Linguistics — University of Amsterdam

Bachelor Linguistics

University of Amsterdam

Feature economy vs. logical complexity: Predicting the distribution of plosive inventories across the world

Floor van de Leur

August 10, 2022

Supervisors: Dr. Klaas Seinhorst & Prof. Dr. Paul Boersma

Abstract

Every sound in a language can be categorized by a number of articulatory features. These features can be combined in various ways to represent the sound systems of human languages, some of which may be harder to learn than others. A number of tools have been proposed in order to measure the complexity of such systems, among which are feature economy and logical complexity. The feature economy of a phoneme inventory is obtained by dividing the number of phonemes in that language by the total possible number of phonemes – considering all contrasts that the language uses in its phoneme space. Logical complexity is calculated by counting the number of literals in a minimal formula that describes all the phonemes from a language. This thesis examines the plosive inventories of 317 languages from all over the world, in order to determine whether feature economy or logical complexity is a better predictor for the distribution of plosive inventories across languages. When a certain sound is absent in a language, but all features that make up that sound are used in different combinations for other sounds in the language, we call the missing sound a gap. Our analysis shows that more than half of the examined languages have no gaps in their plosive inventories. Furthermore, we conclude that logical complexity is a better predictor for the distribution of plosive inventories than feature economy. This conclusion holds even when we take the relatedness of languages within a family into account.

Contents

CHAPTER 1

Introduction

Language learning has been a hot topic ever since the Tower of Babel divided mankind – or, in more accurate scientific terms, ever since Homo sapiens started speaking millions of years ago[1]. At the end of 2020, the popular language learning platform Duolingo had around 40 million monthly active users from every single country in the world[2]. But while learning a new language has become more accessible than ever, there still is a lot to be discovered about how and why languages evolve. And although it is impossible to go back in time and research long gone languages, we can learn a thing or two by taking a closer look at the languages that are spoken in the world today.

This thesis examines 317 languages from all across the world, many of which are still spoken today, while some others have gone extinct. We employ two different measures to describe their complexity, and we aim to determine which one of those measures most accurately predicts the distribution of plosives in languages across the world.

Chapter 2 explains some of the important terminology used throughout this text and introduces the notions of feature economy and logical complexity. Chapter 3 outlines the research method used in our typological study. Chapter 4 describes the results from the analysis of our 317 languages. In Chapter 5 we draw some careful conclusions and discuss some possible concerns. The two appendices show the results of the study in more detailed graphs and tables.

[1] Balter, M. (2015). Human language may have evolved to help our ancestors make tools. Science (American Association for the Advancement of Science). `https://doi.org/10.1126/science.aaa6332`.

[2] Cindy Blanco (2020). 2020 Duolingo Language Report: Global Overview. `https://blog.duolingo.com/global-language-report-2020`.

CHAPTER 2

Feature economy and logical complexity

2.1 Terminology

Every language consists of *phonemes*, which can be described using a set of articulatory features. For example, we may describe the English sound [b] as a bilabial voiced plosive. The main information source for this thesis, Maddieson (1984), uses these articulatory features to place all phonemes of a language into a grid, with the rows and columns representing the different features. Table 2.1 and 2.2 show the consonant grids for German and Tigre, an Afro-Asiatic language spoken in Eritrea.

To begin with, I will explain some of the terminology that will be used frequently throughout this text. First of all, the *phoneme space* of a language is the size of the complete grid, that is, the product of the number of rows and columns. We can see that German has a phoneme space of 12x9, and Tigre has a phoneme space of 15x11. Note that the grids differ per language: a distinctive feature in one language may not be of any importance in another language. The *phoneme inventory* of a language is the number of feature combinations that the language actually uses; thus, the total number of phonemes in a language. Tigre apparently has a phoneme inventory size of 27, while the size of the German phoneme inventory is 22.

Table 2.1: The German consonant inventory. Adapted from Maddieson (1984).

	bilabial	labio-dental	dental/alveolar	alveolar	palato-alveolar	palatal	velar	uvular	variable place
voiceless aspirated plosive	p^h		t^h				k^h		
voiced plosive	b		d				g		
voiceless sibilant affricate			ts						
voiceless nonsibilant affricate		pf							
voiceless nonsibilant fricative		f					x		h
voiced nonsibilant fricative		v							
voiceless sibilant fricative				s	ʃ				
voiced sibilant fricative				z	ʒ				
voiced nasal	m		n				ŋ		
voiced trill								ʀ	
voiced lateral approximant				l					
voiced central approximant						j			

Table 2.2: The Tigre consonant inventory. Adapted from Maddieson (1984).

	bilabial	labio-dental	dental	alveolar	palato-alveolar	palatal	velar	pharyngeal	glottal	variable place	labial-velar
voiceless plosive			t̪				k		ʔ		
voiced plosive	b		d̪				g				
voiceless ejective stop			t̪’				k’				
voiceless sibilant affricate					tʃ						
voiced sibilant affricate					dʒ						
voiceless sibilant ejective affricate				ts’	tʃ’						
voiceless nonsibilant fricative		f						ħ		h	
voiced nonsibilant fricative								ʕ			
voiceless sibilant fricative				s	ʃ						
voiced sibilant fricative				z	ʒ						
voiceless sibilant ejective fricative				s’							
voiced nasal	m		n̪								
voiced trill				r							
voiced lateral approximant				l							
voiced central approximant						j					w

Our typological research only focuses on the plosives of a language. Table 2.3 shows the plosive grids for German and Tigre, with the glottal stop (ʔ) left out. The reason for this choice is explained in the Method section. We now see that the size of the *plosive space* of Tigre is 6, and the size of its *plosive inventory* is 5. The empty cell in the top left corner is called a *gap*, and because of this gap, we say that Tigre is *irregular*. The plosive space of German is completely filled, so both the plosive space and the plosive inventory have size 6. Because the German language fully employs its plosive space, we say that German is *regular*. Note that we consider only plosives in this research, and therefore, we might refer to the plosive space and plosive inventory of a language by phoneme space and phoneme inventory, respectively.

