



Documentary lexicography & Reflections on documenting the Nasal lexicon

Blaine Billings
blainebillings.com

Department of Linguistics
University of Hawai'i at Mānoa

Presentation at Linguistics Friday Colloquium Series
University of North Carolina at Chapel Hill

7 November 2025



Contents

1. Documentary lexicography
2. Nasal & Documenting the lexicon
3. Additional research
4. Conclusion

Documentary lexicography

Abstract

A dictionary is often one of the first resources desired by a language community endeavoring to document their language. In spite of this and the supposed equal value accorded to the elements of the traditional Boasian trilogy (grammar, dictionary, corpus), lexicographic output is underrepresented in linguistic literature. This presentation reemphasizes the need to document and analyze the lexicon and discusses approaches and challenges to lexicographic work in the documentary context by drawing on experiences from documenting the lexicon of Nasal, an endangered Austronesian language spoken in southwest Sumatra, Indonesia. The end of this presentation additionally highlights further research topics that naturally arise from a dictionary project and the insights they can provide both into a language itself and into the historical understanding of the broader region.

Language documentation

Primary goal:

“an accurate and adequate record of a language for posterity”

— Rhodes & Campbell (2018: 107)

Documentary linguistics arose to *adapt* linguistics to documentary context (see Austin 2010)

Boasian trilogy: Grammar, dictionary, corpus (texts, recordings, transcriptions, etc.)

Contemporary grammar:dictionary:corpus ratio of 10:3:1 (Grinevald 2001)

Lexicography

Primary goal:

“the creation of dictionaries”

— Rundell (2012: 63)

“to present a full account of the words of a language, in all their meanings and patterns of use”

— (Kilgarrieff 2007)

Lexicography has long-established tradition

Broad output: learners’ dictionaries, terminology dictionaries, mono-/bi-/multi-lingual dictionaries

Prototypical lexicographic work is that of high-resource languages

Lexicography in lang. doc. (1): Goals

Lexicography and documentary linguistics arose out of different traditions

Traditional and *documentary* lexicography share foundation, beyond...

Goals:

- **Lexicography:** Production of dictionaries
- **Documentary linguistics:** Production of adequate record
- Lexicography ultimately concerned with presentation of lexical data
- Documentary linguistics ultimately concerned with description and preservation of data

“There never was lexicography without word-lists and/or dictionaries, but there were for a long time (and still are) word-lists and/or dictionaries without lexicography”
— Hüllen (1999: 3, emphasis original)

Lexicography in lang. doc. (2): Methods

Methods:

- **Traditional methods:** massive text corpora (Altenberg & Granger 1996, Krishnamurthy 2008, Faaß 2018, and many others)
 - Documentary project corpora often *very* small or have to start from scratch
- **Documentary methods:** rapid word collection (Boerger & Stutzman 2018), semantic domain elicitation (Moe 2003)
- **Lexicographic training:** database development, lexical semantics, user-based design
- **Documentary training:** phonetics, phonology, morphology, syntax, field methods, data collection, archiving

“It is regrettable that so few universities in North America offer training in lexicography as part of their graduate programs in linguistics, especially given the central role that dictionaries play in both the documentation and revitalization of Indigenous languages”

— Lachler & Pankratz (2017: 112)

Lexicography in lang. doc. (3): Outputs

Outputs:

- Documentary linguistics operates across raw and primary data (Himmelman 2012)
- Documentary linguistics focuses on *record*, lexicography on *presentation*
- Landmark grammatical works often unmatched by lexicographic counterpart
- **Lexicographic output:** Specialized dictionaries, learner's dictionaries, tens of thousands of entries, example sentence, grammatical information, etc.
- **Documentary output:** (< 15% lexical materials archived, Paterson III 2015)

“shorter dictionaries [such as] Sapir and Swadesh’s posthumous (1960) dictionary, which was edited by Haas and which contains about 3,000 entries; by lexicographic standard, this is too short to be considered a full dictionary”

— Chelliah & de Reuse (2011: 228)

Lexicography in lang. doc. (4): Value

Value:

- Lexicographic work often goes unpublished for decades (→ iterative archiving, publication)
- As academic field, documentary linguistics does not encourage prioritizing production and publication of lexical documentation

