Methodological Tools for Linguistic Description and Typology

edited by

Aimée Lahaussois

Marine Vuillermet

Language Documentation & Conservation Special Publication No. 16

Table of Contents

8

A proposal for conversational questionnaires

Alexandre François

*Langues, Textes, Traitements Informatiques, Cognition
(CNRS – ENS – Sorbonne nouvelle)
& Australian National University*

This paper proposes a new approach for collecting lexical and grammatical data: one that meets the need to control the features to be elicited, while ensuring a fair level of idiomaticity. The method, called *conversational questionnaires*, consists in eliciting speech not at the level of words or of isolated sentences, but in the form of a chunk of dialogue. Ahead of fieldwork, a number of scripted conversations are written in the area's lingua franca, each anchored in a plausible real-world situation – whether universal or culture-specific. Native speakers are then asked to come up with the most naturalistic utterances that would occur in each context, resulting in a plausible conversation in the target language.

Experience shows that conversational questionnaires provide a number of advantages in linguistic fieldwork, compared to traditional elicitation methods. The anchoring in real-life situations lightens the cognitive burden on consultants, making the fieldwork experience easier for all. The method enables efficient coverage of various linguistic structures at once, from phonetic to pragmatic dimensions, from morphosyntax to phraseology. The tight-knit structure of each dialogue makes it an effective tool for cross-linguistic comparison, whether areal, historical or typological. Conversational questionnaires help the linguist make quick progress in language proficiency, which in turn facilitates further stages of data collection. Finally, these stories can serve as learning resources for language teaching and revitalization.

Five dialogue samples are provided here as examples of such questionnaires. Every linguist is encouraged to write their own dialogues, adapted to a region's linguistic and cultural profile. Ideally, a set of such texts could be developed and made standard among linguists, so as to create comparable or parallel corpora across languages – a mine of data for typological comparison.

Keywords: Linguistic fieldwork; methodology; elicitation; idiomaticity; grammar; lexicon; conversation; spontaneous speech; parallel corpora; language typology and comparison.

1.   Presentation: A new type of questionnaire

The field linguist setting out to describe a language can resort to a variety of methods for gathering primary data (Himmelmann 1998, Lüpke 2010). These methods range between the two extremes of a continuum: on one end, highly-controlled data, which tend to be unnatural; on the other end, spontaneous speech that can be idiomatic, yet difficult to control for linguistic features.

After a short overview of these methods, I will propose an approach that attempts to combine the two advantages of control and naturalness – namely, *conversational questionnaires*. The short sample in (1) gives the reader an idea of what this kind of questionnaire can look like.

(1)   A sample of a conversational questionnaire:

1.   A –   What is it you're carrying in your basket?
2.   B –   This? Oh, it's just some fruit I picked in the woods.
3.   A –   Where did you find them?
4.   B –   Well, I walked across the river down over there, and climbed up the hill.
5.   There's quite a few trees up there with ripe fruit.
6.   A –   Oh, great! Let me go and see if I can find some too.

The idea is for the fieldworker to read this sort of exchange aloud as in a theatrical performance, and have native speakers render it in their own language. This method constitutes an efficient manner of obtaining naturalistic and well-formed utterances, in a way that is painless to both linguists and consultants. Each sentence is anchored in a fictitious but realistic context, which reduces ambiguity and misunderstandings to a minimum. Yet while such dialogues encourage the production of idiomatic speech, they also allow the fieldworker to keep control of the precise linguistic features they wish to test.

While conversational questionnaires can be relevant at any stage of fieldwork, they are designed to be useful even in the early stages of language analysis. These drills form an efficient preparation for the further stages of data collection – whether recording oral stories and actual conversations, or observing actual communicative events in the field.

If the linguist has gained enough familiarity with the community's environment and cultural habits, they may start using conversational questionnaires very early on during their fieldwork on a new language. The method provides a substitute to elicitation methods based on wordlists or isolated sentences, and leads to quick progress in the exploration of a new language. While collecting lexical data from various semantic fields along the way, conversational questionnaires can help document a fair portion of a language's grammatical system: from phonological processes to sentence intonation, from simple case frames to subordinate structures, from possessive patterns to TAM markers,

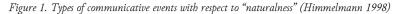from discourse markers to evidentials – to mention just a few.

After presenting the rationale behind conversational questionnaires (Section 2), I will discuss their methodological principles (Section 3). Section 4 will provide and analyze a sample of five possible dialogues, ready to be tested by linguists on virtually any living language. Based on my personal experience, I will highlight the advantages of the method, and propose various ways in which it can contribute to new practices in the domain of language description and typology.

2. The rationale for conversational questionnaires

2.1 The dialectics between control and naturalness

Himmelmann (1998: 185) proposed a typology of communicative events that the linguist may encounter in the field, ranked by degree of "naturalness" (Figure 1). His classification brings to light the crucial correlation between, on the one hand, the degree of control over the data to be analyzed, and on the other hand, their poverty in terms of naturalness.

1. NATURAL COMMUNICATIVE EVENTS
2. OBSERVED COMMUNICATIVE EVENTS
3. STAGED COMMUNICATIVE EVENTS
   ▫ without props
   ▫ with props
4. ELICITATION
   ▫ contextualizing
   ▫ translation
   ▫ judgment

linguistic self-awareness

investigators' control

Figure 1. Types of communicative events with respect to "naturalness" (Himmelmann 1998)

Some elicitation methods have the advantage of keeping a high degree of control over the data collected. This is the case of wordlist elicitation – whether these lists are limited to "basic vocabulary" à la Swadesh, or include richer and more expert vocabulary such as terms for fauna and flora. While wordlists can provide preliminary insights into the lexicon or the phonology of a language, they say little about its actual fine-grained semantics, as they are often distorted by the organization of meaning in the metalanguage (Haviland 2006, Bradley 2007). And obviously, these lists tell us nothing about the language's grammar or phraseology.

Ready-made grammatical questionnaires[1] are often designed for the elicitation of one particular field of grammar. Their unit of observation is generally the sentence, which makes them a preliminary tool for observing what utterances may sound like in the target language. These questionnaires often have the advantage of providing numerous subtle grammatical tests within a clearly circumscribed domain – e.g. possession, quantifiers, subordination, long-distance dependency... But what is convenient for the linguist can be more difficult to process for native speakers. Because they are presented in isolation, the sentences making up these questionnaires are sometimes opaque, or deprived of any meaningful context that could help native speakers make sense of them. Sentences (2)-(4) are taken from existing grammatical questionnaires:

(2) *A cat sleeps.*   (Skopeteas et al. 2006: 229)
(3) *John is continually doing it (on this occasion).*   (Moreno Cabrera 1991: 59)
(4) *Walter thinks that Bill told Harry that Dave respects X.* (X=Walter)
       (Dimitriadis & Everaert 2001: 22)

Admittedly, grammatical questionnaires of this sort are not meant to have every sentence translated directly: they are instead designed to help the linguist think of potential research questions during their own investigation, and meant to be adapted to local contexts. But even if a sentence like (4) has different names substituted in, it will not be much easier to process. Faced with such linguist's constructs as (2)–(4), consultants have no choice but to try and translate them literally, using calques or unusual syntax, at the risk of resorting to unnatural turns of phrase. The contrived sentences that result from such an exercise sound as bad in the target language as they do in English, and fail to reveal the language as it is really spoken. This issue is not just a matter of aesthetics: it raises the profound question of the quality of the data we want to collect – if we are serious about making linguistics an empirical discipline.

Another drawback of isolated sentences is the cognitive cost they impose on speakers. The more isolated an utterance is from any accessible pragmatic context, the more difficult it is to provide a natural, or even correct, translation. A questionnaire consisting of decontextualized sentences can be dull and tiresome to consultants (cf. Chelliah & de Reuse 2011:210). The more difficult the tasks of comprehension and translation, the more likely it is that native speakers will find the whole experience unsettling, as they may feel it exposes the limitations of their linguistic knowledge. They will find greater satisfaction in methodologies that allow them to speak naturally, thereby leading to a smooth experience for all participants involved.

---

[1] See *http://tulquest.huma-num.fr/en* for a collection of such questionnaires; as well as other chapters of this volume.

At the other end of the spectrum, various methods focus on spontaneous speech (see the useful overviews by Lüpke 2010: 67-90, Hellwig 2006: 330-342). Their advantage is to provide natural utterances, with a significant increase in data quality compared to the methods of elicitation mentioned above. This can take the form of "staged communicative events", i.e. semi-controlled experiments such as spoken responses to a short film (e.g. *The Pear story*, Chafe 1980), or to a visual stimulus forming a narrative (e.g. Vuillermet & Desnoyers 2013; Barth & Evans 2017, among many others).

Recording fluent speech using audio or video techniques provides valuable samples of a language, making it possible to capture how it is spontaneously used by native speakers. Whether they consist of narratives taken from the society's oral literature (Finnegan 1992, 2010), in procedural texts, or in conversations, these natural samples constitute the high-quality material most appropriate for the endeavor of language documentation (Himmelmann 1998, Mithun 2001, Woodbury 2011). Once properly transcribed and annotated, these records form a proper empirical corpus that can be mined for examples in a grammatical or lexical description.

Indispensable as they are, high-quality samples of spontaneous speech are more easily collected and analyzed during a later phase of fieldwork, once the linguist has become familiar with the language, enough to understand most exchanges on the fly. My focus here is on an earlier stage of language discovery, when little is known of the language, and it is still too early to start transcribing fluent speech. During this initial phase, the linguist needs to be able to maintain some level of control over the data to be collected.

2.2 The original motivation for conversational questionnaires

If the goal is to hear the language as it is really spoken, the best option is probably to experience full immersion in a community. This was indeed my strategy as I learned to speak the Mwotlap language of Motalava island (Banks Islands, Vanuatu), through a six-month immersion in the field in 1997–98 (François 2001:32–36; 2014:29–37). Such an approach makes it possible to learn phrases and utterances as they occur in day-to-day situations, until one is able to master the language.

Language immersion does not require any ready-made questionnaires; it relies on chance encounters, and on the likelihood of finding oneself in a broad array of real-life situations during the immersion period. As the linguist hears a new utterance, they write it down, and resort to on-the-spot elicitation of whichever new word or pattern they may wish to explore. By conducting various tests, and varying the parameters of the initial utterance (e.g. person or number of participants, type of patient, modality…), one can progressively see the grammatical structures of the language emerge.

Yet this approach based on full immersion can only succeed when spending long periods of time in the field – a luxury that is not always available to linguists.

I felt the need for conversational questionnaires only in 2003, as I was preparing for a second stage of fieldwork discovery. Rather than dedicating many months to the same community, I was planning to spend the next fieldwork seasons exploring linguistic diversity in the Banks and Torres Islands – a particularly multilingual archipelago, with 17 languages for a population of 9,300 (François 2011, 2012). For this exploratory type of survey, I would only be able to dedicate two or three weeks on average to each community.

My main goal was always to collect spontaneous speech in the form of narratives or conversations (§2.1) for each language, which I eventually did.[2] If such an opportunity presented itself, I would record stories, songs or other forms of speech right away, even on day one; but my initial ignorance of each language would mean spending some time extracting the basic structures from the texts themselves, a procedure which was possible yet less than optimal. In order to facilitate my collection and transcription of fluent speech, I felt the need to set up an efficient method for gaining essential knowledge of each language's phonology, morphosyntax and phraseology, ideally in a matter of days.

My personal experience with self-study language books had shown that the most efficient approach to language learning was never through wordlists or isolated sentences, but always through naturalistic snippets of connected conversation.[3] Likewise, my first immersive fieldwork made me realize that linguistic constructions and phraseological strategies come embedded in entrenched routine dialogues, rather than in monologic speech. Face-to-face interaction is central to the life and evolution of linguistic systems; it is the natural context where speakers keep track of each other's referents, negotiate the pragmatics of their utterances, and ensure the success of their speech acts (cf. Levinson 1995; Ochs *et al*. 1996; Chafe 1994, 1997; Mithun 2015).

