



HAL
open science

Language documentation twenty-five years on
Frank Seifart, Nicholas Evans, Harald Hammarström, Stephen Levinson

► **To cite this version:**

Frank Seifart, Nicholas Evans, Harald Hammarström, Stephen Levinson. Language documentation twenty-five years on. *Language*, Journal of the Linguistic Society of America, 2018. hal-01968838

HAL Id: hal-01968838

<https://hal.science/hal-01968838>

Submitted on 3 Jan 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

DISCUSSION

Language documentation twenty-five years on

FRANK SEIFART

*CNRS & Université de Lyon, University of
Amsterdam, and University of Cologne*

NICHOLAS EVANS

*ARC Centre of Excellence for the
Dynamics of Language, The Australian
National University*

HARALD HAMMARSTRÖM

*Uppsala University and Max Planck
Institute for the Science of Human History*

STEPHEN C. LEVINSON

*Max Planck Institute for
Psycholinguistics*

This discussion note reviews responses of the linguistics profession to the grave issues of language endangerment identified a quarter of a century ago in the journal *Language* by Krauss, Hale, England, Craig, and others (Hale et al. 1992). Two and a half decades of worldwide research not only have given us a much more accurate picture of the number, phylogeny, and typological variety of the world's languages, but they have also seen the development of a wide range of new approaches, conceptual and technological, to the problem of documenting them. We review these approaches and the manifold discoveries they have unearthed about the enormous variety of linguistic structures. The reach of our knowledge has increased by about 15% of the world's languages, especially in terms of digitally archived material, with about 500 languages now reasonably documented thanks to such major programs as DoBeS, ELDP, and DEL. But linguists are still falling behind in the race to document the planet's rapidly dwindling linguistic diversity, with around 35–42% of the world's languages still substantially undocumented, and in certain countries (such as the US) the call by Krauss (1992) for a significant professional realignment toward language documentation has only been heeded in a few institutions. Apart from the need for an intensified documentarist push in the face of accelerating language loss, we argue that existing language documentation efforts need to do much more to focus on crosslinguistically comparable data sets, sociolinguistic context, semantics, and interpretation of text material, and on methods for bridging the 'transcription bottleneck', which is creating a huge gap between the amount we can record and the amount in our transcribed corpora.*

Keywords: endangered languages, language documentation, language description, linguistic diversity, history of linguistics

1. LANGUAGE ENDANGERMENT AND THE DOCUMENTATION MOVEMENT. In 1992 *Language* published a multiauthored contribution on endangered languages (Hale et al. 1992), coordinated by Ken Hale, which is number one in the JSTOR download list for *Language*. It was not the first attempt to call attention to language loss—Robins and Uhlenbeck (1991) had published a collection *Endangered languages* the year before, Nancy Dorian and Joshua Fishman had drawn attention to the processes of language obsolescence (Dorian 1989, Fishman 1991), and the *International Journal of the Sociology of Language* and the *Bulletin of the International Committee on Urgent Anthropological Ethnological Research* had repeatedly investigated the issue (e.g. Capell 1962, Stone 1962, Dressler & Wodak-Leodolter 1977, Dow 1987, 1988). But the 1992 publication was a wake-up call to linguists, with powerful essays by especially Ken

* We thank *Language* editor Andries W. Coetzee and two anonymous *Language* referees for comments on earlier versions of this discussion note. We acknowledge support for new ways of carrying out comparative work on underresearched languages by the Volkswagen Foundation, through its DoBeS program, the Arcadia Foundation, through its ELDP program, and the Max Planck Society, through its Institute for Psycholinguistics in Nijmegen, its Institute for Evolutionary Anthropology in Leipzig, and its Institute for the Science of Human History in Jena. N. Evans would also like to thank the Australian Research Council (projects: The Wellsprings of Linguistic Diversity and the ARC Centre of Excellence for the Dynamics of Language).

Hale and Michael Krauss, the first eloquently addressing the human value lost, the second surveying the rates of loss around the world. Krauss called for a serious upgrading of the field's commitment to work on the world's languages, pointing out that '[u]niversities and professional societies have crucial influence in determining research and educational priorities', and asking '[a]re graduate students encouraged to document moribund or endangered languages for their dissertations?' and '[h]ow much encouragement is there to compile a dictionary of one?'. He rounds off his essay by stating: 'Obviously we must do some serious rethinking of our priorities, lest linguistics go down in history as the only science that presided obliviously over the disappearance of 90% of the very field to which it is dedicated' (Krauss 1992:9–10). Twenty-five years after that landmark publication, it is appropriate to ask where we now are, and how effectively and seriously our field has rethought its priorities.

To estimate the current dangers of language loss, we unified data from the three major sources on language endangerment (EiCat, UNESCO, and *Ethnologue*).¹ The result is that around 3,660—that is, more than half of now living languages—are currently threatened, endangered, moribund, or nearly extinct (for details, see Table 1 below). This is consistent with the LSA's guess that a 50% loss by 2100 is guaranteed (Anderson 2010). However, the loss of diversity is much greater in some parts of the world: Simons and Lewis (2013), using the EGIDS scale as a measure,² find radical differences in rate of loss between settler countries (USA, Canada, Australia), with loss rates of 75% from 1950 to now, and nonsettler regions such as sub-Saharan Africa, where the loss rate is less than 10% (for converging results see Harmon & Loh 2010; see Mufwene 2002 on settlement vs. exploitation colonization). Overall, Krauss's (1992) guesstimate that 90% of all languages would be lost or moribund by 2100 seems not unreasonable for many parts of the world, if on the pessimistic side for others.

The last twenty-five years have seen big shifts in the awareness of the general public, major funding agencies, and the linguistics community. Krauss's parallel between endangered biological species and endangered languages has often been invoked since (e.g. Nettle 1999, Nettle & Romaine 2000, Sutherland 2003, Harrison 2007), although speakers of endangered languages often find the comparison with biological species insulting. But language loss is in fact more acute, with recent figures for threatened biological species indicating 22–36% of mammal, 13–14% of bird, and 32–55% of amphibian species, and 30% of plant species (for transdisciplinary research on 'biocultural

¹ Endangerment status categorizations from the *Catalogue of endangered languages* (EiCat; <http://www.endangeredlanguages.com>, accessed December 1, 2016), UNESCO's *Atlas of the world's languages in danger* (UNESCO 2010), and SIL's *Ethnologue* (Lewis et al. 2016) were unified as follows: (i) threatened (used by most adults and some children, negative trend) includes EiCat's and UNESCO's vulnerable and SIL's threatened languages; (ii) shifting (used by many adults but not transmitted to children) includes EiCat's threatened and endangered, UNESCO's definitely endangered, and SIL's shifting languages; (iii) moribund (active use only by grandparent generation) includes EiCat's and UNESCO's severely endangered and SIL's moribund languages; (iv) nearly extinct (hardly used, even by grandparent generation) includes EiCat's and UNESCO's critically endangered and SIL's nearly extinct languages. In case of conflict between sources, we prefer sourced data over nonsourced data, preferring thus EiCat over UNESCO over *Ethnologue* (Hammarström et al. 2018). For problems of 'reducing complex social situations to a number' see Grenoble 2016, and see Dobrin et al. 2007:41 on the danger of a 'commodification' effect when languages are 'prioritised by the weakness of their speaker base and their "degree of endangerment"'. But note that most categorizations reported here are supported by careful scrutinization of a wealth of sources (published and personal communications by experts), in particular in the case of EiCat.

² EGIDS (the Expanded Graded Intergenerational Disruption Scale) is an *Ethnologue* measure of endangered status (Lewis & Simons 2010).

diversity' see Loh & Harmon 2005, Maffi 2005).³ National Geographic and many press outlets took up the story,⁴ and a number of foundations continue to press the parallel between loss of biological and cultural diversity (e.g. Terralingua, Living Tongues Institute). Up to 2006, there were some 114 television and feature films on the subject.⁵ Hale's (1992:36) lament over the 'irretrievable loss of diverse and interesting intellectual wealth, the priceless products of human mental industry' also caught the public interest, a position elaborated in Evans's (2010) book *Dying words: Endangered languages and what they have to tell us* and Harrison's (2007) *When languages die: The extinction of the world's languages and the erosion of human knowledge*. And the notion of 'diversity' has also been invoked to dramatize what language endangerment threatens to destroy by the title of Brenzinger's (2007) continent-by-continent survey, *Language diversity endangered*.

The professional linguistic response to the 1992 issue was not instantaneous. As mentioned, scholars had already been pressing the case before this publication, and a number of national linguistics traditions—notably in Australia and France—had long taken language description for granted as a central part of academic linguistics. But the German Society for Endangered Languages (Deutsche Gesellschaft für bedrohte Sprachen), founded in 1997, pressed the Volkswagen Foundation for funding, and by 2000 an ambitious, international, multimillion-euro documentation program, DoBeS,⁶ was launched with then state-of-the-art digital infrastructure at MPI Nijmegen. This was rapidly followed by the establishment of a parallel enterprise ELDP⁷ by the Arcadia Fund, based at SOAS in London in 2002. These projects offered multiyear documentation grants, allowing the training of Ph.D. students in the new objectives, methods, and digital techniques. DEL (Documentation of Endangered Languages), a joint funding initiative by the National Science Foundation (NSF) and the National Endowment for the Humanities (NEH), began in 2005. DoBeS has funded work on about 130 languages, ELDP on over 300, and DEL on over 100. A number of smaller funds also now exist, including national initiatives in Brazil (PRODOCLIN)⁸ and India. By virtue of the documentation movement perhaps 500 additional languages now have archived primary records with annotations.

Significant new prizes recognizing major achievements in grammar writing and corpus building have also been introduced. Since 2007 the Association for Linguistic Typology has awarded two prizes for descriptive grammars of little-studied languages: one for doctoral dissertations, one for published monographs. Another recent innovation (2016) is the DELAMAN Franz Boas Award for the best-curated multimedia text collection of an underdocumented language.⁹ These give additional recognition to achievements in grammar writing and corpus building, on top of longer-established traditions such as LSA's Bloomfield Award, which occasionally includes publications of this type in their purview.

At the time of the Hale et al. publication, the term 'documentation' was used in a way somewhat different from the more specialized meaning it has acquired since. Krauss

³ Numbers are from the IUCN Red List of Threatened Species™ (<http://www.iucnredlist.org/>, accessed April 1, 2018); see also Pimm et al. 2014.

⁴ See <https://www.nationalgeographic.org/projects/enduring-voices/>, accessed June 20, 2018.

