

Endangered Languages and the Efforts to Preserve Them: Justifications, Issues, Costs

Ernest Davis
davis@cs.nyu.edu

August 17, 2026

It is estimated that there are about 7000 languages spoken in the world currently. It seems likely that a large fraction of these will no longer be spoken within a few decades. Strenuous efforts are being made to preserve or at least record as many of these languages as is possible.¹

Most, though certainly not all, “people like us” — we who consider ourselves well educated, cultured, liberal, enlightened, and culturally sensitive and who value those characteristics in themselves and others — have heard about this impending language die-off, consider it very sad, and consider the activists involved in preservation to be doing God’s work. I myself absolutely share in that point of view.

However, it is not obvious *why* the deaths of endangered languages should be considered tragic or how we can be confident that the benefits of preserving languages outweigh the substantial costs. It seems to me important to examine the arguments that justify our viewpoint. The primary goal of the examination is not bolstering our own convictions; certainly not convincing people who hold the opposite opinion; and absolutely not changing our strongly held opinions if we can’t find strong arguments, since our ethical intuitions on these questions are probably more reliable than our ability to formulate arguments. Rather, the examination is important because there are many questions that arise about what should be done and how it should be done. Hopefully being precise about what we are trying to accomplish will help us reach answer to those questions.

1 Disclaimer

The preservation of endangered language connects both to several different scientific and scholarly disciplines and also to social and political issues that evoke deep feelings. The disciplines include systematic linguistics, historical linguistics, cognitive psychology, history, anthropology, sociology, and to a lesser extent philosophy. The socio-political issues include the history of oppressions and the question of how society should now deal with that history; nationalism; the place of small communities of various kinds in the modern world; and education.

I am neither an activist in language preservation nor at all an expert in any of these areas. Over the years, I have done some haphazard reading, mostly of books and articles addressed to a general audience, and I have spent some time thinking about these issues. All of the information in this essay is drawn from that reading; indeed, large chunks of the essay are direct quotations. The writing

¹(Crystal 2000) lists 16 organizations focused on preserving endangered languages generally, or large categories of endangered languages. The Wikipedia article “Endangered Languages” lists 14. There are many more organizations focused on individual languages.

here is often choppy and sometimes repetitive, because there are a lot of interacting issues, and I haven't found a clean way of organizing the discussion overall. The value of the essay, if any, is that it surveys a range of issues that I have not seen brought together in any single source.

The main sources of my information are (Crystal 2000), (Romaine 2011), and (Krishnan 2026); I recommend these enthusiastically. I have also substantially drawn on (Smith Galer, 2026). (Perlin 2013) is a wonderful book that I recommend whole-heartedly, but as it worked out I don't cite that much here. Other works that have provided information are cited in the reference list.

I've organized the central part of this paper by reviewing, first, the major arguments in favor of preservation efforts; second, the choices of goals to pursue and to prioritize in recording and in preserving languages; third, the costs and difficulties that are entailed.

2 What is gained by preserving an endangered language? What is lost when it dies?

2.1 Linguicide

The clearest case is where a more powerful group tyrannically prohibits use of the language of a less powerful group and brutally punishes it. This has been all too common in history and even now; it includes the suppression of Irish Gaelic and Welsh by the English, the suppression of Native American language in the U.S., the suppression of Kurdish in Turkey, and so on. For the details of how this suppression was carried out and resisted, see particularly Smith Galer (2026), who makes this central to her book. I don't see how any 21st century enlightened person of good will can withhold their sympathy from those who are trying to resist this kind of oppression or argue that they do not deserve support.

2.2 Preserving identity

The community that speaks the language, or some part of it, or individuals, feels strongly that they want to preserve their identity and that their language is a large part of their identity. This is, almost always, a *sine qua non* of for a successful preservation effort and, arguably, in itself a sufficient justification for the community to engage in it and for the outside world to support it. If the community feels that preserving their language is important, then that feeling deserves respect.

2.3 Specific psychological, social, and cognitive gains

Some research studies have shown that among members of disadvantaged communities the study and use of the legacy language has social and psychological benefits; it lowers the rates of suicide, substance abuse, unemployment, and social isolation. Smith Galer discusses this at some length.

In some respects this strengthens the previous argument in that it points to important, quantifiable benefits. However, in any specific case, it is more open to question. This kind of effect surely varies widely across different communities and languages; in the case of any specific community and language, one would have to ask to what extent these relations hold. One also has to question the causal direction: Does the study of the legacy language improve the psychological state, or are people who are already in a better state and more eager to engage in studying the language?

There is also evidence that there are general cognitive advantages associated with being bilingual or multilingual, and especially with having multiple native languages, rather than acquiring additional

languages later in life. However that argument doesn't in itself favor the legacy language over any other second language. Moreover, it's not clear whether the advantages of being bilingual are greater than some of the other activities said to improve cognitive ability and social skills, such as learning a musical instrument, singing in a chorus, extensive recreational reading, getting regular exercise, participating in team sports, or communing with nature. One can't do everything, and these alternatives may well be as effective and easier to achieve.

2.4 Scientific interest

Minority languages often have unusual linguistic features that are of interest to linguists or to non-experts with an interest in languages. For an English speaker, the existence of languages with evidential markers² is astonishing, to take a popular example. If there were only one human language — and, in a certain sense, it is purely a historical accident that that is not the case — our understanding of the possibilities open to human language would be much smaller, and no doubt there would be lots of theories about why language could not be otherwise.

This is a strong but limited argument. It makes a case that the language should be carefully recorded and analyzed but not that it should be preserved as a living language in a community of speakers, which is much more costly (see below). Also, it changes the priorities; it assigns greatest value to languages that have unusual features rather than languages whose communities feel strongly about it.