“in nearly all graduate departments of linguistics it has been grammars [...] that have been suitable as doctoral dissertations, whereas dictionaries, or, heaven forbid, text collections—the low end of the hierarchy—have not. The usual justification for this [is] that text collection and even dictionary making are only clerical activities”

— Woodbury (2003)

*“It is not impossible for a linguist to compile a dictionary on the basis of a grammar, text collection, and unpublished lexical field notes, but it is impossible to compile a **comprehensive** grammar on the basis of a dictionary and text collection”*

— Chelliah & de Reuse (2011: 227, emphasis added)

Lexicography in lang. doc (5): A comparison

Table 1: Comparison of *documentary* and *traditional* lexicography.

Documentary	vs.	Traditional
Lexical record		Dictionary
Corpus from scratch, elicitation		Corpus of hundreds of millions of words
Definitions via elicitation		Definitions via corpus analysis
Stakeholders primarily communities		Stakeholders primarily publishers
Funding primarily from government entities or academic institutions		Funding primarily from publishers or private institutions

Lexicography in lang. doc. (6): Final thoughts

The lexicon is a vital aspect of any language

“It would be useful to better bridge general lexicography and dictionary making for small-scale communities in the future. It would be helpful for documentary/descriptive dictionary makers to take note of the insights of general lexicographers, who have a wealth of practical experience. In fact, many issues that documentary dictionary makers struggle with also preoccupy general lexicographers”

— Grimm (2022: 555)

Grammaticographic, lexicographic, and corpus work all crucially depend on one another

“The boundary between lexicon and grammar is a difficult one to draw in a principled way, and [...] the two treatments need to be rigorously compatible and in that sense form part of an integrated larger work” — Evans & Dench (2006: 13)

Final thought: If *documentary lexicography* is not a field in its own right, it is at least a highly specialized function of lexicography, one in which documentary linguists remain both tragically unprepared and strikingly unproductive

Nasal & Documenting the lexicon

Austronesian languages

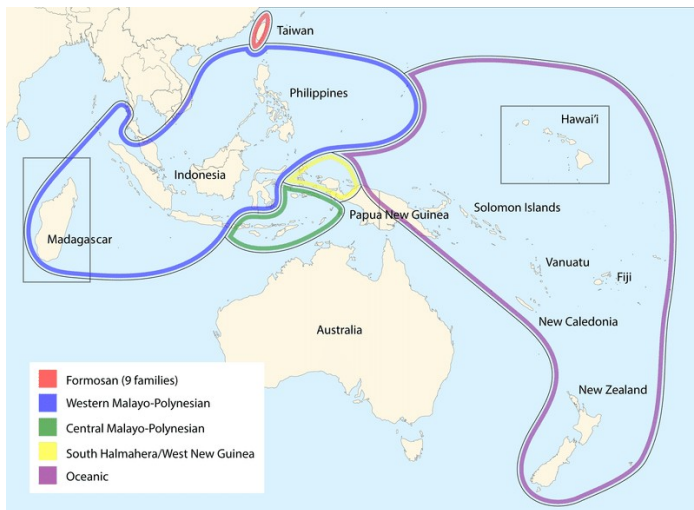


Figure 1: The Austronesian languages.

Nasal people & language

Nasal:

- **Speakers:** ~3,000 people; low intergenerational transmission
- **Region:** Three coastal villages in southwest Sumatra
 - Nasal district, Kaur regency, Bengkulu province, Indonesia
- **Domains:** Only used in private domains (home, friends, market)
 - No use in government, education, media
- **Family:** Sumatran < Malayo-Polynesian < Austronesian (Billings & McDonnell 2024)
- Highly multilingual, diglossic

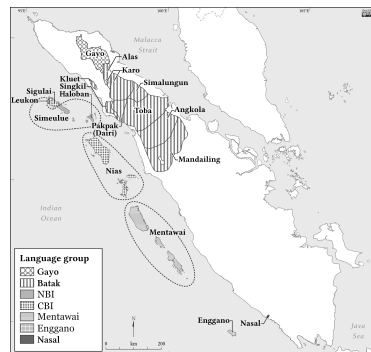


Figure 2: The Sumatran languages.

