Language Resources and Evaluation
https://doi.org/10.1007/s10579-022-09628-x

PROJECT NOTES

CINWA (database of terminology for cultivated plants in indigenous languages of northwestern South America): introducing a resource for research in ethnobiology, anthropology, historical linguistics, and interdisciplinary research on the neolithic transition in South America

Matthias Urban[1] · **Evelyn Michelle Aguilar Panchi**[1] · **Saetbyul Lee**[1] · **Evgenia Brodetsky**[2]

Accepted: 18 November 2022
© The Author(s) 2022

Abstract

This article introduces CINWA, a freely accessible online database of terminology for cultivated plants in indigenous languages of South America based on FAIR principles for scientific data management and stewardship. In the pre-release version we present here, CINWA assembles more than 2700 terms from more than 60 indigenous languages of northwestern South America, and coverage will be continuously expanded. CINWA is primarily designed for use in historical linguistics to explore patterns of lexical borrowing that might be used as a proxy for tracing the pathways by which knowledge of individual cultivated plants and the associated know-how spread from speech community to speech community in pre-Columbian South America. In spite of intensifying research, this is still unclear for most cultivars as the locales of initial cultivation are heterogeneous and spatially diffuse. However, possible uses of the CINWA database are manifold and go beyond this research question. The database can be used as a resource for ethnobiological and comparative anthropological research on South American communities, South American agricultural ecosystems and practices, and for studies in lexical borrowing, language contact, and historical linguistics broadly.

Keywords Lexical database · Cultivated plants · Agriculture · Ethnobiology · Linguistics · Neolithic revolution

✉ Matthias Urban
matthias.urban@uni-tuebingen.de

[1] Center for Advanced Studies 'Words, Bones, Genes, Tools', University of Tübingen, Tübingen, Germany

[2] Department of Linguistics, The Hebrew University of Jerusalem, Jerusalem, Israel

Published online: 13 December 2022

1 Introduction

This article introduces CINWA, a freely accessible online database of terminology for cultivated plants in the languages of South America. Its focus lies on the northern and western parts of the continent. Here, in the early Holocene, people in various locales began experimenting with the tending of plants and the selection of phenotypes that had properties that were beneficial to them (Pearsall, 2015; Piperno, 2011). From these early onsets, South America would become one of the sites of the world that would witness the development of full-blown agriculture and a myriad of agricultural systems that range from large-scale agriculture and sophisticated technologies in the Andes and lowland areas such as the Llanos de Moxos to no less refined agroforestry systems in lowland Amazonia. These have shaped, and continue to shape, the forest ecologies of the lowlands in significant ways (Balée, 2013; Denevan, 2001).

Along with the development of such farming and agroforestry regimes, knowledge systems associated with agricultural techniques, but especially also individual plants suitable for cultivation in different environments, have developed. These knowledge systems are or were enshrined in the many indigenous languages of the South American continent. It originally boasted, according to some estimates, as many as 1000 individual languages and more than 100 independent linguistic lineages (Campbell, 1997; Kaufman, 1990). Many of these languages, and with them the knowledge regarding the physical environment of its speakers, have become dormant already (Harrison, 2007), but a multitude of local languages is still spoken, some severely endangered (Crevels, 2012; Urban, 2021).

The state of documentation of the indigenous languages of South America varies in quality and quantity. For a significant number of languages, including some that were abandoned as a means of daily communication early on, there are wordlists or dictionaries that document the vocabulary and, thereby, often also significant parts of the knowledge of their speakers on the world surrounding them and their daily activities—including, for food producers, the cultivation of plants. Present-day indigenous people of South America know of almost 200 plants that are cultivated or semi-cultivated in different parts of the continent (Denevan, 2001). While some of these are widespread across the entire continent, others are specific to particular climatic and ecological conditions that are needed for the plant to thrive. Some are tied intimately to specific cultural practices, such as the use of annatto, made from the seeds of *Bixa orellana*, which are valued for their bright red color as a dye and raw material for body paint in many cultures of lowland South America.

Documentation of regional indigenous terminology of cultivated plants, whether rudimentary or advanced, is often not readily accessible to interested parties. The terminological distinctions indigenous languages make and the associated knowledge of the speakers are enshrined in sources that are typically far-flung and mostly not available in digital format. In the case of many sources, printed copies are available in only few libraries around the world. So far this makes it next to impossible to obtain a quick and reliable overview of

terminology associated with cultivated plants in the indigenous languages of native South America. In particular, the development of comparative perspectives encounters a hindrance in the fragmented way that relevant information is distributed over different, often hard to get or "grey" literature. In addition, most of the sources present the information utilizing the national languages of present-day South American nation states, Spanish and Portuguese, as metalanguages, and not all interested parties may have a sufficient command of these.

One of the general purposes of CINWA is to facilitate such comparative perspectives for a variety of research purposes. We provide a more detailed yet non-exhaustive list of possible applications of CINWA in Section 7, after having contextualized the database with the CLLD framework which it utilizes and comparable comparative linguistic sources (Sect. 2), providing an overview of CINWA's coverage in terms of languages (Sect. 3), cultivars (Sect. 4), its data collection and representation principles (Sect. 5), and the functionalities of the online interface of the database (Sect. 6).

2 The CLLD framework and comparable work

CINWA utilizes the CLLD (Cross-Linguistic Linked Data) framework (Forkel et al., 2018, 2019), which has established itself over the last years as the quasi-standard for online publication of linguistic data resources. CLLD relies on Linked Data principles for the processing and storage of linguistic data, and has to this end developed CLDF (Cross-Linguistic Data Format), which specifies standards for preparing linguistic data in interrelated .tsv files that can be stored and accessed in the long term.

Developers also furnish a web framework to build and display CLLD data from CLDF-formatted files in an aesthetically uniform way across different CLLD resources, thus ensuring a coherent and intuitive user experience.

CINWA is, to the best of our knowledge, the first online database on names for cultivated plants specifically for any part of the world. However, comparable work in a broader sense does exist. On the one hand, there are databases that reunite lexical and structural information on American languages, also in a CLLD framework: CSD (Rankin et al., 2016) deals with the comparative lexicon of the Siouan languages of North America, while TuLeD (Ferraz Gerardi et al., 2021) provides lexical data for Tupian, one of the largest language families of lowland South America. Both CSD and TuLeD seek to cover information on the vocabulary of the languages of the two families generally, and both also include a few plant terms.[1] Of more general scope is the venerable Intercontinental Dictionary Series (Key & Comrie, 2021), which has been begun in the 1980s and is now also available in CLLD format. The Intercontinental Dictionary Series provides vocabularies of around 1500 words for a wide range of languages of the world, but, reflecting the research interests of the

[1] Ferraz Gerardi et al. (2021, p. 1006) note that "although the current amount of the diverse fauna and flora items in TuLeD is modest when compared to the overall number of concepts, the collection of relevant terms is ongoing and given high priority for the official release."

founding editor Mary Ritchie Key, one particular focus is on South American languages. The local flora and fauna is not well represented in the standardized datasets, however. On the other hand, with Tsammalex (Naumann et al., 2020), there is a database with an outlook and goals that are similar to those of CINWA. Tsammalex is a multilingual lexical database on plants and animals with a highly developed structure that incorporates information on biological classification of the flora and fauna the vocabulary, the world's terrestrial ecoregions it occurs in, and, for some entries, images of the designated species. However, in terms of languages, reflecting the principal interests of its creators, Tsammalex at present focuses on African languages, mostly sub-saharan.

3 Coverage: languages

In terms of coverage, CINWA, at its current state of development that is reflected in the preliminary release version 0.9 that we describe here, focusses on the northern and western quadrant of South America, covering indigenous languages spoken from Colombia in the north to Bolivia in the south, and in all of the three major geographical regions nowadays recognized: coastal lowlands, Andean highlands, and the lowlands of Colombia, Ecuador, Peru and Bolivia to the east of the Andean cordilleras. CINWA pays special attention on this part of the continent at this stage because it harbors the greatest number of sites where, according to current knowledge, the initial cultivation of a significant number of South American cultivars took place (Piperno, 2011). For those researchers interested in the historical linguistics of plant domestication, this is therefore the most relevant region to look into. In further updates to the database, we plan to expand coverage to also include data from languages beyond.

Like other parts of the continent, northwestern South America was and still is eminently multilingual. The preliminary release version of CINWA which we present here has data for 62 languages which belong to as many as 32 distinct language families (including isolates). These languages, together with their genealogical classification, are listed in Appendix A; their locations are shown in Fig. 1, which is a screenshot from CINWA's online interface, with color-symbol type combinations used for different language families. For the genealogical classification, we rely on Glottolog 4.4. (Hammarström et al., 2021). Itself a CLLD resource, Glottolog is the most comprehensive and reliable source of information about the different languages, dialects, and language families of the world, and provides a conservative classification that reflects currently accepted groupings among the experts on the relevant languages.

