Javanese variety. For example, one outcome of these pilot studies was the decision to not include the markers that were identified as evidential markers in Paciran Javanese (*koyoke, ketoke, jekene* 'direct evidentials' and *watake, bonake* 'indirect evidentials'; see Vander Klok 2012 for a description) since it would have made the questionnaire too long.

In the second case, the second author worked in 2015 with three native speakers from Malang, East Java after the questionnaire had been developed and used on Paciran Javanese. To illustrate the implementation as an elicitation set, we focus on Conners' experience.

In Malang, the fieldworker met with each consultant individually, and then had a follow-up session with two of the speakers together where the fieldworker asked a series of clarifying questions. As the questionnaire was originally developed for a different dialect (Paciran Javanese), the fieldworker altered the prompts and discourse contexts to accurately reflect the Malang dialect. During the first elicitation session, these forms were first checked with the native speaker, and corrections were made. This is an important step in adapting a questionnaire for dialectal research, as naturalness beyond simple acceptability could be a confound when dealing with judgments on subtle semantic or pragmatic differences such as with modals.

The elicitations began with a general request for the participant to list and use in a series of sentences all of the modals that they could come up with (whether adverbs, auxiliaries or verbs). This was a fruitful task because each of the participants had had some linguistic training, and so could understand and successfully identify what a modal was. This was also an important preliminary step to ensure that the fieldworker did not proceed under the assumption that Malang has the same inventory of modals as Paciran Javanese. This initial step also allowed for newly discovered modals to be incorporated into the semi-forced choice task. For example, we uncovered the particle *lak,* a term not previously known to the fieldworker, which can be used to express near future certainty. It contrasts in degree with *bakalane*, a morphologically complex expression, previously known to the fieldworker, which also expresses future certainty. What was uncovered is that, according to the speakers, *lak* encodes a greater degree of certainty than *bakalane*, as shown in (4a) vs. (4b):

(4)   Elicitation with speakers of Malang Javanese

    a.   Sesuk   *bakal-an-e*   *udan!*

         tomorrow   **FUT-NMLZ-DEF**   rain

         'It's going to rain tomorrow!' (lit. 'As for tomorrow, the future is rain!')

    b.   *Sesuk*   *lak*   *udan!*

         tomorrow   **FUT.PRT**   rain

         'It's [definitely] going to rain tomorrow!'

After the inventory of modals was collected along with examples, the fieldworker conducted a semi-forced choice task (described in Section 3.3.3) in Malang Javanese, adding new examples gathered from the elicitation.

3.3.2   Questionnaire on modality: Acceptability rating task

For the acceptability rating task, conducted for Paciran Javanese by the first author in 2011, participants were asked to rate the target sentence for acceptability under the given context on a scale from 1-5, where 1 was defined as *cocok 100%* '100% contextually appropriate' and 5, as *gak cocok belas* 'not at all contextually appropriate'. The procedure was as follows: the fieldworker first went over the written instructions verbally with the participants, who also read them. Then the participants completed four practice questions before turning to the main questionnaire. Participants were presented with a context and one target sentence on a laptop screen, and then circled a number between 1 and 5 on a separate piece of paper for the corresponding target sentence. There were 20 participants in total (10 participants for each of the two target sentences per context).[16] This task was not conducted for Malang Javanese.

An example is given in (5), where the Paciran Javanese presentation is illustrated first, followed by the English translation of the context and the glossed target sentence with the results. This example tests whether the modal marker (*paleng* or *mesthi*) is compatible in a context that targets an epistemic possibility reading. The results suggest that the modal *paleng* in Paciran Javanese is compatible with epistemic possibility, with an average rating of 2.3. In comparison, the corresponding target sentence *Kalunge Dewi mesthi ilang* 'Dewi's necklace must be lost' with the modal *mesthi* was rated as incompatible with epistemic possibility, with an average rating of 4.[17]

---

[16] See Vander Klok (2014) for more details on the implementation of the acceptability rating task and the semi-force choice task of the modality questionnaire.

[17] Vander Klok (2013a) analyzes *paleng* as a possibility modal that lexically specifies for epistemic modal flavor and *mesti* as a necessity modal that also lexically specifies for epistemic modal flavor.

(5)   Example of acceptability rating task for the modal questionnaire on Paciran Javanese (Vander Klok 2013a:351):

Dewi ewoh nggoleki kalunge. Dewi gak yakin kalunge iku ilang temenan toh mek lali ndeleh, soale Dewi gak eling nek endi terakhir ndeleh kalunge. Dewi wis nggoleki nek ndhuwure lemari, nek dhuwure tv, nek njero tase, tapi isek durung ketemu. Engko sek! Dewi durung nggoleki nek lemarine adikne…

	Cocok 100% [Appropriate]			*Gak cocok belas* [Inappropriate]	
Kalunge Dewi paleng ilang.	1	2	3	4	5

[**Context in English**: Dewi is looking for her necklace. She's not sure if she lost it or if it is still somewhere in the house because she doesn't remember the last time that she wore the necklace. She looks in her wardrobe and on top of the wardrobe. It's not there. She looks on top of the TV. It's not there. She looks in her backpack; it's not there. Wait! She didn't check her sister's wardrobe yet…]

Kalung-e	*Dewi*	*paleng*	*ilang.*
necklace-DEF	Dewi	EPIS.POS	lose

'Dewi's necklace might be lost.'   Result: 2.3 (average rating score)

3.3.3   Questionnaire on modality: Semi-forced choice task

For the semi-forced choice task, participants were asked to choose the target sentence(s) that was/were most appropriate given the context (and/or offer alternative(s)). This task is most similar to a yes-no task in acceptability judgments, where participants are asked to give a categorical answer (cf. Schütze & Sprouse 2013). The overall procedure for the semi-forced choice task was the same as for the acceptability rating task: participants first went over the instructions with the fieldworker, completed four practice questions, and then turned to the questionnaire. For this task, there were a total of fifteen participants for the Paciran study, and three participants in Malang.