Nasal documentation (1): Past

Unknown to linguists until regional language survey (Anderbeck & Aprilani 2013)

Pre-existing documentation:

- No grammatical description
- No monolingual written sources
- 5 pre-existing “lexical sources”

Table 2: Number of Nasal lexemes by pre-existing print source.

Source	Total
Roos (1890)	7
Helfrich (1891)	5
Stokhof & Almanar (1987)	1,048
Kasim et al. (1987)	209
Anderbeck & Aprilani (2013)	~2,000

Nasal documentation (2): Present



Figure 3: A Nasal dictionary recording.

Ongoing documentation project:

- Started in 2017 (McDonnell 2017, 2019)
- Collaborative with community members
- **Primary goal:** Create lasting record of Nasal language use and resources for linguistic description and classroom teaching of Nasal language
- **A/V Corpus:** 350hrs speech
 - Everyday conversation
 - Local narratives and traditions
 - Elicitation on various linguistic domains
 - Semantic domain discussion
- **Text Corpus:** 80hrs transcribed speech
- **Additional materials:** Child alphabet and counting books, booklets documenting local tradition, dictionary and grammar development in progress

Nasal lexicon

Three layers of Nasal lexicon:

- **Malayic:** Recent loans from Indonesian, Kaur, and South Barisan Malay
- **Lampungic:** Deeper loans, integrated into the lexicon, from historical contact
- **Nasal:** Core vocabulary, shows relationship to Sumatran, relatively small

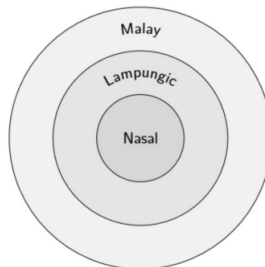


Figure 4: Layers of Nasal lexicon.

Interesting characteristics: Inclusive/exclusive pronoun distinction, tripartite demonstrative system, fossilized infixes, polite register, metaphorical language

2023: Dictionary project began in earnest (i.e., first dictionary recordings)

Approaches (1): Corpus

Corpus development:

- Transcription of pre-existing sources
- Verification of pre-existing material with speakers
- Rest of corpus built from bottom-up

Transcriptions:

- Transcribed using ELAN
- **Tiers:** Speech, Discourse, Code-switching, Translation
- Converted into database tables (ELAN → Python → PostgreSQL)
- Each lexical item tagged with corresponding dictionary entry

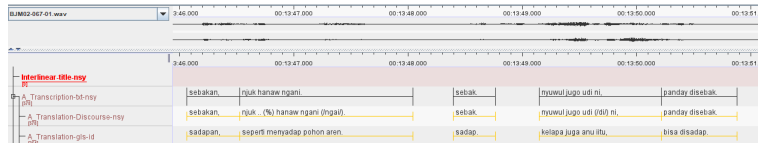


Figure 5: Nasal transcription file.

Approaches (2): Elicitation

Limited range of lexical items in conversation

Additional elicitation needed to surface
specialized vocabulary

Semantic domain elicitation:

- Small groups (3–4 people)
- Monolingual conversation on specific domain
- Guiding questions with open discussion
- Discussed until topic exhausted
- Approximately 1–5hrs recording time
- *Surprisingly* interactive and engaging

Supplemental elicitation:

- Aimed at targeted lexical items
- Definitions and range of meanings
- Collocations, set phrases, morphological paradigms, etc.



Figure 6: Dictionary discussion session.

Approaches (3): Workshops & focus groups

Community workshops:

- Way to keep community informed throughout
- Foster further engagement with eventual outcomes
- Identify ways dictionary can be improved for actual users
- **Topics:** Content, orthography, organization, design



Figure 7: Nasal dictionary workshop.

Focus groups:

- Avoid dynamics that might inhibit discussion at a workshop
- Highlight the finer details of the dictionary output
- Represents broader range of perspectives
- **Topics:** Usability, readability, personal suggestions



Figure 8: Nasal dictionary focus group.

Approaches (4): Miscellaneous

Being the *dictionary person*

Community events as learning opportunities

“my main “research tool”, if it can be called that, was paying attention” — Sutton (1978: xvii)



Figure 9: Photos used to explain Nasal words.