2.3 Creating and using questionnaires in the field

These circumstances and thoughts prompted me to create my own field questionnaire, adapted to the environment and cultural realities of my field area. My intention was to blend two approaches usually kept distinct – namely, questionnaire-based elicitation, and naturalistic conversation.

---

[2] These surveys resulted in 104 hours of recordings, including 50 hours of narratives, in 23 languages (François 2019:282–4). These are archived online, in open access [https://tiny.cc/Francois-archives].

[3] This is, among others, the method famously designed by Alphonse Chérel when he founded the book series *Assimil* in 1929 (Chérel 1929, 1940).

2.3.1 A customized field questionnaire

My field questionnaire itself took the form of a 42-page document, printed on two sides and bound. I made 20 copies – one for each target language – as I left for the field. Each line was printed so as to leave space for answers (see Figures 2 and 3 below).[4]

The core of the document consists of twelve longer dialogues, with an average of 280 words per text: these are the conversational questionnaires proper – the main object of this paper. Besides that main section, my homemade booklet includes shorter sections covering:

- social and sociolinguistic aspects of the language (ethnonyms, glossonyms, number of speakers, vitality, literacy, intermarriage and contact with other languages…)
- IPA chart, phonemic inventories, notes on the phonology
- (blank charts for tabulating) paradigms or personal pronouns and possessive markers
- (blank) kinship charts for consanguines and affines
- detailed maps of the region for collecting toponyms, wind names, geocentric directionals
- month names in the traditional calendar
- numerals and number expressions
- phrases related to time and weather
- eight shorter dialogues (similar to ex. (1); average length: 60 words)
- a list of 25 land animals, with related lexicon (*cat* → 'meow'; *pig* → 'pig's tusk worn as ornament'; *coconut crab* → 'hunt c. crab'; *spider* → 'spiderweb'; *ant* → 'venomous')
- a list of 18 major sea animals, with related lexicon (*octopus* → 'tentacle'; *turtle* → 'shell'; *crab* → 'claws'; *giant clam* → 'it's shutting itself'; *conch* → 'blow the conch'; *shrimp* → 'shrimp trap'). Supplemented with elicitation kit for fish and sea animals.
- a list of 10 birds and flying mammals, with related lexicon (*fowl* → 'eggs', 'chicks', 'cackle', 'rooster', 'rooster feathers'; *pigeon* → 'nest'; *flying-fox* → 'chew fruit'). Supplemented with elicitation kit for birds.
- plant-related terms (*flower* → 'wither'; *fruit* → 'ripen'; *bark* → 'bark a tree')
- a list of 40 trees and plant species, with related lexicon (*taro* → 'taro garden'; *bamboo* → 'bamboo joint'; *coconut* → 'coconut shell', 'c. milk'; *sago* → 'sago thatch'). Supplemented with elicitation kit for flora.
- a list of 52 body parts, with related lexicon (*heart* → 'heart beat'; *head* → 'headache'; *breast* → 'feed a baby'; *blood* → 'bleed'; *voice* → 'loud'…)
- a list of 40 pairs of common adjectives (*high–low*, *deep–shallow*, *rich–poor*, *sharp–blunt*…)
- a list of 100 Proto-Oceanic reconstructions likely to be reflected in modern languages.

---

[4] My field questionnaires are archived online, on the ODSAS platform (François 2013). The one for Dorig (Gaua, Banks, Vanuatu) is provided in open access [https://tiny.cc/AF_Q_Dorig]. Interested readers can consult it as a sample of a complete questionnaire.

Most sections were meant to collect vocabulary and essential phraseology. They helped me learn the language so as to interact with people, and transcribe recordings efficiently; they also provided material for future dictionaries. But the central section of my elicitation kit was a series of dialogues, which together formed a conversational questionnaire. Apart from some vocabulary, the main aim of those texts was to elicit data on morphology, syntax, pragmatics, and phraseology.

2.3.2 Samples of conversational questionnaires

Figure 2 shows the first lines of a dialogue written in Bislama, the lingua franca of Vanuatu. The title reads *Naef blong mi wea?* 'Where is my knife?'. The handwritten translations are in the Lehali language, spoken on Ureparapara island (Banks group, Vanuatu).

Figure 2. Sample of the author's dialogue questionnaire D05 written in Bislama creole, with translations in the Lehali language of Vanuatu (François, pers. data, 2011)

Figure 3 shows the same field questionnaire filled out for Tanema, one of two moribund languages of Vanikoro island (François 2009), in the Solomon Islands. The name in the top-right corner is that of Lainol Nalo, the last speaker of Tanema.

My initial source of inspiration for writing these dialogues were a number of real-life conversations I had experienced a few years earlier, as I was conducting monolingual fieldwork in Mwotlap. My vivid memories of the spontaneous dialogues, consolidated by the handwritten notes I had taken then, had already inspired the linguistic analyses in my doctoral dissertation (François 2001). In that context, some particular situations had struck me as favoring specific constructions and semantic fields. For example, when I got lost one day in a new village, the questions I asked people around me triggered several

strategies for spatial reference, combining deictic and geocentric directions, various motion verbs and spatial prepositions. Later on, as I prepared my exploration of new languages, the memories of such spontaneous exchanges inspired a three-page dialogue, in which someone asks his/her way around a village, and receives directions based on various spatial strategies.
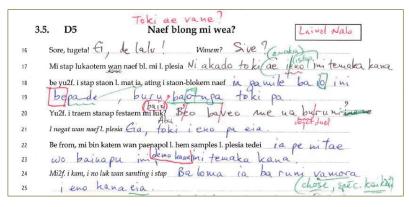

Figure 3. Sample of dialogue D05, with translation in the Tanema language of the Solomon Is. (François, pers. data, 2005)

2.3.3   Choosing the source language

In writing up my conversational questionnaires, I often remembered utterances and turns of phrase initially heard in Mwotlap, and I translated them into the country's national language, Bislama. This English-based pidgin/creole[5] has been adopted since the beginning of the 20th century as the country's lingua franca (Tryon & Charpentier 2004).

I could have used Mwotlap as a lingua franca, considering the number of people across the Torres and Banks islands who can understand it (François 2012:99, 102); yet this would have run the risk of inducing too many calques in the translation between close vernaculars. Admittedly, the structures of Bislama are also very close to its Melanesian substrates, yet they are less strongly anchored in the northern islands, and find their origins in a much broader region (Camden 1979, Tryon & Charpentier 2004). In that sense, this creole can be considered equally distant from all languages in the region I was

---

[5] The term "pidgincreole" proposed by Bakker (2008:138) suits well the status of Bislama. In rural areas of Vanuatu, it is used as a lingua franca – like pidgins; yet in urban settings, it has become the first language of most speakers, which makes it a creole. In 2009, 33.7 percent of Vanuatu families declared using Bislama as their main language of communication at home; 63.2 percent declared using a local vernacular, with Bislama as a second language (François *et al.* 2015:12-13).

studying, and different enough from each variety to reduce the risk of interference. Bislama was here a neutral choice, both sociolinguistically and grammatically.

Using French and English as the source language was, in this case, not a good option. The colonial languages are taught in formal education, yet are hardly ever used in rural areas. Local islanders do not master them sufficiently to have solid intuitions when translating, say, nuances of verbal morphology or lexical semantics. Also, the association of these two languages with the school context would have placed consultant speakers in a situation of linguistic insecurity, and increased their "self-awareness", at the expense of naturalness (cf. Figure 1). The choice of Bislama had the opposite effect. The creole is spoken fluently by everybody in Vanuatu: switching from Bislama to a vernacular is a daily practice in the country, and does not hamper the spontaneity of speech.

For all these reasons, I chose to use Bislama as the source language of my conversational questionnaires.[6]

2.3.4   My typical workflow in the field

As I entered a new language community, I would begin by making friends with the people, socializing with various age groups and families, explaining what my plans were. People were curious about the work of a linguist, and eager to start teaching me their language.

Dialogue D1 (of which an English adaptation is presented in §4.1) was designed to be the first one for each language: it included the most basic greetings ('good morning', 'thank you', 'see you later'), the essential pronouns and verbs, all incorporated in a simple story. Consultants usually enjoyed these dialogues, because they were written so as to combine serious topics with more humorous or lighthearted exchanges (cf. Chelliah & Reuse 2011: 210). Even when I proposed a break after a questionnaire session, oftentimes speakers – especially younger ones – would ask me to "do another story". What I usually did was to use no more than one dialogue per half-a-day session, and alternate those longer texts with shorter sections of the questionnaire (§2.3.1), or to have informal conversations where I'd try practicing what I had just learned.

With an average of two dialogues translated per day, I usually finished collecting the twelve longer texts of my fieldwork questionnaire in less than a week. People of all ages also wished to tell me traditional stories, which I started recording early on – sometimes from day one if they were impatient. As for transcribing those stories, I usually waited until I had finished learning my first "lessons" in the language – that is, translating four or five whole dialogues from my questionnaire. These work sessions alternated with periods

---

[6] In the remainder of this paper, the examples of questionnaires will be given in English.

of rest, walking around the island, chatting with various people in the villages, playing cards or sports, immersing myself in the community in various ways.

Not all dialogues have been translated in every language I worked on. Depending on fieldwork circumstances and priorities, I sometimes skipped some texts, and focused on what was more urgent.[7] For example, my questionnaire for the Lemerig language (Vanua Lava island) only has seven dialogues translated out of twelve, because the few days I was able to spend in 2006 with †Taitus Sërortelsöm – the language's last speaker – were precious: we chose instead to focus on transcribing my recordings from the oral literature.

3. Methodological aspects of conversational questionnaires

I will now expose the methodological principles and advantages of conversational questionnaires. Going beyond the special circumstances of their genesis in my fieldwork experience, this section will take a broader perspective, and examine how this method can be usefully generalized to other field settings. Conversational questionnaires can, in principle, enrich the toolkit of descriptive linguists anywhere in the world, whether they study spoken or signed languages, endangered languages or major tongues, rural or urban cultures.

3.1 Between elicitation and connected speech

3.1.1 Connected speech and idiomaticity

The point of departure of a questionnaire is a naturalistic dialogue representing a plausible real-life situation. A text is prepared ahead of fieldwork, written in English or in the local lingua franca. An English sample is provided in (1) in Section 1 above.

The linguist's work consists in getting consultants to render the dialogue as appropriately as possible in their native language. While sentence-based questionnaires tend to elicit literal, word-by-word translation, and therefore run the risk of sounding unnatural (§2.1), conversational questionnaires are sufficiently tied to daily situations to hopefully bring out the most naturalistic formulations.

The point is not so much to translate a written text as it is to reenact a specific conversation, and elicit the most spontaneous utterances that occur in that context. Thus, if we come back to dialogue (1), the idea is not to translate this line literally:

---

[7] Out of the 204 potential texts to be collected (12 dialogues × 17 languages), we actually translated 144, that is, 70 %.

(1')   **Oh, great**! Let me go and see if I can find some too.

Rather, the linguist's role is to explain the hypothetical situation – which is often self-evident from the dialogue – that speaker A starts her utterance by expressing excitement when she hears about a spot in the woods where fruit is ripe and abundant. How would such excitement be expressed in the target language? It could take the form of a statement like *this is very good* or *I can't believe it*; of an exclamation like *how cool is that!*; or perhaps an interjection such as *hurrah*, or *thank God*. The key advantage of the method is that speakers are not asked to translate words (*How do you say "great" in your language?*) but to mentally tune in to an imaginary dialogue situation, and produce whatever utterance would be most idiomatic.

Because it focuses on naturalistic speech, this approach can generate good-quality data, compared with the translation of isolated sentences. Thus, if I ask a speaker of Lakon – an Oceanic language of Vanuatu (François 2011) – to translate, out of the blue, the sentence "*I do not know*", chances are they will volunteer a literal equivalent (5), which is indeed perfectly grammatical:

(5)   LKN   *Na   tē   roñ   avōh*.
            1SG   NEG₁   hear/know   NEG₂
            'I do not know.'

