2

Linguistic diversity, language documentation and psycholinguistics: The role of stimuli

Birgit Hellwig

University of Cologne

Our psycholinguistic theories tend to be based on empirical data from a biased sample of well-described languages, not doing justice to the enormous linguistic diversity in the world. As Evans and Levinson (2009: 447) put it, a major challenge of our discipline is to harness this linguistic diversity and "to show how the child's mind can learn and the adult's mind can use, with approximately equal ease, any one of this vast range of alternative systems." This paper explores some of the possibilities and limits of how language documentation and description can contribute to taking up this challenge, focusing on the role of both natural data and stimuli in this enterprise.

Keywords: linguistic diversity; language documentation; psycholinguistics; data types; stimuli

1.   Setting the scene: Linguistic diversity

There are different estimates pertaining to the linguistic diversity in this world, but we can probably assume that there are around 7000 languages spoken today (e.g., Simons and Fennig 2017 give a number of 7099). These languages are unequally distributed across the world, with the largest diversity found in the tropical regions of America, Africa, Asia and the Pacific. At the same time, most of the world's population speaks one of the larger languages. The statistics provided by the Ethnologue allow us to estimate that the 10 largest languages are spoken by around half of the world's population, and the 50 largest languages by around 96% of the population. Conversely, the remaining 7000 or so languages are spoken by the remaining 4% of the population, with sometimes very small speaker bases, often in remote and inaccessible regions, and being spoken and acquired in many

different socio-cultural contexts (Evans 2010; Nettle and Romaine 2000; see also Ammon 2015).

Language documentation and description is centrally interested in these remaining 7000 languages, and numerous documentary and/or descriptive projects testify to an enormous diversity on all levels – a diversity that goes well beyond superficial differences (see Evans 2010 for an impressive overview of the attested diversity; and see the references therein for further reading). This diversity poses a challenge to linguistic theory, which is captured in the following quote from Evans and Levinson (2009: 447):

> *"to show how the child's mind can learn and the adult's mind can use, with approximately equal ease, any one of this vast range of alternative systems. [...] [This] calls for a diversified and strategic harnessing of linguistic diversity as the independent variable in studying language acquisition and language processing [...]: Can different systems be acquired by the same learning strategies, are learning rates really equivalent, and are some types of structure in fact easier to use?"*

On one level, their idea is not controversial: in order to substantiate any claims about linguistic and cognitive universals, the diversity of this world needs to be taken into account. Yet, our generalizations about human language and cognition tend to be based on a biased sample of languages from the so-called WEIRD populations (i.e., Western, Educated, Industrialized, Rich, Democratic; a term coined by Henrich et al. 2010), bearing in mind that the extent of the bias is different in different subdisciplines (outlined below). Problematic aspects of this bias are explored in typological research (e.g., Cysouw 2002; 2011; Dahl 1990; Haspelmath 2001; Henrich et al. 2010; Lüpke 2010a), anthropological research (e.g., Schieffelin and Ochs 1986 and, more generally, within the language socialization paradigm) and cross-cultural psycholinguistic research (e.g., Keller 2007; Lieven and Stoll 2009). This research shows that, on the one hand, the languages of our sample are too similar to each other: many of them are related, they share typological features and they are spoken and acquired in similar socio-cultural environments. On the other hand, our sample is probably unusual from a world-wide perspective: it exhibits sometimes unusual typological features, and it is spoken and acquired in unusual socio-cultural environments. More generally, our sample does not reflect the diversity attested in the world, and this has obvious consequences for the validity of our models and theories (for some illustrations of the relevance of cross-linguistic data for theory-building, see, e.g., Slobin and Bowerman 2007 (for language acquisition), or Norcliffe et al. 2015 (for language processing)).

Typology is the subdiscipline that has gone furthest to redress the bias in its databases and to ensure that typological theory is built on a representative sample of the world's languages. There are extensive discussions on issues of language sampling, informed by the overall goal of avoiding genetic and/or areal biases, sometimes also cultural biases (for

overviews of the discussion, see, e.g., Bakker 2010; Perkins 2001). In the end, the typo-
logical sample still depends on the availability of data and descriptions (especially
grammars), but a long tradition of typologically-oriented descriptive research has ensured
that we now have substantial knowledge of the structures of many languages around the
world. The accumulated knowledge can be accessed through numerous grammars and
typological publications (too numerous to cite), but also through on-line typological
databases such as WALS (Dryer and Haspelmath 2013), AUTOTYP (Nichols and
Bickel 2009) or the Universals Archive (Plank and Filimonova 2000). Yet, even within
typology, there remain sampling issues: while genetic and areal biases are carefully
addressed, biases introduced by data types do not receive the same attention. As stated by
Himmelmann (2000: 10),

> *"[t]he data used for well-known and well-documented national languages generally
> represent the normative patterns characteristic of the written standard, which is, at least
> in part, a product of centuries of grammar writing and formal education. These patterns
> are juxtaposed to potentially idiolectal patterns produced by a speaker of a small isolated
> speech community on the occasion of a linguistic interview."*

That is, we tend to typologize over different data sources, thus potentially skewing the
results: we do know that elicited data differs from spontaneous data (for discussions with-
in the language documentation paradigm, see, e.g., Foley 2003; Hellwig 2010; Himmel-
mann 1998; 2006; 2012; Lehmann 2001; 2004; Lüpke 2009; 2010b; Seifart 2008; out-
side of the language documentation discussion, see, e.g., the classic contributions by
Labov 1975 or Schütze 1996), and we also know that there are systematic differences
between oral and literate data types (see, e.g., Chafe 1982; 1985; Maas 2010). We can
take word order typologies as an example, and more specifically, (non-) configurationality
and noun phrase discontinuity. Here, Australian languages are often presented as having
unlimited freedom of word order, but when studying word order variation in natural dis-
course, it emerges that discontinuity is subject to specific information-structural cons-
traints, and hence not unlimited (McGregor 2997; Schultze-Berndt and Simard 2012; see
also Louagie and Verstraete 2016). For Russian, Miller and Weinart (1998: 183) compare
discontinuous noun phrases in spoken and written varieties, and observe that

> *"linguist X working on spoken Russian [...] might well conclude that the language is
> non-configurational; linguist Y working on written Russian [...] would certainly
> conclude that Russian was configurational. The consequences for typology are clear [...]."*

That is, although typology is the subdiscipline that has gone furthest in controlling for
biases, even here, there are remaining sampling issues.

When we move away from typology to psycholinguistics, the limitations of our
databases become strikingly obvious. As testified by a volume of literature too large to

cite, the guiding principle of psycholinguistic research is to discover what is universal and what is language/culture-specific in the acquisition and the processing of human language. There is thus no doubt that linguistic diversity is of central concern to psycholinguistic research. But the actual practice falls short of these intentions. Anand et al. (2011: 3) have conducted a survey of languages researched within psycholinguistics, evaluating data from psycholinguistic journals and conferences, and from existing corpora, and the results are devastating. In terms of experimental data, only 57 languages were represented at all, of which 10 languages accounted for 85% of research (with English alone accounting for just over 30%, and German being a distant second with just over 10% of the research). And in terms of corpus data, five languages accounted for 85% of data sources (again, with English in top position accounting for over 50%; and, this time, the Chinese varieties taken together being a distant second with over 10%).

There are differences across the psycholinguistic subdisciplines, but the overall picture remains the same. In fact, language acquisition is the only subdiscipline that can look back on a tradition of cross-linguistic research,[1] thanks to a number of large-scale initiatives such as Slobin's (1985-1997) classic series *The crosslinguistic study of language acquisition*, or the CHILDES project (MacWhinney 2000). And even here, we have to conclude that

> *"we know something about the acquisition of approximately 70 to 80 languages (i.e., approximately 1% of all the languages spoken today). This 1% of languages also includes languages for which only one acquisition study of a single feature exists [...]"* (Lieven and Stoll 2009: 144).

For all other subdisciplines, the percentage of languages represented for any given topic drops to well below 1%. For example, Norcliffe et al. (2015: 1009) estimate that the empirical foundations for the study of sentence production comprises 0.6% of the world's languages (if counted generously), with our generalizations based on "primarily Germanic and Romance, to a lesser extent Finnish, Hebrew, Chinese, Korean, and Japanese." Conversely, there is hardly any psycholinguistic research on the large number of small, often endangered, languages spoken in remote regions of this world – i.e., on those languages that are of central interest to language documentation.