⁵ <http://www.olestig.dk/endangered-languages/films.html>, accessed June 20, 2018

⁶ DoBeS: Dokumentation bedrohter Sprachen (Documentation of endangered languages)

⁷ ELDP: Endangered Language Documentation Program

⁸ <http://prodoclin.museudoindio.gov.br/>, accessed June 20, 2018

⁹ <http://www.delaman.org/delaman-franz-boas-award/>, accessed June 20, 2018

(1992:9) states: ‘by documentation I mean grammar, lexicon, and corpus of texts’, consistent with England (1992). In modern terminology (e.g. Himmelmann 1998), approaches that carry out grammatical and lexicographical analyses of a range of languages using a relatively theoretically neutral descriptive apparatus fall under ‘descriptive linguistics’, while ‘language documentation’ has since taken on a more specialized meaning (see below), with a focus on how best to produce an optimal and maximally representative corpus of material on an underinvestigated language, sometimes with an explicit rejection of the desirability of devoting time and energy to language description.¹⁰ The coexistence of these broad and narrow senses is probably inevitable now, and our title intends ‘language documentation’ in its broadest sense. However, failure to distinguish ‘language documentation’ and ‘language description’ may have a number of unfortunate effects, including ambiguities about how far grammatical description is expected to flow from a large language documentation project, how sharply the formulation of theoretically interesting questions should be built into documentary research projects, and—in the other direction—how far the real technical advances and increased commitment to data integrity that documentarists have pioneered should be adopted by the entire field of linguistics if it is to recast itself as a discipline with full empirical accountability.¹¹

In fact, the past quarter century has seen the effective establishment and partial institutionalization of a new branch of linguistics dedicated to language documentation, with its own textbooks, journals, educational curricula, and teaching personnel.¹² The new subdiscipline has a number of defining characteristics (Himmelmann 1998)—a focus on primary multimedia data and its archiving, the use of time-aligned digital annotation of recordings, the availability of data in major archives to other scholars for re-analysis, sourcing of all cited language material to searchable and natural primary data, and a preference for interdisciplinary work (Thieberger 2012) with primarily anthropologists, biologists, and ethnomusicologists.

Contributions in Hale et al. 1992 (especially England 1992, Jeanne 1992, Watahomigie & Yamamoto 1992) also initiated thorough reflections on the roles of endangered language speakers and communities in language documentation and the contributions and production of knowledge by indigenous/endangered language scholars (see Sammons & Leonard 2015, Fitzgerald 2017, Leonard 2017 and references therein). One consequence of this is the recognition that different perspectives lead to documentation of different phenomena, different genres, and so forth (Linn 2014). A growing body of literature on community-based language research (Yamada 2007, Czaykowska-Higgins

¹⁰ Himmelmann (2006:23–24) states: ‘with regard to the economy of research resources, it may be more productive to spend more time on expanding the corpus of primary data rather than to use it for writing a descriptive grammar’; and Mosel (2006:307) writes: ‘a thorough analysis [beyond sketch grammar level] can wait until there is time for a specialist investigation’. See Rice 2006, Rhodes et al. 2006:3–4, Hyman 2007, and Evans 2008:346 for arguments against this position.

¹¹ Increasingly, documentary practice requires all example sentences to be sourced, ideally to a naturally occurring text but failing that at least to a point in a numbered field notebook; see Heath 1984 for an early grammar written on these principles, Thieberger 2006 for a hypertext-implemented grammar with links to sound files, and the Austin Principles of Data Citation in Linguistics (Berez-Kroeker et al. 2018). This stands in contrast with the use of unsourced intuited sentences in the generative tradition, whose reliability was subject to critical discussions from early on (e.g. Labov 1975, Dąbrowska 2010, Gibson & Fedorenko 2013).

¹² Handbooks include Gippert et al. 2006 and Filipović & Pütz 2016. Journals include *Language Documentation and Description*, *Language Documentation & Conservation*; conferences include the International Conference on Language Documentation and Conservation and Language Documentation and Linguistic Theory.

2009), decolonizing methodologies (Smith 2012, Leonard 2017, among others), and indigenous research methodologies (Wilson 2008, Kovach 2010) further specifies the complex sociopolitical issues of research with communities speaking endangered languages (for alternative conceptions and critique, see Dobrin 2008, Crippen & Robinson 2013). As concrete examples, we note here methods such as cross-mediated elicitations developed to let teachers benefit as much from the research as the linguist (Léonard & Avilés González 2015) and a recent Ph.D. thesis written in two versions: Portuguese and the object language Kaxinawa (Kaxinawá 2014).

An important aspect of community-based language documentation is training in documentation methods (Jukes 2011), such as the biennial Institute on Collaborative Language Research (CoLang, formerly InField) (Genetti & Siemens 2013) targeting US graduate students, workshops held by SOAS's Endangered Languages Documentation Programme targeting local scholars and documenters, and community-based capacity building (Fitzgerald 2018). In this context, the conceptions of the role of archives and their relationship with language communities are also changing rapidly in important ways (Kung & Sherzer 2013, Linn 2014), and this has been implemented, for instance, in the series of 'Breath of Life/Silent No More' Native California Language Restoration Workshops, in which language archive materials are used to assist the revitalization of moribund or extinct languages (Hinton 2001, Fitzgerald & Linn 2013, Sammons & Leonard 2015).

Regarding the issue of language endangerment itself, Mufwene (2017) claims that linguistics still lacks adequate theorizing of language vitality, although others have pointed out advances in this respect, particularly from research on multilingualism (Bower 2017, Di Carlo & Good 2017, Lüpke 2017) and revitalization (Fitzgerald 2017). Recent language revitalization efforts clearly show that these succeed best when language also serves other community needs (Henderson et al. 2014, Fitzgerald 2017), and when the original documentation takes care to record conversation and speech formulas (Amery 2009). These developments support holistic and inclusive notions of language vitality, placing language documentation in the context of educational and economic processes and well-being (for potential moral incoherence in a preservationist stance, see Mufwene 2002 and Fast 2007).

In summary, the past twenty-five years have seen theoretical advances and lively debates of various facets of language endangerment, the loss of linguistic diversity, and multiple responses. They evidence the recognition of the value of linguistic diversity for 'human intellectual life—not only in the context of scientific linguistic inquiry, but also in relation to ... arts and culture' (Hale 1992:35). In the following sections, we focus on various aspects of the scientific linguistic inquiry of the intellectual wealth embodied in linguistic diversity.

2. HOW DOCUMENTATION EFFORTS HAVE EXPANDED OUR VIEW OF THE POSSIBILITY SPACE FOR LANGUAGES. It might be thought that, after at least half a millennium of linguistic endeavors, we would be approaching the point of diminishing returns, with new languages only rarely offering surprises. But there are few signs of our discoveries flattening:¹³ the documentation of linguistic diversity keeps turning up new phenomena

¹³ Before proceeding, an important caveat is in order, in the light of moves to decolonize research methods and terminology as they are applied to the study of indigenous cultures. The term 'discovery' is fraught in the sense of implying that the categories or phenomena were not previously known even to speakers of the relevant languages, implicitly or explicitly. Finding a less loaded terminology is a task for the field, part of Fishman's (1982:7) vision under which 'each collectivity contributes its own thread to the tapestry of world

that had either been considered impossible or simply had not been contemplated as linguistic categories. Some of these directly rely on ‘documentarist’ approaches, such as new methods of corpus-based work, a new concern with coverage of different speech genres, the availability of field stimuli for structured elicitation, interdisciplinary teamwork, and new technologies of digital video, which make it possible to document language in its local ecology, domestic or ritual, and fully embedded in the context of visual modalities (see e.g. Green 2014). Thus the new knowledge has a depth and richness not captured in the older pencil-and-paper style of fieldwork—although traditional elicitation work remains part of the discipline.

Here is a quick overview of some of the new regions of the design space of human language that the field has discovered since the Hale et al. alert, which remind us that ‘there is nothing to beat actuality, present or past, for clinching possibility’ (Dennett 1995:105). In the realm of linguistic structure, every part of the language system, from phonetics through phonology, morphology, and syntax to semantics, has seen our view of what languages can do greatly expanded. So of necessity the list below is incomplete, selective, and for each phenomenon we can only touch on it superficially (for a more comprehensive treatment see Evans & Levinson 2009).

Beginning with phonetics, the perimeter of tonal contrasts continues to expand. Cruz and Woodbury (2014), for example, recently demonstrated the existence of a fourteen-valued tonal contrast in Chatino. Illustrating the dynamic between documentarist practice and theoretical enquiry, their work teamed up a senior linguist (Woodbury), a native speaker who had recently obtained doctoral training (Cruz), and the added involvement of many native speakers across several varieties:

We also emphasize the continued role [in working out the tone system] of teaching and speaker-training in local contexts. This has been possible ... because of the high levels of interest and appreciation of Chatino in Chatino communities, and the view that writing is a way for the language to receive respect. ... [R]esearch can be more exact when speakers, through linguistic study and through learning to write, become critically aware of the tonological systems of their languages. (Cruz & Woodbury 2014:521)

Next, consider the question of locality (or lack thereof) as it impacts rules for tone assignment and interaction. Work on Giryama has shown that a prefix on the verb can assign a high tone to the penultimate mora of the last word of its phonological phrase, perhaps two words to the right (Volk 2011, Hyman 2016)—a phenomenon that is problematic for autosegmental accounts of spreading which try to find a representation that does not just work descriptively, but also accounts for the phenomenon in a principled way.

As a final phonological example, consider the case of vocal register. Most well-known languages have just one register (phonation type), some (like Middle Khmer) have two, a few have three (e.g. Bai; Edmondson & Esling 2006), and just a tiny number have five (e.g. West !Xóǝ; Naumann 2016). The small number of languages with complex register systems means that, in turn, the sub-design space of how register works has been largely ignored, so that register has not been seen as a phonetic dimension in which the temporal dynamics of the production is important, in contrast to tone where phoneticians have long been aware of the importance of pitch contours. Work on the Mon-Khmer language Chong by DiCanio (2009), which distinguishes four regis-

history, and ... each is accepted and respected for making its own contribution’. For the moment, such terms are used here in the sense of ‘becoming integrated into the conceptual panoply of the global linguistics community’, with full respect for the creative and metalinguistic accomplishments of the cultures that developed the phenomena.

ters, clearly demonstrates that time-course through a syllable is an essential differentiator between register types (see also Thongkum 1988, Esposito 2004). This is not simply an expansion of our taxonomy of register types, since it entails that multiple laryngeal representations may be present on a single prosodic unit (breathily changing to tense, within a single syllable, for the fourth register type).