Now, one of my dialogues had an exclamation 'I don't know!' in context.[8] Interestingly, my Lakon consultants rendered it spontaneously, not with the full sentence (5), but with an unanalyzable word *wē'ĕs!* instead. As it turns out, this is an "ignorative" interjection that is equivalent to Eng. *Dunno!* or *No idea!*, and is always uttered with a shrug.[9] In the domain of controlled elicitation, only a dialogue anchored in a plausible pragmatic situation could provide the opportunity for the spontaneous utterance of such an idiomatic turn of phrase.

3.1.2 The role of the linguist: Control and liberty

In my view, the main role of the linguist here is to channel the work of translation so as to respect two principles: [*a*] faithfulness to the meaning of each utterance in the dialogue; [*b*] freedom to depart from a literal translation, if this favors idiomaticity. These two principles can come into conflict, and it is precisely the role of the linguist to navigate between these two prerogatives.

Let us consider, for example, sentence (1') above *Let me go and see if I can find some too*. The principle of faithfulness implies that the linguist has identified, for this utter-

---

[8] The context was: 'The public phone is down again, when will it be fixed? – I don't know!' [AF.q.Telefon:06].

[9] The neighboring language Mwotlap has a similar ignorative interjection *isi!* (François 2001:1011).

ance, its main contribution to the narrative, as well as its relevance for grammatical elicitation. In the case of (1'), character A expresses her desire to follow the steps of B, and check whether she can also succeed in finding ripe fruit. In pragmatic terms, this sentence is a closure to a short dialogue, as A announces her intention to leave the scene. As far as the grammatical enquiry is concerned, the sentence can potentially provide information on the existence of verb serialization (*let me go and see*), on the expression of a 1sg imperative or hortative (*let me go*, *let me see*), on complement clauses (*see if I can find*), on quantifiers (*find some*), etc.

The linguist must make sure that the consultant is sufficiently comfortable with the translation exercise that they will not try to translate the sentence literally. And indeed, the final sentence does not necessarily need to have a word meaning *go*, or *see*, or *if*, and may well be rendered idiomatically in quite a different way.

Conversely, it may happen that the speaker misses the point of the utterance, and volunteers an inaccurate or incomplete translation (e.g. *I'm going now*, or *I want some too!*) that is too distant from the source version. In that case, the fieldworker should feel free to repeat the initial prompt, and ask the consultant for confirmation that this is the best translation they can think of. While some negotiation is perfectly legitimate, the linguist should refrain from insisting too much, which can be perceived as pressure to produce a literal translation. In my experience, the mere repetition of the stimulus sentence, reenacted by the linguist with a natural tone, is enough to encourage the consultant to provide the best translation possible. A good option here is to work not with one, but at least two speakers, who can complement each other's replies in order to get as close as possible to an idiomatic dialogue in their language.

At the beginning of each work session, I usually took the time to read aloud the whole dialogue in the source language, so as to have the consultants envision the scenario in its entirety; this helped them prepare mentally to render the scene in their language. Likewise, after we went through the whole process of translation, I ended sessions by reading aloud the dialogue we had prepared together, this time in the target language. This was often an occasion for consultants to evaluate the degree of naturalness of the final text. Most often, they nodded in approval, saying "Yes, that's exactly how we'd say it in our language!", which was satisfactory for all. In a few spots, the reading was interrupted by a consultant who came up with an alternative translation that they judged to be more appropriate. That moment of verification was always a very useful way to conclude the session.

In order to assess the consultants' replies as accurate and idiomatic enough, the linguist may need to already have some intuitions about the target language. This is not always possible, especially if the investigation is at an early stage. A solution can be to work out a first translation together, and to give it another go a few weeks later, once the

linguist has made progress in mastering the language, and the consultants are better accustomed to the art of naturalistic translation. This second run provides an opportunity to fine-tune the initial version, and achieve a rendition of the dialogue that is fully satisfactory to everyone.

In principle, the questionnaire can be conducted by a non-linguist – whether a researcher with training in different disciplines, or a member of the community. This could even be seen as an important advantage of conversational questionnaires: they are so intuitive and self-evident that – contrary to most grammatical questionnaires – one does not need linguistics training to use them with consultants. While this is indeed true, some training and practice remain useful to find the proper balance between naturalness of speech, and faithfulness to the function of each utterance.

3.1.3   A new niche in fieldwork methods

Within the typology of communicative events proposed by Himmelmann (Figure 1), conversational questionnaires arguably occupy a new niche in fieldwork methods, halfway between ELICITATION and STAGED COMMUNICATIVE EVENTS.

Conversational questionnaires aim to address the flaws which have often been noted in sentence-based elicitation. Mithun (2001: 45) notes that grammatical questionnaires are heavily dependent on a linguist's expectations of what structures a language may or may not have; and yet, certain language characteristics might only surface in connected speech:

> *"[a]n obvious value of the documentation of natural connected speech is that it permits us to notice distinctions and patterns that we might not know enough to elicit, and that might not even be sufficiently accessible to the consciousness of speakers to be volunteered or retrievable under direct questioning. This material is in many ways the most important and exciting of all."*

Mithun's statement was meant to describe "natural connected speech", but applies equally well to conversational questionnaires. The ignorative interjection which my Lakon speakers volunteered spontaneously when rendering a dialogue is a good example of a linguistic category whose existence was hardly expected, and would probably not have been elicited using classic elicitation methods.

Entire domains fall under the radar of sentence-based elicitation, and only surface in connected conversation. This is particularly true of evidentials (Mithun 2001: 45–8), discourse markers and other devices indicating pragmatic stance or speech acts. To quote Silverstein (1979: 234, cited by Chelliah 2001: 156), these linguistic dimensions are low on the "hierarchy of elicitability". Unless specific elicitation techniques are designed for them (Turnbull 2001), pragmatic strategies are often the neglected garden of fieldwork

elicitation (Grenoble 2007). The very design of conversational questionnaires is meant to address these important flaws inherent to traditional methods of elicitation. This observation will be amply illustrated with the dialogues in Section 4 (see in particular §4.4.2).

The principle of pragmatically-anchored utterances also makes them a potential tool for eliciting data on INTONATION (see Himmelmann 2006). While a given sentence, taken in isolation, can be translated using a default, neutral intonation, a dialogue creates a meaningful context in which a certain prosodic contour naturally comes to mind – whether it encodes surprise, amusement, anger, or any other emotion suitable to the given dialogue situation. Of course, for studying prosody, just like for anything else, no method beats the naturalness of spontaneous conversation; but the conversation-based method provides a combination of idiomaticity *and* control that can be useful to a preliminary observation of intonation.

For similar reasons, conversational questionnaires constitute an interesting option for the description of SIGN LANGUAGES, taking into account all their dimensions – including prosody and facial expressions (Cormier et al. 2010, Dachkovsky et al. 2013, Padden 2015:150) – much better than what is possible under sentence-based elicitation.

3.2 A tool for cross-linguistic comparison

3.2.1 Etic grid and language typology

There are also deeper, theoretical reasons why a dialogue in context should be favored over isolated textbook sentences or wordlists.

A given word may take on different meanings depending on context. Thus in many target languages, an English verb like *carry* has a variety of possible translations, depending on the type of object (carry a baby vs. a bag vs. firewood…), on the exact manner of carrying (carry in both arms vs. on shoulder vs. around neck vs. on head…), on directionality (carry-and-come, carry-and-go, carry upwards…). Obviously, a poorly designed vocabulary list with a "simple" entry *carry* does not provide enough clues to isolate a specific sense (Haviland 2006:153). The common practice is for the linguist to explore the various lexemes volunteered by consultants, by improvising various tests on the spot, so as to understand how that particular field (e.g. verbs of carrying) is organized.

In comparison, a conversational questionnaire targets one specific sense – or a specific "frame" (Rakhilina & Reznikova 2016) – of a polysemous word. Thus, the first sentence of example (1), reproduced here as (6), uses *carry* in a particular sense:

(6)   What is it you're **carrying** in your basket?

In order to explore the semantic domain more broadly, the linguist can legitimately pause the translation of the dialogue, and ask a few questions on the side: can the same verb be used with other forms of objects? with different manners of carrying? If the case is simple, it can be answered right away, and the excursus will be limited. Sometimes, the domain is so rich that it warrants a separate session.

The other senses of an ambiguous English word can also, in principle, be embedded in further dialogues. For example, while (6) illustrates a particular sense of *carry*, another sense can be found in dialogue D5 (§4.5), line #25, shown in (7):

(7)   She's **carrying** her child on her back.

Carry fruit in a basket vs. *carry a child on one's back* describe two separate actions, and these two sentences (6) and (7) thus constitute two distinct data points. The fact that English happens to colexify these two situations using a single word is an interesting property of English, but this should not be taken for granted for other languages.

The same reasoning applies to grammatical morphemes. Many morphemes of English are ambiguous between different possible meanings, and a questionnaire based on isolated sentences remains ambiguous in this respect. A conversational questionnaire, on the other hand, will help the consultant pinpoint a specific interpretation of a given morpheme, without too much effort.

For example, the isolated sentence (2) *A cat sleeps* will be difficult to translate in many languages, due to the ambiguity of the indefinite article in English (specific or not?), and of the general present. Is this a gnomic statement on cats, perhaps truncated (*A cat sleeps longer than a dog*)? Is it supposed to be a habitual action performed by a specific referent (*A cat sleeps on my porch every night*)? Languages may render these two interpretations with different TAM markers, or different noun articles. In the absence of additional context, (2) is an odd sentence, impossible to translate with certainty. Such ambiguity is automatically resolved in a conversational questionnaire.

Thus, to return to the short dialogue in (1), the quantifier *some* in (#2) (*it's just some fruit I picked in the woods*) appears in a [+realis] context, and hence has [+extensional] interpretation (Montague 1970, Moltmann 1997, Zimmermann 2001). This particular token of SOME is functionally distinct from the [-realis] clause (#6) (*Let me go and see if I can find SOME too*), a [-extensional] use which other languages might well translate differently. Eliciting these two sentences in a dialogue can tell us precisely which languages translate (like English) the quantifier *some* with the same form in these two sentences, and which ones draw a distinction between quantifiers in [+realis] vs. [-realis] contexts. For example, Araki – an Oceanic language of Vanuatu –encodes these two uses of Eng. *some* with different morphemes (François 2002:59–69).

In sum, the method of conversational questionnaires paves the way for cross-linguistic comparison, by providing an ETIC GRID from which emic categories can be observed (cf. Evans 2010:509). Each occurrence of an ambiguous lexeme (e.g. *carry*) or morpheme (e.g. *some*), once embedded in a specific context, pinpoints a particular meaning that is defined irrespective of language-specific categories. The onomasiological approach adopted by these questionnaires can then be the starting point for typological comparison.

3.2.2 Conversational questionnaires as parallel corpora

The potential of conversational questionnaires for cross-linguistic comparison is illustrated in Figure 4. This is a sample of my fieldwork database, laid out using SIL's Toolbox program.

Figure 4. Sample of the author's database of conversational questionnaires: a Bislama sentence is translated into 21 languages of Island Melanesia (François, pers. data, 2017)

All the sentences presented in Figure 4 translate the same Bislama sentence, whose English equivalent would be "What for?".[10] This is part of the following exchange in D1 (see an English version in §4.1.1):

(8) 7. A – Well, we're walking down to the river, over there.
    8. B – **What for**? Are you going to be bathing?

The elicitation language, here Bislama [§2.3.3], has a sentence *Blong mekwanem?*, which parses as ⟨PURP do.what⟩ – using an interrogative verb *mekwanem* 'do what'. That elliptic sentence is translated here in 21 Oceanic languages from Island Melanesia: four languages from the Solomon Islands (Temotu province), and 17 from Vanuatu (Banks and Torres Isles), ranked in geographical order from NW to SE, and transcribed in the local orthographies. I created this database simply by compiling my handwritten field notes for individual languages (illustrated earlier in Figure 2 and 3) into a single electronic, searchable text file.