Challenges (1): Academic considerations

Time & money

- Transcription bottleneck
- Definition writing *very* time-consuming
- Funding for lexicographic work is scarce
 - Necessary both for documentation project and for academic progress



Figure 10: Sources of funding for documentary linguistic work.

Publish or perish:

- How and where to publish lexicographic work
- Producing iterative output
- Archiving and associated costs

Lexicographic training for documentary linguists

Challenges (2): Practical considerations

Documentary context:

- Fostering (and managing) community involvement
- Archiving, availability, and accessibility
- Data ownership and sovereignty
- Balancing educational and academic goals

Situational challenges:

- Social dynamics (age, gender, religion, etc.)
- Competing ideas of project direction
- Local government and administration
- Language purism and disagreements



Figure 11: Nasal meeting at local village head office.

Addressing challenges

Intentional community relationships and open communication

Computer assisted methods:

- Automated segmentation
- Automatic speech recognition (Whisper)
- Machine translation (Google-T5)
- Morphological parsing
- Recent developments in computer-assisted lexicography

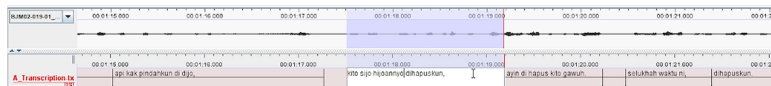


Figure 12: Process of correcting an ASR-generated transcription file.

Academia: Adjacent projects to result in research output driven by lexicographic data (→)

Additional research

Historical linguistics: Sound change

Many Sumatran languages exhibit high proportion of loan vocabulary

- E.g., ~70% in a sample from Gayo (Shorto 1976)

Identifying loan from native phonology can be *very* difficult

Larger body of lexical data allows more informed analysis

Rare **doublets** can reveal native phonological changes

Table 3: Some Nasal doublets exhibiting unique Nasal sound changes.

PMP	Nasal		
	Inherited	Loan	Source
*barəqan 'molar'	səmbəlahum 'molar'	bihum 'cheek'	< Lampungic <i>bihum</i>
*qapur 'lime'	hapul 'lime (betel)'	kapuχ 'lime (construction)'	< Malayic <i>kapur</i>
*wahir 'water'	wail 'water'	ajaχ 'water (in compounds)'	< Malayic <i>air</i>
*wari 'day'	wali 'day'	aχi 'day (in compounds)'	< Malayic (<i>h</i>)ari
*buka 'to.open'	buko 'to.open'	buka? 'to.break.a.fast'	< Malayic <i>buka(?)</i>

Historical linguistics: Language relationship

Proper identification of reflexes informs language subgrouping

Table 4: Identifying PMP reflexes in Nasal.

PMP	PSUM	Nasal		Malayic
		Native	Loan	
*j	*g	/g/~/h/~/k/	/d/	/d/
*pəjəs ‘painful’	*pəgəs	<i>pəgus</i>	—	<i>pədas</i>
*pajaj ‘rice’	*pagaj	<i>pəhaj</i>	<i>padi</i>	<i>padi</i>

*z	*d	/d/~/χ/	/d͡ʒ/	/d͡ʒ/
*pəzəm ‘close.the.eyes’	*pədəm	<i>pədum</i>	—	<i>pəjam</i>
*zauq ‘far’	*dauh	—	<i>d͡ʒauh</i>	<i>d͡ʒauh</i>

*q	*h	/h/	/h/~/Ø	/h/~/Ø
*quləj ‘worm’	*huləg	<i>huluk</i>	—	<i>ulat</i>
*qutək ‘brain’	*hutək	—	<i>uta?</i>	<i>otak</i>

*#w	*w	/w/	Ø	/h/~/Ø
*wahir ‘water’	*wair	<i>wail</i>	<i>ajaχ</i>	<i>air</i>
*wari ‘day’	*wari	<i>wali</i>	<i>aχi</i>	<i>hari</i>

Lexical retention:

- Nasal *χili* < PMP *hadiri (only *regular* Sumatran witnesses)
- Nasal *sihum* < PMP *sijəm (one of two witnesses on mainland Sumatra)
- Nasal *hasal* < PSUM *hasar (only witness outside of north Sumatra)
- Nasal *təluh* ‘village’ < PSUM *təruh ‘low’ (retention with semantic innovation)
 - Possible shared semantic innovation with Lampungic *tiuh* ‘village’?