Giving the above guiding principle, psycholinguistics thus has a central interest in extending its research on acquisition, production and comprehension to a more representative sample of the world's languages (see especially Norcliffe et al. 2015 for a

---

[1] Note that research on child language is not only conducted within psycholinguistics, but also within anthropology, especially within the very influential paradigm of language socialization (see, e.g., Schieffelin and Ochs 1986). While this paper focuses on the relationship between language documentation and psycholinguistics, it should be kept in mind that this does not exclude collaboration with other disciplines. In particular, research on child language necessarily includes collaboration with anthropology.

brief but excellent overview of the history of cross-linguistic psycholinguistic research). But such an extension is faced with numerous methodological and ethical challenges – challenges that are familiar to any language documentation project. They range from simple issues such as working in a difficult climate or with limited access to power supplies, to more complex issues such as negotiating cultural expectations or establishing a collaborative research environment. Over the last decades, language documentation has developed considerable expertise in addressing these and other challenges, and it is thus well placed to collaborate with psycholinguistics in conducting psycholinguistic research (from all subdisciplines) under fieldwork conditions. Such collaboration does indeed exist among individual researchers, but not on a larger scale, and discussions about the methodological and ethical implications have only just begun. As such, this situation is very different from, e.g., the collaboration between language documentation and anthropology. Close ties already exist between the two disciplines, and many documentation projects include anthropological components and/or collaboration between linguists and anthropologists. Both disciplines value qualitative data types and the use of minimally invasive and culturally-sensitive data collection techniques, and both have an interest in the interplay of language and culture. This is in contrast to psycholinguistics, which – as a whole – values quantitative data types and relies on more invasive data collection techniques appropriate to a Western academic setting.

And this brings me to the purpose of this paper: to explore potential points of intersection between language documentation and psycholinguistics, with a focus on the role of semi-structured methods of data collection, such as the use of stimuli.[2] This paper does not adopt the view that psycholinguistic methods should necessarily be transferred 1-on-1 to fieldwork context, i.e., it does not see the role of language documentation as helping psycholinguistics to overcome the many 'obstacles' in the field. Instead it takes the view that any collaboration between language documentation and psycholinguistics has to be placed on an equal footing, ensuring that the standards of *both* disciplines are maintained.[3] The starting point for any such endeavor can only be at the points of intersection,

---

[2] Note that the psycholinguistic literature more commonly uses the terms 'semi-structured', 'semi-experimental' or 'broad-spectrum' methods instead of 'stimuli-based' methods (as is more common in the language documentation literature), thus distinguishing them from experimental methods (e.g., Eisenbeiß 2010). In this paper, I mostly use the term 'stimuli-based'.

[3] For example, I consider it unrealistic that we will be able to extend any of the more invasive psycholinguistic methods (such as neurolinguistic techniques involving, e.g., EEG or fMRI) to fieldwork contexts in the near future, at least not on a larger scale. More generally, I have doubts about all experimental methods that severely constrain the participants' responses (that, e.g., measure reaction times; or that involve tasks that are perceived as particularly unnatural by participants, e.g., tasks involving novel words). I allow (of course!) for the possibility that they can be made to work in specific contexts, but I doubt that they will make it into the general toolkit of a language documentation project

of mutual interest. Notwithstanding the hope that we may be able to push the boundaries of what is possible further in the future.

Two disclaimers are in order. First, the goal of this paper is to explore points of intersection between language documentation and psycholinguistics, with a focus on methodologies and irrespective of the psycholinguistic subdiscipline or the specific research question. And second, the intersection between the two disciplines is explored with a view to facilitating psycholinguistic research in the field, discussing the possibilities (and limits). The paper does not pursue the converse perspective, i.e., it does not address issues of how psycholinguistic research can be of benefit to language documentation. Instead, it refers the reader to the (admittedly) small emerging literature on this topic. For example, Eisenbeiß (2006) makes a very good case for the benefits of such a cooperation to our descriptive and documentary efforts as well as to a community's maintenance and revitalization efforts. Similarly, the Child Language Research and Revitalization Working Group (2017) highlights the relevance of child language documentation to maintenance and revitalization efforts. And Hellwig and Jung (submitted) discuss how the study of child-directed speech can enhance our understanding of the adult language.

The remainder of this paper is structured as follows: section 2 focuses on two central challenges faced by psycholinguistics under fieldwork conditions (challenges that arise from our limited knowledge of the language and the population; and challenges that arise from cultural expectations and norms); section 3 pursues possibilities of addressing the challenges (focusing on the role of stimuli and the role of natural data); and section 4 concludes this paper.

2.   Challenges

The different psycholinguistic subdisciplines have well-established research procedures and best practice models that are often not easily reconcilable with the realities of field-work-based research. For example, describing the challenges to language acquisition research in under-documented languages, Kelly et al. (2015: 287) conclude that "[t]hese conditions often make it difficult to follow the best-practice approaches to data collection which are commonly assumed in lab-based FLA [First Language Acquisition; BH] research." And with respect to psycholinguistic research more generally, Whalen and McDonough (2015: 3) weigh the issues and summarize that "[s]till, specialized studies

---

– which is the central concern of this paper. I might be too pessimistic, but the actual practice does speak for itself: the number of such studies under fieldwork conditions is negligible.

are perhaps best done in larger, less endangered language communities, especially given that many larger, unendangered language communities are also understudied."

The above two quotes paint a fairly pessimistic picture of the possibilities. Indeed, there is a multitude of challenges to any fieldwork-based research, including psycholinguistic research in the field. Not least of all, of course, practical and logistic challenges, which often come to mind first – including factors such as variable levels of literacy and technological skills, missing infrastructure (e.g., access to electricity, workspaces or internet facilities), or a challenging environment and socio-economy (e.g., tropical rains, the inaccessibility of a region or the high mobility of a population). The exact nature of these practical and logistic challenges differs, but it is likely that any fieldwork-based research project will face them to varying degrees, including projects with a psycholinguistic component. For example, a community's limited experience with literacy and technology is not only a challenge for the more obvious reasons (e.g., when planning to conduct a reading experiment), but impinges on the entire process of data collection and processing: when negotiating informed consent and research ethics; when training research assistants to help construct and present the test items for an experiment, to record the data, or to transcribe and translate the results; or when disseminating the results to the community. More seriously, our Western research methods with their reliance on the written medium and sophisticated technological setups contribute to creating an unfamiliar, stressful, research environment (see below for a discussion on cultural expectations and societal norms). While such challenges should not be underestimated, this paper focuses on another type of challenge: challenges that arise from the nature of the research procedure itself.

Psycholinguistics values quantitative data types, with an exceptionally strong expertise in experimental research and a focus on controlled experiments. In addition, there are also more semi-structured experimental techniques (that are similar to stimuli-based research in the field) as well as the construction of spontaneous speech corpora (especially in the context of language acquisition research) – but the focus is again on their quantitative analysis. Given this focus, methodological concerns centrally revolve around the careful construction of experiments and corpora, with issues covering the selection of test items (to exclude potential confounds that might, e.g., arise from frequency effects or from semantic or morphological relationships between test items), the selection of participants (to ensure a representative sample that allows for making generalizations about the overall population) or the setup of a controlled environment that excludes (or at least minimizes) unwanted distractions (such as bystanders interfering in an experiment) and/or regulates the recording intervals (as is important for a longitudinal study) (see, e.g., Blom and Unsworth 2010; Eisenbeiß 2006; 2010). From a fieldwork perspective, these requirements on data collection constitute a tall order. In particular, there are

two issues that emerge, pertaining on the one hand to our knowledge about the language and the population, and on the other hand, to cultural expectations and norms.