It may be noted that often what is most distinctive about a particular language or dialect, especially in languages that have close relatives, is the phonology: the collection of basic sounds included and the way that speech is vocalized. This can serve as a shibboleth (quite literally) to mark native speakers, favorably or unfavorably. But it is presumably of less philosophical interest than other linguistic features, and, I would guess, of less interest to the general public, except in settings like dramatic performances.

2.5 Preserving mankind's cultural heritage

Each language is in itself a cultural treasure. It has been developed by the combined efforts of generation of speakers and is one of the great creations of the human mind and spirit. It therefore deserves to be preserved like other great cultural treasure — literature, music, art, architecture — and its loss is a tragedy for our collective cultural heritage, like the loss of a great work of art.

This is an appealing argument, but it is vitiated by the fact that a language is a living thing that inherently changes over time, though usually slowly. The English spoken in Chaucer's time was a cultural treasure, and was as different from any current language as many languages that are the object of preservation, but should we weep for its passing? Should we invest effort in reviving it? (Those are actual questions, not purely rhetorical ones.)

2.6 Preserving specific artistic and cultural creations

There are often artistic or cultural elements — sagas, poems, songs, inscriptions, rites etc. — that use the language that will be lost or seriously weakened if the language disappears. The strength of this argument in any particular case depends, both on how much these cultural elements are valued, and to what extent they can survive translation into another language. If they can survive

²In some language, the grammar of a declarative sentence always indicates how the speaker knows the fact — did they observe it, were they told it, did they infer it etc.

translation, than that may be a more effective way of preserving them than preserving the language, particularly if they would have value for speakers of other languages.

2.7 Languages encode folk knowledge

Endangered languages reflect folk knowledge otherwise unknown. Thurman (2016) makes strong claims for this:

In Samoa, Cox discovered that Polynesian herbal doctors had an extensive nomenclature for endemic diseases and a separate one for those introduced by Europeans. Their sophistication is not unique. The taxonomies of endangered languages often distinguish hundreds more types of flora and fauna than are known to Western science. The Haunóo, a tribe of swidden farmers on Mindoro, an island in the Philippines, have forty expressions for types of soil. In Southeast Asia, forest-dwelling healers have identified the medicinal properties of some sixty-five hundred species. In the nineteen-fifties, drug researchers for Eli Lilly and Company, working on several continents, studied folk remedies for diabetes based on the rosy periwinkle, and isolated an active ingredient—vinblastine—that is used in chemotherapy for Hodgkin’s disease. (The healers who led the researchers to their discoveries never saw any of the profits. Such “bio-prospecting” by pharmaceutical companies is a controversial practice that was largely unregulated until 1993.) Quinine, aspirin, codeine, ipecac, and pseudoephedrine are among the common remedies that, according to Cox and Balick, we owe to ethnobotanists guided and informed by indigenous peoples.

Smith Galer (2026) writes: “We lose far more than a bunch of grammar rules and words when we lose a language. We also lose the wellspring of information that flows from it — its sayings, its means of identifying and perceiving the world, its character. Languages aren’t dictionaries, but encyclopedias, containing entire worlds of often irreplaceable information.” She proceeds to recount three specific examples of useful scientific information that Western scientists had learned from indigenous peoples.

Krishnan (2026) is more skeptical:

We need no exotic examples to see that the folk category of “fish” (in English as in many other languages) — as something that, understandably enough, included whales, dolphins, and porpoises — was genuinely improved upon by the scientific term “mammal” and its associated taxonomies.

Nor is factual error the only thing a language can preserve. Languages also encode values — gender hierarchies baked into pronouns, honorifics that perpetuate caste distinctions — that liberal-minded activists may not be eager to preserve as valuable pieces of linguistic heritage.

This argument clearly only applies to a limited class of languages. If you discover a hitherto unknown tribe speaking an unusual language in an unexplored part of Amazonia or Borneo, then their language might well capture some knowledge about the local natural history as yet unknown to Western science. That’s a lot less likely for a local dialect in a well-travelled part of Europe.

2.8 Preserving a distinctive world view: Whorfianism

The Whorfian theory claims that your native language affects your world-view. Therefore, when a language dies, one way of conceptualizing the world dies with it.

There are in fact two versions of the Whorfian theory, strong and weak. The strong version claims that the differences in world-view are profound. Romaine (2001) expresses this view forcefully.

The extent to which the world view underlying traditional varieties is being replaced or ‘colonized’ by a western conceptual outlook, is a concern. Instead of partially Hawaiinizing *sulufura*, Hawaiians could have chosen the preexisting native compound *kūkae pele*, ‘sulphur, match, brimstone’, literally ‘Pele’s excrement’. The expression fits into an extensive conceptual network derived from *kūkae*, ‘excrement’. ...

Such terms unite native beliefs and traditions and articulate the conceptual unity of the Hawaiian perspective in which people are intimately connected and genealogically related to the natural environment and spiritual world. The weedy creepy grass native to Hawaii called *kūkae pu’a’a*, literally ‘pig dung’, is the plant form of Kamapu’a’a, the pig demigod and suitor of Pele, the goddess of Kilauea volcano on the island of Hawaii. ... These cultural connections are absent from *sulufura*.

If that’s true, then that would give significant support to preserving an endangered language that incorporates a distinctive world view. The problem is, though certainly different languages reflect different beliefs, it is not clear to what degree these kinds of beliefs add up to anything more than shallow differences in conceptual framework. Do the beliefs that Romaine discusses in the passage quoted above amount to a different view of the world that it is important to preserve, or are they just colorful myths that can be told just as well in English? The specific claims made by Edward Sapir and Benjamin Whorf that Native American languages express a radically different view of time and other fundamental categories were discredited long ago.

A major difficulty in answering this question is that it is very hard to know what is meant by a “world view” and even harder to detect difference in world view, except in terms of how speakers express themselves in language, which is a circular argument. Of course, our inability to design experiments to detect a difference in world view or to define “world view” with any precision does not mean that there is no such thing as a world view or that there are no systematic difference between groups of people. So the issue is not entirely resolved, and may never be.