Table 2.3: The plosive inventories of German (a) and Tigre (b).

p^h	t^h	k^h
b	d	g

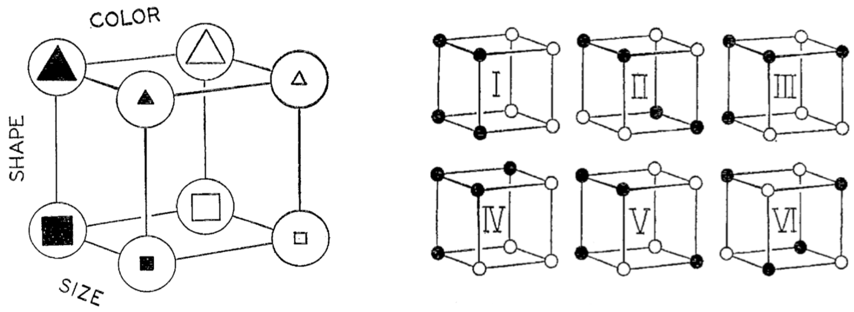
(a)

	t̪	k
b	d̪	g

(b)

2.2 Inventory types

Shepard, Hovland, and Jenkins (1961) conducted a study investigating the learnability of different patterns of feature combinations. Their visual stimuli consisted of three distinct features (shape, size, and color), which could each take on two different values, thus creating a three-dimensional *stimulus space* (Figure 2.1a). Every possible combination of four stimuli belongs to one of six types, as shown in Figure 2.1b. Shepard et al. found that some types are significantly easier to learn than others. The order of difficulty is I < II < [III, IV, V] < VI, where I is the easiest to learn, VI is the hardest, and III, IV, and V have approximately the same difficulty.



(a) (b)

Figure 2.1: (a) The stimulus space that can be constructed from the three binary features. (b) Every combination of four different stimuli can be rotated or reflected to match one of these six types. Reprinted from Shepard, Hovland, and Jenkins (1961).

Seinhorst (2016) conducted a similar experiment to discover the differences in learnability between different phoneme grid types. He used six hand shapes which can be placed in a 2x3 grid, to mimic a basic plosive space such as that of German and Tigre. The stimuli sets consisted of either 3, 4, 5, or 6 stimuli, since most languages with a plosive space of 6 have a plosive inventory of 3 to 6 phonemes (Seinhorst 2017). These combinations of stimuli all belong to one of the eight inventory types shown in Figure 2.2.

In Seinhorst's experiment, each participant saw a subset of the six possible hand shapes, and each hand shape was shown multiple times. After all stimuli were presented, the participants were asked to indicate which hand shapes they had seen. Interestingly, participants never omitted hand shapes that they had indeed seen, but they did incorrectly indicate hand shapes that were not part of the subset shown. Put differently: no participant created an extra gap in the inventory, but almost 15% of the participants filled one or more gaps that were present in their subset. This observation leads us to think that more regular inventory types are easier to learn.

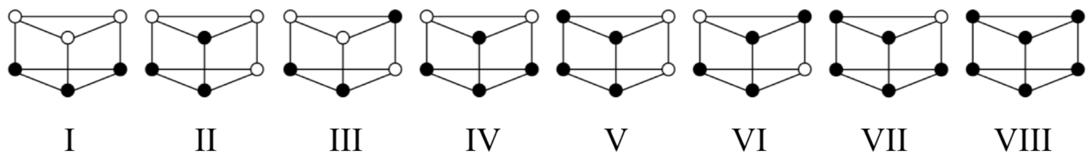


Figure 2.2: Each combination of 3, 4, 5, or 6 stimuli in a 2x3 stimulus space matches one of these eight types. Reprinted from Seinhorst (2016).

2.3 Feature economy

The question arises whether the bias to fill gaps rather than create them will be reflected in the way languages change over time. Martinet (in Clements 2003) showed that gaps indeed tend to disappear, either because they get filled up, or because another phoneme evolves in such a way that the gap is no longer part of the phoneme space.

As opposed to this diachronic approach, our typological research focusses on the synchronic analysis of inventory types. From the results of Shepard et al. and Martinet, we predict that inventory types which are less complex regarding their gaps will be more prevalent in languages across the world. If we want to be able to predict the prevalence of different inventory types, we need a tool to measure their complexity. In the past decades many tools have been proposed: Clements lists for example parsimony, symmetry, pattern congruity, representational economy, and *feature economy*. He argues that the latter seems to be the most accurate measure of complexity.

The principle of feature economy states that “languages tend to maximize the combinatory possibilities of features across the inventory of speech sounds: features used once in a system tend to be used again” (Clements 2003, p. 287). If we apply this to the plosive spaces of German and Tigre in Table 2.3, we see that Tigre has a lower feature economy than German, because Tigre uses the bilabial feature only once. To be able to compare different feature economies accurately, we adopt the following calculation in this thesis: we divide the inventory size of a language by its space size, resulting in a number between 0 and 1, where a higher number represents a higher feature economy. The feature economy of German is then 1.00, while the feature economy of Tigre is $5/6 = 0.83$.

2.4 Logical complexity

Another measure that might be of interest to us is Boolean complexity, which we will refer to as *logical complexity.* Feldman (2000) showed that this measure accurately predicts the outcome of the 1961 Shepard et al. experiment, and was thereby the first to adequately explain those results.

Table 2.4 repeats the plosive spaces of German and Tigre as introduced in Table 2.3, but now with letters indicating the rows and columns. The top left cell is thus denoted by ab and the cell next to it by ab’. The plosive inventory of Tigre can be described by the the formula ab’ + ab” + a’b + a’b’ + a’b”. This is called the *disjunctive normal form* of the Tigre plosive space. This disjunctive normal form can be collapsed into a *minimal formula.* Since Tigre uses all plosives in the second row, we can replace a’b + a’b’ + a’b” with just a’. In the same way, the second and third column can be denoted by b’ and b” respectively. The minimal formula a’ + b’ + b” thus captures all and only those plosives that belong to Tigre’s plosive inventory. Since German uses all cells in its plosive space, we say that its minimal formula is A, which stands for ’all’.