4 Coverage: cultivars

We have relied on Denevan (2001, pp. 307–320, Appendix 1A) for a near-exhaustive list of cultivated plants of South America. The table in this appendix, which provides a common name, other names, a botanical classification, and comments on location

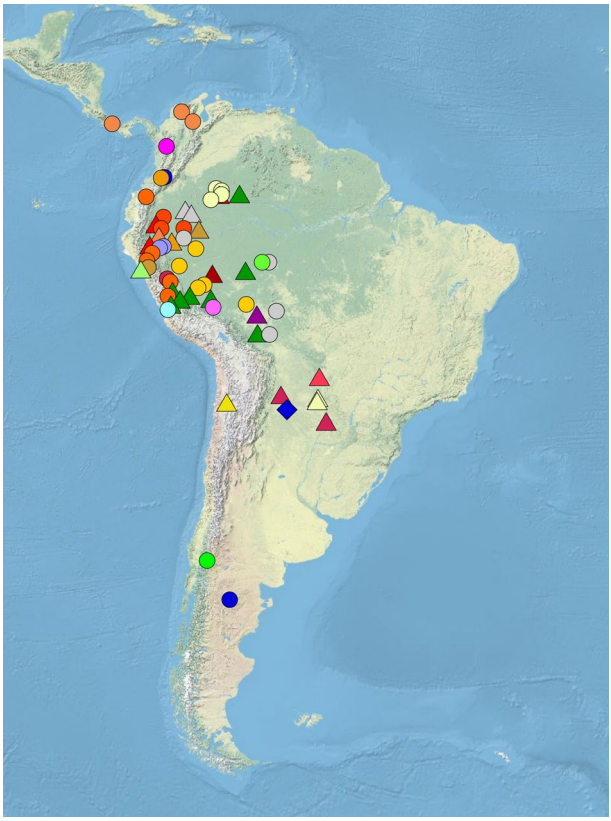

Fig. 1 Map of languages represented in CINWA 0.9

and use of individual cultivars, forms the basis of our work with primary sources for CINWA as described in Sect. 5. Denevan's table is itself based on prior sources, and the author informs that his compilation includes "Domesticates, semi-domesticates, and other selected planted and protected (cultivated) plants, native to the Americas, and prehistoric or probably prehistoric in South America" and that species were included in the table if "[l]isted by at least two sources as domesticated, semi-domesticated, or cultivated" (Denevan, 2001, p. 320, table note a). While surely an independent perusal of botanical and ethnobotanical literature would have resulted in a different, perhaps expanded list, this would have been a research undertaking of such an extent that would have exceeded available resources by far, and would also have engendered its own perils given that we are a team of linguists rather than (ethno)botanists. Our impression was that Denevan's compilation comes as close to definite as any with scope over South America as a whole, so that we found it preferable to adopt his compilation—highlighted as particularly useful in reviews (Newson, 2002; Stadel, 2005)—for our purposes. As adapted for perusal in CINWA, the list features 189 cultivated plants, along with additional information on each one,

including English names, local vernacular names, binomial nomenclature, author citation, plant family, the region in which the plant grows, and its principal use or uses. Pertinent data are in Appendix B; we emphasize again that this list is not our work, but relies heavily on Denevan (2001).

5 Data collection and representation principles

As primary data sources, we draw on published dictionaries, wordlists, and sometimes, where available to us, also dedicated literature that deals with the ethnobotanical knowledge and practices of particular South American speech communities. We used our own expertise in South American languages to identify possible sources, but we additionally also relied on Glottolog's extensive bibliography on the world's languages. Generally, we preferred modern dictionaries that have been compiled by lexicographers with training in linguistics and that utilize phonemic transcription principles. However, such high-quality sources are not available for all languages within the scope of CINWA as delimited in Sect. 3. Given that language loss is a phenomenon that has been going on since the European invasion and that is even accelerating (Crevels, 2012), for some languages it has not been possible to compile such sources. For a subset of languages that are no longer spoken, shorter and less complete documentary material may nevertheless be available, even though such sources are often problematic for representing the sound structure of the described language poorly. Such sources also carry an increased risk of presenting inaccurate or vague information on the meaning of the recollected words. Cultivars may therefore not always be correctly identified. These problematic aspects notwithstanding, we have chosen to incorporate data from such sources since, in spite of their shortcomings, they are the best documentary material available for many a language that may harbor crucial evidence for historical processes related to plant cultivation in South America. Nevertheless, we urge CINWA users to treat such data critically and to be cognizant of the nature of the material and its particular characteristics that reflect the time of its creation.

We have not made it a policy to rely on only one source for data per language. For some languages, more than one source was consulted to increase coverage.

We have extracted the terminology of cultivated plants that the sources document manually by going through all relevant entries. In bilingual dictionaries, i.e. those that feature a section in which an English, Spanish, or Portuguese lexical item is translated to an indigenous language, we have used that section to identify the relevant entries in the other complementary section that translates lemmata in the indigenous language to the European-based vernacular. As a matter of principle, we have always included forms of items and all additional information (part of speech, gloss, etc.) from that section. The reason is that in our experience this section often provides either identical or more complete information vis-à-vis the section that translates from a national language of European origin or English to the indigenous language, which for some of the sources we used is no more than a mere finder list. Where dictionaries only feature one section that is lemmatized according to native language forms, we have had to look through the entire dictionary in search for cultivar terminology.

A significant issue is that the name of a particular cultivar in the local variety of Spanish, and to a minor extent also Portuguese, often differs from region to region. Sometimes this is a matter of distinctions related to the boundaries of modern national states. But the problem trickles down further to the regional level, and sometimes cross-cuts national boundaries. Such terminological variation in the former colonial European languages is a significant obstacle to the precise identification of dictionary or wordlist entries. We have chosen to depart from the names mentioned in the column "Other Names" in Denevan's (2001) appendix on which we also based our list of cultivated plants in Appendix B. To be maximally inclusive and to reduce chances of overlooking data, we have made it our policy to always look up all the local names mentioned in all dictionaries even in cases where we had some prior knowledge on what the name in the local variety of Spanish of the plant might be in each individual case.

One important principle that CINWA follows, not least in response to the above issue, is rich annotation: for each entry, CINWA provides the original entry found in the utilized primary source, including a part of speech code and the gloss, where available, relevant annotations on the characteristics of the plant, cultural associations that prevail in the respective speech communities, and other information. These entries have only been abbreviated if they were excessively long and contained information that was not judged directly relevant to the name of the cultivar and its use in the speech community. Experience shows that such information, while unsystematic across sources, can contain crucial information in different research contexts; therefore, to fulfil its aim to facilitate comparative perspectives on the terminology for cultivated plants for a variety of research purposes, we considered it a high priority to (partially) include the original dictionary gloss.

All lexical items are represented in CINWA in the orthography of the respective sources. We provide an additional phonological representation where the source itself does so, but otherwise do not attempt to infer standardized phonological representations for our datasets—while this would be unproblematic for some sources, for others it would be an enterprise fraught with difficulties that would require a significant amount of philological preparation and even then be associated with considerable uncertainty. We also adopt the citation forms as given in the sources rather than trying to reduce these to stems or roots, but we have not taken over alternate forms that sources provide (for instance, different genders or noun-classifier combinations).

The sources that we utilized to create CINWA rely on a multitude of different orthographies for representing the sound structure of the indigenous languages. Early sources tend to simply utilize the Latin alphabet in an uncritical and problematic way that potentially obscures or distorts the phonology of the described language. We nevertheless include such sources, as per the general principles above, in situations where we have no more adequate modern sources that uses an orthography based on phonemic principles. Such more recent sources typically use an alphabet that is adapted to the phonology of the described languages, and may contain special characters that are not part of the standard Latin alphabet. Often, diacritics are used; in other cases, special phonetic characters from the International Phonetic Alphabet are used alongside standard characters; combinations are also frequent. One problem that arises is that in Unicode, the same character can be created in multiple

ways. For instance, a <o> with a dot below and a circumflex diacritic, <ộ>, can be created in at least three ways: on the one hand, the character is actually available precomposed by Unicode (U+1ED9). But there are alternatives: the same result could be achieved by using the precomposed o with dot below <ọ> (U+1ECD) and adding an acute as a combining character (U+1ECD+0301), by using the precomposed o with acute <ó> (U+00F3) and adding a dot below as a combining character (U+00F3+0323), or by building it up on the basis of <o> without diacritics (U+006F) and adding both acute and dot below through combining characters (U+006F+0301+0323). Depending on the font used, the complex characters thus created may be displayed in slightly different ways, but, more importantly, the characters thus created might behave differently in searches in data processing software, so that a search for either of the characters thus created might fail to find those created by a different strategy. We therefore have prepared a list of characters with diacritics that are used in CINWA and the way they were generated. Our workflow involved a step in which coders, every time they encountered a special character in a source they were about to begin to evaluate for CINWA, checked whether this particular special character was already used in one of the datasets that already had been extracted. If so, they proceeded to create the special character in the same manner. If the particular character in question had not been used before in the data for any other language, they added it to the list along with its name in Unicode and the way they created it so that it can be created in the same way for future datasets. While we placed most emphasis on the uniformity in the way each single character was created, we have usually made it a habit to work with precomposed characters that are made available by Unicode as much as possible (approximating Unicode NFC, see Moran & Cysouw, 2018, p. 28), and to employ combining characters only where this is not the case.

6 Features and functionalities of the CINWA online interface

As is true of CLLD databases generally, CINWA allows its users to access and visualize the data in different ways that cater to the anticipated heterogeneous needs of its future users. We describe the main manners of accessing the database here. From the start page, which provides a general short description of the database, some basic statistics and a citation recommendation, the user is able to navigate to three main sections: one that represents the data per language, one that represents the data per plant, and bibliographic data.

