Some principles for language names

Martin Haspelmath

MPI-SHH Jena & Leipzig University

Linguists are sometimes confronted with choices concerning language names. For example, one and the same language may be referred to as *Persian* or *Farsi*. This short paper discusses some principles that one might use for making decisions when there are variant forms in use, or when one feels that none of the existing names is appropriate. The principles discussed here arose from work on Glottolog, an English-language database of the world's languages (Glottolog.org), where each language has a single primary English name (though variant forms are of course included), and where the goal is to choose the best variant form as the primary name of the language. Whenever the question arises which variant name form to choose, the Glottolog editors are guided by these principles, so they are formulated in a prescriptive way, but with explanation and justification for each principle. It seems that the general issue is also quite important for language documenters, because the names of non-major languages are often not fully established yet, and naming decisions have to be made.

1. Language names (like city names) are loanwords, not code-switches   The first principle is that a language name is a conventional word, i.e. when using a language name in English, we use an English word. (And when using a language name in Chinese, we use a Chinese word, and so on.) Of course, language names are typically borrowed words (not formed originally in English or Chinese, but adopted from another language), but they share this with a large number of other English/Chinese words that are nonetheless fully integrated. For example, names of well-known languages like *Russian, Spanish, Greek, Chinese*, or *Indonesian* are easily recognized as English words (fully integrated loanwords, specifically adjectives which also have a nominal use) by their spelling and pronunciation, and often also by their morphology. The same applies to Russian language names ending in -*skij* such as *anglijskij, ispanskij, grečeskij, kitajskij,* or *indonezijskij*, and to language names in other languages. (For simplicity, I generally talk about English language names in this paper, but most of the principles are valid in the same way for all other languages. So when I say "in English," it can generally be taken as a simpler way of saying "in the given metalanguage, with necessary changes made.")

In this respect, language names are like city names, e.g. *Rome, Athens, Cairo, Moscow,* or *Tokyo*. These have a conventional form in English, and the name forms may be different in different languages (for example, Polish has *Rzym, Ateny, Kair, Moskwa,* and *Tokio* for these five cities). That cities can have multiple names in multi-

ple languages is particularly clear in cases of officially multilingual cities like Brussels (French *Bruxelles*, Dutch *Brussel*) or Cape Town (Afrikaans *Kaapstad*, Xhosa *iKapa*).

Language names are therefore not code-switches. In this regard, they are different from person names, which are unique and not supposed to vary from language to language. Thus, the French president François Hollande has a French name with no English counterpart (one could not call him *Francis Holland*), and one generally tries to approximate the pronunciation in the original language. Educated English speakers are quite likely to pronounce his surname as [ɔlãd], even though English phonology lacks the sound [ã].[1]

Thus, a language name in English is often different from the autoglottonym (or endonym), and sometimes very different, e.g. *German* vs. *Deutsch*, *Finnish* vs. *Suomen kieli*, *Japanese* vs. *nihongo*, *Navajo* vs. *Diné bizaad*. This is not a problem, and there is no general expectation that language names should be similar across languages.

2. Names of non-major languages are not treated differently from names of major languages
That language names are English words is true for all languages, i.e. not just the major national languages. The language names *Zulu, Quechua, Cree, Malayalam, Kalkatungu*, and *Mauwake* are thus conventional English words as well.

Of course, in practice, there is less cross-linguistic variation in these names, but some languages have more stringent requirements for loanword adaptation than English (e.g. the first four languages are called *Zulų, Kečujų, Kri*, and *Malajalių* in Lithuanian), and some languages do not use the Latin alphabet and are thus forced also by the script to adapt the language name in some way (thus, Japanese has *Zuuruu-go, Kechua-go, Kurii-go*, and *Marayaaramu-go* for these four languages).

That names of non-major languages should be treated like the names of major languages may not seem evident. Conceivably one might argue that names of major languages should be treated like person names of monarchs, while names of minor languages should be treated like names of ordinary people (cf. note 1). However, it is unclear how to draw the line between "major" and "minor" languages. And of course, linguists generally treat all languages equally, so Glottolog's policy has been to treat the primary names of all languages as English words.