The weak version, sometimes called “neo-Whorfianism” says that your native language sometimes has a measurable influence on cognitive processes. This is experimentally well established but the influences are mostly quite minor. For example: If you give Spanish speakers a list of objects with personal names to memorize, they find it easier to do when the names agree in gender with the object — e.g. an apple (“*manzana*”) named “Patricia” and a bridge (“*punto*”) named Claudio — than with the opposite pairing. A second example: Russian has no word corresponding to the English “blue”; the word “*goluboj*” means light blue, and the word “*siniy*” means dark blue. Experimental subjects were shown a sequence of tableaux consisting of one blue square on top and two blue squares on the bottom and were instructed to see which of the bottom squares were the identical color as the top square. Russian speakers were faster, by 124 milliseconds on average, at carrying out the task when the non-matching pair crossed the boundary between *goluboj* and *siniy*; English speakers showed no such effect.³ However, it is hard to argue that these small differences in cognitive abilities justify the enormous labors involved in preserving a language.

³Two very readable accounts of this kind of study as of a decade ago are (Deutscher 2010) and (McWhorter 2014). Despite the diametrically opposed titles, the two authors end up at very much the same conclusions, though with different emphases.

3 Choices and priorities

3.1 What are *bona fide* languages that are important to record or preserve?

The ideal case of an endangered language that would be important to preserve, in the sense that it checks all of the boxes enumerated above, would be one with unusual or unique linguistic features, that has no close linguistic relatives and little or no contamination from other languages, that evolved naturally in a community with a strong sense of communal identity, that has a distinctive artistic and cultural heritage in which the language plays a central role, and that is now under threat from one or more dominant languages of quite alien cultures. Such languages certainly exist. However, the impulse toward language preservation also extends or could plausibly be extended to languages that are less ideal candidates for preservation, in one respect or another; and one has to consider to what extent it is reasonable to spend resources on preserving these.

There are two main categories here. The first is languages whose history or current situation is different from the ideal case: Invented languages, pidgins and creoles, and extinct languages. The second, much more difficult issue, is sublanguages, dialects, and variants.

I consider here only languages that can be used for general conversation. I omit forms of expression that are more limited, such as mathematical and scientific notations and jargon, programming languages, and non-verbal music. The arguments for preserving those, if any, are mostly quite different.

3.1.1 Invented languages

Is it important to preserve invented languages such as Esperanto, or is it only important to preserve languages that have evolved naturally? Levi (2001) claims that it is intuitively obvious that no one has strong feelings about preserving invented languages; and concludes from that intuition that “we hold [languages] to be valuable only so long as they are the product of a certain kind of history.” Also, most invented languages are very well recorded; indeed they mostly are, completely or largely, *defined* by some canonical lexicon and grammar; so the only possible question would be the preservation of a community of speakers. But it might be unwise to accept Levi’s claim *tout court* without due thought. There are contrary considerations. Any invented language is a significant intellectual creation of the human mind. Esperanto specifically has existed for 140 years, which is short by the standards of natural languages, but not all that short; there are native speakers of Esperanto; the Esperantists have a distinctive culture (Okrent, 2009); and there is a literature of significant size that was originally written in Esperanto. Languages like LogLan (Okrent 2009) and Ithkuil (Foer 2012) are outliers at least in their semantic structure, and thus perhaps deserve the attention of linguists for that reason.

Sign languages such as American Sign Language and Old French Sign Language were, to a large extent, invented by specific individuals who can be identified historically; but even if, hypothetically, medical advances were to eliminate deafness and hearing disabilities⁴ one might well hesitate before deciding that the disappearance of these languages would be no matter of regret.

Language activists engaged in preserving or reviving a language are often, to some extent, inventing a new language (Romaine 2011). This happens particularly when a standardized version of a language is constructed by merging dialects such as the construction of N’ko by unifying a number of West African Manding languages (Perlin 2024) or the construction of Rumantsch Grischun (RG) as a unified form of the five dialects of Romansh (Akam 2025; discussed further below). In a different

⁴Leaving aside the fraught question of whether that would be desirable if it were medically feasible.

direction, Speaker of Modern Hebrew in 1920 were speaking a language large parts of which had been recently invented by Eliezer ben Yehuda and collaborators.

3.1.2 Pidgins and Creoles

Pidgins and creoles span a wide spectrum. At one end is a language like Haitian Creole. Haitian Creole emerged 300 to 400 years ago from a combination of French and some of the Atlantic-Congo languages with admixtures of several other languages but it is entirely distinct from any other language. The vocabulary is rich and (as languages go) stable; the grammar is sophisticated and stable. It has 13 million native speakers, including the vast majority of Haiti. There is a large literature originally written in the language. The language is thoroughly documented with a large online presence; the Wikipedia website has 72,000 articles. The spoken language is not, at all, endangered; but if it were it would unquestionably be a strong candidate for a large preservation effort.

At the other end of the spectrum are pidgins that have only existed for a few years, with a limited and unstable vocabulary that extends to only a limited subject matter (e.g. trade), with an unfixed and elementary grammar, with a small number of speakers and no native speakers. Pidgins that don't die off soon often evolve into creoles, but sometimes, as in the case of Fanagalo (a lingua franca in South Africa and some other nearby countries), they become so-called "expanded" pidgin languages — richer and more stable than an emerging pidgin, but less structured than a true creole, and with few or no native speakers.

Linguists generally distinguish creoles from pidgins in that creoles have a consistent grammar, a large stable vocabulary, and native speakers (obviously these criteria are somewhat vague and do not align perfectly), and one might propose that creoles, thus characterized, are worth preserving while pidgins can be let go. But it would be hard to argue for drawing a line at that point in terms of the justifications for language preservations that were enumerated in section 2.