The "Languages" view allows the user to access the list of languages covered in CINWA together with associated metadata as also listed in Appendix A. Each language name leads to a full list of entries available for that language, so that researchers interested in the nomenclature of a certain language have the possibility to access all names for cultivated plants that are represented for that language. The resulting list –exemplified in the screenshot in Fig. 2 for Jebero, a language of the Peruvian Amazon, with data from Valenzuela Bismarck (2012)—features, as described in Sect. 5, an (abbreviated) representation of the original gloss as found in the sources we consulted as well as a link to the bibliographic data of that source.

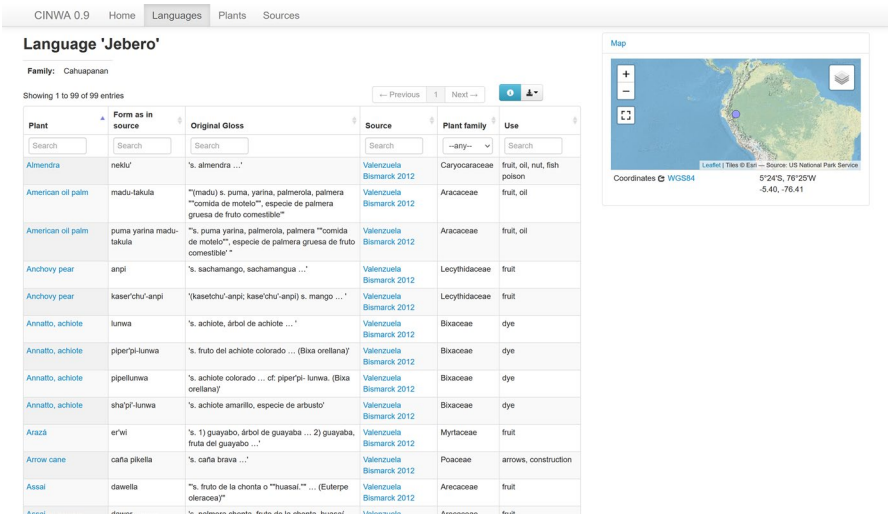

Fig. 2 CINWA 0.9 data for the Jebero language of the Upper Amazon (extracted from the primary source Valenzuela Bismarck, 2012)

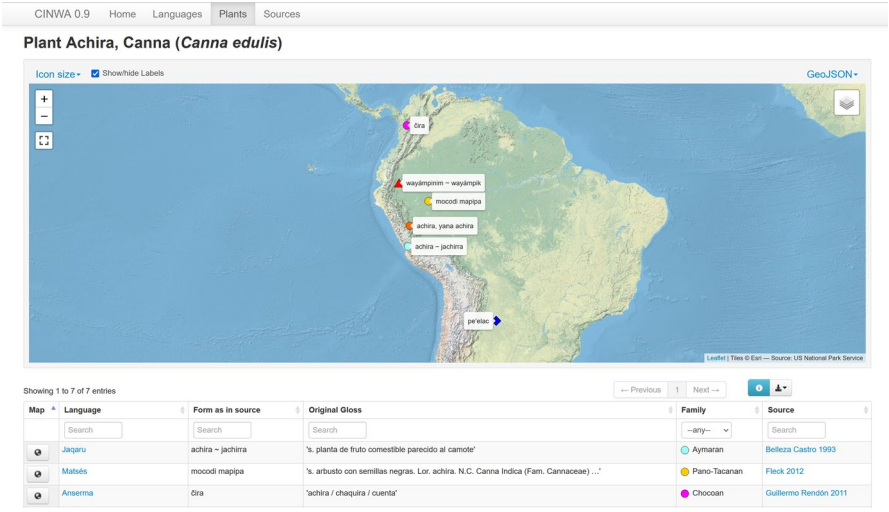

Fig. 3 CINWA 0.9 data for the arrowroot (*Canna edulis*)

In turn, this view also has the English names of available plants linked to a view in which users can conveniently access all names for that particular plant across languages covered in CINWA. This is exemplified in the screenshot in Fig. 3 for the arrowroot (*Canna edulis*).

Together with the visual representation of the distribution of the terms and the languages in which they are found on the associated map, this facilitates initial comparison for similarities, which is one of the prime intended uses of the database.

7 Possible applications

CINWA's possible uses are manifold. Most generally, it is a resource for scholars interested in ethnobotanics and its relation to linguistics and culture (Maffi, 2001). The database may serve as an initial source of information for researchers interested in the cultivated plants and associated names of a particular indigenous community or a particular part of South America. However, given its inherently comparative nature, we expect CINWA to be of greatest use for researchers interest in comparative perspectives on cultivar terminology.

These perspectives might be different ones: since names are the "foundation of linguistic ethnobiology" (Hunn & Brown, 2011, p. 319) ethnobiologists might be interested in the relationships between communities and the cultivars they tend as reflected in the lexical distinctions they enshrine. We imagine that our principle of rich annotation will be particularly useful here, as this will allow researchers to assess the polysemy structure associated with individual cultivars in individual languages that may reveal cultural associations, but also reflect on the diachronic development of the system (Brown, 1986).

There are also a multitude of possible application for more properly historical linguistic questions. One such topic are so-called marking reversals (Witkowski & Brown, 1983), a process of linguistic acculturation in which a native word for a known referent is recruited to designate a newly introduced referent –typically, flora and fauna– together with a modifier. As this new referent grows in importance to such an extent that the original referent ends up with a modifier, the original term comes to refer to the item of acculturation exclusively.

Such questions often relate intimately to South American prehistory. In fact, answering them might make significant contributions to elucidate the vectors of spread of cultivars that can usefully complement the archaeological record that is still sparse in many regions, not least because of rapid decomposition of organic material in hot humid environments. One pertains to the spread of agriculture and the cultivation of particular plants in pre-Columbian South America, and this is the main goal which we had in mind when creating CINWA. For instance, it is known that the coastal valleys of present northern-Peru harbor some of the earliest evidence for cultivars whose initial focus of cultivation is thought to lie to the east of the Andes (Dillehay et al., 2007), and the Huancabamba depression, a sector of the Andes characterized by reduced altitude and width that makes east–west transversal movement relatively easy, could have served as a convenient corridor for the introduction of cultivars from the east (Kaulicke, 2020). Data such as that found in CINWA, in fact, furnishes evidence that this is the case: in the upper Amazon region of northern Peru, indigenous languages know the peanut butter fruit tree (*Bunchosia armeniaca*) by names like *oshon* in Chayahuita (Hart, 1988), and this term has been borrowed into the lowland Quechua variety of San Martín as *ushun* (Park et al., 1976). In this case, we can be very confident about borrowing and its directionality not just because of the general formal similarity, but also because of the fact that the form was adapted to the threevocalic vowel system in Quechua by replacement of mid vowels (provided that Chayahuita was the direct source of borrowing), and especially because cognates of the word are absent in any other Quechua varieties

(and indeed, the tree would not grow in the regions where these are spoken). But there is more: in the extinct Mochica language of coastal Peru on the other side of the Andes, Brüning (2004) documented the form <ošórre> in work with the last speakers of the language. Consistent with a more remote borrowing history, this form shows greater distance from the present-day Upper Amazonian forms (though the final <e> is probably Spanish-induced, reducing the difference basically to the quality of the sonorant). In this sense, CINWA data can also contribute to resolving questions of the origin and spread of individual cultivars that have already amply been debated, but not answered conclusively yet.

Some related questions have occupied researchers from different disciplines already for a long time, and CINWA is eminently suited to contribute to resolving these. One of these is whether the introduction of the banana to South America pre- or postdates European contact (Nordenskiöld, 1922; Jeffreys, 1963; Langdon, 1993; Balée, 2013, pp. 94–95), and linguistic arguments have played a role in that debate. Another open question that should be of renewed interest since Ioannidis et al.'s (2020) finding of a genomic signature of Native American-Polynesian contact events is the origin and geographical extension of use of South American terms for the sweet potato in Polynesia. Polynesian designations are so strikingly similar in some South American languages, prominently of the Quechuan family, that they have been understood as a signature of precisely such contact events (Adelaar, 1998). One issue is that the distribution of the relevant term is hitherto not known to have extended to those regions of South America in coastal Colombia that are most likely implicated in the contact events. Thanks to its scope, CINWA will allow a more rigorous assessment than possible before.

8 Anticipated future development

Here we present a preliminary yet fully functional pre-release version (0.9) of CINWA. We plan to expand CINWA in the following phase of development in two ways: on the one hand, while we already have achieved a fairly dense coverage that allows to look into research questions on the basis of a principled empirical dataset, we aim to extend coverage further with the ultimate long-term goal of full complete coverage of indigenous languages of northwestern South America. On the other hand, we plan to implement a system of coding for morphological structure of the items as certain types of morphological structure, in particular what in English would be endocentric compounds, are often of importance in ethnobiological classification (in ethnobiological nomenclature, so-called "composite labels", "binomial terms" or "secondary lexemes," Hunn & Brown, 2011).

Finally, we plan to implement a system of assessing similarities of forms between families that is similar to cognate coding in databases that cover lexical variation within language families (e.g. Dellert et al., 2020; Ferraz Gerardi et al., 2021). In our case, the goal would, however, rather be to establish a rating system of formal similarities between terms that might hint at prehistoric borrowing of designations across languages and language families which occurs sometimes repeatedly and then gives

rise to so-called "Wanderwörter"—widespread formally similar terms that reflect different patterns of repeated borrowing (Haynie et al., 2014). With this aim in mind, we plan to develop a rating scheme that would take into account the known phonological histories of the involved language families and principles of loanword adaptation (e.g. Uffmann, 2015) to derive hypotheses as to borrowing pathways and networks.