Table 2.4: The plosive inventories of German (a) and Tigre (b), with letters indicating the rows and columns.

	b	**b’**	**b”**
a	p^h	t^h	k^h
a’	b	d	g

(a)

	b	**b’**	**b”**
a		t̪	k
a’	b	d̪	g

(b)

The logical complexity of an inventory type is the number of literals in its minimal formula. Since the minimal formula for the German plosive inventory consists of one literal, A, its logical complexity is 1. The plosive inventory of Tigre has just one gap, but because of that its minimal formula contains three literals, giving Tigre a logical complexity of 3. A higher number thus represents a more complex inventory type.

We now have two possible measures for system complexity, but which one predicts the prevalence of inventory types best? No comparison of these two measures seems to have been carried out in past typological research. Therefore, the main question this thesis seeks to answer is whether feature economy or logical complexity is a better predictor for the prevalence of different inventory types in the languages across the world.

CHAPTER 3

Research method

3.1 The sample

As Maddieson (1984) points out, it is impossible to draw a truly random sample from the population of all languages. First of all, there are many languages in the world about which we have no data, or at most incomplete and inadequate data. Most of these languages are non-western, which makes it almost impossible to get rid of the western bias when drawing a sample. Second, "languages of the world" is not a clearly defined population. Languages are subject to change: new languages emerge, others evolve or go extinct. There is not, and will never be, a satisfactory definition of what precisely a language is.

Because the boundaries of our population are unclear, and a lot of members cannot be part of the sample due to lack of information, there is no good basis for drawing a truly random sample. Our research therefore includes all 317 languages documented in Maddieson, but great care must be taken when drawing conclusions from this sample.

3.2 Analyzing plosive inventories

While there are many different types of phonemes, one thing that all languages in our sample have in common is that they use plosives. Therefore, this research focuses only on plosive inventories. For every language, we listed the features that are necessary to create a grid that fits all plosives of that particular language. We chose to divide these features into four categories:

- The features that make up the laryngeal system of the language: voiced, voiceless, aspirated, laryngealized, prenasalized, et cetera;
- The places of articulation: labial, dental, alveolar, retroflex, palato-alveolar, palatal, velar, and uvular;
- The secondary articulatory features, if distinctive: labialized, palatalized, pharyngealized;
- Any other distinctive features (which turned out to be only the distinction between long and short phonemes).

As stated in the previous chapter, we chose to leave the glottal stop (ʔ) out of our analysis. A glottal stop is produced by a complete closure of the glottis, the space between the vocal folds (Ladefoged and Disner 2012). Because the vocal folds need to be able to vibrate in order to produce a voiced sound, the voiced glottal stop has been judged impossible by the International Phonetic Association[1]. Including that phoneme in the analysis would thus leave many languages

[1] IPA Chart, `http://www.internationalphoneticassociation.org/content/ipa-chart`, available under a Creative Commons Attribution-Sharealike 3.0 Unported License. Copyright © 2018 International Phonetic Association.

Table 3.1: Laryngeal features on a continuum (a) versus in a grid (b).

voiced	voiceless	voiceless aspirated

(a)

	voiced	voiceless
	X	voiceless aspirated

(b)

with a gap that could never be filled. The same goes for pharyngeal plosives. While the voiceless pharyngeal plosive is extremely rare – and in our sample only found in Iraqw – it is physically possible, whereas the voiced pharyngeal plosive is not, according to the International Phonetic Association. We have therefore left the Iraqw pharyngeal voiceless plosive out of the analysis.

While analyzing our sample of languages, we encountered a problem concerning laryngeal systems. If a language has voiced, voiceless and voiceless aspirated plosives, should that be considered a continuum (Table 3.1a) or a two-by-two grid with a gap at the place of the X (Table 3.1b)? This issue arose not only with aspiration, but also with features that are harder to place on a continuum. For example: Selepet, a language of the Indo-Pacific family, has only voiced prenasalized and voiceless aspirated plosives. Treating this system as a continuum does not seem right, since nasalization and aspiration are not two sides of the same coin. Treating the system as a two-by-two grid with two empty cells would imply that this language is highly uneconomical in filling its phoneme space. Since neither of these approaches seems justified, we chose to treat every feature within the laryngeal system of a language as a separate entity, not connected to the other features by a continuum or a grid. This approach provides us with more flexibility while analyzing laryngeal systems.

After determining which features the plosive inventory of a language employs, we listed the size of the plosive space and the size of the inventory. We then calculated both the feature economy and the logical complexity of each language.

3.3 Calculating feature economy

As an example, let us consider the plosive system of Siona, a Tucanoan language spoken in Colombia and Ecuador. Table 3.2 shows the 8 plosives that Siona employs. The laryngeal system consists of two features: voiceless and laryngealized voiceless. Upon closer inspection it becomes clear that the five columns are made up of only four different places of articulation: bilabial, dental, retroflex and velar. The velar column can be divided into "plain" and labialized plosives. Thus, Siona distinguishes two secondary articulatory features. We can now conclude that every place of articulation has space for 2 laryngeal features x 2 secondary features = 4 plosives. Since Siona has 4 places of articulation, its total space size is 4 x 4 = 16, even though Siona has only 8 plosives. This means that Siona has a feature economy of 8 / 16 = 0.50. Table 3.3 shows the complete plosive space of the Siona language.