In the realm of inflection, there has been a steady expansion in our stock of core semantic categories. This has been particularly evident at the meeting point of social cognition and grammar, thanks to the exploration of languages where much of the intersubjective coordination, through the negotiation and adjustment of mutual knowledge, is accomplished by core grammatical categories. These include an elaborated conception of how evidential systems work, such as the discovery that ‘evidence events’ in languages such as Matsigenka (Fleck 2003, 2007) can be specified for tense independently of the tense of the reported event, bringing in the need for a ‘perception event’ to mediate between the reported event and the speech event. And evidential systems in some languages include values for (presumed or questioned) addressee perceptions, such as Duna ‘addressee potential evidentials’ (San Roque 2015) with meanings like ‘as you should be/should have been able to see’. We have likewise vastly expanded our understanding of ‘egophoricity’, that is, grammatical systems for flagging the privileged access (Hargreaves 2005) of a particular speech-act participant to personal knowledge, experience, or a conscious self (Floyd et al. 2018). And languages have been described in which core grammatical systems, such as verb inflections in Kogi (Bergqvist 2016) or auxiliary bases in Andoke (Landaburu 2007), encode different values of ‘multiple perspectives’ (Evans 2006)—that is, multivalued paradigms partitioning the degrees to which a state of affairs is presumed to be accessible to speaker and/or hearer (Evans et al. 2018a,b). All of these developments indicate that the modeling of mutual awareness and accessibility between speaker and addressee needs to be assigned a much more central position than it has been in morphosyntax and semantics, at least for some languages, and cannot simply be relegated to the study of discourse.

The many new morphosyntactic phenomena uncovered by recent fieldwork include the Papuan language Walman’s achieving coordination by transitive verbs (Brown & Dryer 2008) (so roughly *Moru she-adds-him Kilip they-read-it* for ‘Moru and Kilu read it’) and the existence of long-distance agreement of verbs in the Daghestanian language Tsez with topic constituents (Polinsky & Potsdam 2001).

At the morphosyntax-discourse interface, recent work by Haig and Schnell (2016) and Brickell and Schnell (2017) have refuted the long-standing ‘preferred argument structure’ explanation (Du Bois 1987) that attributes ergative alignment to the favoring of intransitive subjects and objects for introducing new referents. This new work extends the language sample by adding the Austronesian languages Teop and Vera’a as well as Northern Kurdish, stemming directly from language documentation projects, to previously reported corpora, and it employs larger corpora that can test new questions (e.g. determining the proportion of \pm human among the overall rare lexical A arguments). This work raises standards of replicability and falsification by making all texts available with sound files, translations, and crosslinguistically comparable syntactic annotation to facilitate meaningful comparison—a salutary example of how a combination of broader crosslinguistic sampling and more accountable corpora can lead to revisions of received wisdom in the field.

Within the field of semantics, numerous examples can likewise be found of how beliefs about language articulated before 1992 have proved too narrow. Our knowledge of the semantics of space has exploded, replacing the view that humans are egocentric by

nature in their spatial reckoning with a more underdetermined ontogeny that allows for a range of conceptual systems, from egocentric ‘left-right’ reckoning to the use of compass coordinates, according to the language learned (e.g. Levinson 2003, Levinson & Wilkins 2006, Levinson et al. 2018). Likewise, we now know vastly more about the semantics of quantification (e.g. Matthewson 2008, Paperno & Keenan 2017), leading to interesting revisions to our view of how quantification should best be represented: Matthewson (2001) argues that determiner quantifiers like *every* are not formed directly through combination of a quantifier with an NP predicate, but involve a two-step process where the NP predicate is first combined with a determiner-like element. The broader interest of work like this is that structures from a seemingly exotic language, St’át’imcets (Lillooet Salish), are used to reveal purportedly universal structures more directly than can be seen in English.

Whole semantic domains have long been thought to be ‘ineffable’, beyond the reach of language, for example the odor domain (Levinson & Majid 2014). Regarded since Kant (2006 [1798]:50) as ‘the most dispensable’ of the senses, as ‘of extremely slight service’ to humans by Darwin (1874:17), and ‘vestigial’ by Pinker (1997), the consequence was the prediction that ‘there is no semantic field of smells’ (Sperber 1974:115–16). Yet we now know that some languages have elaborate abstract odor lexica (e.g. Aslian languages of Southeast Asia), which have been investigated using new methods that combine olfactory stimuli and psychological methods for studying categorization and recall. Thus languages like Jahai and Maniq have vocabularies of fifteen to twenty abstract smell terms, embedded in cultures that place great emphasis on smell (Majid & Burenholt 2014, Wnuk & Majid 2014, O’Meara & Majid 2016).

Or take reciprocals, focusing here particularly on the two cases of a ‘strong reciprocal’ like ‘everyone in this room knows each other’ and a ‘chained reciprocal’ like ‘the graduating students followed each other onto the stage’, in which the leading student follows no one and the last student is followed by no one. Twenty years ago it was believed (Dalrymple et al. 1998) that crosslinguistically these meanings would always be conflated, under a single polyadic quantifier that applies over a range of specific cases including both of these and some others. Yet we now know this claim to be false: Indo-Pakistani Sign Language has a specific construction for the latter ‘chaining’ type, exploiting the four-dimensional modality afforded by sign language to make semantic distinctions not commonly made in spoken languages (Evans et al. 2011, Zeshan & Panda 2011). Here, then, it was not until the triple conjunction of a wider sample, the inclusion of a sign language, and the targeted use of stimuli permuting a range of situation types that we could discover this new type of semantic distinction.

Moving beyond core structural concerns, work with little-studied languages is pushing our knowledge of how language is learned, processed, socially organized, aesthetically extended, and how it evolves (e.g. mixed languages, emergent sign languages). Among recent studies in this vein we mention:

- Studies of how ergativity is learned across a wide range of languages (Bavin & Stoll 2013), how children grapple with the challenge of learning polysynthetic morphology in Inuit (Allen 2017), Mayan languages (Mateo Pedro 2015, Pye 2017), Murrinhpatha (Forshaw et al. 2017), and Chintang (Stoll et al. 2017), and how complex morphosyntax is acquired by adults in language revitalization (Morgan 2017);
- Processing studies of languages with unusual typological features (Norcliffe et al. 2015, Sauppe 2017), which demonstrate that basic word-order differences, or different voice systems, entail substantially different processing routines;

- Crosslinguistic patterns in local speech-rate variation, as measured in language documentation corpora (Seifart, Strunk, et al. 2018);
- Studies of previously unappreciated ways in which sociolinguistic variation can be organized for social signaling, such as patrilects in Sui (Stanford 2009);
- Investigations of special forms of signaling that are interdependent with structural properties of speech, such as drum signaling in the Northwestern Amazon that adopts rhythmic patterns from speech (Seifart, Meyer, et al. 2018), the ‘talking xylophone’ of the Sambla in Burkina Faso and its relation to Seenku phonology (McPherson 2016), and repair organization in Chinantec whistled speech (Sicoli 2016);
- Studies that identify key moments in the emergence of structure in evolving languages, such as the sign language phonology with only emerging double-articulation in Al-Sayyid Bedouin Sign Language (Sandler et al. 2011) or the merging of Kriol and Warlpiri elements to form new auxiliary structures in Light Warlpiri (O’Shannessy 2013).

3. HOW MUCH HAS BEEN ACHIEVED? Probably the single biggest technical revolution in the language sciences since 1992 has been the growth of digital technology for the recording, storage, and handling of multimedia records. This has made possible the building of large-scale digital archives and, more important, the means of storing and analyzing time-aligned transcription and annotation. Especially the ELAN software¹⁴ and interoperable corpus tools such as Praat (Boersma & Weenink 2018) and Toolbox¹⁵ could then be integrated, making it possible to produce highly annotated text for further analysis and enrichment. Older archives that held media material on diverse languages could then be transformed into digital archives with potential worldwide online availability, including access to the donor communities themselves.¹⁶ TLA (The Language Archive based at the MPI for Psycholinguistics, which includes the DoBeS archive) now holds about 80 terabytes of well-described resources, amounting to c. 20,000 hours of digitized audio/video recordings with 5 million annotations in over 200 languages; similar amounts of resources are held at the ELDP’s ELAR archive, as well as PARADISEC and AILLA, with foci on the Pacific and the Latin America, respectively.¹⁷ These archives face a problem of constant maintenance, since the underlying technology is constantly changing, while funding bodies have little interest in contributing to the maintenance costs, and neither academic institutions nor the professional associations take long-term responsibility for them. This makes the archives themselves potentially endangered.

Traditional description in the sense of grammars and dictionaries has gathered pace in parallel with documentation. We estimate, based on *Glottolog* entries (Hammarström et al. 2017), that we can now find basic (Swadesh-type) word lists for 89% of languages, and description of substantial elements of the grammar for around 60% of all living languages. Table 1 shows how descriptive efforts in the past twenty-five years

¹⁴ <https://tla.mpi.nl/tools/tla-tools/elan/>

¹⁵ <https://software.sil.org/toolbox>

¹⁶ These developments were overseen by a few individuals who have indirectly had a substantial impact on the field. A nonexhaustive list would include Peter Wittenburg, Nick Thieberger, and Brian McWhinney.

¹⁷ PARADISEC holds 5,900 hours of archived audio materials representing more than 1,060 languages from around sixty-seven countries in some 13.5 terabytes. AILLA holds 7,500 hours of archived audio materials, representing more than 300 languages from at least twenty-eight countries. ELAR holds 12,240 hours of audio and 4,161 hours of video representing more than 580 languages (with more substantial data on about 300) from at least ninety-one countries, comprising over 21 terabytes.

contributed to this, in particular regarding endangered languages, on which 574 first-time descriptions have been produced since 1992.

STATUS NOW	DESCRIBED	DESCRIBED	IN 2016 STILL
	UP TO 1992	1992–2016	UNDESCRIBED
Not endangered (3,179 languages)	1,459 (45.9%)	385 (12.1%)	1,335 (42.0%)
Threatened (953 languages)	469 (49.2%)	142 (14.9%)	342 (35.9%)
Shifting (1,544 languages)	678 (43.9%)	235 (15.2%)	631 (40.9%)
Moribund (564 languages)	261 (46.3%)	91 (16.1%)	212 (37.6%)
Nearly extinct (600 languages)	262 (43.7%)	106 (17.7%)	232 (38.7%)
<i>Total endangered (3,661 languages)</i>	<i>1,670 (45.6%)</i>	<i>574 (15.7%)</i>	<i>1,417 (38.7%)</i>
Extinct (899 languages)	345 (38.4%)	82 (9.1%)	472 (52.5%)

TABLE 1. Description (in form of at least a grammar sketch, in either published form or unpublished but accessible M.A. and Ph.D. theses) and current endangerment status of the world's 6,840 living L1 languages from the language inventory of Hammarström et al. 2017. Endangerment status categorization is explained in

n. 1. Languages categorized here as not endangered include 155 languages on which no endangerment information is available.