9 Conclusions

In this article, we describe and officially release CINWA 0.9, a new database and resource which aims to cover the terminology for cultivated plants in South American indigenous languages, focusing on the Northwest of the continent, and have described the rationale, layout, and structure of CINWA and described possible applications. We have explained the CLLD framework on which the database draws and discussed comparable work (Sect. 2), and discussed CINWA's coverage in terms of languages (Sect. 3) and cultivated plants (Sect. 4). The preliminary release version 0.9 provides around 2700 such terms, covering 62 languages from 32 families for a subset of around 189 cultivated plants known to the indigenous societies of South America. We have described the principles of data collection and curation, emphasizing our rich annotation approach (Sect. 5). We have described the functionalities of the CINWA online interface (Sect. 6) and the possible applications it aims to facilitate (Sect. 7). In the coming years, we aim to expand CINWA further in terms of data coverage and functionality (Sect. 8).

Appendix A: languages in CINWA 0.9 with associated metadata

Name	Classification	Latitude	Longitude	Glottocode	ISO 639–3 code
Achuar-Shiwiar	Chicham	−2.83	−77.26	achu1248	acu
Aguaruna	Chicham	−5.3	−77.92	agua1253	agr
Amahuaca	Pano-Tacanan	−10.2	−72.32	amah1246	amc
Amarakaeri	Harakmbut	−12.49	−70.55	amar1274	amr
Andoa	Zaparoan	−3.3	−76.61	ando1255	anb
Angaité	Lengua-Mascoy	−22.85	−58.45	anga1316	aqt
Anserma	Chocoan	6.25	−76.05	anse1238	ans
Apurinã	Arawakan	−8.22	−66.77	apur1254	apu
Barasana-Eduria	Tucanoan	0.02	−70.81	bara1380	bsn
Barí	Chibchan	9.08	−72.94	bari1297	mot
Cajamarca Quechua	Quechuan	−7.09	−78.32	caja1238	qvc
Candoshi-Shapra	Candoshi-Shapra	−4.24	−76.93	cand1248	cbu
Cayubaba	Cayubaba	−13.25	−65.48	cayu1262	cyb
Chachapoyas Quechua	Quechuan	−6.27	−77.66	chac1250	quk
Chácobo	Pano-Tacanan	−12.13	−66.72	chac1251	cao
Chimila	Chibchan	10.19	−74.26	chim1309	cbg
Cholón	Hibito-Cholon	−9.23	−75.96	chol1284	cht

Name	Classification	Latitude	Longitude	Glottocode	ISO 639–3 code
Cocama-Cocamilla	Tupian	−4.5	−74	coca1259	cod
Cubeo	Tucanoan	1.32	−70.19	cube1242	cub
Culina	Arawan	−8.55	−70.66	culi1244	cul
Culli	Culli	−7.88	−78.12	cull1235	NA
Curripaco	Arawakan	0.7	−67.5	curr1243	kpc
Huallaga Huánuco Quechua	Quechuan	−9.57	−75.59	hual1241	qub
Hup	Naduhup	0.59	−69.84	hupd1244	jup
Imbabura Highland Quichua	Quechuan	0.32	−78.37	imba1240	qvi
Iquito	Zaparoan	−3.31	−73.97	iqui1243	iqu
Isconahua	Pano-Tacanan	−7.75	−74.52	isco1239	isc
Iyo'wujwa Chorote	Matacoan	−22.21	−62.65	iyow1239	crq
Jaqaru	Aymaran	−12.74	−75.81	jaqa1244	jqr
Jebero	Cahuapanan	−5.4	−76.41	jebe1250	jeb
Júma	Tupian	−7.28	−64.1	juma1249	jua
Katawixi	Katukinan	−7.25	−64.83	kata1270	xat
Kotiria	Tucanoan	1.08	−69.57	guan1269	gvc
Kunza	Kunza	−23	−69	kunz1244	kuz
Maca	Matacoan	−25.01	−57.37	maca1260	mca
Machiguenga	Arawakan	−12.13	−75.5	mach1267	mcb
Mapudungun	Araucanian	−38.74	−71.28	mapu1245	arn
Mashco Piro	Arawakan	−11.47	−70.8	mash1270	cuj
Matsés	Pano-Tacanan	−5.74	−72.63	mats1244	mcf
Minica Huitoto	Huitotoan	−1.55	−73.11	mini1256	hto
Mochica	Mochica	−8.02	−79	moch1259	omc
Murui Huitoto	Huitotoan	−1.14	−73.83	muru1274	huu
Ngäbere	Chibchan	8.83	−82.39	ngab1239	gym
Nomatsiguenga	Arawakan	−11.53	−74.44	noma1263	not
North Junín Quechua	Quechuan	−11.21	−75.81	nort2980	qvn
Omurano	Omurano	−4.82	−75.4	omur1241	omu
Páez	Paezan	2.62	−76.31	paez1247	pbb
Payagua	Payagua	−20.21	−58.16	paya1236	NA
Puelche	Chonan	−42.22	−68.67	puel1244	pue
Sanapaná	Lengua-Mascoy	−22.49	−58.38	sana1298	spn
Shawi	Cahuapanan	−5.64	−76.83	chay1248	cbt
Sirionó	Tupian	−15.53	−63.97	siri1273	srq
Toba	Guaicuruan	−23.77	−61.96	toba1269	tob
Totoro	Barbacoan	2.55	−76.64	toto1306	ttk
Trinitario	Arawakan	−15.33	−65.41	trin1274	trn
Tucano	Tucanoan	0.58	−69.58	tuca1252	tuo
Warázu	Tupian	−12.86	−63.19	paus1244	psm
Yagua	Peba-Yagua	−3.43	−72.22	yagu1244	yad
Yaminahua	Pano-Tacanan	−9.88	−71.6	yami1256	yaa
Yanesha	Arawakan	−10.55	−75.31	yane1238	ame

Name	Classification	Latitude	Longitude	Glottocode	ISO 639–3 code
Yine	Arawakan	− 11.11	− 73.31	yine1238	pib
Záparo	Zaparoan	− 2	− 76.36	zapa1253	zro

Appendix B: cultivated plants of South America. Source: based on Denevan (2001, pp. 307–320, Appendix 1A)

English name	Local names	Linnean name	Author citation	Alternative Linnean name	Alternative author citation	Family	Region	Use
Achira; Canna	Capacho	Canna edulis	Ker-Gawl.			Cannaceae	Lower Andes	Root
Achocha	Cáygua; achokcha	Cylanthera pedata	(L.) Schrad			Cucurbitaceae	Andes	Vegetable
Agave; maguey	Pakpa; cabuya	Agave americana	L.			Amaryllidaceae	Andes	Fiber
Ahipa	Ajipa; huitoto	Pachyrhizus ahipa	(Wedd.) Parodi			Fabaceae	Lower Andes; Pacific coast	Root
Algarrobo		Prosopis spp.				Fabaceae	Pacific coast	Fruit
Almendra	Almendro; sawari; piqui; piquía	Caryocar spp.				Caryocaraceae	Tropics	Fruit; oil; nut; fish poison
Amaranth	Kiwicha; achita; bledo; millmi; chaquillón	Amaranthus caudatus	L.	Amaranthus edulis; Amaranthus mantegazzianus	Speg.; Passer	Amaranthaceae	Andes; Pacific coast	Cereal
Amaranth	Bledo; jacato	Amaranthus cruentus	L.			Amaranthaceae	Andes	Cereal
Amaranth	Kiwillo; sangorache	Amaranthus quitensis	Kunth			Amaranthaceae	Andes	Cereal
American oil palm	Corozo; puma yarina; caiaué; nolí	Elaeis oleifera	Kunth			Arecaceae	Northern Amazonia	Fruit; oil
Anchovy pear	Sacha mango	Grias peruviana; Grias spp.	Miers			Lecythidaceae	Western Amazonia	Fruit

English name	Local names	Linnean name	Author citation	Alternative Linnean name	Alternative author citation	Family	Region	Use
Annatto; achiote	Mantur; urucú	Bixa orellana	L.			Bixaceae	Tropics	Dye
Anona; sweetsop; sugar apple	Annona; fruta de conde	Annona squamosa	L.			Annonaceae	Tropics	Fruit
Añu	Isaño; mashua	Tropaeolum tuberosum	R. & P			Tropaeolaceae	Andes	Root
Arazá	Guayaba brasileira; araçá-boi	Eugenia stipitata	McVaugh			Myrtaceae	Western Amazonia	Fruit
Arracacha	Apio; raccacha; lakacha	Arracacia xanthorriza	Bancr			Umbelliferae	Andes	Root
Arrayan		Myrcianthes foliosa	(Kunth) Mc Vaugh	Eugenia foliosa	(Kunth) DC	Myrtaceae	Andes	Fruit
Arrow cane	Flecha; lata; caña brava; giant cane	Gynerium sagittatum	(Aubl.) Beauv			Poaceae	Tropics	Arrows; construction
Arrowroot; maranta	Guapo; sagu; ereu	Maranta arundinacea	L.			Marantaceae	Tropics	Root
Assai	Açaí; palmito; chonta; murrapo	Euterpe oleracea	Mart			Arecaceae	North Tropics	Fruit
Avocado	Palta; aguacate; abacate	Persea gratissima	Mill			Lauraceae	Tropics	Fruit
Ayahuasca	Caapi; yagé; pindé; timbo branco; napi	Banisteriopsis caapi	(Spruce ex Griesb.) Morton	Banisteria caapi	Spruce ex Griesb	Lauraceae	Tropics	Fruit
Ayapana		Eupatorium ayapana Vent	Vent			Asteraceae	Amazonia	Condiment
Babaco	Papayo calentano	Carica pentagona	Heilborn			Caricaceae	Southern Amazonia	Fruit; oil
Babassu	Babaçu	Orbignya phalerata	Mart			Arecaceae	Southern Amazonia	Fruit; oil