Starting with our knowledge about the language and the population, psycholinguistic research presupposes a good knowledge of both. In a Western context, any project investigating aspects of language acquisition or processing will have access to a comprehensive knowledge about the structure, lexicon and usage of the language (including not only reference grammars and dictionaries, but especially databases that give detailed information on frequencies, on familiarities, and/or on formal and semantic relationships between words). And they will have access to information on the composition and size of the overall population, and to supplementary information on the selected participants (covering basic socio-linguistic and economic data, but increasingly also very detailed information on their performances in working memory tests or IQ tests, or, when working with children, information on measures of linguistic development such as mean length of utterance, or the results from the MacArthur-Bates Communicative Development Inventory). All this is, of course, information that does not exist for many languages of the world, nor can it be easily generated. With respect to our knowledge of the grammatical structure, Valentine (2001: xxxi) nicely outlines the magnitude of the challenge, remarking that

> "[a] language is a natural object with a beauty and a capacity to inspire awe on the order of Niagara Falls or Lake Superior, if we take the time to appreciate it. Writing a reference grammar provides the enjoyment of thousands of hours of careful scrutiny, though at the same time one realizes acutely the truth of Michael Krauss's statement that a hundred linguists working for a hundred years could never get to the bottom of a single language."

Such reservations not only hold for our knowledge of the grammar, but for our knowledge of most aspects of the language and the population. In fieldwork-based projects, there is a clear limit to the amount and kind of knowledge that is available or that can realistically become available, and even large-scale collaborative documentation projects will not be able to reach the level and standard taken for granted in the better-documented languages.

In this context, Kelly and Nordlinger (2014: 189) report on an interesting and revealing comment by an anonymous reviewer, who "asks why anyone would begin a documentation project [on child language; BH] without having a good knowledge of the target language." From a Western psycholinguistic perspective, this is an obvious and legitimate question – and yet it misses a crucial point: to wait until such information may eventually become available will continue to exclude the majority of languages from psycholinguistic research, and will perpetuate the bias of our empirical databases indefinitely. I would thus prefer to turn the question around and ask instead: what kind of

psycholinguistic research is realistically possible under fieldwork conditions? Which existing research procedures and best practice models can we adapt so that they are transferable to fieldwork contexts without compromising the quality and standards of either field? And this is a second point of importance: it is not only a question of adapting psycholinguistic methods such that they can be made to work under fieldwork conditions, but of reconciling two very different research fields – guarding against compromising the standards of psycholinguistics, but also against compromising the standards of language documentation. And this takes the discussion to the second issue raised above, that of cultural expectations and norms.

Psycholinguistic methods rely on specific cultural expectations and societal norms that are not necessarily present outside of Western contexts (or, more specifically, outside of Western academic contexts). As Anand et al. (2011: 2) phrase it,

> *"[a] more serious challenge, not often recognized, is that the experimental method is heavily culturally circumscribed. It relies upon specific societal norms: the importance of test-taking, willingness to maintain exclusive focus on unnatural tasks, and an abstract social contract with the experimenter. Additionally, most experimental tasks are solitary, and require responses to linguistic material presented out of context, often by a machine."*

The authors here deal with experimental methods in a more narrow sense, but, to a lesser degree, this discussion also applies to stimuli-based methods.

For example, child language research often involves stimulated events. A typical activity is to provide toys and games (which are designed in such a way that they generate the linguistic expressions under investigation) to the caregiver (often the mother) and the child, asking them to play together (Eisenbeiß 2009; 2010). In Western contexts, this procedure works well: the caregiver and the child maintain joint attention on the objects and they very naturally engage in linguistic activities over them, such as the caregiver labeling objects for the child and in turn eliciting object names from the child. But this kind of engagement is not characteristic of many other societies, and – where it does occur – it is not necessarily accompanied by the same linguistic events (see especially Keller 2007; Mastin and Vogt 2016; Vogt et al. 2015). Mastin and Vogt (2016: 259), e.g., observe for rural Mozambique that

> *"rural infants' Coordinated-JA [joint attention; BH] interactions are often silent, but when speech does occur there is little naming of objects, and when caregivers do name objects, they often do not use gestures to provide deictic information that could help acquire the appropriate association. So, the more time infants spend in Coordinated-JA, the fewer opportunities they have to learn from the utterances addressed to them, since infant-directed utterances rarely contain object labels."*

From a psycholinguistic perspective, this finding has important consequences for our theories of learning: in communities like the rural Mozambiquan community mentioned

above, language learning does not only (and maybe not even primarily) take place in contexts of coordinated joint attention; instead, contexts such as overheard speech and observation were found to play important roles. From a fieldwork perspective, this has consequences for our data collection methods. Using the established psycholinguistic methods creates situations that are unusual and even potentially stressful to the participants: e.g., the play context outlined above is unusual in many parts of the world. This procedure thus has ethical implications, which we discuss in a separate publication (Hellwig and Eisenbeiß submitted). And it has implications on the ecological validity of the collected data. Within language documentation and description, there are some discussions on the validity of stimuli data, e.g., it is reported that Frog Story narratives trigger more referential expressions than are found in traditional narratives (Berez-Kroeker 2018; Foley 2003). This does not necessarily render the stimuli data invalid, but it strongly speaks for complementing experimental and stimuli-based methods with other data types, as the validity of the data can only be evaluated against a corpus of natural data. Another ensuing issue is the question of comparability: can we compare such data to data collected with the same methods in a Western context – a context where such procedures do reflect natural language use to a greater extent? As for Frog Story research, the stimulus was originally used by Bamberg (1987) to investigate the development of narrative skills in German-speaking children, inspiring many follow-up studies of this kind (see especially Berman and Slobin 1994). This research capitalizes on a Western practice of story-telling where a caregiver and a child interact over a picture book, thus ensuring that the experimental setup has some familiarity to the child. But, interestingly, even within the Western world there are reports that "studies conducted in experimental and semi-experimental settings [...] systematically underestimate the potential narrative capabilities of young children" (Nicolopoulou 1996: 387). Using this stimulus in a non-Western context where children are socialized into very different story-telling practices would only exacerbate this issue (see also Hellwig to appear-a; to appear-b).

Finally, in addition to issues of evaluating their ecological validity, these methods pose a more basic challenge: they have a tendency to objectify participants, assigning them the passive role of respondents to stimulus material. As such, these methods challenge a fundamental tenet of documentary research: documentary research explicitly strives towards collaborative research models that recognize the community's right to control research on their language; to make decisions on project design, research goals and methodologies; and to assume an active role as participants during data collection. This shift in perspective lies at the heart of language documentation as a discipline, having been foreshadowed by Hale et al. (1992), and elaborated further over the last two decades (see, e.g., Cameron et al. 1992; 1993; Dobrin 2005; Grinevald 2003; Hill 2002; Leonard & Haynes 2010; Mosel 2006; Rice 2011; Smith 1999; Stebbins 2012; Wilkins 2000; Yamada 2007). For

example, just looking at the titles of the contributions in Newman and Ratliff (2001), we can immediately see the value placed on collaboration: 'The give and take of fieldwork' (McLaughlin and Sall 2001), 'Who shapes the record: the speaker and the linguist' (Mithun 2001), or 'Learning as one goes' (Rice 2001).

Given this shift in perspective, it is not surprising that fieldwork-based studies favor the use of interactive and dialogic methodologies that allow participants some freedom in shaping the project and influencing its direction. Psycholinguistic studies, by contrast, favor methodologies that maintain control over the task and the responses of the participants. As such, the favored methodologies of the two disciplines cannot easily be reconciled. This is also the case when it comes to their research questions: psycholinguistic research questions of any kind are not likely to be the top priority of a community, which is concerned about, e.g., prestigious genres of their language becoming endangered, preferring to invest scarce resources into their documentation and preservation. As such, communities may be happy to participate in lexical elicitation that documents semantic domains considered important or prestigious, but they may be reluctant when it comes to experiments that are designed to investigate the organization of the mental lexicon. Furthermore, at the more naturalistic end of the continuum, they may be happy to record elders narrating the history of their community or a poet's oral performance, but they may have their doubts about the value of the more mundane aspects of daily life, such as a group of adults chatting (or even gossiping) or the language of children. At the same time, there are equally good reasons for communities to invest in psycholinguistic research, as such research is likely to be of long-term benefit, e.g., in the Western world, psycholinguistic research was and is an important (albeit not the only) prerequisite for the development of materials for teaching or for speech therapy (Eisenbeiß 2006); also, research into language acquisition and socialization is set to inform language maintenance and revitalization efforts, as well as educational policies (Child Language Research and Revitalization Working Group 2017). There is thus no reason why a psycholinguistic project cannot be in the interest of the community and hence no reason not to design and shape it jointly with the community – but this, in turn, requires a commitment to taking the community members' view into account, of thinking through the dangers and benefits of this research together, of weighing up scarce resources and setting priorities, and of jointly determining and developing study design and methods. And while such considerations are an integral part of all language documentation research, their importance may be underestimated within psycholinguistics, and may even conflict with psycholinguistic standards that strive towards maintaining objectivity and independence of research, and hence not involving participants in research design.