While these figures are impressive, the accumulation of grammatical descriptions has not shown the kind of rapid transformation that might have been hoped for. On the one hand, the proportion of descriptive efforts devoted to endangered languages since 1992 has only slightly increased, from 58% of all descriptions prior to 1992 being on languages that are endangered or now extinct to 63% since 1992. On the other hand, the situation since 1992 has not improved substantially in terms of the commitment by leading universities to grammar writing as a central activity in the language sciences, and just a few centers have been the major academic producers.¹⁸ Many renowned linguistics programs (e.g. Stanford, MIT) do not accept descriptive grammars of underdocumented languages as a doctoral topic (although MIT, for instance, does offer a master's program in Indigenous Languages, with Jessie Little Doe Baird notable as one of the graduates), and the most prolific producers of grammatical descriptions barely appear in the top league of world rankings. If we track the top dozen linguistics programs as ranked by the QS rankings since 2013, only two of the grammar-writing powerhouses make the cut: UC Berkeley in four years out of five (#10 in 2014, #11 in 2015, #9 in 2016, and #9 in 2017) and the Australian National University in two years out of five (#9 in 2013 and #11 in 2014). Without more professional recognition of the scientific value of grammar writing, as already urged by Krauss in 1992, the field will find it difficult to confront the huge challenge of obtaining good grammatical coverage of the about 2,750 of the world's languages that are still undescribed, among which more than 1,400 are endangered (Table 1).

However, there have been major developments in terms of online language catalogues and typological surveys. SIL has produced *Ethnologue*, a substantial catalogue of languages, since 1958, online since 2005 (Simons & Fennig 2017), with demographic, bibliographic, and phylogenetic information, and ISO 639-3 three-letter codes for now about 7,500 languages (some extinct). Unfortunately, SIL has now put this valuable resource behind a paywall, but meanwhile Glottolog makes available online

¹⁸ Taking the total number of Ph.D. grammars between 1954 and 2016 as a measure, the ranking is #1 U. Leiden (75), #2 UC Berkeley (74), #3 Deccan College India (74), #4 Australian National U. (71), #5 Sorbonne Nouvelle (Paris 3) (65), #6 UCLA (49), #7 Indiana U. (47), #8 UT Austin (45), #9 University of London (41), #10 U. Hawai'i (39). Taking a more recent measure, the number of grammars between 2000 and 2016, the ranking is: #1 U. Leiden (49), #2 UT Austin (28), #3 U. E. Campinas (21), #4 LaTrobe (31), #5 U. Oregon (20), #6 U. Lumière Lyon (20), #7 Australian National U. (19), SUNY Buffalo (19), Radboud U. Nijmegen (19), #10 U. Hawai'i (17).

constantly updated, full bibliographic information with much more conservative, academically accepted, phylogenies (Hammarström et al. 2017). These catalogues allow for the first time in history a systematic overview of the languages of the world. They make clear the enormity of the descriptive problem in at least two respects. First, we can now safely estimate the number of living languages at around 7,000. Curiously, this number has actually gone up, as some languages were prematurely assumed defunct, and others were added that had previously passed under the radar, including Cuba (McGill & Blench 2012), Baazəm (Lovegren 2012), Ulterior Mixe (Wichmann 2008), and Yilan Creole Japanese (Yuehchen & Shinji 2010), plus many village-level sign languages, such as Kata Kolok (Marsaja 2008), Adamorobe (Nyst 2007), and San Juan Quiahije Chatino (Hou 2016). Second, among 428 Glottolog-recognized language families, close to half (188) are isolates, which have special importance as the last survivors of presumed whole lineages. Taken together, these isolates represent more than 40% of the 430-odd identifiable independent trajectories in language evolution that linguists can use, for example, to test models of how language structures arise.

We also have some kind of overview of the range of language diversity thanks to other more recent developments, the growth of lexical databases (ASJP with basic lexicon on over 5,000 languages; see Wichmann et al. 2018), and typological surveys or language databases with systematic feature inventories. The first, the *World atlas of language structures (WALS)*, released 2005 as an atlas, since 2008 online (Dryer & Haspelmath 2013), has had substantial impact, with many papers based on its data in interdisciplinary journals like *Science* and *PNAS*; another is AUTOTYP (Bickel et al. 2017). Yet other databases have aimed at more systematic coverage of features per language, affording uses such as structural phylogenetics for historical purposes (Dunn et al. 2005) or to show the causal linkage across features (Dunn et al. 2011). Under construction now is a family of improved databases (Glottobank; see <http://glottobank.org/>) for morphosyntax (Grambank), basic vocabulary (Lexibank), cardinal numbers (Numeralbank), and paradigm structure (Parabank). PHOIBLE collects phoneme inventories, now including 1,700 languages (Moran et al. 2014), similar in coverage to the World Phonotactics Database (Donohue et al. 2013).

With the caveat that all such databases are only as good as the underlying source material and inherit the limitations of Procrustean coding that attempts to fit the diversity of language-specific categories into crosslinguistically applicable definitions, they nonetheless offer us unparalleled insight into which phenomena are rare, the distribution of features in phylogenetic and geographic space, and the prehistory of linguistic diversity. The parallel development of powerful data-mining and reconstruction tools (such as phylogenetics, phylogeography, causal graphs) enables linguistics to explore on a global scale the parameters that have shaped global language diversity—biases of various types from preference in the neurophysiology of processing (Bickel et al. 2015), variable influences of the environment on speech sounds (Maddieson & Coupé 2015), possible effects of different genetic profiles on the evolvability of tone (Dediu & Ladd 2007), the impact of population size on phonemic inventories (Moran et al. 2012) and on morphological typology (Nettle 2012), and patterns of sound symbolism (Blasi et al. 2016). Thus in 2017 we are in a far better position not just to investigate, measure, and appreciate the typological and cultural diversity of languages than we were in 1992, but also to model how this diversity reflects a testable set of evolutionary biases which engender a wide range of design solutions while still exhibiting orderliness in the frequency distribution of structural outcomes. For the first time linguistics has the sense of cumulative science and global overview that will allow it to partner with the larger,

more modern sciences in answering fundamental questions about the nature and origins of our preeminent ability.

4. THE SCOPE FOR IMPROVEMENT. Although much has been achieved in the last twenty-five years, there are also opportunities that have been squandered. The first is the failure to collect closely controlled material that would allow coordinated comparison. If researchers could agree to undertake a small battery of tasks on every language (e.g. standardized word lists, recording of minimal pairs in carrier phrases, inventories of grammatical categories, semantic field stimuli—although some of this may not be possible when working with severely threatened languages of which only elderly speakers remain), in addition to their own major descriptive goals, it would be possible to reconcile the requirements of systematic comparison with the search for the unique genius of specific languages. These standard instruments support the construction of phylogenies and the search for relationships across language families, the characterization of the phonetic design space, the search for implicational universals in morphosyntax or semantics, and so forth, thus greatly enhancing the value of the individual language data points.

Second, there has been a striking neglect of sociolinguistics within the documentarist program, despite the gathering of much primary usage data. We actually know little about the sociolinguistics of the small, rural, usually unwritten languages typically targeted by this research (for exceptions see Childs et al. 2014, Hildebrandt et al. 2017; see also Hill 2006)—yet these were the societies in which the current linguistic diversity largely evolved. Without this, we remain ignorant about the key engines of linguistic diversity throughout most of human history. Multilingualism in its myriad forms is often largely bypassed by language documentation projects, even though it may structure language change (McConvell & Meakins 2005, O’Shannessy 2013, Epps 2019) or resist it (Singer & Harris 2016).

Third, the semantic side of language documentation has been relatively neglected, which risks leaving the interpretation of much recorded material unclear. This systematic neglect (Evans & Sasse 2007) reflects both the relative difficulty of acquiring insight and of providing sufficient context in time-aligned digital formats. Addressing the former, advanced methods for documenting semantics and pragmatics have recently been developed (Matthewson 2004, Bochnak & Matthewson 2015, Murray 2017). Regarding the latter, new technologies like the Aikuma app (Bird et al. 2014, Adda et al. 2016, Gauthier et al. 2016) allow interpolation of oral commentary, helping fieldworkers collect the necessary interpretative material.

Fourth, language documentation projects typically manifest a yawning gap between the amount of material recorded and archived and the amount of data that is minimally annotated (transcribed and translated), let alone more thoroughly analyzed (e.g. morphologically segmented and glossed). As a result, the corpus size resulting from most language documentation programs is much smaller than what is required to answer many questions, severely diminishing its utility for future generations. This gap reflects the ‘transcription bottleneck’: an hour of recorded material can take between 40 and 100 hours to transcribe. A concerted attack on this problem needs to involve three steps: (i) optimizing transcription software like ELAN for more ergonomic use, (ii) developing machine-learning tools that can spot recurrent sound-word patterns on the basis of transcribed training samples to suggest transcriptions for the commonest chunks of language, and (iii) field linguists modifying their practice so as to assist the task of machine learning. This latter may simply take the form of good recording practice like

using head-mounted microphones and standardized transcription. But another way of enhancing the success of machine-assisted transcription is to make multiple parallel or semi-parallel recordings, so as to have a robust envelope of phonetic variation across speakers that assists in generalizing sound-transcription matching from one speaker to another (Ćavar et al. 2016, Adams et al. 2018). Here the push to bring in smart machine-learning tools to open up the transcription bottleneck meets the imperative to record more sociolinguistic variation, as outlined above.

5. CONCLUSIONS. A quarter of a century on, we are in a different, much improved place regarding our understanding of language endangerment, academic and community responses, and the scientific knowledge about linguistic diversity, which has been the focus of this discussion note. The data that has been flowing in from the twenty-five years of intensified language documentation does not just enrich our conception of the structural design space. It is also beginning to give a more nuanced and variable picture of the constraints on language processing and learning which themselves sculpt language evolution as structural candidates through the ‘now or never bottleneck’ of being used, learned, absorbed into speech-community norms, and transmitted onward.