Such a comparative database, based on a shared questionnaire, can be put to various uses. Comparison may involve related languages (as in Figure 4) and serve to showcase, and potentially measure, the internal diversity within a given subfamily. For example, Figure 4 here shows that all the languages tested have a synchronically unanalyzable interrogative verb meaning 'do what' (TEA *mikae*, LVN *wo*, TNM *jive*, TIK *oa*, HIW *tave*…) – except three. Lehali, Vera'a and Olrat, like English, simply combine the verb 'do' with their interrogative word for 'what'.

Such comparison can be conducted in the perspective of historical linguistics, helping the linguist identify shared retentions and shared innovations. A larger database could also involve genealogically diverse languages, and function in the spirit of *parallel corpora* advocated by some typological linguists (Cysouw & Wälchli 2007, Dahl 2007). For example, the sentence in Figure 4 can serve to elicit the existence of interrogative verbs (Hagège 2008) across the world's languages.

3.2.3 Towards a universal thesaurus of speech motifs

Phoneticians and dialectologists have been using Aesop's tale "the North Wind and the Sun" (International Phonetic Association 1912; Boula De Mareüil *et al.* 2017), or the parable of the Prodigal son (Bec 1986; Heeringa & Nerbonne 1999), as a standard text to build parallel corpora across dialects and languages. In the same spirit, linguists could agree on a set of dialogues as the basis for empirical cross-linguistic comparison. The

---

[10] The identifier \rf *d01.Rot:08* indexes the eighth sentence in dialogue D1 [see §4.1.1]. The Bislama word *rot* (<Eng. *road*) echoes the title of D1 in my questionnaire, TRIFELA I MIT LONG ROT "Three people met on the road".

conversational questionnaires designed by linguists would constitute a new genre among elicitation tools, each one with its own purposes and advantages.

If every utterance of every dialogue were assigned an identifier – such as a DOI – it could be cited in a systematic way. For example, a study comparing interrogative verbs across languages could refer to the sort of [+dynamic] use found in {CQ.D01.08}, i.e. *Conversational questionnaire D01, line #08* (cf. Figure 4). In the event that other linguists were to use the same standard dialogue for their own fieldwork, they could then confront that cross-linguistic study with their own data. If we adopt standard questionnaires that are used and reused by fieldworkers around the world, our discipline could come closer to an empirical science that promotes the citability and reproducibility of its research results – a valuable objective in itself (Berez-Kroeker et al. 2018).

The granularity of the database could even be refined, so as to point not to a whole sentence in a dialogue, but to the functional components within it. In the same way that the catalog of Aarne–Thompson–Uther (Uther 2011) provides a reference grid for *folktale motifs* across the world's narrative traditions, likewise a universal thesaurus could index what may be called "*speech motifs*", referring to specific meanings, pragmatic values, speech acts, that can be encoded in languages. For example, the speech motif found in Figure 4 above can be defined as follows:

(9) Example of a SPEECH MOTIF:
     { A tells B that A is going somewhere, and B asks A: 'WHAT FOR?' }.

The work of typologists would then be to list the various strategies used by the world's languages to express that speech motif: a full sentence (*what will you do?*); an elliptic sentence (as in English *what for?*); an interrogative verb (like in Vanuatu languages); an interrogative adverb (*why?*); a prepositional phrase (*for what purpose?*)...

The number of "speech motifs" is potentially infinite, but a comparative database based on existing dialogues and corpora could provide a good start. Such an onomasiological thesaurus would be similar to some cross-linguistic tools used in studies of the lexicon, e.g. the *Concepticon* (List et al. 2016, 2018).

3.2.4 Culture-specific or universal?

Conversational questionnaires mimic conversational routines that can be assumed to take place in the actual use of most languages.

Admittedly, the very nature of conversational routines varies from culture to culture. Not all languages have an equivalent of *How do you do?* or *Bon appétit!*, and one should not assume that verbal greetings, requests, apologies or other speech acts, are carried out in the same way everywhere (Wierzbicka 2003; Trosborg 1994, 2010). One option for the linguist is to design dialogues in such a way that they are specifically tailored to the

typical interactions attested in a given area. In many parts of the world, a phrase like *Where are you going?* is a more standard greeting than *Good morning* (Gil 2015:280, 354), and a dialogue must be able to adapt to this. The same applies to cultural references: a story in Vanuatu may concern the drinking of kava or the preparation of breadfruit, while one for Basque speakers might mention sheep milk cheese or a pelota competition.

Such a process of dialogue customization requires some level of familiarity with the local culture. Depending on circumstances, this may be best carried out by the speakers themselves, or by the linguist – provided the latter has acquired sufficient experience with the local mores. While the dialogues themselves can be prepared in such a way as to be usable on day 1 of fieldwork without requiring any previous knowledge of the language, it is in fact a preferable situation for the linguist to be already well-acquainted with the material culture of the local population, as well as their social practices and pragmatic routines.[11]

Adapting conversational questionnaires to local cultures and practices is highly desirable, and should by all means be pursued. This practice has the added value of combining linguistic documentation with the preservation of cultural knowledge. That said, a drawback of this option is the difficulty of using the same dialogue across different cultural areas – an objective that a typologist might want to pursue [§3.2.3]. A possible compromise is to combine both types of dialogues in a fieldwork project: some stories anchored in local referents, and others written in such a way as to be reusable, and comparable, across continents, with minimal local adaptation. In all cases, the linguist should provide consultants with some leeway to adapt the text to local referents and pragmatics.

3.3 An efficient tool for language learning

The fluency I acquired in Mwotlap during my initial immersion (§2.2) provided me with enough linguistic background to quickly assimilate other Oceanic languages with a similar typological profile – even though their lexical diversity is superior to that of the Romance or Germanic families (François 2011: 203 – cf. Figure 4 above). The format of my custom-made questionnaire, thanks to the density of linguistic information, allowed me to learn the essential vocabulary, constructions and grammatical structures of each new language within the first week of my stay. This efficiency, in turn, helped me socialize with community members; this created a virtuous circle whereby I quickly enriched my

---

[11] The context of a regional linguistic survey, in which the linguist is already familiar with the area's cultural practices yet new to each language (see §2.2), is a situation where such questionnaires can serve in the very early days of visiting a new community.

knowledge not only of the language, but also of the community's oral traditions, songs and poems, cultural tastes, and social dynamics.

Conversational dialogues make for an efficient language-teaching resource not only for the linguist, but also for future learners of the language. These may be native speakers learning literacy, semi-speakers in a context of language revitalization, or L2 students wishing to discover a new language. The pedagogic power of such dialogues is indeed to be expected, considering that the approach was initially inspired by the principle of self-teaching methods (see fn.3). Conversational questionnaires are thus easily converted to pedagogic resources, in the spirit of the teach-yourself grammars that exist already.[12]

A number of conversations elicited using these questionnaires were included in some of the literacy materials I created for Vanuatu languages (François 2004–15; 2019:286-9). After hearing about my questionnaires, Aimée Lahaussois (p.c., July 2018) reports that she, together with Guillaume Jacques, successfully tested the method in her fieldwork on Khaling (Nepal): community members created new dialogues typical of their daily interactions, and used them as materials for language revitalization with the younger generations.

This sort of application of conversational questionnaires can have language learners perform short "drama" scenes based on the scripted dialogues – thereby meeting the proposals made by Nathan & Fang (2009: 155–7). As these authors note, this sort of practice helps bridge language description and language pedagogy, a valuable objective especially in case of language endangerment.

In sum, conversational questionnaires belong to several genres at once. They provide raw data for the linguist aiming to describe a given language. They can form the basis of parallel corpora for language comparison, whether historical or typological. And they can be easily adapted into teaching resources, for language learners of various profiles.

4.   Five conversational questionnaires

In the remainder of this paper, I will present a selection of five conversational questionnaires, as an illustration of the method, and inspiration for future linguistic elicitation. Each text is followed by instructions on how best to use that specific dialogue, and comments highlighting how it can contribute to the exploration of certain linguistic domains.

---

[12] In the Pacific domain with which I'm more familiar, my inspiration for dialogue-based questionnaires came from several pedagogic resources I had personally used – including Dutton & Voorhoeve (1974) for Hiri Motu; Dutton & Thomas (1985) for Tok Pisin; Lercari *et al.* (2001) for Drehu; Paia & Vernaudon (2003) for Tahitian.

The first dialogue D1 is an English adaptation of the Bislama script that was part of my own field questionnaire in Vanuatu. The other dialogues presented here (D2–D5) were created especially for the present publication.

4.1   Dialogue D1 – We're going fishing

In principle, D1 can serve as the first dialogue to be used in the field, because it covers the essentials of a basic conversation. If the linguist prefers to start with a shorter session, they can achieve a less elaborate version of D1 by omitting lines #21–30, or even #11–30. A rendering of D1 in the Oceanic language Mwotlap (Vanuatu) is provided in the Appendix.

4.1.1   The text

TITLE: *We're going fishing*.
CONTEXT: A man [A] and his wife run into a woman [B] on the road.

1.   A –   Good morning!
2.   B –   Hello!
3.   A –   How are you$_{SG}$?
4.   B –   I'm fine, and you$_{DU}$?
5.   A –   Oh, we$_{EXC:DU}$ are fine too.
6.   B –   Where are you$_{DU}$ going?
7.   A –   Well, we$_{EXC:DU}$'re walking down to the river, over there.
8.   B –   What for? Are you$_{DU}$ going to be bathing?
9.   A –   No, no! We$_{EXC:DU}$ won't be bathing.
10.   We$_{EXC:DU}$'re going fishing.
11.   We'll try to catch some river fish for dinner.
12.   B –   Oh I see. Is there a celebration in the village, or something?
13.   I didn't know.
14.   A –   No, there's no celebration.
15.   We$_{EXC:DU}$ just need food for our$_{EXC:PL}$ family.
16.   My wife had bought a chicken at the market the other day,
17.   but our$_{EXC:DU}$ children$_{DU}$ ate it all already:
18.   now we$_{EXC:PL}$ have nothing left at home!
19.   Besides, we$_{EXC:PL}$ don't have enough money any more.
20.   We$_{EXC:DU}$ really have to go fishing today.
21.   B –   I understand.
22.   I'm sorry that you$_{PL}$ have no food left.

23.   I can help you_PL, my friends.

24.   My husband and I, we_EXC:DU have lots of vegetables from our_EXC:DU garden,

25.   which he harvested yesterday, and brought back home.

26.   If you_DU want, we_EXC:DU can give you_PL some.

27. A – Oh really? Thank you_DU very much!

28. B – You_DU're welcome.

29. A – You_DU'll give us_EXC:PL some vegetables, and we_EXC:DU'll give you_DU some fish.

30.   Tonight we_INC:PL will all eat together.

31. B – Great! Good bye!

32. A – See you_SG later!

4.1.2   About this dialogue

The story is meant to be applicable to many cultures, as river fishing tends to be a widespread activity in the world – unlike ocean fishing, for example, which is restricted to coastal cultures. The attempt to make this dialogue as universal as possible entails that no specific species are named: #11 mentions 'river fish', #24 'vegetables'. Linguists and language consultants should feel free to adapt these sentences to local fauna and flora, and replace generic 'vegetables' with the name of whatever the staple food is in the local culture. The selection of the proper verb in #25 will depend on this choice, whether 'harvest' or 'pick', and so on. If some elements of the text (river fishing, food gardening, market, money…) are absent from the local culture, an adapted dialogue can be proposed.

4.1.3   Linguistic notes

Dialogue D1 allows the linguist to efficiently collect a large quantity of data on the language, at several levels of analysis. The following provides an overview of the type of material that can be contributed by this dialogue.