3. Taking up the challenge

Given the many challenges, it is not surprising that collaboration between language documentation and psycholinguistics is still in its infancy. There do not (yet) exist standards and best practice models, and we do not (yet) have a good overview over all the potential issues that may arise. It is thus by no means clear to what extent the requirements and standards of the two fields can be reconciled. Assuming that such a reconciliation is desirable (see especially the arguments in Eisenbeiß 2006), there are two promising points of intersection between the two disciplines: stimuli-based research, and corpora of natural language use. The potential of both are explored in this section.

3.1 The role of stimuli

Psycholinguistic research revolves to a large extent around sophisticated experimental methods, which are, indeed, not easily transferable to fieldwork conditions, since their design and implementation presuppose considerable prior knowledge about the language and the speaker population, as well as a certain overall familiarity and understanding of the participants with the research procedure. That is, controlled experiments in a narrow sense are likely to continue to be of limited applicability within language documentation – potentially becoming more feasible in the long run, as documentation and description progress. However, experiments come in different types, including the semi-structured technique of using stimuli that allow some control over the variables, but still strive to allow participants freedom in shaping their responses and to keep the situation as natural as possible.

The use of stimuli has long been a mainstay of linguistic fieldwork. When looking at some of our earliest fieldwork guides (e.g., Bouquiaux & Thomas 1976; Samarin 1967), we see that they include very detailed information, not only on the topics to cover in elicitation, but also – and especially – very detailed suggestions about which questions to ask, frequently including picture stimuli as visual aids. Their content and form was inspired by the linguistic and typological knowledge of the time, and the goal of using them was descriptive: to ensure a better description of the language under investigation. Their focus was not on generating comparable data, i.e., data that can be compared across languages (with the purpose of developing typologies) and/or across speakers (with the purpose of ensuring representativeness of the data and/or to detect variation). Following on from these early fieldwork guides, our newer fieldwork guides usually no longer include such detailed stimuli (e.g., Bowern 2006; Chelliah & de Reuse 2011; Crowley 2007; Everett & Sakel 2012; Newman & Ratliff 2001). On the one hand, their inclusion

is no longer possible: the number of available stimuli has increased tremendously, and the guides instead point to external sites where fieldworkers can find stimuli and questionnaires on specific topics and/or get inspiration for developing their own (including guides to the available guides, e.g., Majid 2012; Mosel 2012). Some important sites are, e.g., the *L&C Field Manuals and Stimulus Materials* site at the Max Planck Institute for Psycholinguistics (http://fieldmanuals.mpi.nl/) or the *TulQuest* site of the French Federation of Linguistic Typology and Universals (http://tulquest.huma-num.fr/). A more detailed discussion of such stimuli will follow further below. And on the other hand, we observe a shift in focus towards increasingly sophisticated metadiscussions on the advantages and disadvantages of various methods, on the kinds of data each method generates or does not generate, and what kinds of challenges they pose for creation, implementation and analysis.

For language documentation, Himmelmann (1998: 185-186) introduced the by-now commonly accepted basic types of communicative events: observed,[4] staged and elicited. They differ in their naturalness, i.e., in the degree to which the event would have taken place even without researchers asking for it. And they differ in the amount of control researchers exercise over it, shaping and manipulating its structure and content. Stimuli fall under the heading of staged communicative events, i.e., events that are enacted for the purpose of the recording session: some are closer to the 'natural' end in that they delimit the context but then allow speakers to talk freely within the given context (e.g., asking a speaker to narrate a story based on a picture book or film stimulus); and others are closer to the 'controlled' end in that they control for the parameters of interest and require speakers to respond to specific questions (e.g., asking a speaker to look at pictures and answer specific questions about each picture).

There is an overall consensus that documentation should attend to all three types of events, as the data generated by them complement each other. That is, staged communicative events (including stimuli) cannot replace observed and elicited events, but they do occupy an important place within language documentation, as the many contributions to this special issue testify: they report on a good number of case studies, detailing both the challenges and the rewards that arise in stimuli-based research.

The contributors to this special issue repeatedly mention one central challenge: to design good stimuli presupposes considerable knowledge about the domain of interest (e.g., topological relations) and the possible variation within it (e.g., relations of support, attachment, containment etc.), i.e., it presupposes knowledge about the so-called 'etic grid' that allows us to delimit the domain in a sensible way and to select the relevant

---

[4] He also includes 'natural communicative events', which, however, cannot be documented, as the known presence of an outside observer (a researcher and/or a recording device) may impact on the self-awareness of the speakers, thus influencing their linguistic behavior.

variables to be manipulated. As such, stimuli are time-consuming to create and require a pre-testing phase. Crucially, though, their use does not depend on a *comprehensive* prior knowledge of the given domain in the research language. That is, stimuli can be used in an exploratory way to develop a first understanding of the categorization patterns in a given domain and of the linguistic means of encoding them. Further challenges are raised by questions about the cultural appropriateness of the data collection method itself (see especially Du Bois 1980 for a discussion of issues that can arise when implementing stimuli-based research, here when showing the Pear Story stimulus). This question in turn impacts the roles of researchers and research participants and their respective agency. Similar issues also arise in elicitation, and documentation projects have gained considerable experience in striking a balance between data collection methods that invest more or less agency in participants, i.e., balancing events where participants assume a more controlling role (e.g., the recording of an observed communicative event) with those where the researcher assumes a more controlling role (e.g., an elicitation session on a grammatical topic). Stimuli-based research and cooperative research can thus be reconciled, and documentation outlets such as *Language Documentation and Conserva-tion* regularly publish discussions of such research (see, e.g., the recent special publication edited by Barth and Evans 2017, based on a stimulus introduced in San Roque et al. 2012). And a final challenge concerns the question of the validity of the generated data (see, e.g., Foley 2003, who compares narratives generated by the Frog Story stimulus with traditional narratives) – again, this is a danger that arises from all elicitation, and that is circumvented by creating a varied corpus that is *not* restricted to stimulus data.

These challenges notwithstanding, there are good reasons for using stimuli, as discussed by the contributors to this special issue. Similar to elicitation, stimuli-based research allows for control (i.e., to delimit the field and to systematically manipulate the parameters of interest). But different from elicitation, it takes steps to minimize the linguistic self-awareness of the speaker by providing a context (in the form of the stimulus, often in visual form). That is, speakers give their response based on a specific context (which is known to both the speaker and the researcher), and they do not have to imagine a context. This procedure in turn reduces the risk of misunderstanding inherent to all forms of elicitation. At the same time, this focus on responding to a given context distracts attention away from the linguistic structure, thereby eliciting more spontaneous responses and less prescriptive language use. Stimuli furthermore generate a large number of relevant expressions – i.e., they generate enough data points to investigate even low frequency phenomena. And they usually allow room for follow-up discussions with speakers, including discussions about expressions that cannot be used in a given context, thus producing negative evidence.

Importantly, stimuli-based research provides a partial solution to the issue of generalizability raised in section 1. Language documentation has come a long way towards ensuring a more diversified database, and the field as a whole is moving away from the practice of working with a handful of speakers and towards that of including a broader variety of participants. As a result, variation is starting to become an integral part of our documentations and descriptions, while at the same time introducing checks and balances that ensure that our findings do not just represent the idiolect of a single speaker (see, e.g., a recent special publication edited by Hildebrandt et al. 2017). But there is a limit to what single linguists, and even teams working collaboratively, can achieve: even with the best of efforts, we will not be able to construct a corpus that is carefully balanced for the different variables, and hence generalizability will always remain an issue. This is where stimuli-based research is set to play a major role: at least for specific domains, it is possible to ensure generalizability – to collect comparable data from a larger sample of the population, detecting variation within a language, and, conversely, allowing for more robust generalizations of language-specific patterns. At the same time, results can be compared across dialects and languages, forming the basis for statements about cross-linguistic patterns.