For non-major languages that came to the attention of English speakers only recently, one may ask how they got their name. This is not a question I am concerned with here, though it should be studied more by historians of linguistics (see Aussant (ed.) 2009 for some relevant work).[2] For the purposes of this paper, I am assuming that every language already has at least one English name, or at least a name in one of the major contact languages (Spanish, Indonesian, etc.).

---

[1]Names of monarchs and of some ancient personalities are different, and do behave like city names, e.g. Michelangelo is *Michel-Ange* in French, and Pope Francis is *Franciszek* in Polish.

[2]In many cases, the English name comes from a glottonym, though language names are also often taken from ethnonyms and toponyms (e.g. river names), and not uncommonly they were made up by linguists on the basis of material from the object language (e.g. taking the word for 'person,' or 'our language').

3. Each language has a unique name   Just as cities never have multiple names within the same language (and variation is exclusively handled by renaming, e.g. *Leningrad > St. Petersburg*, or *Madras > Chennai*), languages do not have multiple names (i.e. synonyms are not acceptable). Of course, sometimes multiple names are in current use, but this cannot be a general solution. When Madras was renamed to *Chennai*, the old form *Madras* was used for a while, but it became old-fashioned and is being used less and less. The situation is analogous with changes of language names; thus, when the language *Papago* was renamed to *O'odham*, there was never a period during which both names were acceptable to the community of linguists, or to speakers of American English. As soon as the new name was introduced, the old one became obsolete. This does not necessarily imply that all members of the community obey the same norm and change their behavior at the same time, of course, but the nature of the norm is that only one name is possible.

In most parts of the world there are no official organizations that fix the (English) names of languages, and there may be disagreements over which is the correct name (or the best name) for a language. Nevertheless, the general expectation exists that language names are unique, and when someone proposes a new name for a language (e.g. Stenzel (2013: 1), who says that the name *Kotiria* should be used for the language that was previously widely known as *Guanano*), they expect that this usage will be followed by others.[3]

Sometimes linguists use terms such as *Hindi/Urdu* and *Serbian/Croatian*. These are not multiple alternative names for the same language, but pairs of languages that are so close to each other that linguists sometimes treat them as a single language system. If the claim is that they ARE a single language, the spelling has to be *Hindi-Urdu* in Glottolog.

Ideally, language names would also be unique in the sense that a language name is not used for two different languages. This is generally the case within a particular region, but it is sometimes violated on a world-wide scale. Thus, there is a language called *Buru* in Nigeria and a language called *Buru* in Indonesia. This homonymy is acceptable, but in Glottolog, we add further information in parentheses to such language names ("Buru (Nigeria)"; if two languages in the same country have the same name, some other solution needs to be found).

Linguists sometimes use a different name in parentheses, e.g. "Sakha (Yakut)" (Pakendorf 2009), "Erromangan (Sye)" (Crowley 1998), or "Camling (Chamling)" (Ebert 1997). This seems to be due to the desire to make the title of a book or paper more transparent, but if the idea is that the name of the language should be changed (as is the case for Sakha, whose speakers no longer accept the older name *Yakut*), it defeats the purpose. In any event, a string such as "Erromangan (Sye)" cannot be

---

[3]A reviewer notes that in his experience from India and Myanmar, there are a number of languages which have different names in the two countries, and there is no unification going on. But I doubt that people in India and Myanmar have different views concerning the uniqueness of language names. The lack of unification is more likely to be due to different traditions and a low level of communication between the elites of both countries (and also the fact that in Myanmar the English language does not have the same official role). Note also that the Constitution of Myanmar says that the name of the major language should be *Myanmar* in English, not *Burmese* – thus, the attitude to the name of the language is exactly the same as that described for Kotiria in the text.

seen as a single name, but must be interpreted as something like "the Erromangan language, which was formerly known as Sye."