3.1.3 Extinct languages

If there are no living speakers left, is it worthwhile trying to revive the spoken language from records that have been made?

Thurman (2015) features an admiring portrait of Joubert Janten Gomez, then twenty-one years old, who had taught himself to speak Selk'nam, and is trying to keep knowledge of the language alive. Selk'nam was an indigenous language in Chile, and the language of Gomez's maternal ancestor. The last native speaker had died forty years before, and, judging from Thurman's account, there does not seem to be any significant surviving community that identifies with the language. Is Gomez's project admirable and worthy of support, or merely Quixotic antiquarianism?

The Cornish language had become extinct, or nearly extinct, by the end of the 19th century, but was revived largely due to the efforts of Henry Jenner at the start of the 20th century. Cornish now has about 500 fluent speakers, with a small number of people who speak it as a first language. It is a recognized minority language in the U.K. There are novels, children's books, songs, and films being created in Cornish.

At an extreme end, there is a Living Latin movement, which advocates the revival of Latin as an active, spoken language (Djinis 2023).

Arnold Toynbee (1920) wrote contemptuously of such efforts:

When the spirit of archaism is moved to express itself in the field of language and literature, the supreme *tour de force* to which it can address itself is to bring a dead

language back to life by putting it back into circulation as a living vernacular; and such an attempt is being made today in several parts of our Westernized World. The impulse toward this perverse undertaking has come from the nationalistic craze for distinctiveness and cultural self-sufficiency.

Toynbee named “the Norwegians, the Irish, the Ottoman Turks, the Greeks, and the Zionist Jews” as nations “engaged in producing a distinctive national language of their own by putting back into circulation some language which has long ceased to be current in any but an academic sphere,” and noted that “none of them is a chip of the original block of Western Christianity”.

Though 2026 readers will mostly entirely reject Toynbee’s Eurocentrism, and his cocksure sense of cultural superiority, the general question of the value of reviving an extinct language remains knotty.

3.1.4 Closely related languages, dialects, variants

Finally, there is the question of how finely perservationists should discriminate languages and dialects and how narrow and specific are the language variants / sublanguages / dialects that should be preserved. This is often the most difficult question of “language legitimacy” to be decided, and it often gives rise to exceptionally bitter conflicts in the speaker community (further discussed in section 5.2). A few examples:

Romansh has been one of the four official languages of Switzerland since 1938. In 2019 there were about 40,000 Swiss residents who spoke it as their primary language. Romansh is divided into five dialects: Sursilvan, with more than 50% of the speakers; and Vallader, Puter, Surmiran, and Sutsilvan with the rest. The dialects each has a standardized written form and are not always mutually comprehensible; speakers of different dialects often prefer to use Swiss German when speaking to one another. A standardized form, Rumantsch Grischun (RG), which combined these but was largely based on Sursilvan, was introduced in 1982, but was met with hostility by members of the smaller dialects (Akam 2025). Romansh as a whole is considered an endangered language; *a fortiori*, each of the five dialects individually is. How should preservation efforts be divided between RG and the individual dialects?

The linguistic complexities in Switzerland pale beside those in Italy. Smith Galer devotes a chapter to the local languages and dialects of Italy, centered around her personal engagement with her mother’s and grandmother’s native language, Piacentine. Piacentine spoken in the province of Piacenza in northwest Italy. According to Linguasphere Observatory, Piacentine is one of eleven dialects within the Emilian family of dialects⁵. The entire Emilian family is considered “definitely endangered.”

The Wikipedia article “Languages of Italy” lists five languages spoken in Italy as “severely endangered”, twenty-one as “definitely endangered”, and seven as “vulnerable”. Note that that enumeration counts Emilian as a *single* language, not as an umbrella term for eleven separate dialects; presumably some or most of the other languages in the list likewise consists of multiple dialects. All of these are Indo-European and all but a few are basically Romance languages. Where does one even start? Italy legally recognizes twelve minority languages: Franch, Occitan, Franco-Provençal, German, Ladin, Friulian, Slovene, Sardinian, Catalan, Arberesh, Greek, and Croatian. Quite a few additional local languages have some kind of official recognition at the level of provincial law; but not Emilian or any of its dialects. Smith Galer is indignant about that; does she do well to be angry? (I am asking seriously, not rhetorically.)

In the Ethnologue⁶ online catalogue of languages, Sursilvan and Piacentino are considered dialects

⁵Wikipedia article, “Emilian language”

⁶<https://www.ethnologue.com/>

of Romansh and Emilian respectively, rather than full-fledged languages. If these and comparable dialects are considered as languages, then the standard estimate of 7000 living languages might be a substantial underestimate.

Different considerations apply to Ebonics or African-American Vernacular English, the sociolect spoken by many African-Americans. It has linguistic features that are stable and noticeably different from Standard American English, including a number of features derived from African languages. It has been extensively studied and recorded by linguists. It is not endangered as a spoken language, but if it were, would it be important to preserve it?

Is it important to preserve Appalachian English? or the Liverpoolian accent? Should we mourn the disappearance of Victorian London Cockney?

From a mathematician's standpoint, the fundamental problem can be viewed as one of individuation. That is, whether two dialects are "the same language" is exceptionally indeterminate; much more so than, for instance, the problem of individuating biological species. The standard criteria that two languages are the same if they are mutually comprehensible does not work, because it is not a transitive relation; that is, if speaker B understands both speakers A and C, it is not necessarily the case that A can understand C.

3.2 Bilingualism with a dominant language

Should children being taught the legacy language be brought up bilingual in the dominant language? Or is it preferable that they should be monolingual, at least at first, and learn the dominant language only as a second language?