Much of contemporary stimuli-based research takes place at the intersection of language and cognition: studying different semantic domains and their categorization patterns, and exploring the possibilities and limits of variation across languages. By now, there is an impressive number of cross-linguistic typologies that result from such stimuli-based research, employing stimuli of many different types, e.g., questionnaires that serve as the basis for the real-life re-enactment of scenarios (as in the Demonstrative Questionnaire by Wilkins 1999; see Levinson et al. 2018 for a typology based on this stimulus); pictures, photos or short video clips that serve as the basis for a response (as in the Topological Relations Picture Series by Bowerman and Pederson 1993; see, e.g., Levinson and Meira 2003); picture books or video clips that depict stories to be narrated (as in the Frog Story by Mayer 1969; see, e.g., Berman and Slobin 1994); or interactive games that require two or more speakers to interact in order to construct a story (as in the Family Problems Picture Task by Carrol et al. 2009; see San Roque et al. 2012) or negotiate a problem (as in the Men and Tree stimulus by Levinson et al. 1992; see Pederson et al. 1996); as well as stimuli that are designed to explore non-linguistic cognition (as in the Animals in a Row stimulus; see again Pederson et al. 1996).

The small sample of existing stimuli mentioned above was selected for two reasons: partly because each one exemplifies a different type of stimulus, and partly because they are well-known and successful examples of stimuli-based research – successful on the level of language documentation and description (i.e., increasing the quality of our documentation and descriptions of individual languages), and successful on the level of theory-building (i.e., giving us a clearer understanding about universality and variation in their

respective domains, and thereby having an impact on our theories of human language and human cognition). That is, they exemplify the major reasons we go through the considerable effort of designing good stimuli: advancing our understanding of individual languages and informing our understanding of human language and cognition. Specifically, they enable us to bring under-described and under-documented languages to speak to our linguistic and psycholinguistic theories: they generate comparable data that form the basis for our typologies, they show that a pattern in a specific under-documented language is not 'exotic' or 'idiosyncratic' but representative of a type – patterns that need to be accounted for by our theories.

In all these typologies, the comparability of the data is ensured through the use of the same stimulus material and the same setup of the procedure. That is, the data is specifically collected for the typology. This procedure, in turn, has consequences for the language sample – it presupposes the availability and collaboration of a fieldworking linguist, making it almost impossible to avoid areal or genetic biases. The result is often

> *"an opportunistic sample, which has arisen from the chance that the authors have had to work closely together, and thus produce closely matched descriptions of the languages in which they are expert" (Levinson and Wilkins 2006: 6).*

In the above studies, the opportunistic sampling method has succeeded in the sense that it has unearthed patterns that are different from those attested in the better-documented Western languages. That is, even on the basis of an opportunistic sample including under-documented languages, chances are that previously unknown patterns will emerge. In the long run, though, the question of sampling will become more important, as it will have consequences for the generalizability of the results (see e.g. Stoll and Bickel 2013 for a proposal in the context of researching language acquisition). Levinson's and Wilkins' quote above highlights a further issue: the role of the language expert, and their extensive overall knowledge of a language and culture, which are indispensable prerequisites for collecting, processing and interpreting the stimuli data. And this, in turn, means that stimuli-based research can never be done in isolation, without access to other data types – and the most important such resource is a corpus of natural data.

3.2.   The role of corpora

While psycholinguistics has a strong focus on experimental research, natural data constitutes another important resource, thereby giving us a second point of intersection between psycholinguistics and language documentation. The principles underlying the construction of corpora suitable for psycholinguistic research are not necessarily the same as those underlying the construction of language documentation corpora – usually the

former place more constraints on the selection of participants, sampling intervals and amounts of data (for discussions of such issues in the construction of child language corpora, see e.g. Behrens 2008; Demuth 1996; 2008; Eisenbeiß 2006; 2010; Tomasello and Stahl 2004). Such constraints originate from the need to be able to quantitatively analyze corpora, and they thus often play only a secondary role in language documentation with its focus on qualitative analyses. But while these considerations do not play a primary role in constructing language documentation corpora, they do tie in with discussions and more recent developments in our field. On the one hand, language documentation has always been concerned with the question of ensuring that the corpus constitutes "a comprehensive and representative sample of communicative events as natural as possible" (Himmelmann 1998: 168). And on the other hand, recent years have seen a move towards cross-corpus typologies based on quantitative analyses across a number of language documentation corpora (e.g., Haig and Schnell 2016; Seifart et al. 2018).

That is, their different methods and goals notwithstanding, both disciplines share an interest in natural data and in principles of corpus construction. The main obstacle here is likely to be a practical issue. From a psycholinguistic perspective, language documentation corpora remain small and often do not reach the masses of data necessary for psycholinguistic research. For example, language acquisition research crucially depends on the availability of longitudinal studies, where a number of children (the more, the better) are recorded regularly (e.g. weekly) over the course of a year or more – resulting in large amounts of data that need to be transcribed, translated and annotated, going well beyond the capabilities of an under-resourced documentation project on an under-documented language. Nevertheless, the mutual interest in natural data provides common ground for discussions and cooperation, and recent years have seen a number of longitudinal child language corpora emerging within the language documentation context (such as on the Tibeto-Burman language Chintang within a DoBeS-funded project, see, e.g., Stoll and Bickel 2013b; Stoll et al. 2012; or on the East Papuan language Qaqet, see Hellwig to appear-a).

Within language documentation, corpora assume a central role because they allow our documentation to be a "lasting, multipurpose record of a language" (Himmelmann 2006: 1): corpora are not only useful for a specific narrow research question at a specific time by a specific researcher, but they can be used to explore and pursue multiple different questions by multiple different users. In the same way, they play a crucial role in providing supplementary information needed for more controlled research, both experimental and stimuli-based. In section 2, I argued that a major obstacle to psycholinguistic research on under-documented languages is limited knowledge about the language; corpora of natural language can give us access to at least some of this knowledge. For example, when re-

searching child language, it is necessary to know what the children are actually hearing from the adults and children around them – and this information is unlikely to be found in grammatical descriptions, or in adult-to-adult language. A corpus that focuses on children in their natural interactions, however, is going to contain this kind of information as a matter of course.[5] And this information, in turn, often sheds light on the adult language itself. For example, Hellwig and Jung (submitted) show how some properties of child-directed speech allow us insights into the metalinguistic knowledge of adult speakers, thereby enriching our overall understanding of the language. Similarly, Demuth and Ellis (2009: 95-96) reflect on the crucial role of Demuth's Sesotho child corpus in their research (one of the first longitudinal corpora of a non-Western language), concluding that "[t]he existence of larger acquisition corpora [...] has also made it possible to examine more closely the input that children hear. [...] This process also often leads to a better understanding of the structure of the target language." This comment is especially interesting in light of the fact that Bantu languages (such as Sesotho) are fairly well understood – and yet the efforts invested into constructing an acquisition corpus has paid off and resulted in a better understanding of the adult language, both of its usage and its grammatical structures. More generally, corpora of natural data can be used for accessing different kinds of supplementary information on, e.g., frequencies of words and constructions, distribution of lexical items or common collocations, or they allow for the calculation of measures of linguistic development such as the mean length of utterance of a child (see also Eisenbeiß 2010: 13-14). As always, what kind of information can be extracted depends on the kind of corpus. At the risk of stating the obvious, a corpus designed for studying language acquisition may give us information on frequencies in child-directed speech or may allow us to calculate the mean length of utterance for a child, while a corpus designed for studying the adult language is unlikely to give us this kind of information. Instead, it would give us information on, e.g., frequencies in adult-directed speech (which may or may not be similar to frequencies in child-directed speech).

All of this supplementary information is an important prerequisite for being able to conduct more targeted investigations, both experimental and stimuli-based: patterns emerging in the natural data allow for the development of hypotheses that motivate specific targeted studies, they feed into the design of the study (e.g., enabling the selection of appropriate test or stimulus items) and they make possible the interpretation of the results (e.g., assessing the naturalness of the results). As discussed in section 3.1, experimental and stimuli-based studies differ in the amount of control the researcher exercises

---

[5] There are considerable ethical challenges involved when conducting psycholinguistic research under fieldwork conditions, and especially any research involving children is set to raise issues such as, e.g., questions of legal authorization. This paper does not focus on such challenges (but see Hellwig and Eisenbeiß submitted, for a discussion).

over the responses of the participants (see also Himmelmann 1998): the more controlled the study, the more important it becomes to control for the numerous confounds – and many of our language documentation corpora will remain too small to give us insights into all such confounds. But for less controlled studies, such as stimuli-based research, they do provide enough information. And conversely, stimuli allow us to explore smaller domains in more in-depth ways, e.g., by investigating the semantic relationships that hold in a specific lexical fields – with the results feeding back into corpus construction and enriching the annotation of our corpora.