4.1.3.1 *Lexicon*

D1 focuses on basic morphosyntax, and puts little emphasis on vocabulary. As a result, many sentences deliberately employ the same simple lexemes, so as to yield contrastive clauses that focus on the morphology; thus compare lines 4 with 5; 8 with 9; 10 with 20; etc.

In spite of the apparent lexical simplicity, D1 does bring forth a number of lexical items, which constitute a good start for a first text:[13]

---

[13] Items in parentheses are words whose translation equivalent is likely to employ different strategies depending on the language. Thus, while 'walk' will probably be translated by a verb everywhere, only a minority of languages possess a verb 'have'; its negation may be a separate verb meaning 'lack', or may

- VERBS: go, walk, bathe, fish, catch, buy; (dine), eat_INTR, eat_TR; harvest; bring; (have), (lack); know, (not.know), understand; (try), need, want; help, give
- NOUNS: food, river, fish, chicken, vegetables, market, family, house/home, village, celebration/party, money; husband, wife, children, friends, garden
- QUANTIFIERS: all_INAN, all_ANIM, enough, some, nothing
- ADVERBS: today, tonight, yesterday, the other day; already; together; (really)
- LOCATIVES: down; over there; (to the river); (home)

As it is, D1 provides only some of the morphological forms of these lexemes: 'village' (12) appears in a locative phrase; 'know' (13) is found exclusively in a negative clause; 'children' (17) only appears as an agentive phrase in the dual, etc. Depending on the profile of the language, it may be easy or complex to infer from these surface forms the other elements of the paradigm – e.g. 'child' in the singular, 'children' in other cases… This investigation can be carried out by the linguist through elicitation, either as a follow-up to D1, or through separate dialogues.

4.1.3.2 *Phraseology*

Besides individual lexemes, D1 collects a good deal of common phraseology:

- GREETINGS   good morning; hello; how are you; good bye; see you later…
- INTERJECTIONS   thanks; you're welcome; great!; no, no!
- DISCOURSE PARTICLES oh; well; besides; really; but
- VOCATIVE   my friends

Some formulations, which in English take the form of a constructed clause, may be rendered in the language by different strategies (§3.1.1). For example, what is given in (12) as "*Oh I see*" might be translated in some languages by the 1SG form of a verb 'see' (or 'understand' or 'know') like in English; but in other languages, it might be more idiomatic to use an interjection (something like 'Alright!' or 'True!' or 'Yes'), or perhaps just a vocal gesture (e.g. [ʔoːˑ], or a click).

Likewise, in (3) "*How are you?*" of course must not be translated literally: its natural equivalent might be something like '*Are things good?*', while its answer in (4) could be a formula such as '*Thank God*' or '*Peace only*'… In (27), some languages lacking a *Thank you* interjection might still have an idiomatic way to express gratitude in that particular situation: e.g. *You are very kind* or *God bless you*. Some cultures might rather opt for a mere gesture or a facial expression. The more idiomatic the formulation, the better.

---

resort to a negative existential for example. The equivalent of Eng. 'try' may be a conative affix or an adverb; 'really' in (20) may be encoded by prosodic strategies; etc.

4.1.3.3 Personal pronouns, number, functions

While it would be difficult, and unnatural, to create a story featuring all possible personal pronouns, D1 can help elicit a fair number of them, for various persons and functions. Table 1 lists the various places in D1 which enable the elicitation of specific pronominal forms.

	A	O	S	possessor
1sg	[12], 13, 21, 26	26	22	16, 24
2sg		[32]	[3]	
3sg	[16], 25	17	[31]	
1exc:du	11, 15		[5], 7, 9, 10, 20	17, 24
1inc:du				
2du	26, 29	29	6, 8	
3du	[17]			
1exc:pl	[18, 19]	29		15, [18, 19]
1inc:pl			30	
2pl	[22]	23, 26		[22]
3pl				

Table 1. Personal pronouns potentially elicited by Dialogue D1

For presentational purposes, the default assumption here is a language with a contrast of clusivity (Filimonova 2005), and three numbers (singular, dual, plural) – but of course other systems are possible. Considering the ambiguity of Eng. pronouns *we* or *you*, each form is tagged for its status: e.g. "*If you$_{DU}$ want, we$_{EXC:DU}$ can give you$_{PL}$ some*". The situation depicted in the dialogue should make it clear which number applies in each case: singular, dual, trial, paucal, or plural (Corbett 2000). The dialogue suggests that character A and his wife have two children (cf. dual on 'children' in (#17)), hence the plural forms of (#15-29) refer to a family of four. Character B is married, but does not seem to have children. The plural form in (#30) refers to a group of six people.

Apart from number and person, Table 1 sorts the various pronominal occurrences into four general semantic roles: A (most agent-like argument in a transitive clause), O (most patient-like argument in a transitive clause), S (sole argument in an intransitive clause); and possessor.[14] This sorting is merely indicative, and does not preempt language-

---

[14] Numbers in square brackets correspond to sentences for which the presence of a person index will most likely depend on language-specific phraseology. For example, (19) *We don't have money* may be translated in some languages by means of a verb 'have' (or 'lack') with a 1pl pronoun encoded as an A argument; but other languages may resort to a possessive structure of the type *Our money is lacking*, where 1pl has the role of a possessor (cf. Heine 1997:58).

specific constructions. It simply provides the fieldworker with a checklist of which persons are attested for which syntactic function in D1, and indicates where the potential gaps are in the data.

The blanks in Table 1 may be filled by the linguist later on, either through elicitation or through new dialogues. Should a story be written to complement it, it will ideally pick a situation favoring the presence of inclusive pronouns (whether 1INC:DU or 1INC:PL) as well as third person pronouns, since these are the ones lacking most from D1. Dialogue D2 (§4.2) will partly fill that gap.

4.1.3.4 Verbal morphosyntax

Apart from personal pronouns, other potential aspects of a language's morphology are likely to be highlighted in dialogue D1 – particularly verbal morphology.

Negation is found in numerous sentences, including: negative interjection *No, no!* (#9); negation of a realis verb (#13), of an irrealis verb (#9); negative existential (#14, 19).

D1 brings together various categories of tense, aspect and modality: e.g. past or perfect (#16, 17, 25); stative (#5, 15, 22); progressive (#6, 7); future (#8, 11, 29, 30); deontic (#20); potential (#23, 26); conditional (#26).

The text also has the potential to elicit serial verbs, or patterns of associated motion if they exist: e.g. *go fishing* (#10), *he harvested [it] and brought back home* (#25).

In terms of case frames, D1 has samples of:

- monovalent verbs, both stative-patientive (*be fine*) and agentive (*go, walk, bathe, celebrate*);
- bivalent verbs: *buy, eat, catch, harvest, bring, help, want, understand*;
- trivalent verbs: *give*.

This first text should provide enough data to determine the main patterns of syntactic alignment in the target language – both primary alignment for transitive verbs, and secondary alignment for ditransitives (Haspelmath 2005; Malchukov et al. 2010).

4.2. Dialogue D2 – Preparing for the New Year

4.2.1 The text

TITLE: *Preparing for the New Year*
CONTEXT: Two adults, A and B, are speaking together.

1. A – Hey, you know what? In two months' time, we'll have celebrations for the New Year.
2. B – Oh, that's right. This time, the people from village X

    will all be coming to our community:

3.		men and women, children, old people, entire families...
4.	A –	We should welcome them in a friendly way,
5.		just like they did to us last year.
6.	B –	Yes of course, like we always do.
7.	A –	Our leaders will start with a welcome speech.
8.		Then, people will pray in the morning.
9.		After that, we'll all share lunch together.
10.	B –	In the afternoon, I hope we can have songs and dances.
11.		The people of X will sing their songs, dance their dances;
12.		we too, we shall sing our own songs, and do our own dances...
13.	A –	Yes, good idea. Everyone loves music and dance.
14.	B –	As we find some rest, our elders will be able to tell stories from the olden times, for the young to hear.
15.	A –	This will be great. They know so many stories.
16.	B –	And then, in the late afternoon, our two communities will part again.
17.		That will be the end of the day of celebration.
18.	A –	Remember that you and I are in charge of the organization this year.
19.	B –	Perhaps we_INC:DU should call a meeting tomorrow morning in the community
20.		and tell our people what they should do.
21.	A –	Some of us can clean the village area,
22.		set up the place for the celebration.
23.	B –	Other people can make the costumes for the dances.
24.		We must choose which dances to showcase;
25.		and we must rehearse them!
26.	A –	Also, don't forget: we'll have to prepare food,
27.		enough food for two hundred people.
28.	B –	Oh dear, that will be a lot of work for us all.
29.	A –	A lot of work indeed. Let us start today!

4.2.2   About this dialogue

D2 is meant to be as universally applicable as possible. This is not easy, due to the culture-specific components typically associated with collective celebrations.

    The "New Year" was chosen as one of the few sorts of celebrations that are widespread across religions and cultures; of course, this does not necessarily coincide with the New Year of the Western calendar, and corresponds to a more general notion of annual cycle which may be accompanied by some form of celebrations. This story evokes a particular kind of event, with one community visiting another one, performing dances and so on; while this was inspired by New Year celebrations at my Vanuatu field site, it is possible

that in other civilizations, the closest equivalent to such a scenario may be linked to a moment distinct from the New Year, such as a wedding or a funeral. In that case, the translation may choose to adapt the text in this respect.

    Line 8 uses a generic verb *pray*, which could in principle apply to different religions and worshipping traditions. As per the principles suggested in §3.2.4, the text could be somewhat adapted to local mores – e.g. *pray* could be replaced by *go to the mosque*, or *make offerings*, or by terms related to other similar collective rituals.

    The story leaves in blank the name of village X; this is meant to be filled by whichever real village or community would fit the description in the target context. This sort of local customization of stories tends to be well received by local communities, and makes the questionnaire a less impersonal exercise.

    The details of the dialogue can indeed be adapted to local contexts – with the caveat that the closer it remains to the original, the easier it will be to later compare the story across languages.

4.2.3   Linguistic notes

The story in D2 supplements D1 with respect to pronouns. The default interpretation of 1st person pronouns throughout the text is inclusive plural (*we'll have celebrations; we should welcome them; they did to us; we'll all eat together; work for us all...*). The text has also 1st inclusive dual (*you and I are in charge...; we should call a meeting*). Many verbs represent 3rd person plural (*they did to us; our leaders will start; people will pray; the elders will be able to tell stories..., for the young to hear; everyone loves music*).

    The dialogue features various quantifiers: *we'll all have lunch; everyone loves; two hundred people; some of us; a lot of work*. A language distinguishing paucal from plural (cf. Corbett 2000: 23) might well exploit that distinction to render some contrasts in this story, e.g. between the smaller groups within the village (*some of us can clean...*) and the larger groups made up of the two communities (*we'll all have lunch together...*).

    The text elicits numerous time expressions: *in two months' time; this time; last year; then; after that; in the afternoon; tomorrow morning; today.*

    The modality is mostly irrealis, including:

•	predictive future:	*we'll have celebrations; people will pray; we'll all have lunch; that will be the end; it will be a lot of work,*
•	deontic:	*we should welcome them; we should call a meeting; we must choose; we must rehearse...,*
•	potential:	*our elders will be able to tell stories* (i.e. 'have the opportunity'); *some of us can clean...,*
•	hortative:	*let us start today,*
•	imperative:	*remember,*
•	prohibitive:	*don't forget.*

The dialogue elicits a number of subordinate structures: *welcome them just like they did to us*; *tell stories for the young to hear*; *remember that we are in charge*; *tell our people what they should do*; *choose which dances to showcase…*

4.3 Dialogue D3 – Seeing the doctor

4.3.1 The text

TITLE: ***Seeing the doctor***

CONTEXT: A patient [A] visits the doctor [B].