English name	Local names	Linnean name	Author citation	Alternative Linnean name	Alternative author citation	Family	Region	Use
Bacuripari	Baçu; pacuri	Rheedia brasiliensis	Planch. & Triana			Clusiaceae	Eastern Amazonia	Fruit
Banana passion fruit	Curuba; tasco	Passiflora mollissima	(Kunth) Bailey	Tacsonia mollissima	Kunth	Passifloraceae	Andes	Fruit
Barbados cherry	Cerezo; acerola; cerejeira do Pará	Malgiphia punicifolia; Malgiphia spp.	L.			Malpighiaceae	Tropics	Fruit
Barbasco; rotenone	Cube; timbó; nicou	Lonchocarpus nicou; Lonchocarpus spp.	(Aubl.) DC			Fabaceae	Tropics	Fish poison; arrow poison
Barbasco	Ciruelito; para-para; timbó	Phyllanthus piscatorum; Phyllanthus spp.	Kunth			Euphorbiaceae	Northern Amazonia	Fish poison; medicinal
Barbasco	Cunumi; huaco	Clibadium asperum (Aubl.); Clibadium spp.	(Aubl.) DC			Asteraceae	Amazonia	Fish poison
Barbasco	Cube; kumu; timbó	Tephrosia sinapou; Tephrosia spp.	(Buch.) A Chev	Tephrosia toxicaria	Sw. Pers	Fabaceae	Amazonia	Fish poison; medicinal
Basul	Sachaporoto; pajuro	Erythrina edulis; Erythrina spp.	Triana			Fabaceae	Andes	Seed; living fences
Begonia		Begonia spp.				Begoniaceae	Amazonia	Medicinal; ornamental
Biribá	Anona	Rollinia mucosa	(Jacq.) Baill	Annona mucosa	Jacq	Annonaceae	Tropics	Fruit
Blackberry	Zarzamora; mora de Castilla	Rubus spp.				Rosaceae	Andes	Fruit
Borojoa	Purui grande; borjó	Borojoa sorbilis	(Ducke) Cuatr	Aliberta sorbilis; Thieleodoxa sorbilis	Ducke; Ducke	Rubiaceae	Amazonia	Fruit
Brazil nut	Castaña; castanha do Pará	Bertholletia excelsa	H. & B			Lecythidaceae	Tropics	Nut; medicinal

English name	Local names	Linnean name	Author citation	Alternative Linnean name	Alternative author citation	Family	Region	Use
Breadnut; ramón	Machinga	Brosimum alicastrum	Sw			Moraceae	Tropics	Fruit; seed
Bullock's heart; custard apple	Sacha anona	Annona reticulata	L.			Annonaceae	Tropics	Fruit
Cabuya	Pia floja	Furcraea cabuya	Trel			Amaryllidaceae	Andes	Fiber
Cacao; cocoa	Cacau	Theobroma cacao	L.			Sterculiaceae	Tropics	Fruit; medicinal; stimulant
Cacaui	Cacau sacha; choclatillo	Theobroma speciosum	Willd			Sterculiaceae	Amazonia	Fruit
Cache	Petaco; matatropa; toda la vida	Phaseolus polyanthus	Greenman			Fabaceae	Northern Andes	Legume; seed
Caimito	Abíu; caimo; cauje; abieiro; sapote	Pouteria caimito	(R. & P.) Radlk	Lucuma caimito	Roem. Schult	Sapotaceae	Tropics	Fruit
Calabash; tree gourd	Calabaza; higuera; jicaro; pate; pati; pamuco; tutuma; cuia; ancara	Crescentia cujete	L.			Bignoniaceae	Tropics	Fruit; gourd
Canarana	Mishquipanga; huitillo	Renalmia spp.				Zingiberaceae	Amazonia	Fruit; dye
Cañihua	Kaniwa; cañahua	Chenopodium pallidicaule	Aellen			Chenopodiaceae	Andes	Cereal
Canistel	Lucma; cutité	Pouteria macrophylla	(Lam.) Eyma	Lucuma spp.; Aridella spp.; Zitellaria spp.		Sapotaceae	Amazonia; Andes	Fruit
Casabanana	Secana; curua; cajuba; curuba	Sicana odorifera	(Vell.) Naud			Cucurbitaceae	Tropics	Vegetable
Cashew	Caju; merey; marañon	Anacardium occidentale	L.			Anacardiaceae	Tropics	Nut; fruit

English name	Local names	Linnean name	Author citation	Alternative Linnean name	Alternative author citation	Family	Region	Use
Cayaponia	Munteká	Cayaponia ophtalmica	R.E. Schultes			Cucurbitaceae	Colombian Amazonia	Vine; medicinal
Chambira	Fiber palm; tucum; cumare	Astrocaryum chamira; Astrocaryum spp.	Burret			Arecaceae	Northwest Amazonia	Fruit; medicinal; seed; fiber
Chayote	Christophine; chuchu	Sechium edule	(Jacq.) Sw			Cucurbitaceae	Tropics	Fruit; vine
Cherimoya; custard apple	Chirimoya; masa; cherimolia	Annona cherimola	Mill			Annonaceae	Tropics; Andes	Fruit
Chilean mango; mangu	Tecu	Bromus mango	E. Desv			Poaceae	central Chile; Argentina	Cereal; seed
Chilean strawberry	Quellen; frutilla; morango de Chile	Fragaria chiloensis	Georgi			Rosaceae	Southern Andes	Fruit
Chili pepper; pepper; ají	Chile; rocoto; pimenta; uchu; tabasco; panka; maiagueta	Capiscum annuum; Capsicum baccatum; Capsicum chinense; Capsicum frutescens; Capsicum pubescens; Capsicum spp.	L.; L.; Jacq.; L.; R. & P			Solanaceae	Andes; Tropics	Condiment
Coca		Erytroxylon coca var. coca	Lam			Erythroxylaceae	Lower Andes; coastal Peru	Narcotic; medicinal
Coca Colombina		Erythroxylon novogranatense	(Morris) Hieron			Erythroxylaceae	Northern Andes; coastal Peru	Narcotic; medicinal
Coca ipadu	Ipadú	Erythroxylon coca var. Ipadu	Plowman			Erythroxylaceae	Amazonia	Narcotic; medicinal
Coconilla	Coconilla colorado	Solanum coconilla	Huber			Solanaceae	Tropics	Fruit
Coconut	Cóco; coqueiro	Cocos nucifera	L.			Arecaceae	Tropics	Fruit; oil

English name	Local names	Linnean name	Author citation	Alternative Linnean name	Alternative author citation	Family	Region	Use
Cocopalm	Mucajá; macaúba	Acrocomia sclerocarpa	Mart			Arecaceae	Central Amazonia	Fruit; oil; palm heart
Cocoyam; yauita	Huitina; mangareto otó; malanga; onko; uncucha; ocumo; tannia; taioba; tegüe	Xanthosoma sagittifolium; Xanthosoma spp.	Schott			Araceae	Tropics	Root
Colombian passion fruit		Passiflora antioquiensis	Karst			Passifloraceae	Northern Andes	Fruit
Coriander	Culantro de monte	Eryngium foetidum	L.			Apiaceae	Tropics	Spice; medicinal
Cotton	Algodón; algodão	Gossypium barbadense; Gossypium hirsutum	L.; L:			Malvaceae	Tropics; Subtropics	Fiber
Cow tree	Sorveira; sorva	Couma utilis	(Mart.) Muel. Arg			Apocynaceae	Amazonia	Fruit; latex; gum
Cupá	Kupá; cipó-babão	Cissus gongyloides	Burch			Vitaceae	Brazilian Amazonia	Vegetable
Cupuassu	Cupuaçu	Theobroma grandiflorum	(Willd. Ex Spreng) Schum			Sterculiaceae	Eastern Amazonia	Fruit
Datura	Nongué	Datura spp.				Solanaceae	Tropics	Medicinal; narcotic
Flat sedge	Piripiri; pripreoca	Cyperus spp.				Cyperaceae	Western Amazonia	Sedge; condiment; medicinal
Floripondo; angels trumpet	Floripondio; toé; toatoé; maricahua	Brugmansia aurea; Brugmansia insignis; Bruckmansia suaveolens	Langerhein; (Barb. Rodr.) Lockwood; H. & B. ex Willd			Solanaceae	Amazonia; lower Andes	Medicinal; hallucinogen