4. Conclusion

The challenges for conducting psycholinguistic research in the field seem forbidding and insurmountable, and the practice shows that we still have a long way to go to reach the goal of "a diversified and strategic harnessing of linguistic diversity as the independent variable in studying language acquisition and language processing" (Evans and Levinson 2009: 447). Such a goal can only be reached through cooperation between language documentation and psycholinguistics. And while there are no established best practice models for such a cooperation, there are initiatives that explore ways of adapting psycholinguistic methods to the realities of fieldwork outside a lab environment – without compromising the standards of either discipline, and by playing to the strengths of both.

This paper argues for seeking the collaboration over the two data types that are valued by both disciplines: natural data and stimuli-based (or semi-experimental) data. These two data types complement each other, and we cannot rely on one to the exclusion of the other. This is, of course, a well-known fact within the language documentation paradigm, but it extends equally to psycholinguistic research in the field: given our limited knowledge of many languages, psycholinguistic research cannot start with experimental methods, but has to incrementally build up the relevant knowledge. This need not, and indeed should not, be a sequential process: if we wait until our knowledge of a language has reached a level comparable to that of the better-described languages, the inclusion of linguistic diversity into our psycholinguistic theories would forever remain a very distant prospect. Instead, we should strive to build up this knowledge in parallel, and, as argued in section 3, both corpus construction and stimuli-based research will help us in this endeavor.

It is true that, from a psycholinguistic perspective, our corpora of natural data often remain small, and they may not provide sufficient information for constructing controlled experiments – but many of them will provide sufficient information for more open-ended, semi-structured, research methods, i.e., stimuli-based methods. And despite

the practical challenges, there are examples of extending language documentation corpora to meet psycholinguistic standards, through cooperation between researchers from language documentation and from psycholinguistics (e.g., the Chintang or Qaqet child language corpora mentioned in section 3.2). The patterns found in natural corpora then feed into the development of hypotheses, the design of our stimuli and the interpretation of their results. And the stimuli, in turn, have a significant role to play, as they address the issue of generalizability raised in section 1: the stimulus is kept constant and controlled, i.e., the same stimulus can be run with multiple speakers and in multiple languages, thus generating data that can be compared across speakers (capturing variation in a single language, thus going beyond describing idiolects of single speakers) and across languages (capturing variation across languages, thus showing that a language does not just represent an exotic or outlier pattern, but represents a larger type).

When we design our stimuli, we should keep this goal in mind: our collective aim as a discipline should be to develop stimuli that can be used in comparing languages, address the issue of generalizability, and provide a framework that helps individual researchers to collect data that feed into our typologies. This will enable us to harness the linguistic diversity of this world to advance our understanding of human language and cognition – and not leave the field to the larger and better-described Western languages.

Acknowledgements

References

Ammon, Ulrich. 2015. *Die Stellung der deutschen Sprache in der Welt*. Berlin: Mouton de Gruyter.

Anand, Pranav & Chung, Sandra & Wagers, Matthew. 2011. Widening the net: Challenges for gathering linguistic data in the digital age. *NSF SBE 2020. Rebuilding the mosaic: Future research in the social, behavioral and economic sciences at the National Science Foundation in the next decade*.

(http://www.nsf.gov/sbe/sbe_2020/submission_detail.cfm?upld_id=121; accessed 2017-09-04.)

Bakker, Dik. 2010. Language sampling. In Song, Jae Jung (ed.), *The Oxford handbook of linguistic typology*, 100–127. Oxford: Oxford University Press.

Bamberg, Michael G. W. 1987. *The acquisition of narratives: Learning to use language*. Berlin: Mouton de Gruyter.

Barth, Danielle & Evans, Nicholas (eds.). 2017. The Social Cognition Parallax Interview Corpus (SCOPIC). *Language Documentation and Conservation*, special publication 12.

Behrens, Heike (ed.). 2008. *Corpora in language acquisition research: History, methods, perspectives*. Amsterdam: Benjamins.

Berez, Andrea L. 2003. Speakers' attention to path, location and narrative tasks: Beyond frog stories in Ahtna. Manuscript.

Berez-Kroeker, Andrea L. 2018. Directional reference in discourse and narrative: Comparing indigenous and non- indigenous genres in Ahtna. In Holton, Gary & Thornton, Tom (eds), *Language, landscape and toponymy in Alaska and beyond*. Fairbanks: Alaska Native Language Center.

Berman, Ruth A. & Slobin, Dan I. 1994. *Relating events in narrative: A cross-linguistic developmental study*. Mahwah, NJ: Lawrence Erlbaum.

Blom, Elma & Unsworth, Sharon (eds.). 2010. *Experimental methods in language acquisition research*. Amsterdam: Benjamins.

Bouquiaux, Luc & Thomas, Jacqueline M. C. 1976. *Enquête et description des languages à tradition orale*. Paris: Selaf.

Bowerman, Melissa & Pederson, Eric. 1992. Topological relations picture series. In Levinson, Stephen C. (ed.), *Space stimuli kit 1.2*, 40–51. Nijmegen: Max Planck Institute for Psycholinguistics.

Bowern, Claire. 2006. *Linguistic fieldwork. A practical guide*. New York: Palgrave MacMillan.

Cameron, Deborah & Frazer, Elizabeth & Harvey, Penelope & Rampton, Ben & Richardson, Kay. 1992. *Researching language: Issues of power and method*. London: Routledge.

Cameron, Deborah & Frazer, Elizabeth & Harvey, Penelope & Rampton, Ben & Richardson, Kay. 1993. Ethics, advocacy and empowerment: Issues of method in researching language. *Language and Communication* 13(2). 81–94.

Carroll, Alice & Evans, Nicholas & Hoenigman, Darja & San Roque, Lila. 2009. *The family problems picture task. Designed for use by the Social Cognition and Language Project*. A collaboration of the Australian National University, Griffith University, University of Melbourne and the Max Planck Institute for Psycholinguistics.

Chafe, Wallace L. 1982. Integration and involvement in speaking, writing and oral literature. In Tannen, Deborah (ed.), *Spoken and written language: Exploring orality and literacy*, 35–53. Norwood, NJ: Ablex.

Chafe, Wallace L. 1985. Linguistic differences produced by differences between speaking and writing. In Olson, David R. & Torrance, Nancy & Hildyard, Angela (eds.), *Literacy, language and learning: The nature and consequences of reading and writing*, 105–123. Cambridge: Cambridge University Press.

Chelliah, Shobana L. & de Reuse, Willem J. 2011. *Handbook of descriptive linguistic fieldwork*. Dordrecht: Springer.

Child Language Research and Revitalization Working Group. 2017. *Language documentation, revitalization, and reclamation: Supporting young learners and their communities*. Waltham, MA: EDC.
(https://www.edc.org/sites/default/files/uploads/RouvierWhitePaperFinal.pdf: acc. 2018-06-20.)

Crowley, Terry. 2007. *Field linguistics: A beginner's guide*. (Edited and prepared for publication by Nicholas Thieberger.) Oxford: Oxford University Press.

Cysouw, Michael. 2002. Interpreting typological clusters. *Linguistic Typology* 6. 49–93.

Cysouw, Michael. 2011. Quantitative explorations of the worldwide distribution of rare characteristics, or: The exceptionality of northwestern European languages. In Simon, Horst J. & Wiese, Heike (eds.), *Expecting the unexpected: Exceptions in grammar*, 411–432. Berlin: Mouton de Gruyter.

Dahl, Östen. 1990. Standard Avarage European as an exotic language. In Bechert, Johannes & Bernini, Giuliano & Buridant, Claude (eds.), *Towards a typology of European languages*, 3–8. Berlin: Mouton de Gruyter.

Demuth, Katherine. 1996. Collecting spontaneous production data. In McDaniel, Dana & McKee, Cecile & Smith Cairns, Helen (eds.), *Methods of assessing children's syntax*, 3–22 Cambridge, MA: The MIT Press.