Crystal (2000), who to my mind generally speaks with the voice of practical good sense, argues that minority groups should also learn the dominant language. Not to do so puts them at a substantial disadvantage in terms of their education, their employment, and their social status. Besides, as mentioned, being bilingual has general cognitive advantages. Fishman (2000), who characterizes his approach to language preservation in the phrase "Reversing Language Shift" argues that the goals of language preservation will be best advanced if speakers speak only the legacy language natively.⁷

3.3 Language police

Whether one says 'Among men' by saying *inter homines* or by saying *inter hominibus* does not affect the person considering things rather than signs. In the same way, what else is a barbarism except a word pronounced with letters or sounds different from those which those who spoke Latin before us were accustomed to use? Whether *ignoscere* is spoken with a long or short third syllable makes little difference to a man asking God to forgive his sins, in whatever way he can pronounce the word. What then is integrity of expression, except the preservation of the customs of others, confirmed by the authority of ancient speakers?

The more men are offended by these things, the weaker they are. And they are weaker in that they wish to seem learned, not in the knowledge of things, by which we are truly instructed, but in the knowledge of signs, in which it is very difficult not to be proud.

— Augustine, *On Christian Doctrine*. Trans. D.W. Robertson, Jr.

Should language preservations be purists, opposing change, particularly contamination by dominant languages? Or, once the language has been rescued, should preservationists step back contentedly and watch the language evolve as the spirit moves it?

⁷I may be misreading Fishman. I find his jargon difficult to get through.

There are arguments on both sides. In favor of letting the language evolve, it may be argued that language change is inevitable, natural, and healthy; only dead languages remain static. Organizations like the Académie Française or the Academy for the Hebrew Language, which try to preserve the historic characteristics of the language and in particular to prevent the language from being flooded by English loan words, are widely scorned as over-educated elitists, on the wrong side both of linguistic science and of history. Besides, if the preservation effort succeeds in creating a large language group, then the efforts of the purists are probably doomed to failure. Certainly Modern Hebrew, which is in some ways the most successful attempt at reviving a language, has changed rapidly and, in some respects, radically in the last one hundred years, despite the efforts of the Academy. As discussed below (section 5.2) it is indeed often the case that purist preservationists are from educated elites, which can give rise to conflicts.

In favor of the purist approach⁸ one may ask, What is the point of rescuing a legacy language if you then calmly allow it to be turned into some kind of mongrel of the dominant language? All language change or stasis is a result of choices made by humans; what is the rationale for allowing “organic” choices and prohibiting deliberate choices? (As an analogy: no one advocates that human beings or organizations be discouraged from either deliberately changing themselves or deliberately trying to prevent change.) And, in particular, what possible sense does a preference for organic change make in the context of an effort at language preservation, which, in most cases, is a huge, deliberate effort to resist what is happening organically? The purist attitude is no less “historically valid” than the organic approach. As the above quote from Augustine indicates, language police with a snooty attitude about “correctness” go back at least 1600 years, and probably much longer. Like them or not, they are part of history.

A more extreme version is retroactive purism: the claim that preservationists should try to expel “contamination” from other languages such as loan words that has already taken place. For example, the New Zealand commission Te Taura Whiri i te Reo Māori has advocated this for the Māori language, and has pursued it with some success.

3.4 What priority should be given to language preservation among other community needs?

Crystal (2000) recounts one particular debate:

The Academy of Mayan Languages was criticised in 1992 for devoting a great deal of effort to a project to deal with reforestation. Critics argued that an economic programme of this kind was so far removed from language — and even from culture — that it was an inappropriate use of resources. The limited funds should have been focused more directly on language-related activities which met the immediate demand for help. Supporters argued that economic measures of this kind would eventually lead to an increase in the community’s power and prestige, thereby placing its language on a stronger footing. There is plainly truth on both sides.

3.5 Responsibility of the dominant society toward a minority language

What are the moral responsibilities of a majority speaking a dominant language toward a minority speaking an legacy language?

⁸I believe that Fishman (2001) advocates against allowing change, indeed for reversing change where it has occurred. However, again I may be misreading him.

For examples: What are the responsibilities of the US government, or US society at large, or particular US-based institutions (e.g. the universities or the media) toward Native American languages? What are their responsibilities toward Spanish? Toward Mandarin or Cantonese? Toward Vietnamese? Toward Pashto? What are the responsibilities of Canada toward Québécois? What are the responsibilities of the UK toward Welsh? Toward Irish? Toward Cornish? Toward Hindi? What are the moral responsibilities of Russia toward Tatar? Toward Tajik? Toward Chechen? Toward Armenian? Toward German? What are the responsibilities of Japan toward the Ryukyuan languages? Toward Ainu? Toward Korean?

I'm not going to attempt to answer any of those, but I will point out some of the different considerations that might arise.

Does the majority group have a history of trying to suppress the endangered language or of oppressing the minority group?

Is the minority disadvantaged, socially, politically, or economically, relative to the majority?

Do most of the speakers of the minority language currently live within this society e.g. Ainu in Japan? Or is the language the dominant language in some other countries e.g. Spanish in the US? Or are the speakers of the language dispersed among many countries and a majority in none e.g. Romani or Kurdish?

Is this the ancestral home of the minority language community or is this a diaspora community? (I am not sure logically why this would matter, but perhaps it might.)

Can the language be studied elsewhere? It is one thing to suggest to an Italian-American who wants to improve their knowledge of Italian that they might want to spend a year or two in Pisa. It would be quite another thing to suggest to a Uyghur-American that they spend a year or two in Xinjiang.

As a rather extreme (it seems to me) view on this, Smith Galer writes about the fact that the UK school system did not give her the opportunity to study Italian, "In one of their essays on *linguicide* Tove Skutnabb-Kangas and Robert Phillipson talk about this. . . . These linguists do not mince their words,. The lack of 'immigrant minority education in these countries' equates to 'linguistic genocide, as defined by the UN'. They do not blame people like my mum or my nonna for not passing Italian on. They blame our government." Apparently Smith Galer considers, not merely that the UK government has a moral obligation to provide the children of the 170,000 native speakers of Italian in the UK with access to Italian instruction in the public schools, but that if it fails to do so, it is guilty of *linguicide*.