This task differed from the acceptability rating task in that participants were presented with two target sentences per context, as in (6) for Paciran Javanese. They could then choose the target sentence (a) or (b), both, neither, and/or give an alternative sentence that is contextually appropriate. The example in (6) first gives what was presented in the original questionnaire, followed by the English translation and glossed target sentences with the results in (7). The context targets an epistemic reading; thus, of the two target sentences, the one with *mesthi* 'EPIS.NEC' is clearly compatible (13/15 participants), while *kudu* 'ROOT.NEC' is infelicitous in this context (with 0/15 participants having chosen

this sentence). Additionally, two participants independently and separately provided a third sentence with the modal *paleng* 'EPIS.POS' shown in (7c), suggesting that this modal is also felicitous in epistemic contexts.[18]

(6)   Example of a semi-forced choice task for the modality questionnaire on Paciran Javanese (Vander Klok 2013a:360):

Sirahmu ngelu gak wara-waras. Terus awakmu reng dokter. Wes diprekso tapekne gak ono penyakit opo-opo. Dadi….

   a.   *Iku mesthi kakean pikiran.*
   b.   *Iku kudu kakean pikiran.*

(7)   [**Context in English** (inspired by Rullmann et al. 2008:321): You have a headache that won't go away, so you go to the doctor. You were examined but no sickness whatsoever is revealed. So…]

   a.   *Iku*   ***mesthi***   *k-ake-an*   *pikir-an.*
        DEM   **EPIS.NEC**   KE-many-NMLZ   think-NMLZ
        'It must be from stress.'   (chosen by 13/15 participants)

   b.   *Iku*   ***kudu***   *k-ake-an*   *pikir-an.*
        DEM   **ROOT.NEC**   KE-many-NMLZ   think-NMLZ
        'It has to be from stress.'   (chosen by 0/15 participants)

   c.   *Iku*   ***paleng***   *k-ake-an*   *pikir-an.*
        DEM   **EPIS.POS**   KE-many-NMLZ   think-NMLZ
        'It might be from stress.'   (offered by 2/15 participants)

Example (8), also Paciran Javanese, presents the semi-forced choice analog to the acceptability rating task illustrated in (5). This context, which targets an epistemic possibility interpretation, tests whether either *mesthi* or *paleng* (or both) is/are compatible with this type of modal force since both are compatible with epistemic modality. The results of the semi-forced choice questionnaire show that 14/15 participants chose the sentence with *paleng* and 1/15 participants chose sentence with *mesthi*. Additionally, one participant who chose the target sentence with *paleng* also offered a sentence with *durung mesthi* 'not.yet EPIS.NEC', demonstrating an alternative way to express epistemic possibility.

---

[18] See Vander Klok (2013a) that *paleng* also only lexically specifies for epistemic modal flavor like *mesthi* in Paciran Javanese.

(8) Context as in (5). Two target sentences in (a) and (b); (c) offered by a participant.

    a. *Kalung-e*   *Dewi*   **paleng**   *ilang.*
         necklace-DEF   Dewi   EPIS.POS   lose
         'Dewi's necklace might be lost.'   (chosen by 14/15 participants)

    b. *Kalung-e*   *Dewi*   **mesthi**   *ilang.*
         necklace-DEF   Dewi   EPIS.NEC   lose
         'Dewi's necklace must be lost.'   (chosen by 1/15 participant)

    c. *Kalung-e*   *Dewi*   **durung**   **mesthi**   *ilang.*
         necklace-DEF   Dewi   not.yet   EPIS.NEC   lose
         'It's not certain yet whether Dewi's necklace is lost.' (offered by 1 participant)

Figure 2. Participant working on the modality questionnaire in 2011.
(used with permission)

3.4   Second case study: Questionnaire on polar questions

The second case study concerns a questionnaire on yes-no (or polar) question strategies (Vander Klok et al. 2017). Four Javanese varieties were investigated in total using this questionnaire: one Central Javanese variety as spoken in the city of Semarang, and three East Javanese varieties as spoken in Montong (Tuban Regency), Weru, and Blimbing (Lamongan Regency) villages. Due to the different locations, the first author engaged two

research assistants[19] who were from those areas to administer the questionnaire.

This questionnaire was implemented as an acceptability rating task across all four varieties with 70 items for Montong, Blimbing, Weru, and 88 items for Semarang (partially because there was an additional strategy in Semarang Javanese, further discussed below). Because this questionnaire served as an initial exploration of the various strategies for yes-no questions across Javanese varieties, there were no contexts or fillers. There were 10 participants for each location and we aimed for gender parity.

The methodology for this questionnaire was the following: participants first went over the instructions in their local variety with the research assistant (explained in (9) below). They were instructed to rate each question from between 1 ('completely natural') to 5 ('completely unnatural'). They then completed four practice questions and asked any additional questions about the process before the start of the actual questionnaire. Once the participants completed the rating task, they were invited to give any additional written comments as feedback using a space designated for that purpose.

There were two versions of the questionnaire instructions. Specifically, we used one set of instructions for Semarang Javanese, a Central Javanese dialect, and another set for the three East Javanese varieties (all located along the north shore of Java), as illustrated in (9a-b) and the English version in (9c). It was determined beforehand through consultation with native speakers and the research assistant that the single set of instructions was clear for each of the East Javanese varieties under study.[20]

(9)   Dialectal group-specific instructions for the acceptability rating task on Javanese polar questions.

   a.   Semarang Javanese version (Central Javanese variety)

         Wenehana biji marang ukara ngisor iki kanthi mbunderi angka 1-5 sing maksute 1 = lumrah/natural banget, dene 5 = ora lumrah/natural banget.

   b.   Montong, Blimbing, Weru version (East Javanese varieties)

         Tulung keki biji nek kalimat tanya ngisor iki carane kluwengi pilih salah sijine ongko 1 sampek 5 sing maksute 1 = wes biasa, nek 5 = ora biasane.

   c.   English version (not used in the implementation)

         Please rate each of the following by circling a number between 1-5 where 1 = completely natural and 5 = completely unnatural.

---

[19] Wuri Sayekti conducted the questionnaire in Semarang, and Finatty Ahsanah in the three locations in East Java.

[20] Please see the Appendix for a partial example of the questionnaire for Semarang Javanese.

Javanese polar questions can be created by three different typologically attested strategies: intonation, word order, and particles (Vander Klok 2017). Since the questionnaire was written, it involved the latter two strategies. The items were carefully designed to minimize the lexical differences across dialects and maximize the comparison of the phenomenon under study. For instance, the test items used event predicates that were the same across dialects, such as *nyekel* 'catch', *mbayar* 'pay', *ketemu* 'meet', or *tuku* 'buy'. We avoided lexical items which were different across dialects, such as *cublok* 'fall down' (Central Javanese) vs. *cicir* (some East Javanese varieties). Some differences were unavoidable, such as the proximate demonstrative *kuwi* in Central Javanese and *iku* in East Javanese, but we aimed to use the definite suffix *–(n)e* instead, which is the same form across both dialectal groups.