As far as endangerment is concerned, however, little has changed: language loss is accelerating under the steamrollers of nation-state consolidation, fast development, habitat erosion, and globalization. There is thus the same or greater urgency than was expressed in 1992, with the very same things at stake: the eons of cultural and cognitive adaptations to vanishing environments, and the human emotional response to them, are being destroyed or abandoned on a worldwide scale. As a concrete example of what we are losing, consider the following: of the 122 compounds derived from plants in the modern pharmacopeia, 80% were derived from compounds used in traditional healing, that is, from ethnomedicine. This reflects the difficulty of finding effective biologic agents simply by direct screening of the over quarter of a million plants on the planet at c. 230 million US dollars per shot (Fabricant & Farnsworth 2001). This is just one token of everything else we are losing when we lose the traditional knowledge of an ethnic group, much of it bound up in language: knowledge of food and water resources and local agricultural practices, ways of managing ecologies for sustainability, kinship and social systems that embody variants of how to live together to mutual benefit, how to rear self-confident and happy children (Diamond 2012), visions of humans’ place in nature, and so much else, both material and spiritual.

REFERENCES

- ADAMS, OLIVER; TREVOR COHN; GRAHAM NEUBIG; HILARIA CRUZ; STEVEN BIRD; and ALEXIS MICHAUD. 2018. Evaluating phonemic transcription of low-resource tonal languages for language documentation. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan, 3356–65. Online: <http://www.lrec-conf.org/proceedings/lrec2018/pdf/490.pdf>.
- ADDA, GILLES; SEBASTIAN STÜKER; MARTINE ADDA-DECKER; ODETTE AMBOUROUE; LAURENT BESACIER; DAVID BLACHON; HÉLÈNE BONNEAU-MAYNARD; et al. 2016. Breaking the unwritten language barrier: The BULB project. *Procedia Computer Science* (Special issue: 5th Workshop on Spoken Language Technologies for Under-resourced Languages (SLTU-2016), Yogyakarta, Indonesia) 81.8–14. DOI: 10.1016/j.procs.2016.04.023.
- ALLEN, SHANLEY E. M. 2017. Polysynthesis in the acquisition of Inuit languages. *The Oxford handbook of polysynthesis*, ed. by Michael Fortescue, Marianne Mithun, and Nicholas Evans, 449–72. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199683208.013.25.

- AMERY, ROB. 2009. Phoenix or relic? Documentation of languages with revitalization in mind. *Language Documentation & Conservation* 3.138–48. DOI: 10.1255/4436.
- ANDERSON, STEPHEN R. 2010. *How many languages are there in the world?* (Linguistic Society of America brochure series: Frequently asked questions.) Washington, DC: Linguistic Society of America. Online: <http://www.linguisticsociety.org/content/how-many-languages-are-there-world>.
- BAVIN, EDITH L., and SABINE STOLL (eds.) 2013. *The acquisition of ergativity*. (Trends in language acquisition research 9.) Amsterdam: John Benjamins. DOI: 10.1075/tilar.9.
- BEREZ-KROEKER, ANDREA L.; HELENE N. ANDREASSEN; LAUREN GAWNE; GARY HOLTON; SUSAN SMYTHE KUNG; PETER PULSIFER; LAUREN B. COLLISTER; THE DATA CITATION AND ATTRIBUTION IN LINGUISTICS GROUP; and THE LINGUISTICS DATA INTEREST GROUP. 2018. The Austin Principles of data citation in linguistics. Version 1.0. Online: <http://site.uit.no/linguisticsdatacitation/austinprinciples/>.
- BERGQVIST, HENRIK. 2016. Complex epistemic perspective in Kogi (Arwako). *International Journal of American Linguistics* 82.1–34. DOI: 10.1086/684422.
- BICKEL, BALTHASAR; JOHANNA NICHOLS; TARAS ZAKHARKO; ALENA WITZLACK-MAKAREVICH; KRISTINE HILDEBRANDT; MICHAEL RIESSLER; LENNART BIERKANDT; FERNANDO ZÚÑIGA; and JOHN B. LOWE. 2017. The AUTOTYP typological databases. Version 0.1.0. Online: <https://github.com/autotyp/autotyp-data/tree/0.1.0>.
- BICKEL, BALTHASAR; ALENA WITZLACK-MAKAREVICH; KAMAL K. CHOUDHARY; MATTHIAS SCHLESEWSKY; and INA BORNKESSEL-SCHLESEWSKY. 2015. The neurophysiology of language processing shapes the evolution of grammar: Evidence from case marking. *PLoS ONE* 10:e0132819. DOI: 10.1371/journal.pone.0132819.
- BIRD, STEPHEN; FLORIAN R. HANKE; OLIVER ADAMS; and HAEJOONG LEE. 2014. Aikuma: A mobile app for collaborative language documentation. *Proceedings of the 2014 Workshop on the Use of Computational Methods in the Study of Endangered Languages*, Baltimore, 1–5.
- BLASI, DAMIÁN E.; SØREN WICHMANN; HARALD HAMMARSTRÖM; PETER F. STADLER; and MORTEN H. CHRISTIANSEN. 2016. Sound–meaning association biases evidenced across thousands of languages. *Proceedings of the National Academy of Sciences of the United States of America* 113.10818–23. DOI: 10.1073/pnas.1605782113.
- BOCHNAK, M. RYAN, and LISA MATTHEWSON (eds.) 2015. *Methodologies in semantic fieldwork*. Oxford: Oxford University Press.
- BOERSMA, PAUL, and DAVID WEENINK. 2018. Praat: Doing phonetics by computer. Online: <http://www.praat.org/>.
- BOWERN, CLAIRE. 2017. Language vitality: Theorizing language loss, shift, and reclamation (Response to Mufwene). *Language* 93.e243–e253. DOI: 10.1353/lan.2017.0068.
- BRENZINGER, MATTHIAS (ed.) 2007. *Language diversity endangered*. (Trends in linguistics: Studies and monographs 181.) Berlin: Mouton de Gruyter.
- BRICKELL, TIMOTHY C., and STEFAN SCHNELL. 2017. Do grammatical relations reflect information status? Reassessing preferred argument structure theory against discourse data from Tondano. *Linguistic Typology* 21.177–208. DOI: 10.1515/lingty-2017-0005.
- BROWN, LEA, and MATTHEW S. DRYER. 2008. The verbs for ‘and’ in Walman, a Torricelli language of Papua New Guinea. *Language* 84.528–65. DOI: 10.1353/lan.0.0044.
- CAPELL, ARTHUR. 1962. Linguistic research needed in Australia. *Bulletin of the International Committee on Urgent Anthropological Ethnological Research* 5.23–28.
- ĆAVAR, MAŁGORZATA; DAMIR ĆAVAR; and HILARIA CRUZ. 2016. Endangered language documentation: Bootstrapping a Chatino speech corpus, forced aligner, ASR. *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, Paris, 23–28. Online: http://www.lrec-conf.org/proceedings/lrec2016/pdf/1006_Paper.pdf.
- CHILDS, TUCKER; JEFF GOOD; and ALICE MITCHELL. 2014. Beyond the ancestral code: Towards a model for sociolinguistic language documentation. *Language Documentation & Conservation* 8.168–91. DOI: 10.1255/4436.
- CRIPPEN, JAMES A., and LAURA C. ROBINSON. 2013. In defense of the lone wolf: Collaboration in language documentation. *Language Documentation & Conservation* 7.123–35. DOI: 10.1255/4577.

- CRUZ, EMILIANA, and ANTHONY C. WOODBURY. 2014. Finding a way into a family of tone languages: The story and methods of the Chatino Language Documentation Project. *Language Documentation & Conservation* 8.490–524. DOI: 10.125/24615.
- CZAYKOWSKA-HIGGINS, EWA. 2009. Research models, community engagement, and linguistic fieldwork: Reflections on working within Canadian indigenous communities. *Language Documentation & Conservation* 3.15–50. DOI: 10.125/4423.
- DĄBROWSKA, EWA. 2010. Naive v. expert intuitions: An empirical study of acceptability judgments. *The Linguistic Review* 27.1–23. DOI: 10.1515/tlir.2010.001.
- DALRYMPLE, MARY; MAKOTO KANAZAWA; YOOKYUNG KIM; SAM MCHOMBO; and STANLEY PETERS. 1998. Reciprocal expressions and the concept of reciprocity. *Linguistics and Philosophy* 21.159–210. Online: <https://www.jstor.org/stable/25001700>.
- DARWIN, CHARLES. 1874. *The descent of man and selection in relation to sex*. 2nd edn. London: John Murray.
- DEDIU, DAN, and D. ROBERT LADD. 2007. Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin. *Proceedings of the National Academy of Sciences of the United States of America* 104. 10944–49. DOI: 10.1073/pnas.0610848104.
- DENNETT, DANIEL C. 1995. *Darwin's dangerous idea: Evolution and the meanings of life*. New York: Simon & Schuster.
- DI CARLO, PIERPAOLO, and JEFF GOOD. 2017. The vitality and diversity of multilingual repertoires: Commentary on Mufwene. *Language* 93.e254–e262. DOI: 10.1353/lan.2017.0069.
- DIAMOND, JARED. 2012. *The world until yesterday: What can we learn from traditional societies?* New York: Viking.
- DI CANIO, CHRISTIAN T. 2009. The phonetics of register in Takhian Thong Chong. *Journal of the International Phonetic Association* 39.162–88. DOI: 10.1017/S0025100309003879.
- DOBRIN, LISE M. 2008. From linguistic elicitation to eliciting the linguist: Lessons in community empowerment from Melanesia. *Language* 84.300–324. DOI: 10.1353/lan.0.0009.
- DOBRIN, LISE M.; PETER K. AUSTIN; and DAVID NATHAN. 2007. Dying to be counted: The commodification of endangered languages in documentary linguistics. *Language Documentation and Description* 6.37–52. Online: <http://www.elpublishing.org/PID/070>.
- DONOHUE, MARK; REBECCA HETHERINGTON; JAMES MCELVENNY; and VIRGINIA DAWSON. 2013. *World phonotactics database*. Canberra: Department of Linguistics, The Australian National University. Online: <http://phonotactics.anu.edu.au>.
- DORIAN, NANCY C. (ed.) 1989. *Investigating obsolescence: Studies in language contraction and death*. (Studies in the social and cultural foundations of language 7.) Cambridge: Cambridge University Press.
- DOW, JAMES R. (ed.) 1987. *New perspectives on language maintenance and language shift I*. (Special issue of *International Journal of the Sociology of Language* 68.) Berlin: Mouton de Gruyter. DOI: 10.1515/ijsl.1987.68.0.
- DOW, JAMES R. (ed.) 1988. *New perspectives on language maintenance and language shift II*. (Special issue of *International Journal of the Sociology of Language* 69.) Berlin: Mouton de Gruyter. DOI: 10.1515/ijsl.1988.69.0.
- DRESSLER, WOLFGANG U., and RUTH WODAK-LEODOLTER (eds.) 1977. *Language death*. (Special issue of *International Journal of the Sociology of Language* 12.) Berlin: Mouton de Gruyter. DOI: 10.1515/ijsl.1977.12.0.
- DRYER, MATTHEW S., and MARTIN HASPELMATH (eds.) 2013. *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Online: <http://wals.info/>.
- DU BOIS, JOHN W. 1987. The discourse basis of ergativity. *Language* 63.805–53. DOI: 10.2307/415719.
- DUNN, MICHAEL; SIMON J. GREENHILL; STEPHEN C. LEVINSON; and RUSSELL D. GRAY. 2011. Evolved structure of language shows lineage-specific trends in word-order universals. *Nature* 473.79–82. DOI: 10.1038/nature09923.
- DUNN, MICHAEL; ANGELA TERRILL; GER REESINK; ROBERT A. FOLEY; and STEPHEN C. LEVINSON. 2005. Structural phylogenetics and the reconstruction of ancient language history. *Science* 309.2072–75. DOI: 10.1126/science.1114615.