Historical linguistics: Language contact

Intra-Austronesian contact:

- High proportion of Lampung vocabulary:
 - Nasal *piuh* ‘wring’ < Lampungic *piuh* (expected **pəluh* < PMP **pərəq*)
 - Nasal *biah* ‘taro’ < Lampungic *biah* (expected **bilah* < PMP **biraq*)
 - Nasal *bahun* ‘house’ < Lampungic *bahun* (innovation)
- High portion of Malayic vocabulary:
 - Nasal *hudʒan* ‘rain’ < Malayic *hudʒan* (expected **huʒan* < PMP **quzan*)
 - Nasal *dʒəxuʔ* ‘noose.trap’ < Malayic *dʒərat* (expected **χəluʔ* < PMP **zərək*)
 - Nasal *əmpədu* ‘bile’ < Malayic *əmpədu* (expected **həmpəgu* < PMP **qapəju*)
- No other apparent Austronesian contact

Extra-Austronesian contact:

- Unique vocabulary not in neighboring languages (direct rather than mediated contact?)
 - Nasal *spənuʔ* (< Dutch?) vs. Malayic *ʔʃəlana pəndʒan* ‘pants’ (< Hindi)
 - Nasal *ləmbudʒo* (< Sanskrit) vs. Malayic *səmaŋka* ‘watermelon’ (< native?)
- Differing phonological forms (possible direct rather than mediated contact?)
 - Nasal *dʒəntəxo* vs. Malayic *təntara* (< Sanskrit)
 - Nasal *χəbəʔo* vs. Malayic *rabu* (< Arabic)
- High proportion of nativized Arabic vocabulary

Historical linguistics: Fossilized affixes

Non-productive <əm> ‘AV’ infix:

Table 5: Nasal <əm> infix.

Root	<əm> ‘AV’	Gloss
<i>kail</i>	<i>k<əm>ail</i>	‘harvest’
<i>kusi</i>	<i>k<əm>usi</i>	‘clear.away’
<i>talāh</i>	<i>t<əm>alah</i>	‘chop’
<i>səlaw</i>	<i>s<əm>əlaw</i>	‘cry’
<i>suah</i>	<i>s<əm>uah</i>	‘burn’

tək- allomorph of *tə*- ‘NVOL’ prefix:

Table 6: Nasal *tə(k)*- prefix.

Root	<i>tə(k)</i> - ‘NVOL’	N- ‘AV’	Gloss
<i>həluŋ</i>	<i>tə-həluŋ</i>	<i>ŋə-həluŋ</i>	‘hear’
<i>ili?</i>	<i>tə-ili?</i>	<i>ŋ-ili?</i>	‘thresh.with.feet’
<i>ulaŋ</i>	<i>tə-ulaŋ</i>	<i>ŋ-ulaŋ</i>	‘return (something)’
<i>hədun</i>	<i>tə-hədun</i>	<i>m-ədun</i>	‘sit’
<i>hətu?</i>	<i>tək-ətu?</i>	<i>m-ətu?</i>	‘fart’
<i>iuḥ</i>	<i>tək-iuḥ</i>	<i>m-iuḥ</i>	‘urine’
<i>isiŋ</i>	<i>tək-isiŋ</i>	<i>m-isiŋ</i>	‘feces’
<i>utah</i>	<i>tək-utah</i>	<i>m-utah</i>	‘vomit’

Frequency: Genre patterns

Linked dictionary: Analysis of various frequency effects

Which genres most contribute to new entries?

Is there a point in which conversational recordings stop “out-performing” dictionary discussions?

Table 7: Lexemes in transcriptions of Nasal recordings by genre.

Genre	Total	Excluded	Unique		Hapax	
	Lexemes	Lexemes	Lexeme	Root	Lexeme	Root
Song	3,875	169	462	372	79	27
Map Task	19,342	1,264	783	565	197	69
Dictionary	58,146	2,220	3,004	1,845	1,661	746
Conversation	59,457	4,808	3,451	2,118	1,999	934
Overall	140,820	8,461	5,462	2,999	3,936	1,776

Frequency: Speaker patterns

Table 8: Nasal speaker representation in tagged data.