3.6 The moral responsibility of outsiders to advocate in the speaker community for language preservation

Crystal (2000) urges this strongly.

It is crucial that people become more widely aware of this sequence of events [involved in language death], simply because there are hundreds of cases already known where a community, some way along the road toward cultural assimilation . . . has come to hold strongly negative views towards its ethnic language. Its members need to realize the effects of these attitudes on their descendants, who will not be in any position to choose. While therefore affirming the right of the local community to make its own decisions, outsiders still have a valuable role to play in ensuring that it is an informed decision. It must all be very carefully handled, as we are living in an intellectual climate where issues of human rights and self-determination have come to rank highly in any thinking about intervention. But a policy of total non-intervention in indigenous affairs, however

well-intentioned, would be a blatant disregard of history. . . . I see no essential difference in educational programs designed to remove ignorance about, say, water-management or pest-control and awareness-raising designed to remove ignorance about language.

Crystal does admit, however, that the task may sometimes be hopeless:

If an indigenous community believes that all Europeans are inherently evil, acquisitive, or seeking to dominate, or that all proposals about intervention ultimately come from the CIA, then the chances of working on the language as such are distant. The issues are no less problematic when they involve social constraints governing interactions — such as when it is believed that white teachers cannot possibly understand an indigenous situation, or that women cannot be allowed to teach men.

3.7 Defense or retribution against a dominant language

At times, a language group that has the power to do so can work legally or socially against another language that they believe threatens theirs. For example, the hostility to the English language and its speakers in Quebec is well known.

Smith Galer devotes a large fraction of her chapter on Ukrainian to the legal and social suppression of Russian in Ukraine, particularly since the Russian invasion of 2022. This suppression ends up being burdensome even on many patriotic Ukrainians, since many of them, including famously President Zelenskyy, grew up speaking Russian as their first language. Smith Galer's tone here is one of a neutral journalist, so, unlike most of her book, I am not sure what is her attitude toward this. My own feeling, both in this instance and in general, is that, though this attitude is certainly understandable up to a point, if it is taken too far, it becomes unwise and ungenerous.

4 Costs and challenges of recording a language

Crystal (2000) discusses these in detail.

A basic linguistic record of a language involves creating an orthography, if none exists, a lexicon, a grammar, and a corpus of audio or video recordings of speech. This necessarily requires a significant, though not enormous, one-time investment of time, both from the linguist, and from the speakers that they are working with. Either a member of the language community must be trained in linguistic analysis or a trained linguist must learn the language from speakers. The skill set required of an outsider is somewhat off the beaten track of a linguistic education, involving both sensitivity to unexpected characteristics of the language, and the anthropological skills needed to interact with a community whose culture is often very different. Once the information has been collected, a permanent record must be created, either in print or online.

After the language has been recorded, the costs of maintaining it are small, but non-zero. If the language changes, the records must be updated. Printed records, audio and video recordings, and especially online resources all require maintenance by someone who is paying attention or they degrade and vanish.

4.0.1 Regularization

To record a language is, inevitably, to regularize it, to some degree. In the debate between prescriptivism and descriptivism in dictionaries, grammars, etc., descriptivism is always at most an

imperfectly achieved aspiration; there always remains some taint of prescriptivism. The linguist transcribing the speech of native speaker must decide that at times the speaker has switched to a different language; that their choice of wording or details of their pronunciation in a particular instance is a personal aesthetic/stylistic choice rather than a characteristic of the language; that some aspects of their intonation are not part of the language; that some of the sounds or gestures that they make are not speech; that systematic differences between two speakers is purely a matter of their personal idiolect; and so on. There are linguistic guidelines for these kinds of questions, but often in the end, they are difficult matters of taste. If the linguist misjudges them, particularly if they are biased by preconceptions, then the record they are making is flawed.

The same considerations arise, of course, when the language is taught, either formally or informally. The teacher must decide when the student has made a mistake, correct them, and show them the correct way of saying it. They also arise when an orthography is developed. An orthography decides on word boundaries and discretizes sounds into a finite collection of phonemes.

Even in a pure state of nature, or in purely immersive learning — speakers say what they say, hearers interpret what is being said as best they can — although these considerations are much reduced, they do not entirely vanish. Even if you learn a language purely by immersion or you learn a first language, one part of knowing the language inherently entails knowing which aspects of speech are linguistically significant and which are arbitrary.

When there is a deliberate effort to regularize dialects into a standard language, then the dialects become disfavored or lost. As we will discuss (section 5.2) this can lead to bitter conflicts.

5 Costs of preserving the spoken language

Preserving an endangered language as a living, spoken language is, generally, laborious and costly. The costs can be both up-front and ongoing, involving both costs that are quantifiable in hours of labor and dollars and cents, and unquantifiable in terms of creating conflict and distress.

5.1 The costs of teaching

Preserving a language, either as a first or second language, involves devoting some significant amount of school time to teaching the language. Curricula and teaching materials must be created and regularly updated. Teachers must be trained and paid.

If the language is to be learned as a first language, then at least the parents, and preferably other family members, caregivers, and friends, must commit to regularly using it in conversation. Depending on the circumstances, this can be burdensome.