The purpose of this study was mainly exploratory but based on previous research and fieldwork observations. First concerning word order, the test items for subject-auxiliary inversion included a number of auxiliaries which are grouped into two sets across at least three Javanese dialects: Peranakan Javanese (spoken by ethnic Chinese; Cole et al. 2008), Paciran Javanese and Standard Javanese (Vander Klok 2015). One set of auxiliaries, which includes *oleh* 'DEON.POS', can be fronted to form a polar question, while the other set, which includes *ape* 'FUT', cannot, as shown in (10).

(10)   a.   ***oleh***   *aku*   *cicip-i*   *iwak*   *panggang?*   PACIRAN JAVANESE
             DEON.POS   1SG   try-APPL   fish   grilled
             'May I try the grilled fish?'

       b.   ****ape***   *mbak*   *Nunung*   *masak*   *nastar?*
             FUT   Miss   Nunung   AV.cook   cookies
             ('Will mbak Nunung bake cookies?')   (Vander Klok 2015:150)

Based on these previous findings (Cole et al. 2008; Vander Klok 2015), the hypothesis was that the differences in grammaticality between these two sets would be the same across the four Javanese varieties explored in this study. This same result was borne out. While the test items were held constant across the questionnaire variants, known lexical variation of auxiliaries across dialectal groups was accounted for. For example, the Semarang Javanese questionnaire included both *entuk* and *oleh* as possibility deontic modals (with *entuk* as the preferred marker), while only *oleh* for the East Javanese varieties. Additionally, Semarang Javanese included the auxiliary *nate*, the *krama* 'high Javanese' counterpart to *tau* 'PST' in *ngoko* 'low Javanese', as this marker was noted in fieldwork to be used as a *ngoko* 'low Javanese' marker. Finally, Semarang Javanese included two future auxiliaries *arep*, a volitional future, and *bakal*, a non-volitional future, while the East

Javanese varieties only included *ape*, as illustrated in (11).[21] These were all judged as ungrammatical.

(11) Test items for subject-auxiliary inversion with auxiliaries marking the future across Javanese dialect groups

    a.   **Arep*   *Nunung*   *nggawe*   *nastar?*   SEMARANG

    b.   **Ape*   *Nunung*   *nggawe*   *nastar?*   MONTONG, WERU, BLIMBING
           PROSP/FUT Nunung   AV.make   cookie
           'Will Nunung make cookies?'

    c.   **Bakal*   *Pak polisi*   *nyekel*   *maling*   *kuwi?*   SEMARANG
           FUT   Mr. police   AV.catch   thief   DEM
           'Will Mr. Police catch that thief?'

Second, concerning deriving polar questions with particles, we explored sentence-initial and sentence-final particles, as well as some combinations. There is only one sentence-initial particle in Semarang Javanese (*apa* [ɔpɔ]) and across the East Javanese varieties (*opo* [opo]) investigated. Sentence-final particles in polar questions can divided into those expressing 'yes' or 'no' and those dedicated to indicating focus or yes-no questions. In Semarang Javanese, *yo* 'yes' and *rak* 'NEG' are used, while in the East Javanese varieties, *yo* 'yes' and *gak* 'NEG' are used. An example of the other sentence-final particles is given in (12) (all grammatical), showing dialectal differences. Apart from the particles, only the future marker differs between the Central Javanese variety (Semarang) with *arep* and the East Javanese varieties with *ape*; the proper name and predicate are otherwise the same. Further, Semarang Javanese has two sentence-final particles (*to* and *ndak*) used to create a polar question, whereas the other varieties do not use the *ndak* particle.

(12) a.   *Nunung*   *arep*   *nggawe*   *nastar*   *to?*   SEMARANG
     b.   *Nunung*   *arep*   *nggawe*   *nastar*   *ndak?*   SEMARANG
     c.   *Nunung*   *ape*   *nggawe*   *nastar*   *leh?*   MONTONG
     d.   *Nunung*   *ape*   *nggawe*   *nastar*   *tah?*   WERU
     e.   *Nunung*   *ape*   *nggawe*   *nastar*   *tah?*   BLIMBING
         Nunung   PROSP/FUT   AV.make   cookie   PRT
         'Nunung will make cookies, right?'

We also explored combinations of the sentence-initial particle *apa/opo* with sentence-final particles; the 'yes'/'no' particles with the focus-type sentence-final particle; and both 'yes' and 'no' particles together. One interesting discovery revealed in the pilot study was that in the variety of Javanese spoken in Montong, the syntactic order of the combination

---

[21] *Ape* is analyzed as a prospective aspect (Chen et al. 2017), but *arep* behaves differently in that it requires a volitional agent (Vander Klok 2012). To underline that further research is needed, we gloss them in (10) and (11) as PROSP/FUT.

of the sentence-final particle *leh* with negation was *gak* 'NEG' > *leh* 'PRT', while in all other varieties investigated, the order is the opposite (PRT > NEG), as shown in (13). Possible semantic differences related to this word order difference needs to be researched.

(13) a. *Gayus* *kudu* *mbayar* *dendo-ne* **gak** **leh?** MONTONG
          Gayus ROOT.NEC AV.pay fine-DEF NEG PRT
          'Gayus has to pay the fine or not?'

     b. *Gayus* *kudu* *mbayar* *dendo-ne* **pa** **rak?** SEMARANG
          Gayus *kudu* *mbayar* *dendo-ne* **tah** **gak?** WERU
          Gayus *kudu* *mbayar* *dendo-ne* **tah** **gak?** BLIMBING
          Gayus ROOT.NEC AV.pay fine-DEF PRT NEG
          'Gayus has to pay the fine or not?'

Note that the particle used with negation in polar questions in Javanese is the one also used for disjunction (Vander Klok 2017); thus, for Semarang Javanese, it is not with *toh* or *ndak* but with *pa*, a shortened form of *apa*. We did include the combination *toh rak* to test this, and it was judged ungrammatical.

    In sum, this second case study provides a further illustration of how the grammatical differences of polar questions across Javanese varieties played a role in the design and implementation of this questionnaire—instructions were provided in the appropriate dialect, and the elicitation items varied according to strategies available in each variety. Some shortcomings of this study were that not all possible combinations were included in the test items and no filler items were included. However, since the main purpose of this study was exploratory, these shortcomings can be rectified by conducting a follow-up study focusing on specific phenomena.