Table 3.2: The plosive inventory of Siona. Adapted from Maddieson (1984).

	bilabial	dental	retroflex	velar	velar labialized
voiceless plosive	p	t̪		k	k^{w}
laryngealized voiceless plosive	p̰		ʈ̰	k̰	k̰^{w}

Table 3.3: The complete plosive space of Siona.

	bilabial	bilabial labialized	dental	dental labialized	retroflex	retroflex labialized	velar	velar labialized
voiceless plosive	p		t̪				k	kʷ
laryngealized voiceless plosive	p̰				ṭ̰		k̰	k̰ʷ

3.4 Calculating logical complexity

To calculate the logical complexity of the Siona language we place its phonemes in a grid, just as we did with German and Tigre in the previous chapter (Figure 2.4). The main difference between Siona on one hand and German and Tigre on the other, is that the Siona plosive space is three-dimensional rather than two-dimensional. We must accordingly place the Siona plosives in a three-dimensional grid (Figure 3.1). The node containing [p̰] in Figure 3.1 can be denoted by abc, and the node containing [kʷ] by a'''b'c'. The disjunctive normal form of Siona is then abc + abc' + a'bc' + a''bc + a'''bc + a'''b'c + a'''bc' + a'''b'c'. Because Siona uses all its velar plosives, we can replace a'''bc + a'''b'c + a'''bc' + a'''b'c' by simply a'''. Similarly, because Siona uses both bilabial plain plosives, we can replace abc + abc' by ab. This gives the minimal formula ab + a'bc' + a''bc + a'''. The logical complexity of a language is equal to the number of literals (that is, every a, b, and c) in its minimal formula, thus the logical complexity of Siona is 9.

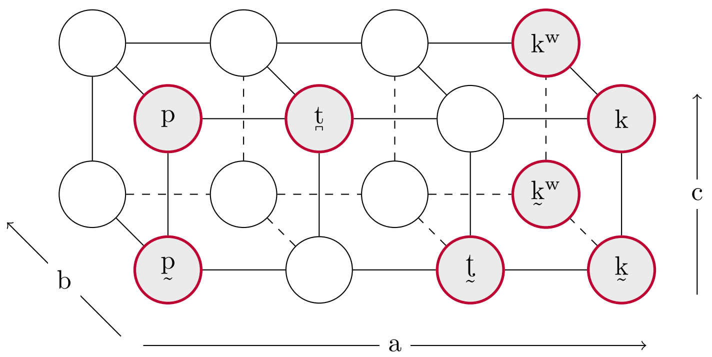


Figure 3.1: A three-dimensional visualization of the Siona plosive system.

Now that we are able to calculate both the feature economy and the logical complexity of every language, we can analyze the suitability of both measures. The next chapter presents the results of this analysis. Appendix A contains the feature economy and logical complexity calculations for all 317 languages in our sample.

CHAPTER 4

Analysis and results

4.1 Regularity

Let us first take a look at the distribution of regular and irregular languages in our sample. As defined in Chapter 2, a regular language fully employs its phoneme space. Thus, every language with a phoneme inventory the same size as its phoneme space is considered regular – note that we measure regularity with regard to plosives only. Out of the 317 languages in our sample, a small majority of 176 languages (55.5%) is regular, as opposed to 141 irregular languages.

Table 4.1 shows all attested combinations of space sizes and inventory sizes, with gray cells representing regular languages. The exceptionless regularity of the languages with some of the smallest plosive spaces (2, 3, and 5) is unavoidable, since a prime number cannot form a grid with more than one row. Put differently, a smaller inventory size would in these cases automatically result in a smaller space size. In contrast, the regularity of languages with space size 4 is not imperative. We expected that most languages with space size 4 would be made up of a two-by-two grid, leaving the possibility for an inventory size of 3, but it turns out that there are no languages with two-by-two grids in our sample. Apparently, languages tend to use all three major articulation places (labial, coronal and dorsal). Only five languages in our entire sample have a plosive inventory with just two places of articulation.

Table 4.1: Each cell contains the number of languages with a certain combination of space size (column) en inventory size (row). A gray cell indicates a regular combination, i.e. a combination where the space size and inventory size are equal.

	2	**3**	**4**	**5**	**6**	**8**	**9**	**10**	**12**	**15**	**16**	**18**	**20**	**24**	**27**	**30**	**32**	**36**	**48**	*Sum*
2	1																			*1*
3		19			1															*20*
4			18		15				1											*34*
5				3	19	3	2		2											*29*
6					83	9	1	1	3		1									*98*
7						10	2	1	6		1									*20*
8						29	6	1	11		1	1								*49*
9							8	4	3		3	2	1							*21*
10								5	2		1	1			1					*10*
11									4		1	1								*6*
12									6	1	2		2	2		1				*14*
13														2						*2*
14										1			1							*2*
15											1									*1*
16											3						1			*4*
17													1						1	*2*
18															1		1			*2*
20																		1		*1*
24														1						*1*
Sum	*1*	*19*	*18*	*3*	*118*	*51*	*19*	*12*	*38*	*2*	*14*	*5*	*5*	*5*	*2*	*1*	*2*	*1*	*1*	

Figure 4.1 shows that bigger spaces also tend to have more phonemes. That seems like a rather obvious observation: a two by two space must have at least two phonemes, while a six by six space has at least six – and most likely more. Figure 4.2 (on the next page) shows that the mean feature economy decreases as the space size increases. A bigger space size thus corresponds to a bigger inventory size, but the correlation is not linear: if two space sizes differ by a factor two, the size of their inventories will on average differ by less than a factor two. The line in Figure 4.1 shows that a space size of 10 corresponds to a mean inventory size of approximately 7.5. Doubling the space size does not double the inventory size: the mean inventory size for a space size of 20 is less than 15.

Whereas the mean feature economy decreases as the space size increases, the logical complexity increases whenever the space size increases (Figure 4.3). Feature economy and logical complexity thus seem to be inversely proportional to each other. The next section will explore the correlation between feature economy and logical complexity further.