The problem that different names are often in use for the same language has given rise to the idea of language codes such as the ISO 639-3 code (assigned by SIL International for ISO) and the Glottocode (assigned by the Glottolog editors). These are very useful for technical purposes, but they do not supplant the general expectation that each language has a single name in English, as they are codes, not names, and are not intended for use outside of technical contexts (somewhat like airport codes, which distinguish Melbourne in Australia (MEL) from Melbourne in Florida (MLB) but are not intended to supplant the city names). It is true that the existence of unique language codes means that language names carry less of a burden, but the current paper is about language names in ordinary use, not about the issue of distinguishing languages in technical contexts.

The following principle (§4) follows immediately from the principle of uniqueness.

4. New language names are not introduced unless none of the existing names is acceptable for some reason   There seems to be a tendency among fieldworkers to use new English names for languages that were previously known by some other name. For example, we find statements such as the following:

> Wubuy (previously known as Nunggubuyu (Heath 1984)) is an Australian language of the Gunwinyguan family. (Baker et al. 2010: 65)

> Anywa…is a Western Nilotic language spoken on both sides of the Ethiopian-Sudanese border…In most of the written literature, the name of both the people and the language is spelled Anywak, Anuak, Anyuak…(Reh 1996: 3)

It may be that the new name appears more appropriate to the fieldworker for some reason, but in order to be adopted more widely, a justification is required. Thus, Stenzel (2013: 1) writes:

> 'Wanano' is a name given by unknown outsiders and its use has been called into question by village leaders and the directors, teachers and students of the indigenous school. In 2006, the group publicly adopted the policy of using exclusively their own traditional name Kotiria 'water people' to refer to themselves and to their language and have requested that the outsiders working with them do the same.

When such a justification is given, Glottolog will normally adopt the new usage, but otherwise the principle of conservatism prevails. Clearly, the introduction of additional names may lead to confusion and may thus reduce the visibility of the language community, so this disadvantage must be outweighed by other advantages.

Even though this paper is primarily concerned with the use of language names by linguists and more specifically in Glottolog, it should perhaps be noted briefly that the principles of §3 (favoring a unique name) and §4 (favoring a stable name) are also beneficial for the languages (and thus by implication for the speakers). Languages that do not have a unique name or whose name changes frequently are less easy to remember or find (e.g. on the internet), so by changing a name or using multiple names, linguists run the risk of contributing to confusion and lack of visibility of a language (and thus of the speaker community).

5. Language names that many speakers object to should not be used   This is the best-known principle of language naming, and many languages have been renamed on its basis. Thus, the language formerly known as "Galla" is now called *Oromo*, the language formerly known as "Lappish" is now called *Saami*, and so on.

While this principle is not controversial among linguists, there may be doubts about how to apply it, because different speakers of the language may have different views about what to call their language in English. Language documenters sometimes operate with a vague concept of "speaker community", but unless the community is organized in a formal way that is widely recognized, reference to the community's views must be interpreted with caution.

It should also be noted that this principle applies only to cases where speakers have an opinion on what outsiders, and here specifically English speakers, call them. It is quite possible that speakers of languages in Russia or China or Indonesia have no particular views on what their name should be in English, e.g. whether the Lezgian language (*lezginskij* in Russian, *lezgi* in Lezgian) should be called *Lezgian* or *Lezgin* or *Lezgi* in English. But of course, if the speakers reject a certain name in Russian or Indonesian for their language, then a close counterpart of the name in English should be avoided as well.

6. Language names in English are written with ordinary English letters, plus some other well-known letters   Since language names are English words, they are generally written with English letters. Ideally, this would be the set of 26 basic Latin characters (from *a* through *z*), but special letters from French (*à, â, é, è, ê, î, ô, ù, û, ç*), from Spanish (*ñ, á, í, ó, ú*), from Portuguese (*ã, õ*), and from German (*ä, ö, ü*) are acceptable as well, in view of the cultural closeness of these languages to English, and the fairly widespread knowledge of some basics of these languages among linguists and other educated speakers of English.[4]

This means that special characters from IPA or other special characters are not acceptable as part of language names. Thus, a language name may not contain a