1.   A –   Good morning Doctor, how are you?
2.   B –   I'm fine, and you?
3.   A –   Well… I'm not feeling well these days.
4.   That's why I came to see you.
5.   B –   What's happening to you? You look sick.
6.   A –   I can't sleep well at night.
7.   I sweat, I have nightmares, and then I wake up in the middle of the night.
8.   Sometimes I feel hot, sometimes I'm cold. I must have fever?
9.   B –   Let me check your forehead… Oh yes, you're hot!
10.   Do you cough?
11.   A –   No, I don't cough.
12.   But every time I wake up, I'm very thirsty; I feel I need to drink.
13.   And also, my belly hurts. It's painful.
14.   B –   Does it hurt during the day? or only at night?
15.   A –   Mostly at night. I don't know why.
16.   Doctor, I'm a bit worried: what is going on?
17.   B –   Did you eat anything particular lately?
18.   A –   Hm, let me remember… No, I don't think so.
19.   Oh wait, actually yes I did!
20.   Last week, my child came back from the forest
             with some strange fruit I had never seen.
21.   He gave them to me, for me to try.
22.   B –   Did you?
23.   A –   Yes I did. Actually I liked it, it was sweet. I ate many of them.
24.   But then, I became sick after that.
25.   B –   I see. It must have been that fruit that made you sick.
26.   If you hadn't eaten so much, you wouldn't have gotten sick like this.
27.   A –   Oh Doctor, you're right. I shouldn't have.
28.   What should I do now?

29.   B –   Don't worry. I'll give you some medicine for you to drink.
30.   You will take it twice a day: once in the morning,
             and once again in the evening, after dinner. Alright?
31.   A –   Alright Doctor. And then I'll be better?
32.   B –   Yes, you should get better soon.
33.   This is efficient medicine against fever and belly-ache.
34.   Also, you must get some rest.
35.   Don't go to work: you need to sleep.
36.   A –   Alright Doctor, I understand. I'll get some rest.
37.   Thank you so much!

4.3.2 Linguistic notes

The reader will be able to see the interest of each new dialogue in the same spirit as D1 and D2 above. I will only provide a few hints here.

Dialogue D3 focuses on the expression of experiencer predicates, physical affects, sensations and feelings: *I sweat*; *I have fever*; *I have nightmares*; *I'm thirsty*; *I'm cold*; *I feel hot*; *I'm not feeling well*; *my belly hurts*; *it's painful*; *I'm worried*; *I liked it*; *you'll get better …*

D3 should help elicit various Tense and aspect meanings:

- habitual stative: *sometimes I feel hot; I'm very thirsty; my belly hurts*
- past stative: *I liked it, it was sweet*
- present stative: *you're hot; I don't know; I'm worried; I understand; you look sick*
- future stative: *I'll be better*
- habitual dynamic: *I wake up; do you cough?*
- past dynamic: *did you eat lately?; my child came back; I ate many of them*
- past inchoative: *I became sick*
- future dynamic: *you will take it*

The text illustrates an array of irrealis modalities:

- deontic: *what should I do now?; you must get some rest; you need to sleep…,*
- promissive: *I'll give you some medicine; I'll get some rest*
- hortative: *let me check, let me remember*
- prohibitive: *don't go to work*
- counterfactual: *if you hadn't eaten…, you wouldn't have gotten sick*

Several clauses illustrate epistemic modality and evidentiality: *you look sick*; *it must have been that fruit*; *you should get better*; *I must have fever*.

Dialogue D3 elicits a number of adverbial phrases:

- time phrases   *these days*; *lately*; *last week*; *at night*; *in the middle of the night*; *during the day*; *twice a day*; *in the morning*; *in the evening*; *after dinner*; *sometimes*; *every time*; *after that*; *never*; *soon*…
- non-temporal adverbs *mostly*; *a bit*; *actually*; *also*.

In the domain of noun phrases, D3 illustrates various forms of possession: kinship (*my child*), body parts (*your forehead*, *my belly*). In those languages that grammatically treat body parts as inalienable, the presence of a possessor is normally obligatory: a "head" or a "belly" will always be mentioned with a specific possessor. In this perspective, sentence #33 is meant to elicit a particular sort of construction, namely the case – quite rare in discourse – when an inalienable noun is used in a generic sense, and thus has no specific possessor. *Medicine against belly-ache* is a phrase which mentions a body part, yet makes it impossible to retrieve a possessor. While many languages may simply use the bare, unpossessed noun (*belly*), others may have to resort to special morphology here. This is especially true of Oceanic languages: thus Mwotlap will have to use here a special suffix *-ge* to fill the possessive slot, while encoding the absence of any referential possessor: *na-tqe-ge* '[s.o.'s] belly' (François 2001:526-545). This sort of construction turns out rarely in a corpus of spontaneous speech, but can be usefully elicited by means of a conversational questionnaire.

These are just some of the linguistic highlights of this dialogue.

4.4.   Dialogue D4 – Where's my notebook?

Dialogue D4 revolves around spatial relations, as two individuals try to locate a lost item. It is based on my field questionnaire D05 illustrated earlier in Figure 2 and 3 (*Where is my knife?*); it is also inspired by one of the lessons of Tahitian proposed in Paia & Vernaudon (2003:92-100).

4.4.1   The text

TITLE: ***Where's my notebook?***
CONTEXT: A young girl [A] is asking her elder brother [B] for her school notebook.

1.   A –   Brother, have you seen my notebook?
2.   I've been looking for it everywhere, but I can't find it!
3.   B –   I don't know, sister. Which notebook?
4.   A –   It's my math book. It's a thick, blue one.
5.   You've seen it already.
6.   I was doing my homework on it last night in the dining room.

7.   B –   Well, you probably left it there, then!
8.   A –   No. I put it away in my schoolbag before I went to sleep. I think.
9.   B –   Hm, did you look in our room, beside your bed?
10.   A –   Yes I did; it's not there.
11.   B –   Or perhaps on Dad's desk, maybe it's hidden under another book?
12.   A –   Wait… No, I can't find it.
13.   B –   Let me look in the kitchen… Hm, not here either.
14.   A –   But I need it for my math class today!
15.   The teacher will be quite angry if I don't have my notebook with me.
16.   He will think I didn't do my homework.
17.   B –   Uh oh… I think I found it!
18.   A –   Really?! where?
19.   B –   Look at our little brother out there in the garden.
20.   What's that he's holding in his hands?
21.   Isn't it your notebook?
22.   A –   Oh my god, yes it is. But he has shredded it into pieces!
23.   B –   It looks like he's been playing with it all morning.
24.   A –   Oh dear, what happened to my homework?!
25.   Now I need to buy a new notebook,
26.   and start my work all over again.
27.   Poor me! What a disaster…

4.4.2   Linguistic notes

Dialogue D4 focuses on some spatial relations, with words such as *where*, *everywhere*, *here*, *there*, *out there*; and various locative phrases (*in my schoolbag*, *in our room*, *in his hands*, *beside your bed*, *on Dad's desk*, *under another book*…).

The text also features several forms of noun modification:
- possession: *my notebook*, *your bed*, *Dad's desk*, *our little brother*…
- qualification: *thick, blue book*; *a new notebook*
- characterization or noun compounding: *my schoolbag*, *my math book*, *my math class*.

D4 has clauses showing various tense–aspect–mood configurations:
- past reference: *have you seen*; *I've been looking for*; *I was doing my homework*; *I put it away*;
- present reference: *I don't know*; *I think*; *he's holding*;
- future reference: *the teacher will be angry*; *he will think*…

Discourse particles and interjections are one of the linguistic domains for which conversational questionnaires surely constitute the best elicitation method. Dialogue D4

includes a wealth of such particles and exclamatives, rendered in English as *Well... then, I think, Hm, Wait, Uh oh, Oh my god, Oh dear, Poor me...* Several of these words are polysemous in English, but take on a certain nuance in the particular context of this dialogue. Obviously not all languages will have equivalent particles, and some languages may add certain interjections where English has none – depending on what is most natural in the flow of the dialogue.

As suggested in §3.1.1, conversational questionnaires can also help control and elicit data on prosody. The sentences in this dialogue express emotions such as annoyance, surprise, amusement, impatience, concern, desperation... Ideally, once the dialogue has been translated, native speakers would manage to impersonate it – similar to a drama – with a natural enough rendering that the audio of their performance can be recorded and analyzed.

4.5. Dialogue D5 – *A family album*

The following dialogue is plausible in many parts of the world, though not everywhere: the existence of photography, and of black and white photo albums, is not attested in all regions. The usual caveat applies, namely, that the text may need some local adaptation.

4.5.1 The text

TITLE: *A family album*

CONTEXT: A person [A] is showing a photo album to a friend [B].

1.   A –   Have you ever seen pictures of my family?
2.   B –   Well, I've met some of your relatives, but I've never seen your pictures.
3.   A –   Here is an old photo album I just found in my parents' room.
4.   B –   Oh, show it to me please!
5.   Who's this on that first photo?
6.   It looks like olden times, it's in black and white.
7.   A –   These are my grandparents, on my mother's side:
         Grandpa here on the left, and Grandma on the right.
8.   The people around them must be their friends, or other relatives.
9.   B –   Was that the day of their wedding?
10.   A –   No, I don't think so. On their wedding, they were older than this.
11.   This must have been the day when they got engaged.
12.   In those times, the day of engagement used to be a major event
         for the whole family, and people would wear beautiful clothes, and all.
13.   B –   Did you know them?

14.   A –   Who? My grandparents?
15.   B –   Yes, your grandparents on the photo.
16.   A –   Actually no. I heard they died before I was born.
         They were born a long time ago.
17.   B –   Oh I see. And who's that on the second photo?
18.   A –   This is my mother, when she was a child.
19.   I guess she was coming back from school: look at her schoolbag.
20.   B –   Oh yeah, nice. And the small boy behind her, who's that?
21.   A –   That's my uncle Teri.
         You've never met him, he lives far away from here.
22.   B –   And what school did they attend then?
23.   A –   That was an old school that doesn't exist any more.
         It was located down the road, to the west, towards the lake.
24.   B –   Let me see this other photo.
25.   This woman is surely your mother again, carrying a child on her back.
26.   Oh, I think I know who this child is. That's you!
27.   A –   That's me indeed! How did you recognize me?
28.   B –   Well, on that photo you're a small child,
         but I recognize your eyes and your smile.
29.   You look mischievous on that picture.
30.   You haven't changed much!
31.   A –   Ha ha. But I'm not a child any more, I've grown up now.
32.   Today my mother would not be able to carry me on her back like that!
33.   I'm taller than her, and heavier too.
34.   B –   Ha, that's true. You've eaten too much!

4.5.2 Linguistic notes

In terms of the lexicon, D5 features some kinship terms: *relatives, family*; *parents*; *mother*; *(maternal) uncle*; *grandpa, grandma*. This elicitation session, incidentally, may provide the fieldworker with an occasion to elicit more kin terms on the side.

D5 will show whether the language's morphosyntax uses the same possessive structures for all nouns, or whether it distinguishes – like many Oceanic languages do, for example – between possession of kin terms (*my grandparents, your mother, their relatives...*), possession of body parts (*her back, your eyes*) and possession of other types of nouns (*their room*; *their friends*; *their wedding*; *her schoolbag*; *your smile*).

Besides possession, D5 elicits complex noun phrases showing a diversity of internal syntax: *the pictures of my family*; *some of your relatives*; *an old photo album I just found*; *the people around them*; *the day of their wedding*; *the day when they got engaged*; *a major event for the whole family*; *your grandparents on the photo*.