3.5   Summary: Questionnaires as a comparative linguistic field research tool

The first case study illustrates how the same tool—a questionnaire on modality—can have different methods of implementation both within one dialect and across dialects. Research on the Paciran dialect revealed significant differences with what had been described for Standard Javanese, primarily in how the different lexical modals carve up the modal space in terms of force and flavor (Vander Klok 2013a, 2015). In Malang Javanese, the questionnaire helped to uncover new terms, and made it clear that in general, lexical modals patterned with Paciran, rather than Standard Javanese. Given that both dialects are spoken in East Java, and Standard Javanese is spoken is Central Java, this

finding is not surprising. There are however modals in Malang Javanese that are not present in Paciran, such as *lak*, which seems to express near certain future.

The second case study was on a questionnaire designed to explore which morphosyntactic strategies of polar questions are available within and across one Central Javanese variety and three varieties of East Javanese. This study highlighted the need to carefully design the target items in order to maximize the comparative aspect, while allowing for linguistic variation (possibly unanticipated) and keeping in mind processing and time constraints. While questionnaires were identified as an advantageous method in both case studies, they still raised some methodological challenges; we address these in the following section.

4.   Challenges of questionnaires

In this section, we describe some challenges of using questionnaires as a tool specific to comparative linguistic field research. Even if a written questionnaire is deemed to be the right tool for the research question and the research population, one can still be confronted with challenges, whether anticipated or not. We use the modality and polar question questionnaires discussed above as case studies to raise these issues. In response, we outline a set of best practices in Section 5, and refer the reader to other resources since these best practices extend beyond using written questionnaires as a fieldwork method.

4.1   Challenge 1: Variation within a written questionnaire for a targeted construction

The first challenge concerns how to conduct a comparative study of a specific linguistic phenomenon—for instance across dialects—while still allowing for (possibly unanticipated) grammatical variation. In creating *written questionnaires for targeted constructions*, a general goal of the (field) linguist is to maximize direct comparison in a context of possible (lexical/phonological/morphosyntactic) variability across languages or dialects, so that examples and forms will be natural and comprehensive for each language/dialect.

For instance, in the second case study, the polar question questionnaire had 88 items for Semarang Javanese but 70 items for the other dialects since only Semarang Javanese uses the particle *ndak* as an additional strategy. In this way, the questionnaire included enough items for cross-dialectal comparison across the four varieties but did not over-generalize. From a processing point of view, omitting the grammatically unavailable

strategies with *ndak* in the dialects spoken in Blimbing, Weru and Montong (East Java) allowed us not to overwhelm participants with ungrammatical examples.

Another example concerns lexical variation: in feedback from the yes-no questionnaire on Semarang Javanese in particular (Wuri Sayekti, p.c.), we learned that some participants rated certain items with a low score for reasons unrelated to grammaticality in some dialects, such as a proper name sounding unfamiliar. For instance, one participant for the Semarang Javanese questionnaire mentioned orally in the practice questions that *Hari* was not a common proper name, and then judged the item lower. We also used the proper name *Kana* [kana] in the rating task for all varieties, only realizing later that this form is homographic to the Standard Javanese distal demonstrative [kɔnɔ], which was a confound for some participants in Semarang Javanese, since both varieties are spoken in Central Java.

4.2   Challenge 2: Addressing (or not) more salient or prestigious variants

A second challenge of implementing questionnaires as a tool for comparative linguistic field research is the different variants of a language. Across Javanese language varieties, this type of challenge takes at least two forms.

First, there is the challenge of dealing with prestigious vs. non-prestigious variants. The variety spoken in the Central Javanese courtly centers of Yogyakarta and Surakarta/Solo is taught in schools and is the standard in formal Javanese writing, even in areas of West and East Java, where regional varieties differ significantly. One of the challenges for the fieldworker is convincing the participants that you want to research their local variety and getting them to express themselves in their own dialect; this is especially challenging in a more formal situation, as with a written questionnaire. Presenting the instructions and target items in the Javanese variety under discussion with representative phonological and lexical differences—possibly the only time someone might see their variety written in a formal setting—helps achieve this. This approach is not without potential issues, of course, as some participants might take the exercise less seriously. For instance, for the Javanese case studies in Section 3, some participants suggested that it is better to study Standard Javanese. However, once we explained that there is more research on that variety and we are interested in the variety they speak, most are sympathetic to this view. We have found that using the local variety in the research materials themselves (including the instructions) is generally perceived as positive, and participants are happy that research is being conducted on local varieties. The other alternative is to sidestep the issue of which language variety to use, and simply use the language of wider communication; in this case,

Indonesian. We did not choose this option in our studies since part of our goal is to recognize and validate colloquial varieties of Javanese. In practice, however, Indonesian was useful as a metalanguage in elicitation to further discuss or explain Javanese language examples.

Second, there is the challenge of ideologies pertaining to Javanese speech levels. As mentioned in Section 3.1, Javanese has a complex linguistic etiquette with various speech levels (*ngoko, madya, krama* 'low, mid, high Javanese') as well as additional humble and honorific vocabulary sets (Poedjosoedarmo 1968, Errington 1988). Knowledge and use of the etiquette system, however, vary extensively across dialects (Smith-Hefner 1989; Conners 2008; Krausse 2017; Vander Klok *to appear*). Even though in some Javanese varieties the speech levels are not as extensive, it is not uncommon for speakers to make sharp distinctions in codes, such as discussed in the elicitation implementation of the modality questionnaire in Section 3.2.1. In both case studies, we focused on *ngoko* 'low Javanese', and tried to set the stage with the instructions in *ngoko* as well. For East Javanese varieties studied in this paper, *ngoko* is the everyday norm and *krama* is not used as extensively as in Central Javanese, and this approach generally did not pose problems. Further, the young age of the participants (most in their early 20s) facilitated the use of *ngoko* since many young Javanese speakers in these areas do not speak *krama* fluently (cf. Setiawan 2012; Vander Klok *to appear*). Overall, there are some exceptions to these sharp code distinctions, which include social media such as SMS, Twitter, or Facebook (Brugman & Conners 2018), or magazines specific to certain varieties, such as *Panjebar Semangat*, a Surabaya weekly; *Djaka Lodang*, a former Yogyakarta weekly; and *Jaya Baya*, a Surabaya weekly that focuses on culture, the arts, and literature.