---

[4]It is an interesting question how we know how language names are pronounced. In English, there is generally a rather loose relationship between spelling and pronunciation, which is also true of language names. Ideally, Glottolog would also contain a pronunciation hint for language names that might present problems. (The situation is different with languages such as Lithuanian or Japanese where the spelling clearly indicates the pronunciation; in such languages, the requirements for the kinds of letters that can be used in language names are much more stringent than in English, i.e. it is even less possible to approximate the autoglottonym).

schwa symbol (ə), or special Vietnamese letters (e.g. *ư*), or a dot below or above a vowel, or a superscript letter, or a macron or an underline, or a digit (such as 7), or a tilde on a letter other than an *a* and an *o*.

This principle is not uncommonly violated by linguists. Table 1 gives a few names of languages for which linguists have used names containing a special letter but where Glottolog adopts a name with a corresponding ordinary letter.

Table 1. Some languages for which names with special letters have been used

Glottolog name	A linguist's name	Reference
Bembele	*Bəmbələ*	Djomeni (2014)
Shixing	*Shĭxīng*	Chirkova (2009)
Muong	*Mừòng*	Barker (1968)
Omaha	*Umoⁿhoⁿ*	Eschenberg (2005)
Newar	*Newār*	Hale (2011)
Lomongo	*Lɔmɔ́ngɔ*	Hulstaert (1961)
Oko	*Ọkọ*	Atoyebi (2009)
Even	*Éven*	Pakendorf (2009)
Squamish	*Sk̲w̲xwú7mesh*	Gillon (2013)

The restriction to ordinary English letters is of eminent practical value, because we want to use language names in the titles of papers and books, and many bibliographical systems are restricted to the ordinary letters. Thus, if one wants to buy Djomeni's (2014) book in an online bookshop, one has to look for "A grammar of Bembele", which means that in effect the language would have two names (which contradicts the principle in §3).

Of course, the reason why the linguists in Table 1 used these unusual letters is that they were aiming for name forms that are very close to the autoglottonym, both phonetically and orthographically. But we saw earlier that such closeness is not a requirement for major languages, so it is not a requirement for non-major languages either (cf. §2).

A related issue is the use of the apostrophe in language names. The apostrophe is of course widely used in English words (e.g. *Mary's, didn't, o'clock*), and also sometimes in names (e.g. *O'Hara*), but it does not have a clear interpretation, and is perhaps most widely interpreted as a sign of omission (because there is a full form *did not* corresponding to *didn't*). In language names, by contrast, it is often used in rather different ways:

 (i) between vowels, perhaps indicating a glottal stop, or at least the beginning of a new syllable (e.g. *Waima'a, Ida'an, Kwara'ae*)

 (ii) after a consonant, apparently indicating a special quality of the consonant (e.g. *Yup'ik, K'abeena, Tz'utujil*)

(iii) after a vowel at the end of the language name, with an unclear interpretation (*Wari', Popti', K'ekchi'*)

These three uses are accepted by Glottolog, though somewhat reluctantly. Especially the second and third uses do not serve a useful purpose for English readers, as the nature of the special quality varies widely anyway and cannot even be reproduced reliably by most linguists. One big disadvantage of the apostrophe is that it is easily overlooked and may thus lead to confusion; another (little known) disadvantage is that there are actually a range of different characters that look very much like the apostrophe and that can be easily confused with it.[5] A third disadvantage is that the apostrophe is interpreted as a word separator by many computer programs. Linguists who find the use of non-English letters in language names acceptable might even want to argue that not the apostrophe, but some other apostrophe-like character should be used in language names (e.g. prime when it means palatalization, half ring when it means glottal stop, and so on), thus creating even more confusion.

There is a fourth and a fifth type of context where the apostrophe sometimes occurs:

(iv) before a consonant (*Ga'dang, Ma'di, Mi'kmaq, C'lela*)

 (v) at the beginning of a language name before a vowel (*'Are'are, 'Auhelawa*)

The fifth type is unacceptable in Glottolog, as the apostrophe gets lost here too easily and searching a language name with an initial glottal stop is too difficult. The fourth type can perhaps also be eliminated; as in (ii) and (iii), it is not possible to interpret it in English and it looks decidedly odd.