Some spatial expressions may be of interest: *in my parents' room*; *on the left, on the right*; *around them*; *behind her*; *on her back*; *from school*; *on the photo*; *far away from here*; *down the road*; *to the west*…

But the main linguistic focus of D5 is arguably tense, aspect, and mood. Sentences include cases of:

- the experiential perfect (*have you ever seen*; *I've never seen*; *I've met*; *you haven't met him*),
- the immediate past (*I just found*)
- an equational predicate in the present (*that's you*; *this woman is your mother*) or in the past (*was that the day*…; *she was a child*)
- a stative in the present (*it looks old*; *he lives far away*; *I know*; *I'm taller*) or in the past (*it was located*; *did you know them*)
- a resultative in the present (*you haven't changed*; *I've grown up now*; *you've eaten too much*)
- a past, semelfactive event (*the day when they got engaged*; *they died*; *they were born*; *I was born*)
- a progressive in the past (*she was coming back from school*)
- a habitual predicate in the past (*used to be a major event*), habitual activity in the past (*they would wear*; *what school did they attend*)
- an imperative (*show it to me*; *look at her schoolbag*)
- a negative counterfactual (*she would not be able to carry me*)

The dialogue also includes various evidential types – indexing ways in which the characters source their statements:

- from hearsay (*I heard they died*; *in those times, people would wear*…)
- from visual cues (*it looks like an old picture*; *I guess she was coming back from school*)
- from reasoning and inference (*the people must be their friends*; *you haven't met him*)
- from firsthand experience (*an album I just found*; *I'm taller than her*)

Several sentences show subordination (*I think I know who this child is*; *an old album I just found*; *the day when they got engaged*; *an old school that doesn't exist any more*; *they died before I was born*…). The structure 'I GUESS she was coming back' may be rendered by a subordinate pattern in some languages – like in English – or by an evidential particle in other languages; the same applies for 'I HEARD they died'.

The dialogue shows a couple of comparative phrases: *they were older than this*…; *I'm taller than her, and heavier too*.

Finally, the text has deictics (*this*, *that*, *here*, *like that*…) as well as noun articles with various meanings – situational, anaphoric, recognitional, generic (cf. Dryer 2014).

5.   Conclusion

While elicitation questionnaires based on isolated sentences can be useful for exploring certain grammatical domains in depth, they lack any retrievable pragmatic context. The resulting ambiguities, lexical or grammatical, can make it difficult for speakers to identify the intended meaning, let alone render it in an idiomatic way. Such basic questionnaires also fail at capturing various ordinary linguistic features such as question–answer pairs, long-range anaphoric reference, or common discourse particles.

Fluent speech is of higher linguistic quality, but makes it difficult for the linguist to control for the presence of specific grammatical features. Due to their size limitations, text corpora collected in the field may lack many constructions useful to the linguist, who can thus wish to elicit them one way or another.

The present paper advocates for a new approach to elicitation, by promoting naturalistic conversation as the key to successful language description. Rather than being based on pure translation, conversational questionnaires consist of a meaningful exchange anchored in a plausible real-life context, and seeking its most natural expression in the target language. While keeping control of the grammatical features to be tested, the linguist camouflages them in a smooth and simple chunk of dialogue that is almost effortless, possibly even pleasant, for consultants.

Five dialogues were presented and analyzed here, as an illustration of the method. Rendering just these five texts in a given language would already provide a wealth of data in a large array of linguistic domains, enough for a solid grammar sketch. Of course, more than five texts are needed to cover the whole array of possible "speech motifs" (§3.2.3) that are encoded in natural languages. Interested linguists are encouraged to take up this endeavor themselves, devising new questionnaires to fill the gaps. These texts can be tailored so as to reflect the cultural peculiarities of a region, or to explore specific areas of interest in a grammatical system.

Ideally, new research projects setting out to study a given domain (kin terms, spatial deixis, reciprocals, discourse particles, emotions, social cognition…) could set up an onomasiological component in their fieldwork kit, in the form of elaborate conversational questionnaires created around their target domain. These cumulative efforts could then be pooled, so other linguists can use these dialogues in their respective field sites. The elaboration of such questionnaires, insofar as it helps the academic community, should be credited as a research output.

Over the years, the typological community could develop a large thesaurus of speech motifs, embodied in naturalistic conversational questionnaires. Once tried across different languages, these dialogues would be the basis for massive multilingual parallel corpora, which would be citable, reproducible, and mutually comparable. This new array of tools will enhance our means to explore empirically the diversity of language patterns around the world.

Appendix A. Translation of D1 in Mwotlap

I provide the translation here of dialogue D1 in Mwotlap, an Oceanic language of northern Vanuatu. Transcription is in IPA. The linguist reader can try and figure out the meaning of each word or morpheme from the English version of D1 presented in §4.1.1.

TITLE: kamjo sɔ van japjap

1.	A –	lɛmtap newe!
2.	B –	lɛmtap newe!
3.	A –	nek itok?
4.	B –	inɔ itok, ᵐba komjo?
5.	A –	ɔ̃ [∨], kamjo itok sɛ.
6.	B –	komjo sɔ van avɛ?
7.	A –	kamjo sɔ van how leᵐbe ɛn, a:: how ʊɻen.
8.	B –	sɔ aktɛʊ̯? a sɔ suwsuw ok?
9.	A –	ɔ̃ɔ̃ɔ̃ [↑↓↗] tatɛh, kamjo tit suwsuw vestɛ kp̄ʷijiʊ̯.
10.		kamjo sɔ van japjap.
11.		kamjo sɔ jap taʊ̯anmɛm momo te leᵐbe.
12.	B –	ɔ hijwe. wo nalavet ae lɛpnɔ?!
13.		nɔk ɛt eʊ̯lal tɛ.
14.	A –	ɔ̃ɔ̃ɔ̃, tatɛh lavet.
15.		kamjo nɛmjos ewe sɔ vel taʊ̯anmɛm ʊ̯enʊ̯en.
16.		iʊ̯nik mewel to nututu vitwaʊ̯ lamaket anejeh ɛ,
17.		ᵐba intimamjo kojo mal ʊ̯en kp̄ʷet,
18.		to:: tatɛh ʊ̯anmɛm hap sɛ leŋm̄ʷ eʊ̯en !
19.		ᵐbasto, tatɛh nɔnmɛm sem sɛ a sɔ hajtejeh.
20.		kamjo tit hiʊ̯ap vestɛ a sɔ van japjap kp̄ʷijiʊ̯.
21.	B –	ɔ, nɔ mɛlɛp.
22.		nɛmʊ̯ajsen a sɔ tatɛh ʊ̯anmi hap sɛ.
23.		nɔ tɛᵐbjiŋ veh kimi, jehe minɔ.
24.		nihnaʊ̯ naʊ̯anmamjo iʊ̯nik ɛ hip ae,
		a nihnaʊ̯ te letkp̄ʷe nɔnmamjo.
25.		ike a miʊ̯iljak to mɛ anɔj to mavan tej mɛ leŋm̄ʷ.
26.		komjo wɔ nɛmjos ɛ, kamjo tɛlɛp taʊ̯anmi veh van.
27.	A –	et, hijwe? vewe komjo a nekeken!
28.	B –	tatɛh, itok.
29.	A –	komjo sɔ lɛp taʊ̯anmɛm hinaʊ̯ mɛ,
		to kamjo sɔ lɛp taʊ̯anmi momo van.
30.		kp̄ʷijiʊ̯ aŋkp̄ʷoŋ ɛ ʊ̯en ⁿdɛl sɔ ʊ̯enʊ̯en tiwaʊ̯.
31.	B –	namnan les. sowle!
32.	A –	ɛt lɔk sɛ nek!

Acknowledgments

References

Ameka, Felix K. & Dench, Alan & Evans, Nicholas (eds.), *Catching language: The standing challenge of grammar writing*. Trends in Linguistics. Studies and Monographs [TiLSM]. Berlin, New York: Mouton de Gruyter.

Bakker, Peter. 2008. Pidgins versus creoles and pidgincreoles. In Kouwenberg, Silvia & Singler, John Victor (eds.), *The handbook of pidgin and creole studies*, 130-157. Oxford: Wiley-Blackwell.

Barth, Danielle & Evans, Nicholas. 2017. SCOPIC design and overview. In Danielle Barth & Nicholas Evans (eds), *Social Cognition Parallax Interview Corpus (SCOPIC)* Language Documentation & Conservation Special Publication 12, 1–21. Honolulu: University of Hawaii. [http://nflrc.hawaii.edu/ldc/sp-12]

Bec, Pierre. 1986. *La langue occitane*. 5th edition. (Que sais-je ? 1059.) Paris: Presses universitaires de France.

Berez-Kroeker, Andrea L. & Gawne, Lauren & Smythe Kung, Susan & Kelly, Barbara F. & Heston, Tyler & Holton, Gary & Pulsifer, Peter & Beaver, David I. & Chelliah, Shobhana & Dubinsky, Stanley & Meier, Richard P. & Thieberger, Nick & Rice, Keren & Woodbury, Anthony C. 2018. Reproducible research in linguistics: A position statement on data citation and attribution in our field. *Linguistics* 56 (1). 1–18.

Boula De Mareüil, Philippe & Vernier, Frédéric & Rilliard, Albert. 2017. Enregistrements et transcriptions pour un atlas sonore des langues régionales de France. *Géolinguistique* 17. 23–48.

Bradley, David. 2007. What elicitation misses: Dominant languages, dominant semantics. *Language Documentation & Description* 4. 136–144.

Camden, William. 1979. Parallels in structure and lexicon and syntax between New Hebrides Bislama and the South Santo language spoken at Tangoa. In Mühlhäusler, Peter (ed.), *Papers in Pidgin and Creole Linguistics, No.2*, 51–117. Pacific Linguistics, A-57. Canberra: Australian National University.

Chafe, Wallace (ed.). 1980. *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*. Norwood, New Jersey: Ablex.   [http://pearstories.org/]

Chafe, Wallace. 1994. *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. University of Chicago Press.

Chafe, Wallace. 1997. Polyphonic topic development. In T. Givón (ed.), *Conversation: Cognitive, communicative and social perspectives*. (Typological Studies In Language 34), 41-54. New York: Benjamins.

Chelliah, Shobhana L. 2001. The role of text collection and elicitation in linguistic fieldwork. In Newman, Paul & Ratliff, Martha (eds.), *Linguistic Fieldwork,* 152–165. Cambridge: Cambridge University Press.

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

Chérel, Alphonse. 1929. *Assimil : L'anglais sans peine*. Paris: Assimil.

Chérel, Alphonse. 1940. *French without toil*. Assimil spare-time daily courses. Paris: Assimil.

Corbett, Greville. 2000. *Number*. Cambridge: Cambridge University Press.

Cormier, Kearsy & Schembri, Adam & Woll, Bencie. 2010. Diversity across sign languages and spoken languages: Implications for language universals. *Lingua* 120 (12). 2664–2667.

Cysouw, Michael, & Wälchli, Bernhard. 2007. Parallel texts: using translational equivalents in linguistic typology. *STUF-Sprachtypologie und Universalienforschung* 60 (2). 95-99.

Dachkovsky, Svetlana & Healy, Christina & Sandler, Wendy. 2013. Visual intonation in two sign languages. *Phonology* 30 (2). 211–252.

Dahl, Östen. 2007. From questionnaires to parallel corpora in typology. *STUF-Sprachtypologie und Universalienforschung* 60 (2). 172-181.

Dimitriadis, Alexis & Everaert, Martin. 2001. *Anaphora typology questionnaire*. Utrecht Institute of Linguistics. Unpublished ms. [http://www.let.uu.nl/~Alexis.Dimitriadis/personal/database/Questions.pdf]

Dryer, Matthew. 2014. Competing methods for uncovering linguistic diversity: The case of definite and indefinite articles (Commentary on Davis, Gillon, & Matthewson). *Language* 90 (4). 232-249.

Dutton, Thomas Edward & Voorhoeve, Clemens Lambertus. 1974. *Beginning Hiri Motu* (Pacific Linguistics D 24). Canberra: Australian National University.

Dutton, Tom. 1985. *A new course in Tok Pisin (New Guinea Pidgin)*. In collaboration with Dicks Thomas (Pacific Linguistics D 67). Canberra: Australian National University.

Evans, Nicholas. 2010. Semantic typology. In Jae Jun Song (ed.), *The Oxford Handbook of Linguistic Typology,* 504–533. Oxford: Oxford University Press.