Demuth, Katherine. 2008. Exploiting corpora for language acquisition research. In Behrens, Heike (ed.), *Corpora in language acquisition research: Finding structure in data*, 199–205. Amsterdam: Benjamins.

Demuth, Katherine & Ellis, David. 2009. Revisiting the acquisition of Sesotho noun class prefixes. In Lieven, Elena & Guo, Jiansheng & Budwig, Nancy & Ervin-Tripp, Susan & Nakamura, Keiko & Özçalikan, Seyda (eds.), *Crosslinguistic approaches to the psychology of language*, 93–104. New York: Psychology Press.

Dobrin, Lise. 2005. When our values conflict with theirs: Linguists and community empowerment in Melanesia. In Austin, Peter K. (ed.), *Language documentation and description*, vol. 3, 42–52. London: SOAS.

Dryer, Matthew S. & Haspelmath, Martin (eds.) 2013. *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at http://wals.info.

Du Bois, John W. 1980. Introduction: The search for a cultural niche: Showing the pear film in a Mayan community. In Chafe, Wallace L. (ed.), *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*, 1–7. Norwood, NJ: ABLEX Publishing Corporation.

Eisenbeiß, Sonja. 2006. Documenting child language. In Austin, Peter K. (ed.), *Language documentation and description*, vol. 3, 106–140. London: SOAS.

Eisenbeiß, Sonja. 2009. Contrast is the name of the game: Contrast-based semi-structured elicitation techniques for studies on children's language acquisition. Essex Research Reports in Linguistics 57(7). 1–27.

Eisenbeiß, Sonja. 2010. Production Methods. In Blom, Elma & Unsworth, Sharon (eds.), *Experimental methods in language acquisition research*, 11–34. Amsterdam: Benjamins.

Evans, Nicholas. 2010. *Dying words: Endangered languages and what they have to tell us*. Chichester: Wiley-Blackwell.

Evans, Nicholas & Levinson, Stephen C. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32. 429–492.

Everett, Daniel L. & Sakel, Jeanette. 2012. *Linguistic field work: A student guide*. Cambridge: Cambridge University Press.

Foley, William A. 2003. Genre, register and language documentation in literate and preliterate communities. In Austin, Peter K. (ed.), *Language documentation and description*, vol. 1, 85–98. London: SOAS.

Grinevald, Colette. 2003. Speakers and documentation of endangered languages. In Austin, Peter K. (ed.), *Language documentation and description*, vol. 1, 52–72. London: SOAS.

Haig, Geoffrey & Schnell, Stefan. 2016. The discourse basis of ergativity revisited. *Language* 92(3). 591–618.

Haspelmath, Martin. 2001. The European linguistic area: Standard Avarage European. In Haspelmath, Martin & König, Ekkehard & Oesterreicher, Wulf & Raible, Wolfgang (eds.), *Language typology and language universals: An international handbook*, 1492–1510. Berlin: Mouton de Gruyter.

Hellwig, Birgit. 2010. Meaning and translation in linguistic fieldwork. *Studies in Language* 34(4). 802–831.

Hellwig, Birgit. To appear-a. Child language documentation: A pilot project in Papua New Guinea.

Hellwig, Birgit. To appear-b. Children's narratives in Papua New Guinea: A case study of Qaqet. In Bongartz, Christiane & Torregrossa, Jacopo (eds.), *What's in a narrative? Variation in story-telling at the interface between language and literacy*. Frankfurt et al.: Peter Lang.

Hellwig, Birgit & Eisenbeiß, Sonja. Submitted. Extending psycholinguistics to under-researched languages and populations: The principle of justice and ethical challenges for new types of linguistic fieldwork.

Hellwig, Birgit & Jung, Dagmar. Submitted. Child-directed speech – and what it can tell us about the adult language.

Henrich, Joseph & Heine, Steven J. & Norenzayan, Ara. 2010. The weirdest people in the world? *Behavioral and Brain Sciences* 33(2/3). 61–83.

Hildebrandt, Kristine A. & Jany, Carmen & Silva, Wilson (eds). 2017. Documenting variation in endangered languages. *Language Documentation and Conservation*, special publication 13.

Hill, Jane H. 2002. "Expert rhetorics" in advocacy for endangered languages: Who is listening, and what do they hear? *Journal of Linguistic Anthropology* 12(2). 119–133.

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

Himmelmann, Nikolaus P. 2000. Towards a typology of typologies. *Sprachtypologie und Universalienforschung* 53(1). 5–12.

Himmelmann, Nikolaus P. 2006. Language documentation: what is it and what is it good for? In Gippert, Jost & Himmelmann, Nikolaus P. & Mosel, Ulrike (eds.), *Essentials of language documentation*, 1–30. Berlin: Mouton de Gruyter.

Himmelmann, Nikolaus P. 2012. Linguistic data types and the interface between language documentation and description. *Language Documentation and Conservation* 6. 187–207.

Keller, Heidi. 2007. *Cultures of infancy*. Mahwah, NJ: Lawrence Erlbaum.

Kelly, Barbara & Forshaw, William & Nordlinger, Rachel & Wiggelsworth, Gillian. 2015. Linguistic diversity in first language acquisition research: Moving beyond the challenges. *First Language* 35(4/5). 286–304.

Kelly, Barbara & Rachel Nordlinger. 2014. Fieldwork and first language acquisition. In Gawne, Lauren & Vaughan, Jill (eds.), *Selected papers from the 44th Conference of the Australian Linguistic Society, 2013*, 178–192. (http://hdl.handle.net/11343/40963) (Accessed 2018-06-20.)

Labov, William. 1975. *What is a Linguistic Fact?* Lisse: de Ridder.

Lehmann, Christian. 2001. Language documentation. A program. In Bisang, Walter (ed.), *Aspects of typology and universals*, 84–97. Berlin: Akademie Verlag.

Lehmann, Christian. 2004. Data in linguistics. *The Linguistic Review* 21. 175–210.

Leonard, Wesley Y. & Haynes, Erin. 2010. Making "collaboration" collaborative: An examination of perspectives that frame linguistic field research. *Language Documentation and Conservation* 4. 268–293.

Leonard, Wesley Y. 2012. Framing language reclamation programmes for everybody's empowerment. *Gender and Language* 6(2). 339–367.

Levinson, Stephen C. & Brown, Penny & Danzinger, Eve & De León, Lourdes & Haviland, John B. & Pederson, Eric & Senft, Gunter. 1992. Man and Tree & Space Games. In Levinson, Stephen C. (ed.), *Space stimuli kit 1.2*, 7–14. Nijmegen: Max Planck Institute for Psycholinguistics.

Levinson, Stephen C. & Cutfield, Sarah & Dunn, Michael & Enfield, Nick & Meira, Sérgio. 2018. *Demonstratives in cross-linguistic perspective*. Cambridge: Cambridge University Press.

Levinson, Stephen C. & Meira, Sérgio. 2003. 'Natural' concepts in the spatial topological domain – adpositional meanings in crosslinguistic perspective: An exercise in semantic typology. *Language* 79(3). 485–516.

Lieven, Elena & Stoll, Sabine. 2009. Language. In Bornstein, Marc H. (ed.). *The handbook of cross-cultural developmental science*, 543–555. Mahwah, NJ: Lawrence Erlbaum.

Louagie, Dana & Verstraete, Jean-Christophe. 2016. Noun phrase constituency in Australian languages: A typological study. *Linguistic Typology* 20. 25–80.

Lüpke, Friederike, 2009. Data collection methods for field-based language documentation. In Austin, Peter K. (ed.), *Language documentation and description*, vol. 6, 53–100. London: SOAS.

Lüpke, Friederike. 2010a. Rare and endangered – languages or features? An African perspective. *Journal of West African Languages* XXXVII. 119–139.

Lüpke, Friederike. 2010b. Research methods in language documentation. In Austin, Peter K. (ed.), *Language documentation and description*, vol. 7, 55–104. London: SOAS.

Maas, Utz. 2010. Literat und orat. Grundbegriffe der Analyse geschriebener und gesprochener Sprache. *Grazer Linguistische Studien* 73. 21–150.

MacWhinney, Brian 2000. *The CHILDES Project: Tools for analysing talk*. Mahwah, NJ: Erlbaum.