7. Highly unusual pronunciation values of English letters are not acceptable   This principle means, for example, that the letter *v* cannot be used to represent a vowel ("Mvskoke" for *Muscogee*, "Rvwang" for *Rawang*), the letter *q* cannot be used for a pharyngeal fricative ("Qafar" for *Afar*), and the letter *x* cannot be used to represent [x] ("Abxaz" for *Abkhaz*).

The English letter y is often pronounced [j], but when it does not occur in front of a vowel, this pronunciation value is not available in English. Thus, a name such as *Anywa* (to render an autonym *āɲwáa*, cf. Reh 1996: 3) is only marginally acceptable, and a name such as *Yidiny* is even worse (Dixon's (1977) original name was *Yidiɲ*, violating the principle of §6, but many people referred to the language as *Yidiny*: more recently Dixon has switched to *Yidiñ*, a fully acceptable variant, cf. Dixon 2015).

Exceptions to the principle are made for language names adopted from Spanish, Portuguese and Chinese, where *x* often has the value of a fricative. The fricative value is widely known from place names such as *México* (pronounced with [x] in Spanish), *Belford Roxo* (pronounced with [ʃ] in Portuguese), and *Xi'an* (pronounced [çian] in Mandarin Chinese).[6]

---

[5]Right single quotation mark (U+2019), modifier letter apostrophe (U+02BC), prime symbol (U+2032), okina (U+02BB), modifier letter half ring (U+02BE), combining comma above (U+0313).

[6]The letter *x* with a special value is also found in the language name *Xhosa*, which is very well established, and whose pronunciation ([kosa]) is widely known.

A marginal case is the use of the letter *q* for the glottal stop. One might argue that [k] is not a worse way to emulate [ʔ] than to omit it, so if one prefers the pronunciation [tokabakita] to [toabaita] (for a language whose autoglottonym is [toʔabaʔita]), one can use the language name *Toqabaqita* (Lichtenberk 2008) instead of *To'aba'ita* (Lichtenberk 1988; though by the principle of §4, the earlier name should not have been abandoned by the author). Another marginal case is the use of the letter *c* for [c] or [tʃ], as in the language name *Camling* (Ebert 1997; but see Rai 2013).

8. Language names must be pronounceable for English speakers   This follows from the first principle: Language names used in an English context are English words, hence they cannot contain segments that are not part of English (let alone tones or special phonation types). When they contain letters that stand for segments that do not exist in English, these are (or at least can be) pronounced with an English value, e.g. *Kazakh* is pronounced [kazak], not (necessarily) [kazax], and *Gavião* can be pronounced [gaviao].

This also means that language names cannot contain click consonants. The language called *Korana* is alternatively sometimes called *!Ora*, but this violates the ordinary letter requirement (§5), and the autoglottonym with the initial click simply cannot be pronounced, even in approximation, by English speakers who do not have special training in phonetics.

Moreover, sound combinations that are unpronounceable in English are best avoided, so *Kutenai* is better than *Ktunaxa* (which also violates the principles of §4 and §6).

As noted at the end of §4, choice of language name sometimes has an indirect negative effect on the community of speakers, as when there is confusion about the name of a language. The same applies to the principles in §6–8, because names that cannot easily be written or pronounced will tend to be avoided (many linguists have probably experienced a situation where a linguist gives a talk about a language but does not mention it in speaking because she or he is unsure how to pronounce it).[7]

9. Language names begin with a capital letter   Like city names, Language names always begin with a capital letter in English. This means that language names such as "rGyalrong", "ut-Ma'in" and "isiZulu" are not acceptable. The name "rGyalrong" is best replaced by the variant form *Gyalrong* (also for the reasons mentioned in §8), "ut-Ma'in" should be replaced by *Ut-Ma'in*, and names of Bantu languages with a prefix such as *isi-* or *chi-* or *ki-* must either have an integrated, capitalized prefix (e.g. *Chichewa, Kikongo*) or leave out the prefix (e.g. *Swahili, Zulu*).