- EDMONDSON, JEROLD A., and JOHN H. ESLING. 2006. The valves of the throat and their functioning in tone, vocal register and stress: Laryngoscopic case studies. *Phonology* 23.157–91. DOI: 10.1017/S095267570600087X.
- ENGLAND, NORA C. 1992. Doing Mayan linguistics in Guatemala. *Language* 68.29–35. DOI: 10.1353/lan.1992.0052.
- EPPS, PATIENCE. 2019. Amazonian linguistic diversity and its sociocultural correlates. *Language dispersal, diversification, and contact: A global perspective*, ed. by Mily Crevels and Pieter Muysken. Oxford: Oxford University Press, to appear.
- ESPOSITO, CHRISTINA MARIE. 2004. Santa Ana del Valle Zapotec phonation. *UCLA Working Papers in Phonetics* 103.71–105. Online: <https://escholarship.org/uc/item/2c70d80m>.
- EVANS, NICHOLAS. 2006. View with a view: Towards a typology of multiple perspective constructions. *Berkeley Linguistics Society* 31.93–120. DOI: 10.3765/bls.v31i1.3429.
- EVANS, NICHOLAS. 2008. Review of Gippert et al. 2006. *Language Documentation & Conservation* 2.340–50. DOI: 10125/4353.
- EVANS, NICHOLAS. 2010. *Dying words: Endangered languages and what they have to tell us*. Oxford: Wiley-Blackwell.
- EVANS, NICHOLAS; HENRIK BERGQVIST; and LILA SAN ROQUE. 2018a. The grammar of engagement I: Framework and initial exemplification. *Language and Cognition* 10.110–40. DOI: 10.1017/langcog.2017.21.
- EVANS, NICHOLAS; HENRIK BERGQVIST; and LILA SAN ROQUE. 2018b. The grammar of engagement II: Typology and diachrony. *Language and Cognition* 10.141–70. DOI: 10.1017/langcog.2017.22.
- EVANS, NICHOLAS; ALICE GABY; STEPHEN C. LEVINSON; and ASIFA MAJID (eds.) 2011. *Reciprocals and semantic typology*. Amsterdam: John Benjamins.
- EVANS, NICHOLAS, and STEPHEN C. LEVINSON. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32.429–92. DOI: 10.1017/S0140525X0999094X.
- EVANS, NICHOLAS, and HANS-JÜRGEN SASSE. 2007. Searching for meaning in the library of Babel: Field semantics and problems of digital archiving. *Archives and Social Studies: A Journal of Interdisciplinary Research* 1.63–123.
- FABRICANT, DANIEL S., and NORMAN R. FARNSWORTH. 2001. The value of plants used in traditional medicine for drug discovery. *Environmental Health Perspectives* (Supplement 1: Reviews in environmental health) 109.69–75. DOI: 10.2307/3434847.
- FAST, ANNICKA. 2007. Moral incoherence in documentary linguistics: Theorizing the interventionist aspect of the field. *Proceedings of the Fifth University of Cambridge Postgraduate Conference in Language Research*, Cambridge, 64–71. Online: http://www.ling.cam.ac.uk/camling/Manuscripts/CamLing2007_Fast.pdf.
- FILIPOVIĆ, LUNA, and MARTIN PÜTZ (eds.) 2016. *Endangered languages and languages in danger: Issues of documentation, policy, and language rights*. (IMPACT: Studies in language and society 42.) Amsterdam: John Benjamins. DOI: 10.1075/impact.42.
- FISHMAN, JOSHUA A. 1982. Whorfianism of the third kind: Ethnolinguistic diversity as a worldwide societal asset (The Whorfian hypothesis: Varieties of validation, confirmation, and disconfirmation II). *Language in Society* 11.1–14. Online: <https://www.jstor.org/stable/4167289>.
- FISHMAN, JOSHUA A. 1991. *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*. (Multilingual matters 76.) Clevedon: Multilingual Matters.
- FITZGERALD, COLLEEN M. 2017. Understanding language vitality and reclamation as resilience: A framework for language endangerment and ‘loss’ (Commentary on Mufwene). *Language* 93.e280–e297. DOI: 10.1353/lan.2017.0072.
- FITZGERALD, COLLEEN M. 2018. Creating sustainable models of language documentation and revitalization. *Insights from practices in community-based research: From theory to practice around the globe* (Trends in linguistics: Studies and monographs 319), ed. by Shannon T. Bischoff and Carmen Jany, 94–111. Berlin: De Gruyter Mouton. DOI: 10.1515/9783110527018-005.
- FITZGERALD, COLLEEN M., and MARY S. LINN. 2013. Training communities, training graduate students: The 2012 Oklahoma Breath of Life Workshop. *Language Documentation & Conservation* 7.185–206. DOI: 10125/4596.

- FLECK, DAVID W. 2003. *A grammar of Matses*. Houston: Rice University dissertation.
- FLECK, DAVID W. 2007. Evidentiality and double tense in Matses. *Language* 83.589–614. DOI: 10.1353/lan.2007.0113.
- FLOYD, SIMEON; ELIZABETH NORCLIFFE; and LILA SAN ROQUE (eds.) 2018. *Egophoricity*. (Typological studies in language 118.) Amsterdam: John Benjamins. DOI: 10.1075/tsl.118.
- FORSYTH, BILL; LUCINDA DAVIDSON; BARBARA KELLY; RACHEL NORDLINGER; GILLIAN WIGGLESWORTH; and JOE BLYTHE. 2017. The acquisition of Murrinhpatha (Northern Australia). *The Oxford handbook of polysynthesis*, ed. by Michael Fortescue, Marianne Mithun, and Nicholas Evans, 473–94. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199683208.013.27.
- GAUTHIER, ELODIE; DAVID BLACHON; LAURENT BESACIER; GUY-NOEL KOUARATA; MARTINE ADDA-DECKER; ANNIE RIALLAND; GILLES ADDA; and GRÉGOIRE BACHMAN. 2016. LIG-AIKUMA: A mobile app to collect parallel speech for under-resourced language studies. *INTERSPEECH 2016: Show & Tell contribution*, San-Francisco. Online: https://www.isca-speech.org/archive/Interspeech_2016/pdfs/2003.PDF.
- GENETTI, CAROL, and REBEKKA SIEMENS. 2013. Training as empowering social action: An ethical response to language endangerment. *Responses to language endangerment: In honor of Mickey Noonan: New directions in language documentation and language revitalization* (Studies in language companion series 142), ed. by Elena Mihás, Bernard Perley, Gabriel Rei-Doval, and Kathleen Wheatley, 59–78. Amsterdam: John Benjamins. DOI: 10.1075/slcs.142.04gen.
- GIBSON, EDWARD, and EVELINA FEDORENKO. 2013. The need for quantitative methods in syntax and semantics research. *Language and Cognitive Processes* 28.88–124. DOI: 10.1080/01690965.2010.515080.
- GIPPERT, JOST; NIKOLAUS P. HIMMELMANN; and ULRIKE MOSEL (eds.) 2006. *Essentials of language documentation*. Berlin: Mouton de Gruyter.
- GREEN, JENNIFER. 2014. *Drawn from the ground: Sound, sign and inscription in Central Australian sand stories*. (Language, culture & cognition 13.) Cambridge: Cambridge University Press.
- GRENOBLE, LENORE A. 2016. A response to ‘Assessing levels of endangerment in the Catalogue of Endangered Languages (ELCat) using the Language Endangerment Index (LEI)’, by Nala Huiying Lee & John Van Way. *Language in Society* 45.293–300. DOI: 10.1017/S0047404515000950.
- HAIG, GEOFFREY, and STEFAN SCHNELL. 2016. The discourse basis of ergativity revisited. *Language* 92.591–618. DOI: 10.1353/lan.2016.0049.
- HALE, KEN. 1992. Language endangerment and the human value of linguistic diversity. *Language* 68.35–42. DOI: 10.1353/lan.1992.0052.
- HALE, KEN; MICHAEL KRAUSS; LUCILLE J. WATAHOMIGIE; AKIRA Y. YAMAMOTO; COLETTE CRAIG; LAVERNE MASAYESVA JEANNE; and NORA C. ENGLAND. 1992. Endangered languages. *Language* 68.1–42. DOI: 10.1353/lan.1992.0052.
- HAMMARSTRÖM, HARALD; SEBASTIAN BANK; ROBERT FORKEL; and MARTIN HASPELMATH (eds.) 2017. *Glottolog 3.1*. Jena: Max Planck Institute for the Science of Human History. Online: <http://glottolog.org/>.
- HAMMARSTRÖM, HARALD; THOM CASTERMANS; ROBERT FORKEL; KEVIN VERBEEK; MICHEL A. WESTENBERG; and BETTINA SPECKMANN. 2018. Simultaneous visualization of language endangerment and language description. *Language Documentation & Conservation* 12.359–92. DOI: 10.125/24792.
- HARGREAVES, DAVID. 2005. Agency and intentional action in Kathmandu Newar. *Himalayan Linguistics* 5.1–48. DOI: 10.5070/H95022977.
- HARMON, DAVID, and JONATHAN LOH. 2010. The index of linguistic diversity: A new quantitative measure of trends in the status of the world’s languages. *Language Documentation & Conservation* 4.97–151. DOI: 10.125/4474.
- HARRISON, DAVID K. 2007. *When languages die: The extinction of the world’s languages and the erosion of human knowledge*. (Oxford studies in sociolinguistics.) Oxford: Oxford University Press.
- HEATH, JEFFREY. 1984. *Functional grammar of Nunggubuyu*. Canberra: Australian Institute of Aboriginal Studies.
- HENDERSON, BRENT; PETER ROHLOFF; and ROBERT HENDERSON. 2014. More than words: Towards a development-based approach to language revitalization. *Language Documentation & Conservation* 8.75–91. DOI: 10.125/4611.