Majid, Asifa. 2012. A guide to stimulus-based elicitation for semantic categories. In Thieberger, Nicholas (ed.), *The Oxford handbook of linguistic fieldwork*, 54–71. Oxford: Oxford University Press.

Mastin, J. Douglas & Vogt, Paul. 2016. Infant engagement and early vocabulary development: a naturalistic observation study on Mozambican infants from 1;2 to 2;1. *Journal of Child Language* 43(2). 235–264.

Mayer, Mercer. 1969. *Frog, where are you?* New York: Dial books for young readers.

McGregor, William B. 1997. Functions of noun phrase discontinuity in Gooniyandi. *Functions of Language* 4. 83–114.

Miller, Jim & Weinert, Regina. 1998. *Spontaneous spoken language. Syntax and Discourse*. Oxford: Clarendon Press.

McLaughlin, Fiona & Sall, Thierno Seydou. 2001. The give and take of fieldwork: Noun classes and other concerns in Fatick. In Newman, Paul & Ratliff, Martha (eds.), *Linguistic fieldwork*, 189–210. Cambridge: Cambridge University Press.

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.

Mosel, Ulrike. 2006. Fieldwork and community language work. In Gippert, Jost & Himmelmann, Nikolaus P. & Mosel, Ulrike (eds.), *Essentials of language documentation*, 67–85. Berlin: Mouton de Gruyter.

Mosel, Ulrike. 2012. Morphosyntactic analysis in the field: A guide to the guides. In Thieberger, Nicholas (ed.), *The Oxford handbook of linguistic fieldwork*, 54–71. Oxford: Oxford University Press.

Nettle, Daniel & Romaine, Suzanne. 2000. *Vanishing voices: The extinction of the world's languages*. New York: Oxford University Press.

Newman, Paul & Ratliff, Martha (eds.). 2001. *Linguistic fieldwork*. Cambridge: Cambridge University Press.

Nichols, Johanna & Balthasar Bickel. 2009. *The AUTOTYP genealogy and geography database: 2009 release*. Electronic database, http://www.uzh.ch/spw/autotyp.

Nicolopoulou, Ageliki. 1996. Narrative development in social context. In Slobin, Dan I. & Gerhardt, Julie & Kyratzis, Amy & Guo, Jiansheng (eds.), *Social interaction, social context, and language: Essays in honor of Susan Ervin-Tripp*, 369–390. Mahwah, NJ: Lawrence Erlbaum.

Norcliffe, Elisabeth & Harris, Alice C. & Jaeger, T. Florian. 2015. Cross-linguistic psycholinguistics and its critical role in theory development: Early beginnings and recent advances. *Language, Cognition and Neuroscience* 30(9). 1009–1032.

Pederson, Eric & Danziger, Eve & Wilkins, David & Levinson, Stephen C. & Kita, Sotaro & Senft, Gunter. 1998. Semantic typology and spatial conceptualization. *Language* 74(3). 557–589.

Perkins, Revere D. 2001. Sampling procedures and statistical methods. In Haspelmath, Martin & König, Ekkehard & Oesterreicher, Wulf & Raible, Wolfgang (eds.), *Language typology and language universals: An international handbook*, 419–434. Berlin: Mouton de Gruyter.

Plank, Frans & Filimonova, Elena. 2000. The Universals Archive: A brief introduction for prospective users. *Sprachtypologie und Universalienforschung* 53. 109–123. Available online at https://typo.uni-konstanz.de/archive/intro/index.php.

Rice, Keren. 2001. Learning as one goes. In Newman, Paul & Ratliff, Martha (eds.), *Linguistic fieldwork*, 230–249. Cambridge: Cambridge University Press.

Rice, Keren. 2011. Documentary linguistics and community relations. *Language Documentation and Conservation* 5. 187–207.

San Roque, Lila, Gawne, Lauren & Hoenigman, Darja & Miller, Julia C. & Rumsey, Alan & Spronck, Stef & Carroll, Alice & Evans, Nicholas. 2012. Getting the story straight: Language fieldwork using a narrative problem-solving task. *Language Documentation and Conservation* 6. 135–174.

Schieffelin, Bambi & Ochs, Elinor (eds.). 1986. *Language socialization across cultures*. Cambridge: Cambridge University Press.

Schultze-Berndt, Eva & Simard, Candide. 2012. Constraints on noun phrase discontinuity in an Australian language: The role of prosody and information structure. *Linguistics* 50(5). 1015–1058.

Schütze, Carson T. 1996. *The empirical base of linguistics*. Chicago: The University of Chicago Press.

Seifart, Frank. 2008. On the representativeness of language documentations. In Austin, Peter K. (ed.), *Language documentation and description*, vol. 5, 60–76. London: SOAS.

Seifart, Frank & Strunk, Jan & Danielsen, Swintha & Hartmann, Iren & Pakendorf, Brigitte & Wichmann, Søren & Witzlack-Makarevich, Alena & de Jong, Nivja H. & Bickel, Balthasar. 2018. Nouns slow down speech across structurally and culturally diverse languages. *Proceedings of the National Academy of Sciences*. (https://doi.org/10.1073/pnas.1800708115) (Accessed 2018-06-20.)

Simons, Gary F. & Fennig, Charles D. (eds.). 2017. *Ethnologue: Languages of the World*, 20th edition. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com.

Slobin, Dan I. 1985-1997. *The crosslinguistic study of language acquisition,* Vol. 1-5. Mahwah, NJ: Erlbaum.

Slobin, Dan I. & Bowerman, Melissa. 2007. Interfaces between linguistic typology and child language research. *Linguistic Typology* 11. 312–226.

Smith, Linda T. 1999. *Decolonizing methodologies: Research and indigenous peoples*. 2nd edn. London: Zed Books.

Stebbins, Tonya N. 2012. On being a linguist and doing linguistics: Negotiating ideology through performativity. *Language Documentation and Conservation* 6. 292–317.

Stoll, Sabine & Bickel, Balthasar. 2013a. Capturing diversity in language acquisition research. In Bickel, Balthasar & Grenoble, Lenore A. & Peterson, David A. & Timberlake, Alan (eds.), *Language typology and historical contingency*, 195–216. Amsterdam: Benjamins.

Stoll, Sabine & Bickel, Balthasar. 2013b. The acquisition of ergative case in Chintang. In Bavin, Edith L. & Stoll, Sabine (eds.), *The acquisition of ergativity*, 183–208. Amsterdam: Benjamins.

Stoll, Sabine & Bickel, Balthasar & Lieven, Elena & Paudyal, Netra P. & Banjade, Goma & Bhatta, Toya N. & Gaenszle, Martin & Pettigrew, Judith & Rai, Ichchha Purna & Rai, Manoj & Rai, Novel Kishore. 2012. Nouns and verbs in Chintang: Children's usage and surrounding adult speech. *Journal of Child Language* 39(2). 284–321.

Tomasello, Michael & Stahl, Daniel. 2004. Sampling children's spontaneous speech: How much is enough? *Journal of Child Language* 31(1). 101–121.

Valentine, J. Randolph. 2001. *Nishnaabemwin reference grammar*. Toronto: University of Toronto Press.

Vogt, Paul & Mastin, J. Douglas & Schots, Diede M. A. 2015. Communicative intentions of child-directed speech in three learning environments: Observations from the Netherlands, and rural and urban Mozambique. *First Language* 35(4/5): 341–358.

Whalen, D. H. & McDonough, Joyce. 2015. Taking the laboratory into the field. *Annual Review of Linguistics* 1. 395–415.

Wilkins, David. 1999. The 1999 demonstrative questionnaire: "This" and "that" in comparative perspective. In Wilkins, David (ed.), *Manual for the 1999 Field Season*, 1–24. Nijmegen: Max Planck Institute for Psycholinguistics.

Wilkins, David. 2000. Even with the best of intentions...: Some pitfalls in the fight for linguistic and cultural survival (one view of the Australian experience). In Queixales, Francisco & Renault-Lescure, Odile (eds.), *As linguas amazônicas hoje* [The Amazonian languages today], 61–81. São Paolo: Instituto socioambiental.

Yamada, Racquel-María. 2007. Collaborative linguistic fieldwork: Practical application of the empowerment model. *Language Documentation and Conservation 1(2)*. 257–282.