This principle also extends to hyphenated names, which consist of two (or occasionally more) "sub-names", e.g. *Pak-Tong, Lesing-Gelimi*. Hyphens that separate

---

[7]A language name can also be very problematic by being identical to another English word in a potentially confusing way (as with the languages *One* and *Libido*), or by being too short (as with the languages *U* and *E*). In these cases, linguists should find a better name, not limiting themselves to borrowing the autoglottonym. (If a language name has unwanted associations in English, this may likewise be a reason to modify it, e.g. to choose *Kisi* rather than *Kissi*, cf. Childs (1995).)

syllables (as in *Nuu-chah-nulth*) are not really acceptable, because they are not normal in English usage.

10. Language names may have a modifier-head structure   Most language names are atomic and contain just a single root in English, e.g. *Finn(-ish), Japan(-ese), Copt(-ic), Tamil*, or *Nunggubuyu*. There are also language names consisting of two written words, such as *Tukang Besi* or *Guugu Yimidhirr* (occasionally one even finds three words, as in *Kala Lagaw Ya*), which are nevertheless morphologically simple (in English; just as a loanword such as *parachute* is not complex in English, although its French source word is a compound).

But language names may also be internally complex and consist of a proposed modifier and a head word, e.g. *Pite Saami* (Wilbur 2015), *Chiapas Zoque* (Faarlund 2012), *Southeastern Tepehuan* (Willett 1988). Such names are generally used when there is a larger linguistic or ethnic grouping that has a well-established name but if there are multiple mutually unintelligible languages within such a grouping.

In this regard, language names are different from city names, which do not have such internal structure. Language names with internal structure are perhaps not ideal either, but in many cases there are no good alternatives. Sometimes the precise delimitation of languages is not clear (e.g. *Chiapas Zoque*, which uses the well-known Mexican state name Chiapas but which represents just one possible way of dividing Zoquean varieties into languages), and in other cases, the modifier by itself is not sufficient to refer to a language (e.g. *Southeastern Tepehuan*, as well as *Yucatec Maya*, where *Yucatec* could perhaps also refer to the Spanish variety of Yucatán).

Ideally, such language names with modifiers would perhaps be marked as such, and indeed this is done in Ethnologue (where Southeastern Tepehuan is "Tepehuan, Southeastern", using an older convention inherited from the time when alphabetic order was important for finding items in lists). It should be noted that in languages other than English, modifiers may occur in other positions (so Southeastern Tepehuan is *tepehuán del sureste* in Spanish, and *Pite Saami* is *same de Pite* in French).[8]

11. The usage of prominent authors is given substantial weight   When there are multiple variant names of a little-studied language spoken in a non-English-speaking environment, it is best to adopt the usage of the most prominent writer on the language, especially if the writer is far more prominent than all others (this will serve the goal of §3). For example, the Daghestanian language Hinuq was previously called "Ginukh" or "Hinukh", but after Forker's (2013) detailed grammatical description, which dwarfs everything written on the language in English by a huge margin, the new name *Hinuq* is best adopted (as was done by Glottolog, but not yet by Ethnologue). Similarly, Jaminjung was previously known as "Djamindjung", but after Schultze-Berndt's (2000) detailed account, the name Jaminjung has better chances of becoming the generally accepted unique name.

---

[8]When names with modifiers are borrowed into English from languages with postposed modifiers, this may result in names such as *Jola-Fonyi* (cf. French *diola-fogny*, a variety of *diola*), or *Hmong Daw* (a variety of Hmong). If one borrows the names in this way, one loses the internal structure in English.

(This does not mean that Forker and Schultze-Berndt should have changed the language name; by the principle of §4, they should have adopted an existing language name, in order not to create confusion. But now their detailed and widely respected work has created new facts. In other words, when prominent authors change the name of a language, others should follow them, but nobody should change language names.)