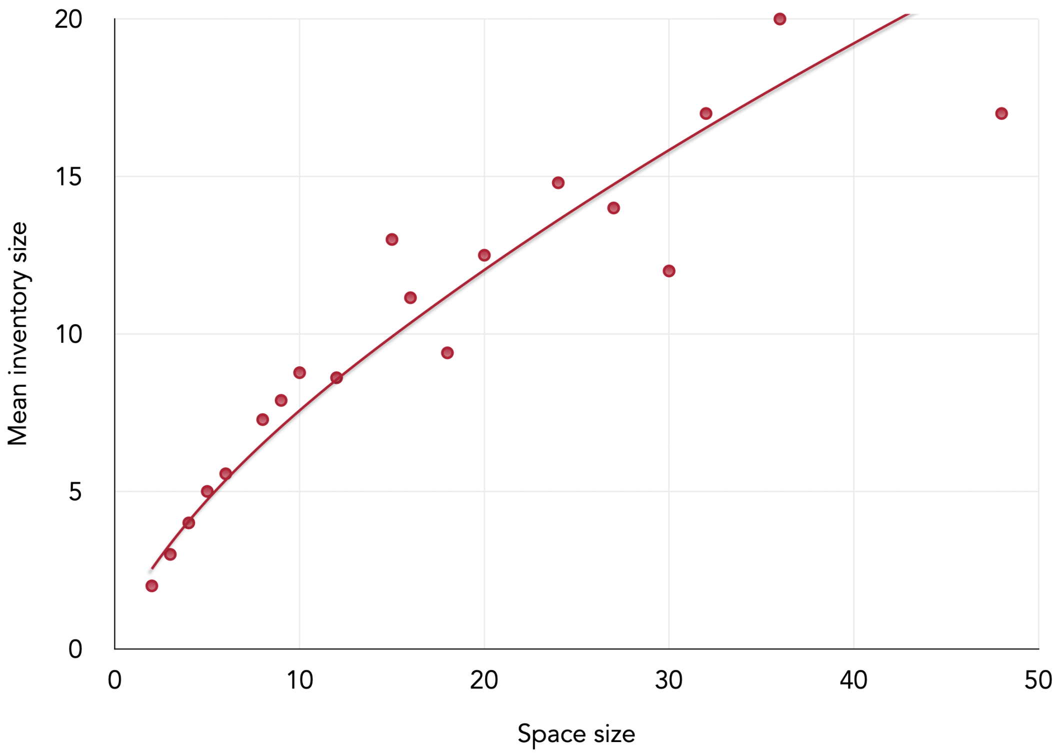


Figure 4.1: Languages with larger plosive spaces generally have larger plosive inventories too.

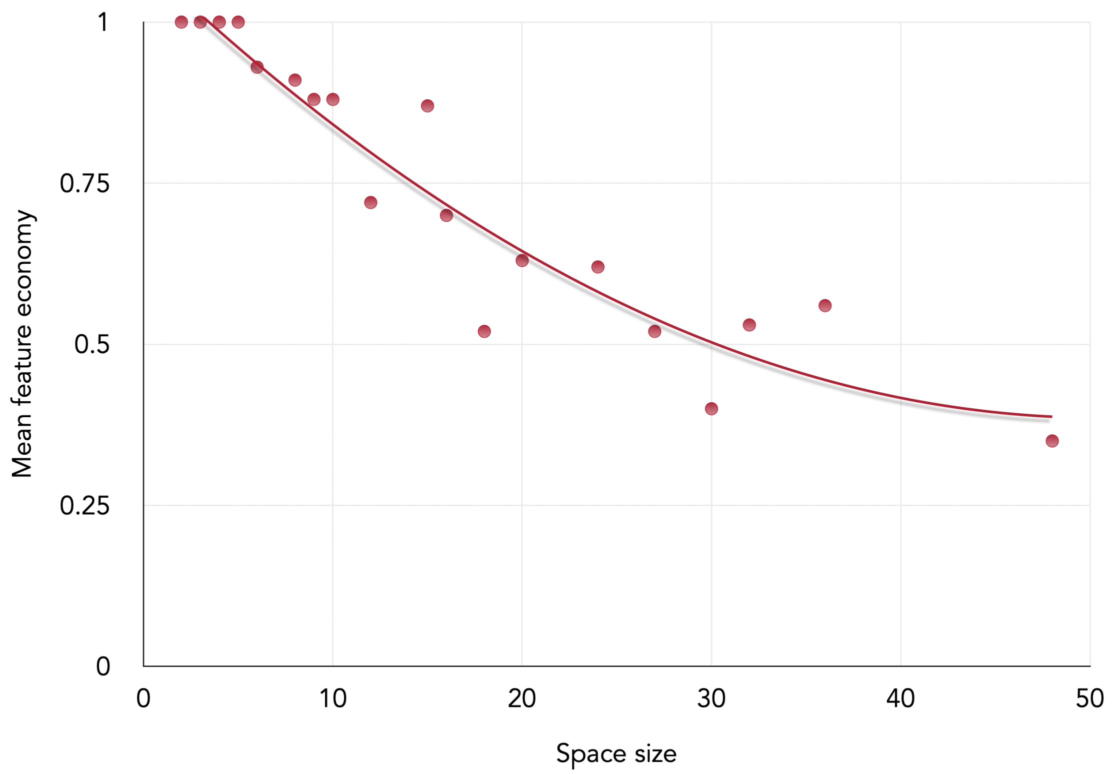


Figure 4.2: When the space size increases, the mean feature economy decreases.