English name	Local names	Linnean name	Author citation	Alternative Linnean name	Alternative author citation	Family	Region	Use
Fragrant granadilla	Maracuja grande	Passiflora alata	Curtis			Passifloraceae	Tropics	Fruit
Galupa	Bejuco; chulup	Passiflora pinnatistipula	Cav			Passifloraceae	Andes	Fruit
Genipa	Genipap; caruto; huito	Genipa americana	L.			Rubiaceae	Tropics	Dye; fruit
Gherkin		Cucumis anguria				Cucurbitaceae	Tropics	Fruit
Giant granadilla; passion fuit	Granadilla; tumbo; maracujá-açu; badea	Passiflora quadrangularis	L.			Passifloraceae	Tropics	Fruit
Golden bell	Campanilla de oro; canaria	Allamanda cathartica	L.			Apocynaceae	Amazonia	Vine; ornamental; medicinal
Goldenberry; cape gooseberry	Capulí; camapun; aguaymanto; topotopo	Physalis peruviana	Mill			Solanaceae	Andes; Amazonia	Fruit
Gourd; bottle gourd	Poro; maté; tapara; camasa; totuma; calabaza; cabaça; cuiera; cuia	Lagenaria siceraria	(Molina) Standley			Cucurbitaceae	Tropics; temperate Andes	Misc
Granadilla; passion fruit		Passiflora spp.				Passifloraceae	Tropics	Fruit
Granadilla de Quijos	Curubejo	Passiflora popenovii	Killip			Passifloraceae	Ecuadorian Amazonia	Fruit
Grape-leaf passion fruit	Granadilla del monte; guate; tumbo	Passiflora vitifolia	Kunth			Passifloraceae	Tropics	Fruit
Green plum; ciruela	Ciruela de fraile; ciruelo; usum	Bunchosia armeniaca	(Cav.) Rich			Malpighiaceae	Andes; Peruvian coast; Amazonia	Fruit

English name	Local names	Linnean name	Author citation	Alternative Linnean name	Alternative author citation	Family	Region	Use
Guaba; pacay	Guabilla; guabo; guamo; guama; inga sipo; pacae; shimbillo	Inga edulis; Inga spp.	Mart			Fabaceae	Tropics	Fruit
Guaná-bana; soursop	Quabana; graviola; masasamba	Annona muricata	L.			Annonaceae	Tropics; temperate Andes	Fruit
Guaraná	Cupana	Paullinia cupana	Kunth	Paullinia sorbilis	Mart	Sapindaceae	Amazo-nia	Stimulant; beverage
Guava; guayaba	Guayabo; goiaba; sawintu; araça	Psidium guajava; Psidium spp.	L.			Myrtaceae	Tropics	Fruit
Guayusa		Ilex guayusa	Loes			Aquifoli-aceae	Western Ama-zonia; lower Andes	Medici-nal; stimu-lant
Guiana orange; bacuri	Bacuri do Paraná	Platonia esculenta	(Arruda) Rick-ett & Stafleu	Platonia insignis	Mart	Clusiaceae	Eastern Ama-zonia; Para-guay	Fruit
Heliconia	Millue	Heliconia hirsuta; Heliconia spp.	L.f			Heliconi-aceae	Tropics	Root; medici-nal; beverage
Hog plum; mombin	Yellow mom-bin; ciruela; cajá; taperebá; jobo; ubos; taperibá	Spondias mombin	L			Anacardi-aceae	Tropics	Fruit
Husk tomato; ground cherry	Capulí	Physalis pubensencs	L.	Spondias lutea	L	Solanaceae	Lower Andes; Amazo-nia	Fruit; medici-nal
Indigo	Añil; mutui; velvet bean	Indigofera suffruticosa	Mill	Indigofera anil	L	Fabaceae	Lower Andes	Dye
Ipecac	Ipecacuana; raicilla; golden root	Psychotria ipecacuana	(Brot.) Stokes	Cephalis ipecacu-anha	(Brot.) Tussac	Rubiaceae	Amazo-nia	Root; medici-nal

English name	Local names	Linnean name	Author citation	Alternative Linnean name	Alternative author citation	Family	Region	Use
Ivory palm	Tagua; yarina; vegetable ivory; ivory nut	Phytelephas macrocarpa	R. & P			Arecaceae	Western Amazonia	Nut; crafts
Jaboticaba	Camboim	Myrciaria caudiflora	McVaugh			Myrtaceae	Brazil; Paraguay	Fruit
Jacaranda	Ashpingo; para-paraúba	Jacaranda copaia	Aubl			Bignoniaceae	Amazonia	Medicinal
Jackbean	Pallar de gentiles; nescafé	Canavalia ensiformis; Canavalia plagiosperma	(L.) DC; Piper			Fabaceae	Tropics; Pacific coast	Legume; seed
Jesus' heart		Caladium bicolor	Ait			Araceae	Amazonia	Ornamental; potions
Kaahée	Sweet herb; stevia	Stevia rebaudiana	Bertoni			Asteraceae	Paraguay	Sweetener; medicinal
Kidney bean; common bean	Frijol; poroto; nuñas; feijão	Phaseolus vulgaris	K			Fabaceae	Tropics; Andes	Legume; seed
Lerén; sweet cornroot	Lirén; alluia; topinambur; dale dale; arruruz; ariã	Calathea allouia	(Aubl.) Lindl	Maranta lutea	Aubl	Marantaceae	Tropics	Root
Lima bean	Pallar; haba; feijão de Lima	Phaseolus lunatus	L.			Fabaceae	Tropics; lower Andes	Legume; seed
Lucuma	Lucmo; lucma	Pouteria lucuma; Pouteria spp.	(R. & P.) Ktze	Lucuma spp.		Sapotaceae	Amazonia; Andes	Fruit
Maca	Maka	Lepidium meyenii	Walp			Cruciferae	Andes	Root
Macambo	Patashte; pataste macambillo; cacau do Perú	Theobroma bicolor	H. & B			Sterculiaceae	Western Amazonia	Fruit; seed

English name	Local names	Linnean name	Author citation	Alternative Linnean name	Alternative author citation	Family	Region	Use
Macoubea	Amapa	Macoubea witotorum	R.E. Schultes			Apocynaceae	Colombian Amazonia	Fruit
Madi	Madivilcum; tarweed	Madia sativa	Molina			Asteraceae	central Chile	Oil; seed
Madrono	Naranjito; bacuripari	Rheedia spp.				Clusiaceae	Amazonia	Fruit
Maize; corn	Maíz; sara; milho	Zea mays	L.			Poaceae	Andes	Cereal
Mamey	Mammey; mammee; abricó	Mammea americana	L.			Clusiaceae	Tropics	Fruit
Mamón	Mamoncillo; maca; pitomba; genip	Melicocca bijuga	L.	Melicocca bijugatus	Jacq	Sapindaceae	Northern Amazonia	Fruit
Manioc	Yuca; cassave; casabe; mandioca; ruma	Manihot esculenta	Crantz			Euphorbiaceae	Tropics	Root
Marimari		Cassia leiandra; Cassia spp.	Benth			Fabaceae	Amazonia	Fruit
Maté	Yierba maté	Ilex paraguariensis	St. Hil			Aquifoliaceae	Subtropics	Beverage
Mauka	Miso	Mirabilis expansa	R. & P. Standl			Nyctaginaceae	Temperate Andes	Root
Melon pear; pepino	Pepino; cachun; kachan	Solanum muricatum	Ait			Solanaceae	Temperate Andes	Fruit
Molle	Arbol de pimienta; mulli	Schinus molle	L.			Anacardiaceae	Andes	Medicinal; beverage; ornamental
Moriche; aguaje	Buriti; miriti	Mauritia flexuosa	L.f			Arecaceae	Amazonia	Fruit; starch; fiber; construction
Mullaca	Bolsa mullaca	Physalis angulata	L.			Solanaceae	Amazonia	Medicinal; narcotic

English name	Local names	Linnean name	Author citation	Alternative Linnean name	Alternative author citation	Family	Region	Use
Nambicuara peanut	Amendoím indigena	Arachis nambyquarae	Hoehne			Fabaceae	Southern Amazonia	Legume; seed
Nance	Muruçi; chaparro	Byrsonima crassifolia	(L.) Kunth			Malpighiaceae	Tropics	Fruit
Naranjilla; Quito orange	Lulo; toronja	Solanum quitoense	Lam			Solanaceae	Tropics	Fruit
Oca	Cavi	Oxalis tuberosa	Molina	Xanthoxolis tuberosa	(Molina) Holub	Oxalidaceae	Andes	Root
Orinoco nut	Inchi; cacay; palo de nuez; tacay nut	Caryodendron orinocense	Karst			Euphorbiaceae	Northern Amazonia	Nut
Paco	Membrillo	Gustavia superba	(Kunth) Berg			Lecythidaceae	Tropics	Fruit
Palillo	Guaribo	Campomanesia lineatifolia	R. & P			Myrtaceae	Western Amazonia; lower Andes	Fruit; medicinal
Panama hat palm	Bonbonaje	Carludovica palmata	R. & P			Cylanthaceae	Tropics	Fiber
Papaya	Lechosa; mamâo; pawpaw	Carica papaya; Carica spp.	L.			Caricaceae	Tropics; lower Andes	Fruit
Papaya (dwarf); mito	Chamburu	Carica candicans	Gray			Caricaceae	Andes	Fruit
Papaya (mountain)	Chamburo	Carica pubescens	Solms			Caricaceae	Lower Andes	Fruit
Pará cress	Berro; botoncillo	Spilanthes acmella	(L.) Murr			Asteraceae	Amazonia	Condiment
Parinari	Supay ocote; umarirana	Couepia subcordata; Couepia spp.	Benth			Chrysobalanaceae	Amazonia	Fruit; nut
Passion fruit; granadilla		Passiflora spp.				Passifloraceae	Tropics	Fruit
Patauá; ungurahui	Bataua; batana; seje	Jessenia bataua	(Mart.) Burret	Oenocarpus bataua	Mart	Arecaceae	Tropics	Fruit