4.3   Challenge 3: Recruiting speakers of a specific variety

A third challenge is that since the questionnaire is designed to be specific (but maximally comparative; see Challenge 1) to a particular language or variety for comparative linguistic research, then the researcher requires participants of that language/variety. We found it difficult to 'define' the language variety, especially in a city setting; this was in part because within the city of Semarang, some speakers identified with a variety of the language pertaining to specific neighborhoods.[22]

Our overall goal in the two case studies was to target the relevant participants through metadata questions, without necessarily excluding those who deviated from the 'ideal' language profile. We thus took a broad approach, targeting speakers who grew up and still

---

[22] See, for example, Samidjan (2013) on potential smaller subdivisions within Semarang dialect. See also Goebel (2002, 2005) on reference to neighborhood varieties.

lived in the same place, and did not exclude those participants whose parents were not from that location or those who had lived elsewhere for a period during their lifetime. Importantly, this information was nonetheless recorded as metadata and we can further subdivide our data into various speaker groups based on these factors.

4.4   Challenge 4: Methodological issues

A fourth challenge, which pertains to any fieldwork or language experiment, is to provide explicit and clear instructions as well as to ensure that the participants fully understand them. For example, we teach participants the task by going through practice questions together and allowing feedback during this stage (this is also possible while they are doing the questionnaire). These two points of the methodological process—clear instructions and ensuring they are understood—are key for participants to properly follow the task.

Some issues that were raised in our case studies were that for the rating task, some participants use only one end of the scale, and some only used 1 and 5. If the fieldworker can recruit a high number of participants, this challenge can be interpreted through relevant statistical tests.[23] Ideally, however, this problem should be resolved through practice questions that define the ends of the scale, before the actual questionnaire is undertaken. Another issue that we encountered was that despite our instructions for the rating task of the questionnaire on yes-no questions, some participants wanted to know or give the answers to the yes-no question (in order to rate the question), showing a lack of understanding of the task.

5.   Some best practices for using questionnaires for comparative linguistic field research

Based on the above challenges, we suggest five practices for implementing written questionnaires for targeted constructions. This paper takes a narrow focus coming from the perspective of comparative linguistic research rather than from using written questionnaires in general (see various ways questionnaires can be classified in Section 2). Despite this focus, we feel many of these points can be useful in the practice of field research and in experimental studies: having a pilot study and practice questions; evaluating which language or variety is the most appropriate at each stage of the questionnaire; gathering

---

[23] It is beyond the scope of this paper to discuss this; see, for instance, Dollinger (2015) and overview chapters in Podesva & Sharma (2013).

metadata on the social background of the participants. In addition to these points, we also advocate for the designation of an appropriate contact person for each language or dialect under study as well as ample space for feedback where the participant is free to use any language or variety. Given that these points can have a wider application than questionnaires, we keep this section brief, but draw from our own experiences and suggest relevant literature that may also be helpful.[24] Overall, this section might be most useful to a scholar embarking on a comparative fieldwork research project.

5.1   Developmental stages

The developmental stages in this section refer to two different types: (i) aiding participants in understanding the task and (ii) creating and refining the test items.

The first point is especially important in certain fieldwork situations where questionnaires are a completely novel task, and participants may not be comfortable with questionnaire tasks. It is crucial to set up a practice task in order to ensure that all participants understand and can perform the task. However, not everyone is capable of every type of task, due to different interests or talents, or perhaps to physical limitations. (For practical advice on working with language consultants, see e.g., Chelliah 2013:51-52; Bowern 2015:148-153.)

Concerning the second point, it is imperative to carry out a pilot test for the questionnaire: this allows the researcher to practice and perhaps train others in conducting the questionnaire, as well as check if there are any issues with the content. Nahhas (2007:84) underlines that it is also important to test the tool in a setting as similar as possible to that in which the experiment will be conducted.

Pilot testing provides a chance to uncover problems with the items and to check both grammaticality and naturalness. When using a questionnaire for comparative linguistic research, different lexical items or grammatical constructions may express similar ideas. The range of meanings or uses encompassed by a lexical item in one dialect is not necessarily co-extensive with that lexical item in another dialect. The questionnaire must therefore be structured broadly enough to include unexpected relevant distinctions or features that may vary.

For some questionnaires, pragmatic loading can also be significant; there can be a certain amount of contextual information provided prior to the experiment. In the section of the modal questionnaire testing the use of two modals for the expression of

---

[24] For practical information on fieldwork, see especially Bowern (2015) and references therein, and Chelliah & de Reuse (2011). For practical information in using more experimental methods in linguistics, see Arunachalam (2013). For practical information for using written questionnaires specifically for social dialectology, see Dollinger (2015).

circumstantial versus epistemic possibility, the original question provided a rich context, and then offered the same carrier sentence that differed only in the selected modal. As described in section 3.2.3 above, the participants were asked to select the sentence(s) that was/were most appropriate given the context, if any, and/or to offer alternatives. A pilot test in Malang Javanese revealed an additional option, incorporating both of the modals with different syntactic distributions, as in (14), and suggesting the second English translation:

(14)   Piloting questions reveals new possibilities to test:

a.		*duk*		*isò*	*thokol*	*ndhék*	*kéné*
b.		*duku*	*paléng*		*thokol*	*ndhék*	*kéné*
c.	*paléng*	*duku*		*isò*	*thokol*	*ndhék*	*kéné*
	EPIS.POS	k.o.fruit	EPIS.POS	CIRC.POS	grow	at	here

'Duku can/might grow here.' ~ 'Duku might be able to grow here.'

Finally, it is important to consider the amount of time needed for this stage, which almost always takes longer than expected. Questionnaires of the type described in this paper are often used to test differences between certain lexical items or particular grammatical constructions. We have found the semi-forced choice implementation useful for this kind of direct comparison. It is necessary to dedicate a enough time to the developmental stages of the questionnaire to ensure that each item is testing the appropriate distinction, so that the results are properly interpretable. It is thus important to accurately gauge the time for both the development of the test items and for ensuring the participant understands the task at hand.

5.2   Contact person and/or research assistant

Ideally, for comparative linguistic field research, the same fieldworker should administer the questionnaire across different languages or varieties for consistency, and there should be a contact person for each variety. By contact person, we mean someone who is not a study participant directly but who can reliably answer questions relating to that particular language/variety and ideally have some linguistic training. We view the contact person as playing a key role in the developmental stages such as creating the corresponding version of the instructions in the local variety and participating in the pilot study.

When the fieldworker is unable to conduct the questionnaire with all study participants, then research assistants can undertake this role. Unlike the contact person, any research assistant must have some training in field methods or data collection. The training can be given by the fieldwork researcher, if needed. (Of course, the contact person can also be the research assistant.) The training process is beneficial for all; it can strengthen

relationships between the researcher and the assistant, and perhaps also with the community as this work may be perceived as important. It also gives the consultant strong transferable skills for other possible employment, such as people skills or basic computer skills. Much of the literature in language documentation advocates for working with the community in language research (e.g. Dwyer 2006 and references therein) and making use of a research assistant can play a role in this endeavor.