The children who are learning the language also often pay a cost, impossible to quantify and hard even to characterize, but often substantial. I am not going to address the eternal, impossible question of how children ideally should spend their time. I accept that children must spend a large part of their time for many years being educated; that society has a moral obligation to ensure this; that the content that the children learn will be largely chosen by adults, and that those choices will be partly sensible, partly traditional, and partly arbitrary; and that, though school can be made more pleasant, inevitably some substantial fraction of the children will find significant parts of their education unpleasant, boring, or frustrating. This has been the case for a long time, and it seems unlikely to me that we will find a much better, very different, alternative any time soon. The only cost I want to point out here is the fact that that time spent studying a legacy language is, inevitably, time not spent doing something else. The time that a child can spend studying and the effort that they can put into studying are finite, and no child can do or learn everything that would be valuable.

5.1.1 Teaching a language in school distorts it.

According to Romaine, when a legacy language is taught in school as a second language, it often loses distinctive features:

More children are learning Basque, Irish, Hawaiian etc. through the education system than learn it at home as a native language. ... This has an enormous effect on their language use, as well as on the languages themselves. Functional restriction of these new varieties to the classroom has entailed structural loss, simplification, and reduction in variability, change rooted in the transmission process; the most complex and irregular parts of the language requiring the longest time to acquire often weaken or disappear as the result of incomplete or imperfect acquisition.

Romaine goes on to adduce example of studies that have demonstrated this loss as regards Welsh, Basque, French in Ontario, Irish, and Hawaiian. She states as a general rule, “It is impossible to revitalize a seriously eroded language without transforming it in the process, often in unexpected ways with unexpected consequences.”

5.2 Internecine conflict

The saddest part of any book on language preservation is the recounting of the bitter fights within the preservationist communities. Romaine (2011) discusses these at length, with many examples.

Even when the speaker population is very small and the issue is something as minor as spelling, the in-fighting can be fierce. Romaine writes,

Different approaches to locating ‘true’ Cornish created rival revivalist factions, each promoting its own version of the language. The rivalry was both acute and unfriendly, with each faction competing vigorously for the loyalty of prospective learners. In 2004, the installation of a welcome mat in Cornish at the Camborne county offices in southwest Cornwall sparked a heated debate over how to spell ‘welcome’. Although the county government tried to diffuse the tension by installing signs using all the different spellings (e.g. *dynnargh* on the welcome mat outside the county offices, but *dynargh* on another sign inside the building), that did not diffuse the tension.

These kinds of fights are extremely common in the preservationist communities of many different languages.⁹ There can also be conflict within an ethnic community between advocates for different languages. The conflict between Hebraists and Yiddishists in the first half of the twentieth century is notorious.

5.2.1 Language preservation meets class conflict

Often these conflicts pit a population of working-class, native language speakers against educated language ideologues. Quoting Romaine again:

The language of Néo-bretonnants, who have learned Breton mainly at university or in evening classes, is based on standard literary Breton and is distinguished on the one

⁹They are also very common among speakers of invented languages. There have been schisms among the enthusiasts for Esperanto, Loglan, Klingon, Tolkien’s Elvish languages and many others (Okrent 2009).

hand by its puristic tendencies (especially vocabulary) but ironically on the other by marked French influence on syntax. Traditional Breton, however, spoken predominantly by working-class people over 50 living in rural areas, displays opposite tendencies, namely more influence from French in vocabulary, but less in syntax. Consider this encounter between an older Breton speaker and a young Néo-breton student ... When the elder speaker used French place names such as *Cornouaille*, the student supplied her with the Néo-breton equivalent *Kernev*. However, the student's pronunciation of the Breton name is influenced by French, making it in effect sound more French than the native Breton's pronunciation of the French equivalent.

Later in her essay she describes a similar dynamic in the revitalization of Galician (the language spoken in Galicia, Spain):

As Galician is appropriated and revalorized in a different political climate, power can now be exercised through Galician, but the same process of social control is reproduced through language.

The former functional split between Galician as home and family language and Spanish as a public language has been transformed into a new one between two clearly differentiated forms of Galician: old rural traditional and urban middle class. In this context, rural Galician, once stigmatized and despised by those now controlling Neo-Galician, has become a symbol of resistance. Rural speakers have reacted negatively to the new standard Galician broadcast on television, because to them Castilianized, bookish, false, artificial and foreign. ...

The interests of those speaking Galician, Cornish, or other revitalized languages for the sake of identity are not identical with those of native speakers.

5.2.2 Merging dialects into a national language

The Marxist historian Eric Hobsbawm (2012) wrote, "National languages are therefore almost always semi-artificial constructs and occasionally, like modern Hebrew, virtually invented. They are the opposite of what nationalist mythology supposes them to be, namely the primordial foundations of national culture and the matrices of the national mind. They are usually attempts to devise a standardized idiom out of a multiplicity of actually spoken idioms, which are thereafter downgraded to dialects"

Nikhil Krishnan (2026) writes with sympathy for both sides:

When French people today complain about the diminished international prestige of their nation and language, and wince when their teens talk about "e-mail" rather than the approved "courriel," they might remember that their ideal de-Englished French was itself a construction built on the graves of the dozens of languages and regional forms of speech which dominated much of rural France well into the nineteenth century. The historian Eugen Weber wrote influentially of how peasants who spoke languages like Occitan, Breton, or Basque became — or, more honestly, were made — Frenchmen. Weber's account does not play down the violence and humiliation of the process. ...

But Weber's account is also a reminder that the project of linguistic homogenization had a republican, and not merely an administrative, justification. During the Revolutionary period, Abbé Grégoire held that a true republic could not flourish if its citizens could not understand one another; in the Third Republic, the push for universal primary education in French explicitly framed it as the condition for civic participation. Kings may have

wanted a single language for the convenience of their scribes and tax collectors, but republicans wanted it for the sake of the citizens.

The French republican project did, in fact, succeed in creating a public sphere where a farmer from Brittany and a lawyer from Provence might read the same newspaper, serve on the same jury, and vote in the same election, without some Parisian smart-ass claiming not to understand what they were saying. What was erected on the site of that linguistic divide could be something of real value.