Speaker	Total	Excluded
Johan Safri	25,734	1,228
Wawan Sahrozi	16,959	658
Bahrul 'Ain	10,406	376
Susila Nurhayati (Elder)	7,644	209
Ahmad Bahuri	6,696	389
Nuzuar	6,574	171
Muliana	6,093	113
Johan Suardi	4,980	160
Iskandar Hamedi	4,853	181
Suherman	4,598	414
Amir Abdullah	4,404	497
Ahmad Baksir	3,925	116
Lili Arsil	3,638	373
Buyung Mahmud	3,250	418
Seni Husnila	3,100	537
Susila Nurhayati	2,840	379
Yogi Aprianda	2,575	96
Mefran Irwan	2,279	105
Lili Katam	2,221	167
Redo Susanto	1,880	82
...		

Demographic effects: Age,
gender, occupation, home
village, time outside village

How does lexeme use and
code-switching vary by
speaker?

Are certain populations
over-represented in the data?

Sociolinguistics: Lexical change

How is language shift reflected in the lexicon?

Can we track rates of lexical replacement?

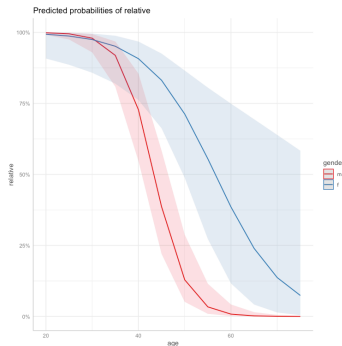


Figure 13: Loan (100%) vs. native (0%) relativizer use by age and gender.

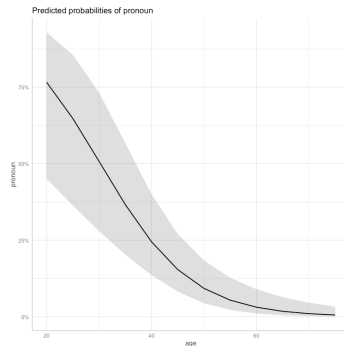


Figure 14: Loan (100%) vs. native (0%) third-person singular pronoun use by age.

Conclusion

Conclusion

Documentation of the lexicon is a *vital part* of documentary linguistic work

Dictionaries are often one of the *first resources* desired by language communities

Analysis of grammar is *dependent* on analysis of lexicon (and vice-versa)

Lexicographic work can be integrated with a variety of other research topics

Acknowledgments

Dr. Bradley McDonnell

Johan Safri & Wawan Sahrozi

Nasal speech community

Funding sources: NSF (BCS-1911641), DSNA, CAORC-AIFIS, Google Cloud



Google Cloud

References (1)

-  Altenberg, Bengt & Sylviane Granger. 1996. *Lexis in contrast: Corpus-based approaches*. Amsterdam: John Benjamins Publishing Company.
-  Anderbeck, Karl & Herdian Aprilani. 2013. *The improbable language: survey report on the Nasal language of Bengkulu, Sumatra*. Dallas: SIL International.
-  Austin, Peter K. 2010. Current issues in language documentation. *Language Documentation and Description* 7. 12–33.
-  Billings, Blaine & Bradley McDonnell. 2024. Sumatran. *Oceanic Linguistics* 63(1). 112–174.
-  Boerger, Brenda H. & Verna Stutzman. 2018. Single-event rapid word collection workshops: Efficient, effective, empowering. *Language Documentation & Conservation* 12. 147–193.
-  Chelliah, Shobhana L. & Willem J. de Reuse. 2011. *Handbook of descriptive linguistic fieldwork*. London: Springer.
-  Evans, Nicholas & Alan Dench. 2006. Introduction: Catching language. In Felix K. Ameka, Alan Dench & Nicholas Evans (eds.), *Catching language: The standing challenge of grammar writing* (Trends in Linguistics: Studies and Monographs 167), 1–40. Berlin: Mouton de Gruyter.
-  Faaß, Gertrud. 2018. Lexicography and corpus linguistics. In Pedro A. Fuertes-Olivera (ed.), *The Routledge handbook of lexicography*, 123–137. London: Routledge.