Filimonova, Elena (ed.). 2005. *Clusivity: Typology and case studies of the inclusive–exclusive distinction*. New York: Benjamins

Finnegan, Ruth. 1992. *Oral traditions and the verbal arts: a guide to research practices*. (ASA Research Methods in Social Anthropology). London: Routledge.

Finnegan, Ruth. 2010. Rewards and issues in studying oral literature: Some personal reflections. In Gunn, Imogen & Turin, Mark (eds.), *Oral Literature and Language Endangerment,* 13–26. Special issue of *Language Documentation & Description*, 8. London: SOAS.

François, Alexandre. 2001. *Contraintes de structures et liberté dans l'organisation du discours. Une description du mwotlap, langue océanienne du Vanuatu*. Doctoral thesis in Linguistics, Université Paris-IV Sorbonne. 3 volumes. [https://tel.archives-ouvertes.fr/tel-00136463/]

François, Alexandre. 2002. *Araki: A disappearing language of Vanuatu*. Pacific Linguistics, 522. Canberra: Australian National University.

François, Alexandre. 2003. *Field questionnaire form northern Vanuatu: Dorig language*. Handwritten ms., digitized and archived online, in open access. Online Digital Sources and Annotation System (ODSAS). [tiny.cc/AF_Q_Dorig]

François, Alexandre. 2004–2015. Materials for vernacular literacy in twelve languages of northern Vanuatu. 14 vol. Self-published. [http://alex.francois.online.fr/AF-literacy-e.htm]

François, Alexandre. 2009. The languages of Vanikoro: Three lexicons and one grammar. In Evans, Bethwyn (ed.), *Discovering history through language: Papers in honour of Malcolm Ross,* 103-126. Canberra: Australian National University.

François, Alexandre. 2011. Social ecology and language history in the northern Vanuatu linkage: A tale of divergence and convergence. *Journal of Historical Linguistics* 1, 175-246.

François, Alexandre. 2012. The dynamics of linguistic diversity: Egalitarian multilingualism and power imbalance among northern Vanuatu languages. *International Journal of the Sociology of Language* 214, 85–110.

François, Alexandre. 2013. Linguistic notes from Island Melanesia: Online archive of digitized fieldwork documents from Vanuatu and the Solomon Islands. Electronic files, 4156 pages. *Online Digital Sources and Annotation System* (ODSAS). Marseille: CREDO–CNRS. [https://www.odsas.net/]

François, Alexandre. 2014. *Structures et dynamiques des systèmes linguistiques: Documentation, description, comparaison des langues océaniennes*. Dissertation for Habilitation à Diriger des recherches. Paris: Institut National des Langues et Civilisations Orientales. [https://tel.archives-ouvertes.fr/tel-02004673]

François, Alexandre. 2019. In search of island treasures: Language documentation in the Pacific. In McDonnell, Bradley & Berez-Kroeker, Andrea L. & Holton, Gary (eds.), *Reflections on Language Documentation 20 years after Himmelmann 1998.* (*Language Documentation & Conservation* special publication, 15.), 276–294.

François, Alexandre & Franjieh, Michael & Lacrampe, Sébastien & Schnell, Stefan. 2015. The exceptional linguistic density of Vanuatu. In François, Alexandre & Lacrampe, Sébastien & Franjieh, Michael & Schnell, Stefan (eds), *The Languages of Vanuatu: Unity and Diversity*. (Studies in the Languages of Island Melanesia, 5). Canberra: Asia-Pacific Linguistics. 1–21.

Gil, David. 2015. The Mekong-Mamberamo linguistic area. In Enfield, Nicholas J. & Comrie, Bernard (eds), *Languages of Mainland Southeast Asia: The state of the art.* (Pacific Linguistics 649), 266-355. Berlin: DeGruyter.

Gippert, Jost & Himmelmann, Nikolaus P. & Mosel, Ulrike (eds.), *Essentials of language documentation* (Trends in Linguistics: Studies and Monographs, 178). Berlin: Mouton de Gruyter.

Grenoble, Lenore. 2007. The importance and challenges of documenting pragmatics. *Language Documentation & Description* 4, 145–162.

Hagège, Claude. 2008. Towards a typology of interrogative verbs. *Linguistic Typology* 12 (1). 1-44.

Haspelmath, Martin. 2005. Argument marking in ditransitive alignment types. *Linguistic Discovery* 3(1). 1–21

Haviland, John B. 2006. Documenting lexical knowledge. In Gippert, Jost & Himmelmann, Nikolaus P. & Mosel, Ulrike (eds.), *Essentials of language documentation* (Trends in Linguistics: Studies and Monographs, 178), 129–162. Berlin: Mouton de Gruyter.

Heine, Bernd. 1997. *Possession: Cognitive sources, forces, and grammaticalization*. (Cambridge Studies in Linguistics, 83.) Cambridge: Cambridge University Press.

Hellwig, Birgit. 2006. Field semantics and grammar-writing: Stimuli-based techniques and the study of locative verbs. In Ameka, Felix K. & Dench, Alan & Evans, Nicholas (eds.), *Catching language: The standing challenge of grammar writing*. Trends in Linguistics. Studies and Monographs [TiLSM], 321–358. Berlin, New York: Mouton de Gruyter.

Heeringa, Wilbert & John Nerbonne. 1999. Change, convergence and divergence among Dutch and Frisian. *Philologia Frisica Anno*. 88–109.

Himmelmann, Nikolaus P. 1998. Documentary and descriptive linguistics. *Linguistics* 36(1). 161–195.

Himmelmann, Nikolaus P. 2006. Prosody in language documentation. In Gippert, Jost & Himmelmann, Nikolaus P. & Mosel, Ulrike (eds.), *Essentials of language documentation* (Trends in Linguistics: Studies and Monographs, 178), 163–181. Berlin: Mouton de Gruyter.

International Phonetic Association. 1912. *The Principles of the International Phonetic Association*. London: Daniel Jones.

Lercari, Claude & Vernaudon, Jacques & Sam, Léonard & Gowé, Marc. 2001. *Qene drehu - Langue de Lifou : Méthode d'initiation* (Langues Kanak: Méthodes et Documents). Nouméa: Centre de Documentation Pédagogique.

Levinson, Stephen C. 1995. Interactional biases in human thinking. In Esther N. Goody (ed.), *Social intelligence and interaction*, 221–260. Cambridge University Press.

List, Johann Mattis & Cysouw, Michael & Forkel, Robert. 2016. Concepticon: A resource for the linking of concept lists. *Proceedings of the Tenth International Conference on Language Resources and Evaluation, May 23-28, 2016, Portorož, Slovenia.*

List, Johann Mattis & Cysouw, Michael & Greenhill, Simon & Forkel, Robert (eds.). 2018. *Concepticon*. Jena: Max Planck Institute for the Science of Human History. https://concepticon.clld.org/.

Lüpke, Friederike. 2010. Research methods in language documentation. *Language Documentation & Description* 7. 55-104.

Malchukov, Andrej & Haspelmath, Martin & Comrie, Bernard. 2010. Ditransitive constructions: A typological overview. In Malchukov, Andrej L. & Haspelmath, Martin & Comrie, Bernard (eds.), *Studies in ditransitive constructions: A comparative handbook*, 1–64. Berlin, New York: Mouton De Gruyter.

Mithun, Marianne. 2001. Who shapes the record: the speaker and the linguist. In Newman, Paul & Ratliff, Martha (eds.), *Linguistic Fieldwork,* 34–54. Cambridge: Cambridge University Press.

Mithun, Marianne. 2015. Discourse and grammar. In Tannen, Deborah & Hamilton, Heidi & Schiffrin, Deborah (eds.), *The handbook of discourse analysis,* 11-41. New York: John Wiley & Sons.

Moltmann, Friederike. 1997. Intensional verbs and quantifiers. *Natural Language Semantics* 5 (1). 1-52.

Montague, Richard. 1970. Pragmatics and intensional logic. *Synthese* 22 (1-2), 68-94.

Moreno Cabrera, Juan Carlos. 1992. Adverbial quantification and other adverbialia. Eurotyp V, 4. Strasbourg: European Science Foundation. [https://www.eva.mpg.de/lingua/tools-at-lingboard/questionnaire/adverbial-quantification_description.php]

Nathan, David & Fang, Meili. 2009. Language documentation and pedagogy for endangered languages: A mutual revitalisation. *Language Documentation and Description* 6, 132–160.

Newman, Paul & Ratliff, Martha (eds.). *Linguistic Fieldwork*. Cambridge: Cambridge University Press.

Ochs, Elinor, Emanuel A Schegloff & Sandra A Thompson. 1996. *Interaction and grammar*. Vol. 13. Cambridge University Press.

Padden, Carol A. 2015. Methods of research on sign language grammars. In Orfanidou, Eleni & Woll, Bencie & Morgan, Gary (eds.), *Research Methods in Sign Language Studies*, 141–155. Chichester, UK: John Wiley.

Paia, Mirose & Vernaudon, Jacques. 2003. *Tahitien: ia ora na. Méthode d'initiation à la langue tahitienne* (Méthodes de langues Bonjour salut). 3rd edn. Paris: Bibliothèque publique d'information Centre Pompidou.

Rakhilina, Ekaterina & Reznikova, Tatiana. 2016. A Frame-based methodology for lexical typology. In Juvonen, Päivi & Koptjevskaja-Tamm, Maria (eds.), *The Lexical Typology of Semantic Shifts,* 95-130. Berlin: de Gruyter.

Roach, P.J. 1989. Report on the 1989 Kiel Convention: International Phonetic Association. *Journal of the International Phonetic Association* 19(2). 67–80.

Silverstein, Michael. 1979. Language structure and linguistic ideology. In Clyne, Paul & Hanks, William F. & Hofbauer, Carol L. (eds.), *The Elements: A parasession on linguistic units and levels*, 193–247. Chicago: Chicago Linguistic Society.

Skopeteas, Stavros & Fiedler, Ines & Hellmuth, Sam & Schwarz, Anne & Stoel, Ruben & Fanselow, Gisbert & Féry, Caroline & Krifka, Manfred. 2006. *Questionnaire on Information Structure: Reference Manual*. ISIS Volume 4. Potsdam: Universitätsverlag Potsdam.

Trosborg, Anna. 1994. *Interlanguage pragmatics: Requests, complaints, and apologies*. (Studies in anthropological linguistics, 7.) Berlin: de Gruyter.

Trosborg, Anna (ed.). 2010. *Pragmatics across languages and cultures*. (Handbooks of Pragmatics, 7.) Berlin: de Gruyter.

Tryon, Darrell & Jean-Michel Charpentier. 2004. *Pacific Pidgins and Creoles: Origins, Growth and Development*. Berlin-New York: Mouton de Gruyter.

Turnbull, William. 2001. An appraisal of pragmatic elicitation techniques for the social psychological study of talk: The case of request refusals. *Pragmatics* 11 (1). 31–61.

Uther, Hans-Jörg. 2011. *The Types of International Folktales: A Classification and Bibliography*. Three volumes. Helsinki: Suomalainen Tiedeakatemia – Academia scientiarum Fennica.

Vuillermet, Marine & Desnoyers, Antoine. 2013. A hunting story – Yendo a cazar: A visual stimulus for eliciting constructions that associate motion with other events. Linguistics Department, UC Berkeley, ms.

Wierzbicka, Anna. 2003. *Cross-Cultural Pragmatics: The Semantics of Human Interaction*. Berlin: deGruyter Mouton.

Woodbury, Anthony C. 2011. Language documentation. In Austin, Peter K. & Sallabank, Julia (eds.), *The Cambridge Handbook of Endangered Languages*, 159-186. Cambridge: Cambridge University Press.

Zimmerman, Thomas E. 2001. Unspecificity and Intensionality. In Féry, Caroline & Sternefeld, Wolfgang (eds.), *Audiatur Vox Sapentiae,* 514–532. Berlin: Akademie Verlag.