12. Not a principle: Closeness of the English name to the autoglottonym Some linguists seem to think that the ideal English name for a language is the closest English counterpart of the autoglottonym. But this cannot be true as a general principle, because many major languages have conventional English names that are quite remote from the autoglottonym (*German, Greek, Finnish, Chinese*, etc.). By the principle of §2, names of non-major languages should not be treated differently from names of major languages, so an English name that is different from the autoglottonym is not per se problematic.[9]

Moreover, many languages do not have an autoglottonym, i.e. the speakers of a language may not have a word for their language in their language itself (see also note 2).

It seems that the choice of language name by some authors can only be explained by an implicit principle that the English name should be identical to the autoglottonym. For example, Gillon's (2013) choice of the name *Skwxwú7mesh* for the language that has long been widely known as Squamish can hardly be explained by a desire of the speakers to abandon the traditional name and use the name *Skwxwú7mesh* in English instead. This can be inferred from the fact that the *Squamish–English Dictionary* was published in 2011, by the Squamish Nation Education Department, surely an authoritative body representing speakers of the language. This case, where a language name is neither writable nor pronounceable in English, would seem to be a particularly striking case of a choice of language name by a linguist that has a negative impact on the language (and thus indirectly on the speakers of the language).

13. Conclusions This paper discussed some principles by which the Glottolog editors are guided, and which are probably more widely useful. They are listed together in Table 2.

Surprisingly, issues of this sort have hardly been discussed earlier. The SIL's *Ethnologue* (Lewis et al. 2015), for example, has no discussion of how it distinguishes between the "main name" of a language and alternates, other than noting that some language names are considered pejorative (cf. §5 above), and there is no discussion

---

[9]On the contrary, one might even say that a language name that sounds more like an older English word is preferable because it makes the language sound more like bigger, more prestigious languages which generally have more English-sounding names, e.g. names including an English suffix. Thus, one might say that *Buginese* is preferred to *Bugis*, because the former includes the suffix *-ese*, which is also found in *Chinese, Japanese, Javanese*, and so on. Thus, Rosén (1977: 16) objects to the name "Ivrit" (which was sometimes used for Israeli Hebrew) on the grounds that it "would place Israeli Hebrew among the non-European speech forms that do not rank as languages of civilization within the European cultural conscience."

Table 2. Eleven principles of language naming

---

1. Language names (like city names) are loanwords not code-switches.
2. Names of non-major languages are not treated differently from names of major languages.
3. Each language has a unique name.
4. New language names are not introduced unless none of the existing names is acceptable for some reason.
5. Language names that many speakers object to should not be used.
6. Language names in English are written with ordinary English letters plus some other well-known letters.
7. Highly unusual pronunciation values of English letters are not acceptable.
8. Language names must be pronounceable for English speakers.
9. Language names begin with a capital letter.
10. Language names may have a modifier-head structure.
11. The usage of prominent authors is given substantial weight.

---

either of what kinds of deviations from normal English words are acceptable and what kinds are not acceptable.

Perhaps the most important point to keep in mind is that the choice of a language name is not primarily a research question. While linguists have the research task to find out what the body part terms or the topicalization constructions of a language are, they do not have the task to find out what the "real name" of the language is. When the current name of a language is not suitable because speakers have strong views against it, or when there are a range of different names in current use, linguists must make a choice, but otherwise there are no reasons to change the name, and in fact good reasons to be conservative.

References

Atoyebi, Joseph Dele. 2010. *A reference grammar of Oko: A West Benue-Congo language of north-central Nigeria*. Köln: Köppe.

Aussant, Émilie (ed.) 2009. La nomination des langues dans l'histoire. *Histoire Épistémologie Langage* 31(2). 5–173.

Baker, Brett & Horrack, Kate & Nordliner, Rachel & Sadler, Louisa. 2010. Putting it all together: Agreement, incorporation, coordination and external possession in Wubuy (Australia). In Butt, Miriam & Holloway King, Tracy (eds.), *Proceedings of the LFG10 Conference*, 64–84. Stanford: CSLI Publications.

Barker, Milton E. 1968. The phonemes of Mường. *Studies in Linguistics* 20. 59–62.