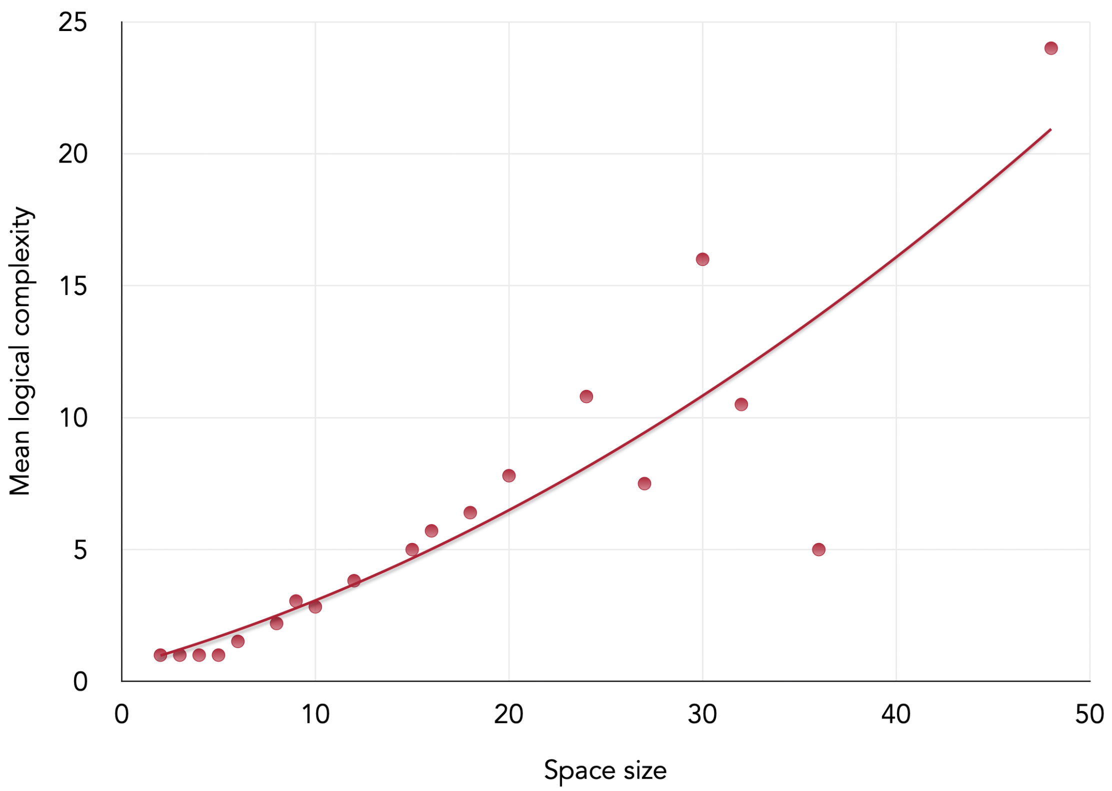


Figure 4.3: When the space size increases, the mean logical complexity increases as well.

4.2 Feature economy versus logical complexity

Figure 4.4 depicts the relationship between feature economy and logical complexity. Every language that fully uses its plosive space has both a feature economy and a logical complexity of 1. We see that a lower feature economy generally corresponds to a higher logical complexity, but the correlation does not seem to be very strong. Consider for example the three languages with the lowest feature economy (< 0.40): their logical complexities range from the highest logical complexity in the sample (24) to a fairly low logical complexity of just 6. Thus, while both measures are correlated, they clearly measure different things.

Let us now consider the frequency of the different feature economies and logical complexities found in the sample. Figure 4.5 (on the next page) plots the frequency of all attested feature economies; Figure 4.6 does the same for logical complexity. Both distributions seem to follow some kind of power law, but a trend line can be fitted much more closely on the distribution of logical complexities ($R^2 = 0.9185$) than on the feature economy distribution ($R^2 = 0.1649$).

Now why would we want our frequency distributions to follow a power law? In other words, why would a measure that closely follows a power law be a better predictor of real-world plosive inventories? It turns out that power laws are found across many fields of study, both in social and natural sciences. To support this claim, let us take a look at some examples.

Perhaps the most well-known example of a power law within the field of linguistics is Zipf's law (Zipf 1949). This empirical law states that word frequencies in natural language utterances follow a power law, where the most frequent word in a corpus is used approximately twice as often as the second most frequent word, three times as often as the third most frequent word, and so on. A more recent observation of a linguistics-related power law is made by Abrams and Strogatz (2003), who showed that the declining number of speakers of endangered languages over the years generally follows a power law.

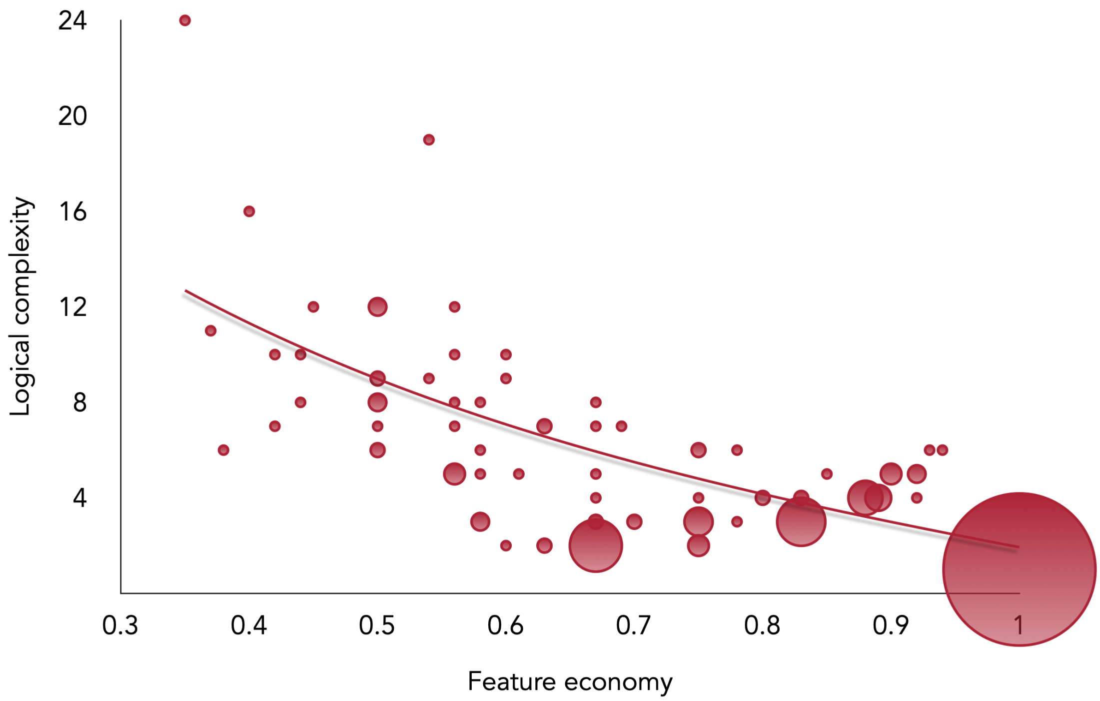


Figure 4.4: The correlation between logical complexity and feature economy. The size of a circle corresponds to the number of languages with that specific combination of feature economy and logical complexity. Generally speaking, a higher feature economy corresponds to a lower logical complexity.