English name	Local names	Linnean name	Author citation	Alternative Linnean name	Alternative author citation	Family	Region	Use
Peach palm	Pejibaye; chonta duro; pijuayo; pupunha	Bactris gasipaes	Kunth	Guilielma speciosa; Guilielma gasipaes	Mart.; Kunth Bailey	Arecaceae	Tropics	Fruit; palm heart
Peach tomato; cocona	Topiro; lulo; cubiu	Solanum sessiliflorum	Dunal	Solanum topiro	H.B. ex Dunal	Solanaceae	Tropics	Fruit
Peanut	Maní; inchis; amendoím	Arachis hypogaea	L.			Fabaceae	Amazonia; Pacific coast	Legume; seed
Penca	Pakpa; chuchao	Furcraea andina	Trel			Amarylli-daceae	Andes	Fiber; ornamental
Pineapple	Piña; abacaxi; ananás; curauá; achupalla	Ananas comosus	(L.) Merr	Ananas sativus; Ananas ananas	Schult.; (L.) Voss	Bromeli-aceae	Tropics	Fruit
Pineapple (wild)		Ananas erectifolius	L.B. Smith			Bromeli-aceae	Amazonia; Orinoco	Fiber
Pitomba		Talisia esculenta	Radlk	Sapindus edulis	St. Hil	Sapindaceae	Tropics	Fruit
Potato	papa; apichu; batata	Solanum tuberosum; Solanum spp.	L.			Solanaceae	Andes	Root
Potato (tropical)	Moshaki; amista; kurahji	Solanum hygrothermicum	Ochoa			Solanaceae	Peruvian Amazonia	Root
Potato bean; jícama	Macucu; nupe	Pachyrhizus tuberosus	(Lam.) Spreng			Fabaceae	Amazonia; Paraguay	Root
Prickly pear	Pata kiska; espino; cacto	Opunita exaltata; Opuntia spp.	Berger			Cactaceae	Andes; Pacific coast	Fencing; misc
Quinoa	Kiuna; trigo Inca	Chenopodium quinoa	Willd			Chenopodi-aceae	Andes	Cereal

English name	Local names	Linnean name	Author citation	Alternative Linnean name	Alternative author citation	Family	Region	Use
Quishuar		Buddleja incana; Buddleja spp.				Buddlejaceae	Andes	Cane; tools; carvings; fuel; leaves for fertilizer and medicine
Quito palm; coquito		Parajubaea cocoides	Burr			Arecaceae	Ecuadorian Andes	Nut; ornamental
Rope plant	Caraua; carúa; caroá; curratow	Neoglaziovia variegata	(Arruda) Mez	Bromelia variegata	Arruda	Bromeliaceae	Tropics	Fiber
Rosy passion fruit	Curuba bogotana; red banana passion fruit	Passiflora cumbalensis	Harms			Passifloraceae	Northern Andes	Fruit
Rubber	Shiringa; seringa; jebe; goma; caucho; seringuiera	Hevea brasiliensis; Hevea spp.	(Willd.) Muell.. Arg			Euphorbiaceae	Amazonia	Latex; seed
Sacaca	Cascarilla	Croton cajucara	Benth			Euphorbiaceae	Amazonia	Medicinal
San Pedro cactus	Gigantón; jahuackollai	Trichocereus cuzcoensis; Trichocereus spp.	Britton & Rose			Cactaceae	Andes; coastal Peru	Hallucinogen; fruit
Sapodilla chiclé	Níspero; zapote; sapoti; naseberry; caimito brasilero	Manikara achras	(Mill.) Fosberg	Manikara zapota; Achras zapota	(L.) Van Royen; L	Sapotaceae	Western Amazonia	Fruit
Sapucaia	Castaña de monte	Lecythis pisonis	Caub			Lecythidaceae	Amazonia	Nut
Scarlet runner bean	Ayocote; fava	Phaseolus coccineus	(L.) DC	Phaseolus multiflorus	Willd	Fabaceae	Subtropics	Legume; seed
Spanish plum; mombin	Cajazeiro; red mombin	Spondias purpurea	L.			Anacardiaceae	Northern South America; Tropics	Fruit

English name	Local names	Linnean name	Author citation	Alternative Linnean name	Alternative author citation	Family	Region	Use
Squash; pumpkin	Zapallo; auyama; jerimum; achocha; abóbora	Cucurbita maxima; Cucurbita ficifolia; Cucurbita moschata; Cucurbita ecuadoriensis	Duchesne; Bouché; (Duchesne) Poir.; Cutler and Whitaker	Pepo ficifolia	(Bouché) Britton	Cucurbitaceae	Pacific coast	
Star apple; cainito	Caimito	Chrysophyllum cainito	L.			Sapotaceae	Tropics	Fruit
Strawberry guava		Psidium catleianum	Sabine			Myrtaceae	Coastal Brazil	Fruit
Sunsapote		Licania platypus	(Helmsley) Fritsch			Chrysobalanaceae	Northern South America	Fruit
Surinam cherry	Pitanga; cereza	Eugenia uniflora	L.	Stenocalyx uniflorus	(L.) Kausel	Myrtaceae	Tropics	Fruit
Sweet calabash	Chulupa; curuba	Passiflora maliformis	L.			Passifloraceae	Colombian Andes	Fruit
Sweet granadilla	Sweet passion fruit; granadilla; tumbo; badea	Passiflora ligularis	Juss			Passifloraceae	temperate Andes	Fruit
Sweet potato	Batata; camote; cumara; aje	Ipomoea batatas	(L.) Lam			Convolvulaceae	Tropics; lower Andes	Root
Tarwi; lupine	Tarhui; chocho	Lupinus mutabilis	Sweet			Fabaceae	Andes	Seed
Tobacco	tobaco; tabaco	Nicotiana rustica; Nicotiana tabacum	L.			Solanaceae	Tropics; lower Andes	Narcotic; medicinal
Tomato	Tomate; paconca	Lycopersicum esculentum	Mill			Solanaceae	Tropics; lower Andes	Fruit
Totai	Mbocayá; coco paraguayo	Acrocomia totai	Mart			Arecaceae	Eastern Bolivia; Paraguay	Fruit; oil; palm heart
Tree tomato	Sacha tomate; tomatillo; tamarillo	Cyphonandra betacea; Cyphomandra spp.	(Cav.) Sendt			Solanaceae	Temperate Andes	Fruit

English name	Local names	Linnean name	Author citation	Alternative Linnean name	Alternative author citation	Family	Region	Use
Trompillo	Guabillo; huitillo; purui	Alibertia edulis	A. Rich. ex DC			Rubiaceae	Tropics	Fruit
Tucumã	Tucumã; chonta	Astrocaryum aculeatum	G. Meyer	Astrocaryum tucuma; Astrocaryum princeps	Mart.; Barb. Rodr	Arecaceae	Amazonia	Fruit; fiber
Tucumão	Curua; cumare; aoura	Astrocaryum vulgare	Mart			Arecaceae	Brazilian Amazonia	Fruit; fiber
Ugni	Murta	Myrtus ugni	Moll	Ugni ugni	Macloskie	Myrtaceae	Chilean Andes	Fruit
Ulluco	Papa lisa; ullucu; melloco	Ullucus tuberosus	Caldas			Basellaceae	Andes	Root
Umarí	Marí	Poraqueiba sericea	Tul	Poraqueiba paraensis	Ducke	Icacinaceae	Western Amazonia; Eastern Amazonia	Fruit; oil
Uvilla; grape tree	Ubilla; puruma	Pourouma cecropiaefolia	Mart			Moraceae	Tropics	Fruit
Vanilla		Vanilla planifolia; Vanilla spp.	Andr			Orchidaceae	Tropics; Andes	Vine; flavoring
Water lemon	Bell apple; yellow granadilla; maracujá; Jamaica honeysuckle	Passiflora laurifolia	L.			Passifloraceae	North Tropics	Fruit
Yacón	Llacón; aricona; aricuma; jíquima	Polymnia sonchifolia	Poeppig & Endlicher			Asteraceae	Temperate Andes	Root
Yampee; American yam; mapuey	Cará; sacha papa; cushcush; ñame	Dioscorea trifida; Dioscorea spp.	L.			Dioscoreaceae	Tropics	Root

English name	Local names	Linnean name	Author citation	Alternative Linnean name	Alternative author citation	Family	Region	Use
Yellow oleander	Flor amarilla	Thevetia peruviana	(Pers.) Schumann	Cerbera peruviana; Cerbera thevetia	Pers.; L	Apocynaceae	Tropics	Medicinal; poison
Yellow passion fruit; purple passion fruit	Yellow granadilla; purple granadilla; maracujá	Passiflora edulis	Sims			Passifloraceae	Tropics; South Brazil	Fruit
Yoco		Paullinia yoco	Schultes & Killip			Sapindaceae	Western Amazonia	Medicinal; hallucinogen
Yopo	Cohoba; paricá	Anadenathera peregrina	(L.) Speg			Fabaceae	Amazonia	Medicinal bark; hallucinogen

Acknowledgements We thank Johannes Dellert, Robert Forkel, Richard Lawrence, Thorsten Trippel, and Alena Witzlack-Makarevich for discussion on technical matters, and we are grateful to Erhard Hinrichs for allowing CINWA to be hosted on the servers of the CLARIN-D infrastructure. Likewise, we acknowledge the helpful comments of the anonymous referees.