While it may not be necessary for the research assistant to have fluency in IPA, he/she must be able to accurately record distinctions that arise during data collection, from the phonetic level through the semantic and discourse levels. If necessary, the data collection can be recorded, so the fieldworker can review the work of the research assistant after the fact. Depending on the skill level of the research assistant, the fieldworker should train or familiarize them with the questionnaire to ensure that they understand both the overall goal and the express intent behind each prompt. Furthermore, the assistant should be instructed in how to go over the practice questions with the goal of having the same experience level across all participants, as mentioned in Section 5.1. Finally, it is important to maintain open and direct contact with those conducting the questionnaire to ensure timely feedback.

5.3 Language/variety selection

Choosing the most appropriate language or variety is important for various aspects of conducting the questionnaire including the language of instruction (both oral and written), the language of the target items of the questionnaire, and the language for asking for feedback. AnderBois & Henderson (2015) underline the importance of reporting the reasoning of this choice in the results of the study – often it is not clear to the reader in which language the study was undertaken, despite potential repercussions on the results, and why this choice was made (it could be for ethnolinguistic, sociolinguistic, or purely linguistic reasons.)

In a comparative linguistic study, it is helpful to make a distinction between dialectal data collection of languages for which there is a published or even notional standard, and those for which there is not. We consider written standard French to be an example of a published standard, while Parisian French (in speech) would be an example of a notional standard. Similarly for Indonesian, while the Indonesian in newspapers or formal speeches can be considered as the explicit standard, Jakartan Indonesian is the notional standard. Javanese presents an additional complexity, as there is no published standard; that said, the variety spoken in and around the royal centers of Yogyakarta and Solo/

Surakarta are considered as the notional standard and circulated as such in teaching materials (see Section 3.1 above).

For languages with no regulated or notional standard, each dialect or variety can be approached neutrally. However, in the context of published or notional standards, care needs to be taken in terms of which language or variety is chosen for each step, taking into consideration the benefits of a shared contact language distinct from the target language, while weighing its possible influence thereupon. For example, in conducting research on Javanese in Indonesia, one benefit of using Indonesian as the language for instructions in administering a questionnaire is to avoid possible influence from 'Standard Javanese' as the notional standard. On the other hand, this choice can be felt to possibly belittle the particular colloquial variety of Javanese under study. Another approach is to create different versions of instructions that correspond to the specific variety being investigated, as we have done for the polar question questionnaire (see Section 3.4). Regardless of the language or variety chosen, instructions should emphasize the interest in studying the variety under discussion. Instructions should be clear and concise, and in a language that is fully comprehensible to the study participant. It can be very helpful to give an example in the instructions. Finally, participants should be free to use the language of their choice in expressing themselves, such as in offering feedback.

5.4   Feedback

An area for feedback from participants should be made available in the written questionnaire, for final comments or alternative target items. As much as we try to control for the hypothesis under study in the questionnaire design, feedback often offers clues as to different ways to better interpret the data.

For example, in one of our items from the modal questionnaire, we provided the following context (adapted from von Fintel & Gillies 2007):

(15)   You are going to visit your friend in the hospital. When you enter into the hospital, you stop at the information desk to inquire what room your friend is in. But the woman at the information desk tells you that you can't visit your friend now because it's already 8pm! She says, "I'm sorry, the hospital regulations say that…" (Visitors MUST leave by 6pm.)

One of our participants noted that this sort of interaction is relatively formal as it takes place in a public setting (the hospital), and between a 'customer' and a paid employee. Further, the employee is quoting or at least paraphrasing an established regulation. Due to all these factors, Indonesian would often be preferred to Javanese. The participant felt that if the 'customer' were a monolingual Javanese speaker and hence likely older,

from a rural area, and less educated, then the employee may well use a variety with elements from the local Javanese dialect – though still in a polite register. This sort of feedback is important when interpreting answers.

5.5   Metadata

In running the questionnaire, detailed metadata on the participants should be collected. While this point seems obvious, it is nonetheless worth taking the time to prepare for this step. For instance, it is important to establish which metadata categories are relevant for the goals of the study. It is also worthwhile to consider how the data might be used in the future, either by the fieldworker herself, or by some other researcher. Collecting too much metadata can be tedious for participants.

If a goal of the research is to be able to make generalizations over different demographic groups, then that demographic information must be collected from all study participants. For a comparative linguistic study, when multiple participants from multiple locations are providing responses, it is difficult to identify all relevant distinctions. For example, if some small subset of participants responds to a prompt in a certain way, it will be useful and interesting to be able to identify whether that group represents a demographic trend.

Beyond general demographic information, we find it useful for work in multilingual places like Java to include questions on which neighborhoods one has lived in for significant periods. It is also important to collect basic data on the language of parents and grandparents, in addition to what language(s) or speech level(s) the participant speaks with them. While collecting this kind of data can be time-consuming, having such information makes it possible to interpret the data with a finer-grained approach than perhaps anticipated.

5.6   Summary of practices

We find it useful to report on the practices and considerations that we successfully deployed in the field in cross-dialectal research. Some of the best practices we highlight above can be summarized as a common-sense approach combined with an awareness of local needs and norms. It is important to pay attention to the developmental stages of research, pilot questions, and determine what sort of metadata is necessary without over-burdening research participants; to be aware of local norms (work with the community, engage a local point of contact, and identify and sensitively navigate different language ideologies and attitudes, especially those that can confuse research results); to be open to

new findings (do not construct your research tools in such a way that they will affect outcomes); and include sufficient space in the process to allow for feedback that could be crucial for interpreting results.

6.   Conclusion

Field linguists have a number of different tools and methodologies at their disposal when conducting fieldwork. The goals of the research project, limitations in time, resources, and project participants, and the scope of collection will all factor into determining the most appropriate method(s) and approach.

The focus of this paper was on the use of *questionnaires for targeted constructions*, and particularly how they can be beneficially used in comparative linguistic field research. We discussed two case studies on Javanese where questionnaires are relevant for collecting data on specific grammatical constructions or features as well as for gathering sociolinguistic information such as language attitudes and use. The case study using the questionnaire on modality showed how a questionnaire can be useful to gather subtle judgments on semantics from multiple respondents in a replicable way. The study also showed how the same questionnaire can be implemented in various ways, depending on the timing and number of participants. The second case study, about a questionnaire designed to explore the morphosyntactic strategies of polar questions across four different Javanese varieties, showed the importance of carefully designing items to maximize direct comparison while allowing for variation.