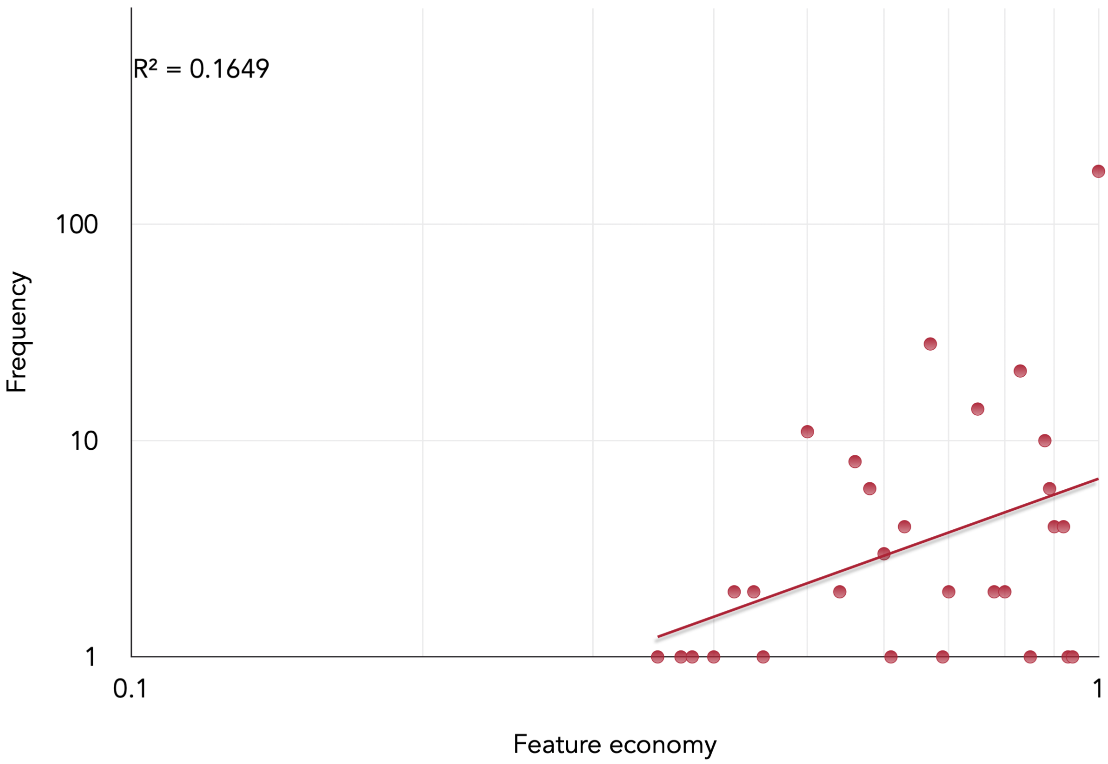


Figure 4.5: The frequency of all feature economies found in the sample, plotted on a horizontally and vertically logarithmic scale.

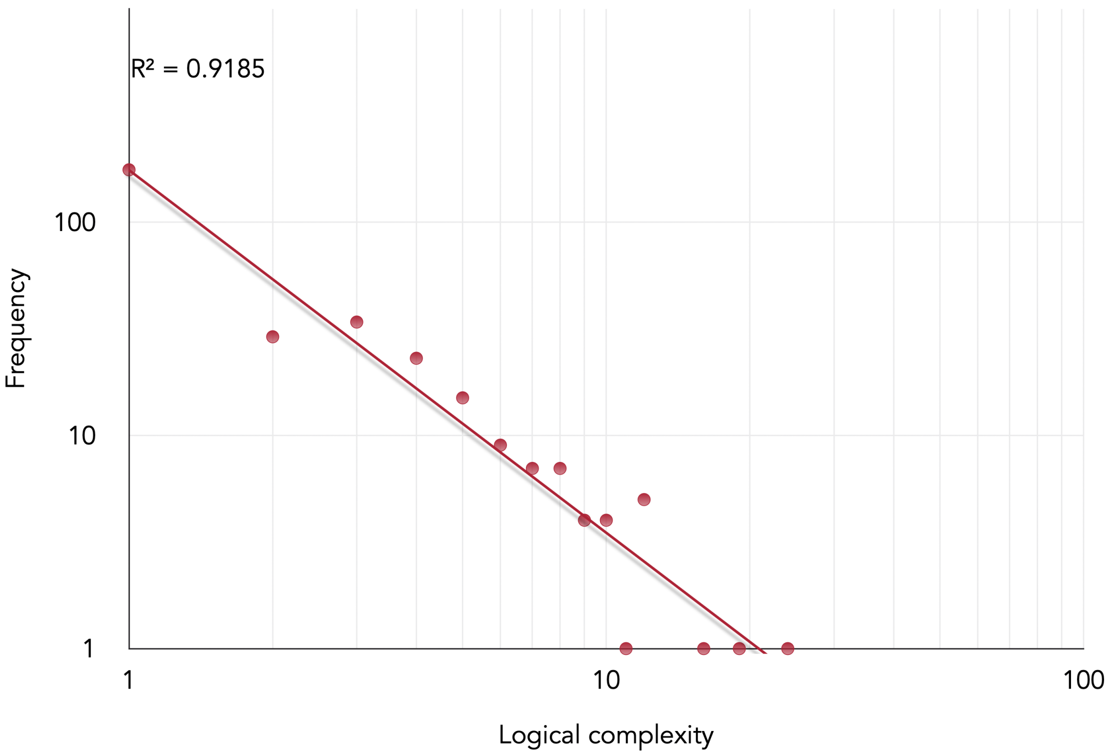


Figure 4.6: The frequency of all logical complexities found in the sample, plotted on a horizontally and vertically logarithmic scale.