6 Dr. Samuel Johnson on the importance of recording languages

The earliest quotation I have found that expresses regret over the death of language and argues the importance to preserve a record is in a letter that Dr. Samuel Johnson wrote in 1766, which Boswell quotes in the *Life of Samuel Johnson*: “I am not very willing that any language should be totally extinguished. The similitude and derivation of languages afford the most indubitable proof of the tradition of nations, and the genealogy of mankind. They add often physical certainty to historical evidence; and often supply the only evidence of ancient migrations, and of the revolutions of ages which left no written monuments behind them.”

Similarly, Boswell, in *Journal of a Tour to the Hebrides*, (September 18, 1773) quotes Johnson as saying, “There is no tracing the connection of ancient nations, but by language; and therefore I am always sorry when any language is lost, because languages are the pedigree of nations.”

7 My personal experience

This paper is not primarily one of advocacy. But since it unquestionably reflects my point of view which is informed by my personal background, perhaps I should describe what that background is.

I am an American Ashkenazi Jew, born in 1956. All of my ancestors that I am aware of were Ashkenazi Jews, and probably all or almost all of them going back several hundred years were Ashkenazi Jews. Six of my eight great-grandparents spoke Yiddish as a native language; the other two were English Jews who spoke no Yiddish. In all likelihood, the large majority of my ancestors going back several hundred years spoke Yiddish as a native language. Many of them would have learned some Hebrew for prayer and for sacred studies, but few of them ever spoke it conversationally. So, if I have a legacy language other than English, it is Yiddish and not Hebrew.

I attended Hebrew Day School from second to eighth grade. Half of the school day was given over to religious study and to the study of the Hebrew language. I continued my study of Hebrew in after-school classes through high school (Davis 2025). Fifty years later, I retain a fairly good passive knowledge of Hebrew and a weak active knowledge. I take satisfaction in that knowledge and have found it a source of pleasure of various kinds over the years. I do not at all regret the many hours I spent studying it.

No one in my school ever spoke Yiddish except a few common words or phrases. No one would have dreamed of suggesting that the school should teach Yiddish. My guess is that, among my ancestors, some fraction would have been pleased to think that their descendant spoke Hebrew, but very few would have been disappointed to learn that he did not speak Yiddish. I do not myself regret that I was not taught Yiddish when I was young.

Of course, Yiddish and Hebrew are exceptional cases. Hebrew was traditionally considered the Holy

Language and has been revitalized as the language of the State of Israel. In contrast Yiddish was not valued as a cultural marker even by those who spoke it until the middle of the nineteenth century. But all languages and all language communities are unique.

8 Bottom Line

At the end of all this discussion, there are only a few things of which I feel at all confident.

1. Deliberate attempts by an external group to suppress a language are immoral and should be resisted.
2. Linguists should work on preserving languages before they vanish.
3. If the community of language speakers wants to try to preserve its language, more power to them.
4. Language activists should try to be realistic in their expectations and generous in their interactions. Conversely, those who interact with language activists should keep in mind that the leaders of successful language preservation efforts are generally fanatics, to some degree, and should make some allowances.

Beyond that, I have no idea what are the answers to any of the questions I've raised here, either in general or in specific cases, or even how one should go about finding answers to the questions.

Bibliography

Simon Akam (2025). "A Very Big Fight over a Very Small Language", *New Yorker*, December 1, 2025.

<https://www.newyorker.com/magazine/2025/12/08/a-very-big-fight-over-a-very-small-language>

David Crystal (2000). *Language Death*. Cambridge University Press.

Ernest Davis (2025). "My Jewish Education in Providence during the 1960s and 1970s," *Rhode Island Jewish Historical Association Notes*, November 2025.

<https://cs.nyu.edu/~davis/personal/MyJewishEducation.pdf>

Guy Deutscher (2010) *Through the Language Glass: Why the World Looks Different in Different Languages*. Metropolitan Boosk

Elizabeth Djinis (2023). "Spoken Latin is Making a Comeback." *Smithsonian Magazine*, February 14, 2023.

<https://www.smithsonianmag.com/history/why-spoken-latin-is-making-a-comeback-180981621/>

Joshua Fishman (2001). "Why is it so hard to save a threatened language?" In Joshua Fishman (ed). *Can Threatened Languages be Saved? Reversing Language Shift, Revisited: A 21st Century Perspective*. Multilingual Matters Ltd

Joshua Foer (2012). "Utopian for Beginners". *New Yorker*, December 16, 2012.

<https://www.newyorker.com/magazine/2012/12/24/utopian-for-beginners>

Nikhil Krishnan (2026). "Sometimes it's OK to Let a Language Die". *New Yorker*, July 20, 2026.

<https://www.newyorker.com/magazine/2026/07/27/how-to-kill-a-language-sophia-smith-galer-book-review>

- Neil Levy (2001). “Why Regret Language Loss”, *Public Affairs Quarterly*, **15**:4 373-384.
- John McWhorter (2014) *The Language Hoax: Why The World Looks the Same in Any Language*. Oxford University Press.
- Arika Okrent (2009) *In the Land of Invented Languages: Esperanto Rock Stars, Klingon Poets, Loglan Lovers, and The Mad Dreamers Who Tried to Build a Perfect Language*. Spiegel and Grau.
- Ross Perlin (2024). *Language City: The Fight to Preserve Mother Tongues in New York City*. Atlantic Monthly Press.
- Suzanne Romaine (2011). “Revitalized Languages as Invented Languages.” In Michael Adams (ed.) *From Elvish to Klingon: Exploring Invented Languages*.
- Sophia Smith Galer (2026). *How to Kill a Language: Power, Resistance, and the Race to Save our Words*. Crown Publishing.
- Judith Thurman (2015). “A Loss for Words”. *New Yorker*, March 23, 2015.
<https://www.newyorker.com/magazine/2015/03/30/a-loss-for-words>