Through these case studies, we also have identified four significant challenges: identifying variation as dialect-internal or cross-dialectal; identifying sociolinguistic variants; delimiting speakers of a variety; and overcoming particular methodological issues. In order to overcome these challenges, we suggested five 'best practices', useful particularly for the scholar embarking on a field research project and using questionnaires for comparative linguistic studies. These practices are to (i) focus on questionnaire development as well as experience leveling of participants; (ii) work with a contact person and/or research assistant for each dialect/variety; (iii) consider which language or variety is most appropriate for instruction, collection, and implementation; (iv) allow for feedback in various ways; and (v) collect appropriate metadata.

Acknowledgements

Appendix A. Example of yes-no questionnaire for Semarang Javanese

Kuesioner   Nomer: _____
Semarang

Wenehana biji marang ukara ngisor iki kanthi mbunderi angka 1 - 5
sing maksute 1 = lumrah/natural banget, dene 5 = ora lumrah/natural banget.

Latihan:

1. Opo Hari nyanyike anake sewengi?	1	2	3	4	5
2. Pelem kuwi mbok wis pangan?	1	2	3	4	5
3. Lili wis 3 taun to?	1	2	3	4	5
4. Lili tau mangan rajungan mau bengi to?	1	2	3	4	5

Kuesioner:

	Lumrah banget				Ora lumrah banget
1. Titin lagek ketemu koncone to?	1	2	3	4	5
2. Opo Bambang njahit celono kuwi ndak an?	1	2	3	4	5
3. Opo Bambang njahit celono kuwi to ning omah?	1	2	3	4	5
4. Opo Bambang njahit celono kuwi?	1	2	3	4	5
5. Siti tau manggon ning Ostrali rak?	1	2	3	4	5
6. Opo Ulum mepe klambine yo?	1	2	3	4	5

References

Abtahian, Maya Ravindranath & Cohn, Abigail C. & Pepinsky, Thomas. 2016. Modeling social factors in language shift. *International Journal of the Sociology of Language* 242:139-179.

AnderBois, Scott & Henderson, Robert. 2015. Linguistically establishing discourse context: Two case studies from Mayan languages. In Bochnak, Ryan M. & Matthewson, Lisa (eds.) *Methodologies in Semantic Fieldwork*, 207-232. Oxford: Oxford University Press.

Arunachalam, Sudha. 2013. Experimental Methods for Linguistics. *Language and Linguistics Compass* 7/4: 221–232, 10.1111/lnc3.12021

Bard, Ellen Gurman & Roberston, Dan & Sorace, Antonella. 1996. Magnitude estimation of linguistic acceptability. *Language* 72, 32-68.

Bornkessel-Schlesewsky, Ina, & Schlesewsky, Matthias. The wolf in sheep's clothing: Against a new judgement-driven imperialism. *Theoretical Linguistics* 33.3 (2007): 319-333.

Bowern, Claire. 2015. *Linguistic Fieldwork: A practical guide*. 2nd edition. New York: Palgrave Macmillan.

Brugman, Claudia & Conners, Thomas J. 2018. Distinguishing properties of SMS and Twitter in Indonesian: a contrastive study. *Digital Scholarship in the Humanities*. fqy028, https://doi.org/10.1093/llc/fqy028

Burton, Strang & Matthewson, Lisa. 2015. Targeted construction storyboards in semantic fieldwork. In Bochnak, M. Ryan & Matthewson, Lisa (eds.) *Methodologies in Semantic Fieldwork*, 135-156. Oxford: Oxford University Press.

Chelliah, Shobhana. 2013. Fieldwork for language description. In Podesva, Robert J. & Sharma, Devyani (eds.), *Research methods in Linguistics*, 51-73. Cambridge: Cambridge University Press.

Chelliah, Shobhana, & de Reuse, Willem J. 2011. *Handbook of Descriptive Linguistic Fieldwork*. Dordrecht: Springer.

Chen, Sihwei & Hohaus, Vera & Laturnus, Rebecca & Louie, Meaghan & Matthewson, Lisa & Rullmann, Hotze & Simchen, Ori & Turner, Claire K. & Vander Klok, Jozina. 2017. 'Past possibility cross-linguistically: Evidence from 13 languages.' In Arregui, Ana & Rivero, Maria-Luisa & Salanova, Andres (eds.), *Modality across syntactic categories*. Oxford: Oxford University Press.

Cole, Peter & Hara, Yurie & Yap, Ngee Thai. 2008. 'Auxiliary Fronting in Peranakan Javanese'. *Linguistics*, 44: 1-43.

Conners, Thomas J. 2010. Death of Linguistics Subsystems: Javanese Speech Levels. (Paper presented at *Humanities of the Lesser Known*, Lund University, Lund, Sweden, 10–11 September 2010.)

Conners, Thomas J. In Press. Javanese Undressed: Isolating Phenomena in 'Peripheral' Dialects. In Gil, David & McWhorter, J. (eds.), *Austronesian Undressed*. Oxford: Oxford University Press.

Conners, Thomas J. & Vander Klok, Jozina. 2016. 'On language documentation of colloquial Javanese varieties.' In Hrics, Lindsay (ed), *Proceedings of 2016 Annual Conference of the Canadian Linguistics Association* (CLA-ACL), 1-12. http://cla-acl.ca/actes-2016-proceedings/

Dollinger, Stefan. 2015. *The Written Questionnaire in Social Dialectology*. Amsterdam: John Benjamins.

Dwyer, Arienne M. 2006. Ethics and practicalities of cooperative fieldwork and analysis. In Gippert, Jost & Himmelmann, Nikolaus P. & Mosel, Ulrike (eds.), *Essentials of Language Documentation* (Trends in Linguistics: Studies and Monographs 178), 31–66. Berlin: Mouton de Gruyter.

Ekowardono, B. Karno & Suprapti & Hartono, Bambang & Setyono. 1999. *Modalitas dalam Bahasa jawa* [Modality in Javanese]. Jakarta: Pusat Pembinaan dan Pengambangan Bahasa, Departemen Pendidikan dan Kudbudayaan.

Errington, J. Joseph. 1988. *Structure and style in Javanese: A semiotic view of linguistic etiquette*. Philadelphia. University of Pennsylvania Press.

Errington. J. Joseph. 1998. *Shifting languages: Interaction and identity in Javanese Indonesia*. Vol. 19. Cambridge: Cambridge University Press.