Childs, G. Tucker. 1995. *A grammar of Kisi: A Southern Atlantic language* (Mouton Grammar Library 16). Berlin: Mouton de Gruyter.

Chirkova, Katia. 2009. Shǐxīng, a Sino-Tibetan language of South-West China: A grammatical sketch with two appended texts. *Linguistics of the Tibeto-Burman Area* 32(1). 1–89.

Crowley, Terry. 1998. *An Erromangan (Sye) grammar*. Honolulu: University of Hawai'i Press.

Dixon, R M. W. 1977. *A grammar of Yidiɲ*. Cambridge: Cambridge University Press.

Dixon, R. M. W. 2015. *Edible gender, mother-in-law style, and other grammatical wonders: Studies in Dyirbal, Yidiñ, and Warrgamay*. Oxford: Oxford University Press.

Djomeni, Gabriel Delmon. 2014. *A descriptive grammar of Bɜ̀mbɜlɜ̀*. München: LINCOM.

Ebert, Karen. 1997. *Camling (Chamling)*. (Languages of the World/Materials 103.) München: Lincom.

Eschenberg, Ardis. 2005. *The article system of Umoⁿhoⁿ*. Buffalo: University at Buffalo. (Doctoral dissertation.)

Faarlund, Jan Terje. 2012. *A grammar of Chiapas Zoque*. New York: Oxford University Press.

Forker, Diana. 2013. *A grammar of Hinuq*. Berlin: De Gruyter Mouton.

Gillon, Carrie. 2013. *The semantics of determiners: Domain restriction in Sḵwx̱wú7mesh*. Newcastle: Cambridge Scholars Publishing.

Hale, Austin. 2011. The modalities of Newār mal. *Himalayan Linguistics* 10(1). 127–150.

Heath, Jeffrey. 1984. *Functional grammar of Nunggubuyu*. Atlantic Highlands, NJ: Humanities Press.

Hulstaert, Gustaaf. 1961. *Grammaire du lɔmɔ́ngɔ. Première partie: Phonologie*. Tervuren: Royal Museum for Central Africa.

Lewis, M. Paul & Simons, Gary F. & Fennig, Charles (eds.). 2015. *Ethnologue: Languages of the world*. 18th edn. Dallas: SIL International. http://www.ethnologue.com/.

Lichtenberk, Frantisek. 1988. The pragmatic nature of nominal anaphora in To'aba'ita. *Studies in Language* 12(2). 299–344.

Lichtenberk, Frantisek. 2008. *A grammar of Toqabaqita*. Berlin: Mouton de Gruyter.

Pakendorf, Brigitte. 2009. Intensive contact and the copying of paradigms: An Ėven dialect in contact with Sakha (Yakut). *Journal of Language Contact* 2. 85–110.

Rai, Vishnu Prasad Singh. 2013. *A grammar of Chamling*. Kathmandu: Sahadev Maharjan.

Reh, Mechthild. 1996. *Anywa language: Description and internal reconstructions*. Köln: Köppe.

Rosén, Haiim B. 1977. *Contemporary Hebrew*. The Hague: Mouton.

Schultze-Berndt, Eva. 2000. *Simple and complex verbs in Jaminjung: A study of event categorization in an Australian language*. Nijmegen: Katholieke Universiteit Nijmegen. (Doctoral dissertation.)

Squamish Nation Education Department & University of Washington. 2011. *Sḵwx̱wú7mesh sníchim - xwelíten sníchim: Sḵexwts = Squamish–English dictionary*. Seattle: University of Washington Press.

Stenzel, Kristine. 2013. *Reference grammar of Kotiria (Wanano)*. Lincoln: University of Nebraska Press.

Wilbur, Joshua. 2015. *A grammar of Pite Saami* (Studies in Diversity Linguistics 5). Berlin: Language Science Press. http://langsci-press.org/catalog/book/17.

Willett, Thomas L. 1988. *A reference grammar of Southeastern Tepehuan*. Buffalo: University at Buffalo. (Doctoral dissertation.)

Martin Haspelmath
haspelmath@shh.mpg.de