References (2)



Grimm, Nadine. 2022. Documentary approaches to lexicography. In Moses Effiong Ekpenyong & Imelda Icheji Udoh (eds.), *Current issues in descriptive linguistics and digital humanities: A festschrift in honor of professor Eno-Abasi Essien Urua*, 551–568. Singapore: Springer.



Grinevald, Colette. 2001. Encounters at the brink: Linguistic fieldwork among speakers of endangered languages. In Osamu Sakiyama (ed.), *Lectures on endangered languages: 2*, 285–314. Osaka: ELPR Publication Series 002.



Helfrich, O. L. 1891. Proeve van een Lampongsch-Hollandsche woordenlijst, bepaaldelijk voor het dialect van Kroë. *Verhandelingen van het Bataviaasch Genootschap van Kusten en Wetenschappen* 45.



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



Hüllen, Werner. 1999. *English dictionaries 800–1700: The topical tradition*. Oxford: Clarendon Press.




Kasim, Yuslina et al. 1987. *Pemetaan bahasa daerah di Sumatra Barat dan Bengkulu*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.




Kilgariff, Adam. 2007. Word sense. In Eneko Agirre & Philip Edmonds (eds.), *Word sense disambiguation: Algorithms and applications* (Text, Speech and Language Technology vol. 33), 29–46. London: Springer.


References (3)

 Krishnamurthy, Ramesh. 2008. Corpus-driven lexicography. *International Journal of Lexicography* 21(3). 231–242.

 Lachler, Jordan & Elizabeth Pankratz. 2017. Moving towards value-added digital repatriation in lexicography for indigenous languages in Canada. In Nicholas Ostler, Vera Ferreira & Chris Moseley (eds.), *Proceedings of the 21st FEL Conference*, 107–114. Hungerford: Foundation for Endangered Languages.

 McDonnell, Bradley. 2017. *Documentation of Nasal: an overlooked Malayo-Polynesian isolate of southwest Sumatra*. Endangered Languages Archive.







 McDonnell, Bradley. 2019. *Documentation of the multilingual linguistic practices of the nasal speech community*. Collection BJM02 at catalog.paradisec.org.au [Open Access].

 Moe, Ronald. 2003. Compiling dictionaries using semantic domains. *Lexikos* 13. 215–223.

 Paterson III, Hugh. 2015. *Lexical dataset archiving: An assessment of practice*. Poster presented at the *4th International Conference on Language Documentation and Conservation*, Honolulu, HI, 26 February – 1 March.

 Rhodes, Richard A. & Lyle Campbell. 2018. The goals of language documentation. In Kenneth L. Rehg & Lyle Campbell (eds.), *The Oxford handbook of endangered languages*, 107–122. Oxford: Oxford University Press.

References (4)

-  Roos, K. H. F. 1890. Aanteekeningen over de afdeeling Kaoer in de residentie Benkoelen. *Tijdschrift voor het binnenlandsch bestuur* 4(1–6). 14–33, 92–112, 145–160.
-  Rundell, Michael. 2012. ‘It works in practice but will it work in theory?’ The uneasy relationship between lexicography and matters theoretical. In Ruth Vatvedt Fjeld & Julie Matilde Torjusen (eds.), *Proceedings of the 15th EURALEX International Congress*, 47–92. Oslo: Department of Linguistics & Scandinavian Studies, University of Oslo.
-  Shorto, H. L. 1976. Gayo consonant correspondences. In Nguyễn Đăng Liem (ed.), *Southeast Asian linguistics studies* (Pacific Linguistics C-94), 87–109. Canberra: The Australian National Univeristy.
-  Stokhof, W. A. L. & Alma E. Almanar. 1987. *Holle lists: Vocabularies in languages of Indonesia, vol. 10/2, Southern Sumatra*. (Pacific Linguistics Series D 75). Canberra: The Australian National University.
-  Sutton, Peter. 1978. *Wik: Aboriginal society, territory and language at Cape Keerweer, Cape York Peninsula, Australia*. Brisbane. (Doctoral dissertation).
-  Woodbury, Anthony C. 2003. Defining documentary linguistics. *Language Documentation and Description* 1. 35–51.

Thank you — Tekhimo kasih

Questions?