Featherston, Sam. 2007. 'Data in generative grammar: The stick and the carrot.' *Theoretical Linguistics* 33.3: 269-318.

von Fintel, Kai, & Gillies, Anthony S. 2007. An opinionated guide to epistemic modality. In *Oxford Studies in Epistemology* 2: 32-62.

Franks, Scott & Gessner, Suzanne. 2013. *A Guide to Language Policy and Planning for B.C. First Nations Communities.* Brentwood Bay, B.C.: First Peoples' Cultural Council.

Goebel, Zane. 2002. Code choice in interethnic interactions in two urban neighborhoods of Central Java, Indonesia. *International Journal of the Sociology of Language* 158:69-87.

Goebel, Zane. 2005. An ethnographic study of code choice in two neighbourhoods of Indonesia. *Australian Journal of Linguistics* 25:85-107.

Hacquard, Valentine. 2011. Modality. In *Semantics: An international handbook of natural language meaning*, eds. C Maienborn, K von Heusinger & Paul Portner, 1484-1515. Berlin: Mouton de Gruyter.

Hatley, Ron. 1984. Mapping cultural regions of Java. In *Other Javas: Away from the kraton*, eds. Ron Hatley, Jim Schiller, Anton Lucas & Barbara Martin-Schiller, 1-32. Clayton, Victoria: Monash University.

Hoogervorst, Tom. 2010. *Describing Surabaya's Linguistic Ecology*. M.A. Thesis, Leiden University.

Krausse, Daniel. 2017. A description of Surabayan Javanese with special reference to its linguistic etiquette. M.A. Thesis: Goethe Universität.

Linn, Mary S. 2004. Handbook 3: Conducting a language survey. *Awakening our languages: ILI Handbook Series.* New Mexico: Indigenous Language Institute.

Matthewson, Lisa. 2004. On the methodology of semantic fieldwork. *International Journal of American Linguistics* 70:369-415.

Nahhas, Ramzi W. 2007. *Steps of language survey: an outline of practical methods.* Ms. Payap University, SIL International. http://li.payap.ac.th/index.php?option=com_content&id=14

Oetomo, Dede. 1990. The Bahasa Indonesia of the middle class. *Prisma* 50: 68-79.

Palmer, Frank Robert. 1986. *Mood and modality*: Cambridge textbooks in linguistics. Cambridge, New York: Cambridge University Press.

Podesva, Robert J. & Devyani Sharma (eds.) 2013. *Research Methods in Linguistics*. Cambridge: Cambridge University Press.

Poedjosoedarmo, Soepomo. 1968. Javanese speech levels. *Indonesia* 6:54-81.

Portner, Paul. 2009. *Modality*. Oxford; New York: Oxford University Press.

Rullmann, Hotze, Lisa Matthewson, & Henry Davis. 2008. Modals as distributive indefinites. *Natural Language Semantics* 16.4: 317-357.

Robson, Stuart. 2014. *Javanese* Grammar *for students, 3rd edition*. Glen Waverley: Monash Papers on Southeast Asia.

Samidjan, Hartono. 2013 *Halah Pokokmen: Kupas tuntas dialek semarangan*. Semarang: Mimbar Media.

Schilling, Natalie. 2013. Surveys and interviews. In Robert J. Podesva & Devyani Sharma (eds.), *Research methods in Linguistics*, 96-115. Cambridge: Cambridge University Press.

Schütze, Carson T. & Sprouse, Jon. Judgment data. In Podesva, Robert J. & Sharma, Devyani (eds.), *Research methods in Linguistics*, 27-50. Cambridge: Cambridge University Press.

Setiawan, Slamet. 2012. *Children's language in a bilingual community in East Java*. The University of Western Australia, Perth Australia. (Doctoral dissertation.)

Smith-Hefner, Nancy. 1989. A Social History of Language Change in Highland East Java. *Journal of Asian Studies* 48(2):257–271.

Vander Klok, Jozina. 2013a. Pure possibility and pure necessity modals in Paciran Javanese. *Oceanic Linguistics*, 52(2), 341-374.

Vander Klok, Jozina. 2013b. 'Questionnaire on modality for cross-linguistic use.' http://www.eva.mpg.de/lingua/tools-at-lingboard/questionnaires.php

Vander Klok, Jozina. 2014. On the use of questionnaires in semantic fieldwork: A case study on modality. In Belkadi, Aicha & Chatsiou, Kakia & Rowan, Kirsty (eds.), *Proceedings of Language Documentation and Linguistic Theory 4*, 1-11. London: SOAS.

Vander Klok, Jozina. 2015. 'The dichotomy of auxiliaries in Javanese: Evidence from two dialects.' *Australian Journal of Linguistics* 35.2, 142-167.

Vander Klok, Jozina. 2017. 'Types of polar questions in Javanese.' *NUSA: Linguistic Studies in and around Indonesia*, 63, 1-44. http://hdl.handle.net/10108/91748

Vander Klok, Jozina. *to appear*. The Javanese language at risk? Perspectives from an East Java village. *Language Documentation & Conservation*.

Vander Klok, Jozina & Matthewson, Lisa. 2015. Distinguishing *already* from perfect aspect: A case study on Javanese *wis*. *Oceanic Linguistics*, 54.1, 172-205.

Vander Klok, Jozina & Ahsanah, Finatty & Sayekti, Wuri. 2017. 'Cross-dialectal variation of Javanese particles in yes-no questions.' (Paper presented at *International Symposium on Languages of Java* (ISLOJ 6), Universitas Dian Nuswantoro, Semarang, Java, Indonesia. May 18-19.)

Wardhaugh, Ronald. 2015. *An Introduction to Sociolinguistics, 7ᵗʰ edition*. London: Lackwell Publishing.

Wedhawati & Nurlina, Wiwin Erni Siti & Setiyanto, Edi & Marsono & Sukesti, Restu & Baryadi, I. Praptomo. 2006. *Tata Bahasa Jawa Mutakhir* [Contemporary Javanese Grammar]. Yogyakarta: Penerbit Kanisius.

Wolff, John U. & Poedjosoedarmo, Soepomo. 1982. *Communicative Codes in Central Java*. Ithaca: Southeast Asia Program.

Zentz, Lauren. 2015. "Love" the local, "use" the national, "study" the foreign: Shifting Javanese Language Ecologies in (Post-)Modernity, Postcoloniality, and Globalization. *Journal of Linguistic Anthropology* 24 (3), 339–359.