- HILDEBRANDT, KRISTINE A.; CARMEN JANY; and WILSON SILVA (eds.) 2017. *Documenting variation in endangered languages*. (Language Documentation & Conservation special publication 13.) Honolulu: University of Hawai'i Press. DOI: 10.125/24754.
- HILL, JANE H. 2006. The ethnography of language and language documentation. In Gippert et al., 113–28.
- HIMMELMANN, NIKOLAUS P. 1998. Documentary and descriptive linguistics. *Linguistics* 36.161–95. DOI: 10.1515/ling.1998.36.1.161.
- HIMMELMANN, NIKOLAUS P. 2006. Language documentation: What is it and what is it good for? In Gippert et al., 1–30.
- HINTON, LEANNE. 2001. The use of linguistic archives in language revitalization: The Native California Language Restoration Workshop. *The green book of language revitalization in practice*, ed. by Leanne Hinton and Kenneth Hale, 419–28. Leiden: Brill.
- HOU, LYNN YONG-SHI. 2016. *Making hands': Family sign languages in the San Juan Quiahije community*. Austin: University of Texas at Austin dissertation. DOI: 10.15781/T2W08WN5V.
- HYMAN, LARRY. 2007. Elicitation as experimental phonology: Thlantland Lai tonology. *Experimental approaches to phonology*, ed. by Maria-Josep Solé, Patrice Speeter Beddor, and Manjari Ohala, 7–24. Oxford: Oxford University Press.
- HYMAN, LARRY M. 2016. Morphological tonal assignments in conflict: Who wins? *Tone and inflection: New facts and new perspectives* (Trends in linguistics: Studies and monographs 296), ed. by Enrique L. Palancar and Jean Léo Léonard, 15–39. Berlin: De Gruyter Mouton. DOI: 10.1515/9783110452754-002.
- JEANNE, LAVERNE MASAYESVA. 1992. An institutional response to language endangerment: A proposal to create a Native American Language Centre. *Language* 68.24–28. DOI: 10.1353/lan.1992.0052.
- JUKES, ANTHONY. 2011. Researcher training and capacity development in language documentation. *The Cambridge handbook of endangered languages*, ed. by Peter K. Austin and Julia Sallabank, 423–45. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511975981.
- KANT, IMMANUEL. 2006 [1798]. *Anthropology from a pragmatic point of view*, trans. by Robert B. Loudon. Cambridge: Cambridge University Press.
- KAXINAWÁ, JOAQUIM PAULO DE LIMA. 2014. *Hãtxa kuñ haska xarabu*. Brasília: Universidade de Brasília dissertation. Online: <http://repositorio.unb.br/handle/10482/18984>.
- KOVACH, MARGARET. 2010. *Indigenous methodologies: Characteristics, conversations and contexts*. Toronto: University of Toronto Press.
- KRAUSS, MICHAEL. 1992. The world's languages in crisis. *Language* 68.4–10. DOI: 10.1353/lan.1992.0075.
- KUNG, SUSAN SMYTHE, and JOEL SHERZER. 2013. The Archive of the Indigenous Languages of Latin America: An overview. *Oral Tradition* 28.379–88.
- LABOV, WILLIAM. 1975. Empirical foundations of linguistic theory. *The scope of American linguistics*, ed. by Robert Austerlitz, 77–113. Lisse: Peter de Ridder.
- LANDABURU, JON. 2007. La modalisation du savoir en langue andoke (Amazonie Colombienne). *L'annonciation médiatisée II : Le traitement épistémologique de l'information*, ed. by Zlatka Guentchéva and Jon Landaburu, 23–47. Louvain: Peeters.
- LÉONARD, JEAN LÉO, and KARLA JANIRÉ AVILÉS GONZÁLEZ (eds.) 2015. *Documentation et revitalisation des langues en danger : Épistémologie et praxis*. Paris: Michel Houdiard.
- LEONARD, WESLEY Y. 2017. Producing language reclamation by decolonising 'language'. *Language Documentation and Description* 14.13–36. Online: <http://www.ejournals.org/PID/150>.
- LEVINSON, STEPHEN C. 2003. *Space in language and cognition: Explorations in cognitive diversity*. Cambridge: Cambridge University Press.
- LEVINSON, STEPHEN C.; SARAH CUTFIELD; MICHAEL DUNN; N. J. ENFIELD; and SÉRGIO MEIRA (eds.) 2018. *Demonstratives in cross-linguistic perspective*. Cambridge: Cambridge University Press.
- LEVINSON, STEPHEN C., and ASIFA MAJID. 2014. Differential ineffability and the senses. *Mind & Language* 29.407–27. DOI: 10.1111/mila.12057.
- LEVINSON, STEPHEN C., and DAVID P. WILKINS (eds.) 2006. *Grammars of space: Explorations in cognitive diversity*. (Language, culture & cognition 6.) Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511486753.
- LEWIS, M. PAUL, and GARY F. SIMONS. 2010. Assessing endangerment: Expanding Fishman's GIDS. *Revue Roumaine de Linguistique* 55.103–20.

- LEWIS, M. PAUL; GARY F. SIMONS; and CHARLES D. FENNIG (eds.) 2016. *Ethnologue: Languages of the world*. 19th edn. Dallas: SIL International. Online: <http://www.ethnologue.com/19/>.
- LINN, MARY S. 2014. Living archives: A community-based language archive model. *Language Documentation and Description* (Special issue on language documentation and archiving) 12.53–67. Online: <http://www.elpublishing.org/PID/137>.
- LOH, JONATHAN, and DAVID HARMON. 2005. A global index of biocultural diversity. *Ecological Indicators* 5.231–41. DOI: 10.1016/j.ecolind.2005.02.005.
- LOVEGREN, JESSE. 2012. Sparse notes on Baazəm. Buffalo: University at Buffalo, ms.
- LÜPKE, FRIEDERIKE. 2017. African(ist) perspectives on vitality: Fluidity, small speaker numbers, and adaptive multilingualism make vibrant ecologies (Response to Mufwene). *Language* 93.e275–e279. DOI: 10.1353/lan.2017.0071.
- MADDIESON, IAN, and CHRISTOPHE COUPÉ. 2015. Human spoken language diversity and the acoustic adaptation hypothesis. *The Journal of the Acoustical Society of America* 138. 1838. DOI: 10.1121/1.4933848.
- MAFFI, LUISA. 2005. Linguistic, cultural, and biological diversity. *Annual Review of Anthropology* 34.599–617. DOI: 10.1146/annurev.anthro.34.081804.120437.
- MAJID, ASIFA, and NICLAS BURENHULT. 2014. Odors are expressible in language, as long as you speak the right language. *Cognition* 130.266–70. DOI: 10.1016/j.cognition.2013.11.004.
- MARSAJA, I. GEDE. 2008. *Desa kolok—A deaf village and its sign language in Bali, Indonesia*. Nijmegen: Ishara Press.
- MATEO PEDRO, PEDRO. 2015. *The acquisition of inflection in Q'anjob'al Maya*. (Trends in language acquisition research 14.) Amsterdam: John Benjamins. DOI: 10.1075/tilar.14.
- MATTHEWSON, LISA. 2001. Quantification and the nature of crosslinguistic variation. *Natural Language Semantics* 9.145–89. DOI: 10.1023/A:1012492911285.
- MATTHEWSON, LISA. 2004. On the methodology of semantic fieldwork. *International Journal of American Linguistics* 70.369–415. DOI: 10.1086/429207.
- MATTHEWSON, LISA. 2008. Pronouns, presuppositions, and semantic variation. *Proceedings of Semantics and Linguistic Theory (SALT)* 18.527–50. DOI: 10.3765/salt.v18i0.2505.
- MC CONVELL, PATRICK, and FELICITY MEAKINS. 2005. Gurindji Kriol: A mixed language emerges from code-switching. *Australian Journal of Linguistics* 25.9–30. DOI: 10.1080/07268600500110456.
- MCGILL, STUART, and ROGER BLENCH. 2012. Documentation, development, and ideology in the northwestern Kainji languages. *Language Documentation and Description* 11. 90–135. Online: <http://www.elpublishing.org/PID/130>.
- MCPHERSON, LAURA. 2016. The talking xylophone of the Sambla: Seenku phonology in a speech surrogate system. Paper presented at Harvard University, April 29, 2016.
- MORAN, STEVEN; DANIEL McCLOY; and RICHARD WRIGHT. 2012. Revisiting population size vs. phoneme inventory size. *Language* 88.877–93. DOI: 10.1353/lan.2012.0087.
- MORAN, STEVEN; DANIEL McCLOY; and RICHARD WRIGHT (eds.) 2014. *PHOIBLE online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Online: <http://phoible.org/>.
- MORGAN, JULIET. 2017. The learner varieties of the Chikasha Academy: Chickasaw adult language acquisition, change, and revitalization. Norman: University of Oklahoma dissertation. Online: <https://shareok.org/handle/11244/50825>.
- MOSEL, ULRIKE. 2006. Sketch grammar. In Gippert et al., 301–9.
- MUFWENE, SALIKOKO S. 2002. Colonisation, globalisation, and the future of languages in the twenty-first century. *International Journal on Multicultural Societies* 4.162–93. Online: www.unesco.org/shs/ijms/vol4/issue2/art2.
- MUFWENE, SALIKOKO S. 2017. Language vitality: The weak theoretical underpinnings of what can be an exciting research area. *Language* 93.e202–e223. DOI: 10.1353/lan.2017.0065.
- MURRAY, SARAH. 2017. *The semantics of evidentials*. Oxford: Oxford University Press.
- NAUMANN, CHRISTFRIED. 2016. The phoneme inventory of Taa (West !Xoon Dialect). *Lone Tree—Scholarship in the service of the Koon: Essays in memory of Anthony T. Traill*, ed. by Rainer Voßen and Wilfrid H. G. Haacke, 311–51. Cologne: Rüdiger Köppe.
- NETTLE, DANIEL. 1999. *Linguistic diversity*. Oxford: Oxford University Press.