Author contributions Conceptualization: MU; Methodology: MU, SL; Programming: SL; Data generation and Investigation: EMAP, EB, MU Writing—original draft preparation: MU; Writing—review and editing: MU, SL, EMAP; Funding acquisition: MU; Supervision: MU.

Funding Open Access funding enabled and organized by Projekt DEAL. The work on CINWA was supported by the Daimler and Benz foundation through a grant to Matthias Urban.

Data availability All data contained in CINWA are available from the online interface at cinwa.org. In addition, raw source files in .tsv format and bibliographic information on the sources used in BibTex format are available for download without restrictions.

Declarations

Conflict of interest The authors have no competing interests to declare that are relevant to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Adelaar, W. F. H. (1998). The name of the sweet potato: A case of pre-conquest contact between South America and the Pacific. In M. Janse & A. Verlinden (Eds.), *Productivity and creativity: Studies in general and descriptive linguistics in honour of E.M. Uhlenbeck* (pp. 403–411). Mouton de Gruyter.

Balée, W. L. (1994). *Footprints of the forest: Ka'apor ethnobotany—the historical ecology of plant utilization by an Amazonian people*. Columbia University Press.

Balée, W. (2013). *Cultural forests of the Amazon: A historical ecology of people and their landscapes*. University of Alabama Press.

Berlin, B. (1992). *Ethnobiological classification: Principles of categorization of plants and animals in traditional societies*. Princeton University Press.

Brown, C. H. (1986). The growth of ethnobiological nomenclature. *Current Anthropology, 27*(1), 1–11.

Brüning, H. H. (2004). *Mochica Wörterbuch. Diccionario mochica. Mochica—castellano/castellano—mochica.* Ed. José Antonio Salas García. Universidad de San Martín de Porres, Escuela Profesional de Turismo y Hotelería.

Campbell, L. (1997). *American Indian languages: The historical linguistics of Native America*. Oxford University Press.

Crevels, M. (2012). Language endangerment in South America: The clock is ticking. In L. Campbell & V. Grondona (Eds.), *The indigenous languages of South America: A comprehensive guide* (pp. 167–234). Walter de Gruyter.

Dellert, J., Daneyko, T., Münch, A., Ladygina, A., Buch, A., Clarius, N., Grigorjew, I., Balabel, M., Boga, H. I., Baysarova, Z., Mühlenbernd, R., Wahle, J., & Jäger, G. (2020). NorthEuraLex: A wide-coverage lexical database of Northern Eurasia. *Language Resources & Evaluation, 54*, 273–301. https://doi.org/10.1007/s10579-019-09480-6

Denevan, W. M. (2001). *Cultivated landscapes of native Amazonia and the Andes*. Oxford University Press.

Dillehay, T. D., Rossen, J., Andres, T. C., & Williams, D. E. (2007). Preceramic adoption of peanut, squash, and cotton in Northern Peru. *Science, 316*, 1890–1893. https://doi.org/10.1126/science.1141395

Ferraz Gerardi, F., Reichert, S., & Coelho Aragon, C. (2021). TuLeD (Tupían lexical database): Introducing a database of a South American language family. *Language Resources & Evaluation, 55*, 997–1015. https://doi.org/10.1007/s10579-020-09521-5

Forkel, R., List, J.-M., Greenhill, S. J., Rzymski, C., Bank, S., Cysouw, M., Hammarström, H., Haspelmath, M., Kaiping, G. A., & Gray, R. D. (2018). Cross-linguistic data formats, advancing data sharing and re-use in comparative linguistics. *Scientific Data, 5*(1), 1–10. https://doi.org/10.1038/sdata.2018.205

Forkel, R., Bank, S., & Rzymski, C. (2019). clld: A toolkit for cross-linguistic databases (v5.0.0). https://doi.org/10.5281/zenodo.3437148

Hammarström, H., Forkel, R., Haspelmath, M., & Bank, S. (2021). *Glottolog 4.4.* Max Planck Institute for Evolutionary Anthropology. https://doi.org/10.5281/zenodo.4761960

Harrison, K. D. (2007). *When languages die: The extinction of the world's languages and the erosion of human knowledge*. Oxford University Press.

Hart, H. (1988). *Diccionario chayahuita–castellano: Canponanquë nisha Nisha nonacaso'*. Ministerio de Educación/Instituto Lingüístico de Verano.

Haynie, H., Bowern, C., Epps, P., Hill, J., & McConvell, P. (2014). Wanderwörter in the languages of the Americas and Australia. *Ampersand, 1*, 1–18. https://doi.org/10.1016/j.amper.2014.10.001

Hunn, E. S., & Brown, C. H. (2011). Linguistic ethnobiology. In E. N. Anderson, D. Pearsall, E. Hunn, & N. Turner (Eds.), *Ethnobiology* (pp. 319–333). Wiley-Blackwell.

Ioannidis, A. G., Blanco-Portillo, J., Sandoval, K., Hagelberg, E., Miquel-Poblete, J. F., Moreno-Mayar, J. F., Rodríguez-Rodríguez, J. E., Quinto-Cortés, C. D., Auckland, K., Parks, T., Robson, K., Hill, A. V. S., Avila-Arcos, M. C., Sockell, A., Homburger, J. R., Wojcik, G. L., Barnes, K. C., Herrera, L., Berríos, S., et al. (2020). Native American gene flow into Polynesia predating Easter Island settlement. *Nature, 583*, 572–577. https://doi.org/10.1038/s41586-020-2487-2

Jeffreys, M. D. W. (1963). The banana in the Americas. *Journal Agriculture Tropicale Et De Botanique Appliquée, 10*(5–7), 189–203. https://doi.org/10.3406/jatba.1963.2705

Kaufman, T. (1990). Language history in South America: What we know and how to know more. In D. L. Payne (Ed.), *Amazonian linguistics: Studies in lowland South American languages* (pp. 13–67). University of Texas Press.

Kaulicke, P. (2020). Early social complexity in northern Peru and its Amazonian connections. In A. J. Pearce, D. G. Beresford-Jones, & P. Heggarty (Eds.), *Rethinking the Andes-Amazonia divide: A cross-disciplinary exploration* (pp. 103–115). UCL Press.

Key, M. R., & Comrie, B. (Eds.). (2021). *The intercontinental dictionary series*. Max Planck Institute for Evolutionary Anthropology. http://ids.clld.org

Langdon, R. (1993). The banana as a key to Early American and Polynesian history. *The Journal of Pacific History, 28*(1), 15–35.

Maffi, L. (Ed.). (2001). *On biocultural diversity: Linking language, knowledge, and the environment*. Smithsonian Institution Press.

Moran, S., & Cysouw, M. (2018). *The unicode cookbook for linguists: Managing writing systems using orthography profiles*. Language Sciences Press.

Muysken, P., Hammarström, H., Krasnoukhova, O., Müller, N., Birchall, J., van de Kerke, S., O'Connor, L., Danielsen, S., van Gijn, R., & Saad, G. (2016). *South American Indigenous Language Structures (SAILS) Online*. https://sails.clld.org

Naumann, C., Güldemann, T., Moran, S., Segerer, G., Fehn, A.-M., & Forkel, R. T. (2020). *Tsammalex: A lexical database on plants and animals*. Max Planck Institute for Evolutionary Anthropology. http://tsammalex.clld.org

Newson, L. A. (2002). Cultivated landscapes of Native Amazonia and the Andes (review). *Hispanic American Historical Review, 82*(4), 788–790.

Nordenskiöld, E. (1922). *Deductions suggested by the geographical distribution of post-Columbian words used by the Indians of S. America*. Elanders Boktryckeri.

Park, M., Weber, N., & Cenepo Sangama, V. (1976). *Diccionario quechua: San Martín*. Ministerio de Educación/Instituto Lingüístico de Verano.

Pearsall, D. M. (2015). Early agriculture in the Americas. In G. Barker & C. Goucher (Eds.), *The Cambridge world history: A world with agriculture, 12, 000 BCE—500 CE* (Vol. 2, pp. 514–538). Cambridge University Press.

Piperno, D. R. (2011). The origin of plant cultivation and domestication in the new world tropics: Patterns, process, and new developments. *Current Anthropology, 52*(S4), S453–S470.

Rankin, R. L., Carter, R. T., Jones, A. W., Koontz, J. E., Rood, D. S., & Hartmann, I. (Eds.). (2015). *Comparative Siouan dictionary*. Max Planck Institute for Evolutionary Anthropology.

Stadel, C. (2005). Cultivated landscapes of Native Amazonia and the Andes (review). *Mountain Research and Development, 25*(4), 388–389. https://doi.org/10.1659/02764741(2005)025[0388:CLONAA]2.0.CO;2

Uffmann, C. (2015). Loanword adaptation. In P. Honeybone & J. Salmons (Eds.), *The Oxford handbook of historical phonology* (pp. 644–666). Oxford University Press.

Urban, M. (2021). Language classification, language contact and Andean prehistory: The North. *Language and Linguistics Compass, 15*(9), e12414. https://doi.org/10.1111/lnc3.12414

Valenzuela Bismarck, P. (2012). *Kirka' lawer'lla'la' ñak: Diccionario shiwilu-castellano, castellano-shiwilu*. Unpublished Manuscript.

Witkowski, S. R., & Brown, C. H. (1983). Marking-reversals and cultural importance. *Language, 59*(3), 569–582.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.