- NETTLE, DANIEL. 2012. Social scale and structural complexity in human languages. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367.1829–36. DOI: 10.1098/rstb.2011.0216.
- NETTLE, DANIEL, and SUZANNE ROMAINE. 2000. *Vanishing voices: The extinction of the world's languages*. Oxford: Oxford University Press.
- NORCLIFFE, ELISABETH; AGNIESZKA E. KONOPKA; PENELOPE BROWN; and STEPHEN C. LEVINSON. 2015. Word order affects the time course of sentence formulation in Tzeltal. *Language, Cognition and Neuroscience* 30.1187–1208. DOI: 10.1080/23273798.2015.1006238.
- NYST, VICTORIA ANNA. 2007. *A descriptive analysis of Adamorobe Sign Language (Ghana)*. Utrecht: LOT. Online: https://www.lotpublications.nl/Documents/151_fulltext.pdf.
- O'MEARA, CAROLYN, and ASIFA MAJID. 2016. How changing lifestyles impact Seri smellscapes and smell language. *Anthropological Linguistics* 58.107–31. DOI: 10.1353/anl.2016.0024.
- O'SHANNESY, CARMEL. 2013. The role of multiple sources in the formation of an innovative auxiliary category in Light Warlpiri, a new Australian mixed language. *Language* 89.328–53. DOI: 10.1353/lan.2013.0025.
- PAPERNO, DENIS, and EDWARD L. KEENAN (eds.) 2017. *Handbook of quantifiers in natural language*, vol. 2. Heidelberg: Springer.
- PIMM, S. L.; C. N. JENKINS; R. ABELL; T. M. BROOKS; J. L. GITTLEMAN; L. N. JOPPA; P. H. RAVEN; C. M. ROBERTS; and J. O. SEXTON. 2014. The biodiversity of species and their rates of extinction, distribution, and protection. *Science* 344.1246752. DOI: 10.1126/science.1246752.
- PINKER, STEVEN. 1997. *How the mind works*. New York: Norton.
- POLINSKY, MARIA, and ERIC POTSDAM. 2001. Long-distance agreement and topic in Tsez. *Natural Language and Linguistic Theory* 19.583–646. DOI: 10.1023/A:1010757806504.
- PYE, CLIFTON. 2017. *The comparative method of language acquisition research*. Chicago: University of Chicago Press.
- RHODES, RICHARD A.; LENORE A. GRENOBLE; ANNA BERGE; and PAULA RADETZKY. 2006. Adequacy of documentation: A preliminary report to the CELP [Committee on Endangered Languages and their Preservation, Linguistic Society of America]. Washington, DC: Linguistic Society of America, ms.
- RICE, KEREN. 2006. The role of linguistic theory in writing grammars. *Catching language: The standing challenge of grammar writing*, ed. by Felix K. Ameka, Alan Charles Dench, and Nicholas Evans, 235–68. Berlin: Mouton de Gruyter.
- ROBINS, ROBERT H., and EUGENIUS M. UHLENBECK. 1991. *Endangered languages*. Oxford: Oxford University Press.
- SAMMONS, OLIVIA, and WESLEY LEONARD. 2015. Breathing new life into Algonquian languages: Lessons from the Breath of Life Archival Institute for Indigenous Languages. *Papers of the Forty-Third Algonquian Conference: Actes du Congrès des Algonquistes*, 207–24. Albany: SUNY Press.
- SAN ROQUE, LILA. 2015. Using you to get to me: Addressee perspective and speaker stance in Duna evidential marking. *STUF—Language Typology and Universals* 68.187–210. DOI: 10.1515/stuf-2015-0010.
- SANDLER, WENDY; MARK ARONOFF; IRIT MEIR; and CAROL PADDEN. 2011. The gradual emergence of phonological form in a new language. *Natural Language and Linguistic Theory* 29.503–43. DOI: 10.1007/s11049-011-9128-2.
- SAUPPE, SEBASTIAN. 2017. Symmetrical and asymmetrical voice systems and processing load: Pupillometric evidence from sentence production in Tagalog and German. *Language* 93.288–313. DOI: 10.1353/lan.2017.0015.
- SEIFART, FRANK; JULIEN MEYER; SVEN GRAWUNDER; and LAURE DENTEL. 2018. Reducing language to rhythm: Amazonian Bora drummed language exploits speech rhythm for long-distance communication. *Royal Society Open Science* 5.170354. DOI: 10.1098/rsos.170354.
- SEIFART, FRANK; JAN STRUNK; SWINTHA DANIELSEN; IREN HARTMANN; BRIGITTE PAKENDORF; SØREN WICHMANN; ALENA WITZLACK-MAKAREVICH; NIVJA H. DE JONG; and BALTHASAR BICKEL. 2018. Nouns slow down speech across structurally and cultur-

- ally diverse languages. *Proceedings of the National Academy of Sciences of the United States of America* 115.5720–25. DOI: 10.1073/pnas.1800708115.
- SICOLI, MARK A. 2016. Repair organization in Chinantec whistled speech. *Language* 92.411–32. DOI: 10.1353/lan.2016.0028.
- SIMONS, GARY F., and CHARLES D. FENNIG. 2017. *Ethnologue: Languages of the world*. 20th edn. Dallas: SIL International. Online: <https://www.ethnologue.com/20/>.
- SIMONS, GARY F., and M. PAUL LEWIS. 2013. The world's languages in crisis: A 20-year update. *Responses to language endangerment in honor of Mickey Noonan: New directions in language documentation and language revitalization*, ed. by Elena Mihas, Bernard Perley, Gabriel Rei-Doval, and Kathleen Wheatley, 3–20. Amsterdam: John Benjamins.
- SINGER, RUTH, and SALOME HARRIS. 2016. What practices and ideologies support small-scale multilingualism? A case study of Waruwi community, Northern Australia. *International Journal of the Sociology of Language* 2016.163–208. DOI: 10.1515/ijsl-2016-0029.
- SMITH, LINDA TUHIWAI. 2012. *Decolonizing methodologies: Research and indigenous peoples*. 2nd edn. New York: Zed Books.
- SPERBER, DAN. 1974. *Rethinking symbolism*, trans. by Alice L. Morton. Cambridge: Cambridge University Press.
- STANFORD, JAMES N. 2009. Clan as a sociolinguistic variable: Three approaches to Sui clans. *Variation in indigenous minority languages (IMPACT: Studies in language and society 25)*, ed. by James N. Stanford and Dennis R. Preston, 463–84. Amsterdam: John Benjamins. DOI: 10.1075/impact.25.23sta.
- STOLL, SABINE; BALTHASAR BICKEL; and JEKATERINA MAŽARA. 2017. The acquisition of polysynthetic verb forms in Chintang. *The Oxford handbook of polysynthesis*, ed. by Michael Fortescue, Marianne Mithun, and Nicholas Evans, 495–515. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199683208.013.28.
- STONE, DORIS. 1962. Urgent tasks of research concerning the cultures and languages of Central American Indian tribes. *Bulletin of the International Committee on Urgent Anthropological Ethnological Research* 5.65–69.
- SUTHERLAND, WILLIAM J. 2003. Parallel extinction risk and global distribution of languages and species. *Nature* 423.276–79. DOI: 10.1038/nature01607.
- THIEBERGER, NICHOLAS. 2006. *A grammar of South Efate: An Oceanic language of Vanuatu*. (Oceanic Linguistics special publication 33.) Honolulu: University of Hawai'i Press. DOI: 11343/31242.
- THIEBERGER, NICHOLAS. 2012. Using language documentation data in a broader context. *Potentials of language documentation: Methods, analyses, and utilization* (Language Documentation & Conservation special publication 3), ed. Frank Seifart, Geoffrey Haig, Nikolaus P. Himmelmann, Dagmar Jung, Anna Margetts, and Paul Trilsbeek, 129–34. Honolulu: University of Hawai'i Press. DOI: 10125/4527.
- THONGKUM, THERAPAN L. 1988. Phonation types in Mon-Khmer languages. *Vocal physiology: Voice production, mechanisms and functions*, ed. by Osamu Fujimura, 319–33. New York: Raven Press.
- UNESCO. 2010. *Atlas of the world's languages in danger*. 3rd edn. Ed. by Christopher Moseley. Paris: UNESCO Publications Office.
- VOLK, EREZ. 2011. *Mijikenda tonology*. Tel Aviv: Tel Aviv University dissertation.
- WATAHOMIGIE, LUCILLE J., and AKIRA Y. YAMAMOTO. 1992. Endangered languages. *Language* 68.10–17. DOI: 10.1353/lan.1992.0052.
- WICHMANN, SØREN. 2008. Om opdagelsen af et grænseoverskridende nyt sprog. *De mange veje til Mesoamerika: Hyldestskrift til Una Canger*, ed. by Jesper Nielsen and Mettelise Fritz Hansen, 63–80. Copenhagen: Afdelingen for Indianske Sprog og Kulturer, Institut for Tværkulturelle og Regionale Studier, Københavns Universitet.
- WICHMANN, SØREN; ERIC W. HOLMAN; and CECIL H. BROWN (eds.) 2018. *The ASJP database (version 18)*. Online: <http://asjp.clld.org/>.
- WILSON, SHAWN. 2008. *Research is ceremony: Indigenous research methods*. Black Point, NS: Fernwood. Online: https://epubs.scu.edu.au/gnibi_pubs/17.
- WNUK, EWELINA, and ASIFA MAJID. 2014. Revisiting the limits of language: The odor lexicon of Maniq. *Cognition* 131.125–38. DOI: 10.1016/j.cognition.2013.12.008.
- YAMADA, RACQUEL-MARIA. 2007. Collaborative linguistic fieldwork: Practical application of the empowerment model. *Language Documentation & Conservation* 1.257–82. DOI: 10125/1717.

- YUEHCHEN, CHIEN, and SANADA SHINJI. 2010. Yilan Creole in Taiwan. *Journal of Pidgin and Creole Languages* 25.350–57. DOI: 10.1075/jpcl.25.2.11yue.
- ZESHAN, ULRIKE, and SIBAJI PANDA. 2011. Reciprocal constructions in Indo-Pakistani Sign Language. *Reciprocals and semantic typology*, ed. by Nicholas Evans, Alice Gaby, Stephen C. Levinson, and Asifa Majid, 91–113. Amsterdam: John Benjamins. DOI: 10.1075/tsl.98.05zes.

[frank.seifart@cns.fr]

[nicholas.evans@anu.edu.au]

[harald.hammarstrom@lingfil.uu.se]

[Stephen.Levinson@mpi.nl]

[Received 22 December 2017;

revision invited 12 March 2018;

revision received 4 June 2018;

accepted pending revisions 17 June 2018;

revision received 4 July 2018;

accepted 5 July 2018]