Power laws are not only found in the field of linguistics, but in many other social sciences as well. For example, the distribution of wealth in society follows a power law (Levy and Solomon 1997), as does the distribution of firm sizes in the United States (Axtell 2001), of war sizes and casualties (Cederman 2003), of the number of criminal acts committed by individuals (Cook, Ormerod, and Cooper 2004), and many more. The same phenomenon can be observed within the natural sciences, where, among other things, the severity of earthquakes, financial crashes, and premature human births follow power laws (Sornette 2002).

Many aspects of human behavior, human characteristics, and natural phenomena thus seem to follow some sort of power law. While this is no proof that the complexity of plosive inventories should exhibit a similar pattern, the well-fitting power law in the distribution of logical complexities does fit in neatly with the widely observed collection of power law patterns within the social and natural sciences. This leads us to the tentative belief that logical complexity may be a better predictor of real-world plosive distributions than feature economy.

4.3 Language families

Before we draw any conclusions, we must consider another relevant aspect of our data set. The languages from our sample are not isolated. Instead, many languages have evolved from the same ancestor and may have therefore influenced each other widely. As a result, we can divide languages into language families of varying sizes. We adopt the division of languages into families as proposed by Maddieson (1984). Our data set contains three isolates: Ainu, Basque and Burushaski. Northern Amerindian is with 51 languages the biggest family in the data set.

The implication of this relatedness is that languages within a family may have certain features not because those features have some kind of advantage to the speakers, but because those languages have evolved from a mutual ancestor with that feature. We must take relatedness into account before we can conclude which of our measures most accurately predicts the direction in which languages may evolve.

Appendix B contains a graph for each of the twenty language families in our sample. Every graph shows the distribution of feature economy and logical complexity in that family. Figures 4.7 and 4.8 (on the next page) contain the graphs for the two largest language families in our sample: Northern Amerindian and Southern Amerindian, respectively. Of the Northern Amerindian languages, 29.4% is regular regarding their plosive inventories; 54.1% of the Southern Amerindian languages is regular. Recall that regular plosive inventories are found in 55.5% of the languages in our sample. Looking at the graphs in Appendix B, we can see that every family with more than two languages contains both regular and irregular languages; their irregular languages have various complexity values, regardless of whether we measure complexity by feature economy or by logical complexity; the Paleo-Siberian family has the lowest concentration of regular languages (25.0%), and the highest concentration of regular languages (89.5%) is found within the Australian languages.

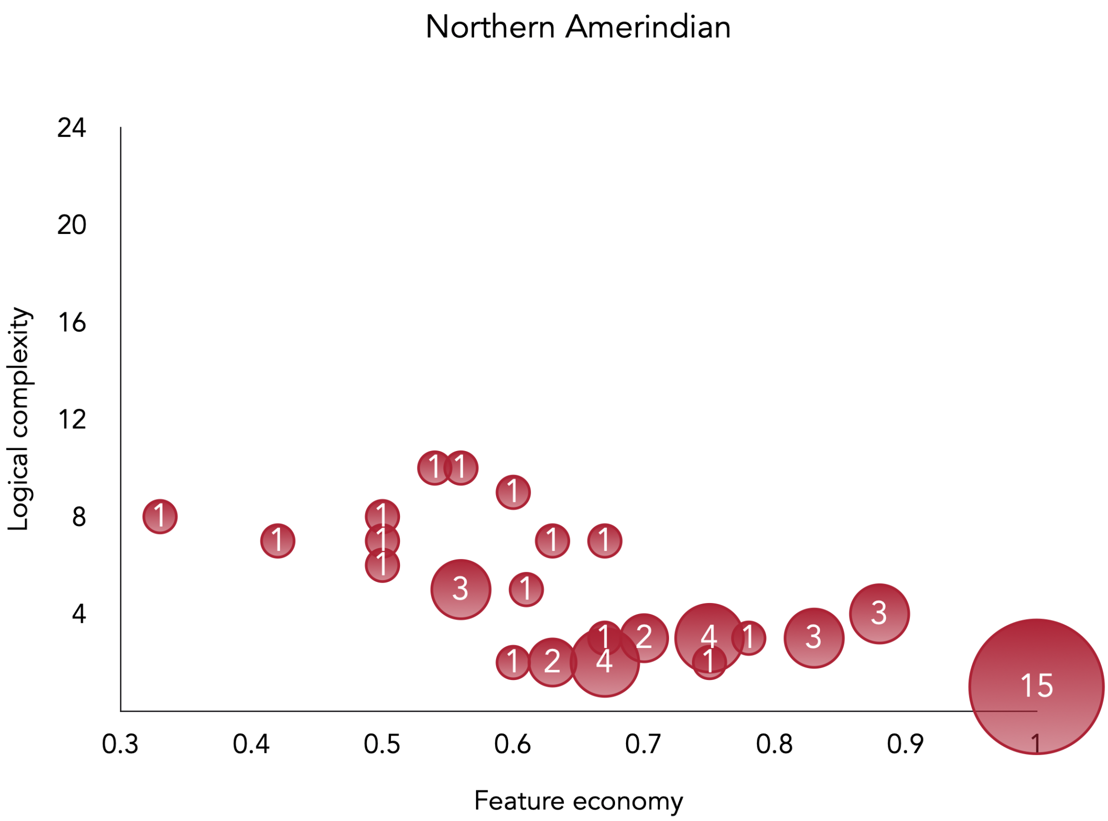


Figure 4.7: The distribution of feature economy and logical complexity in Northern Amerindian languages.

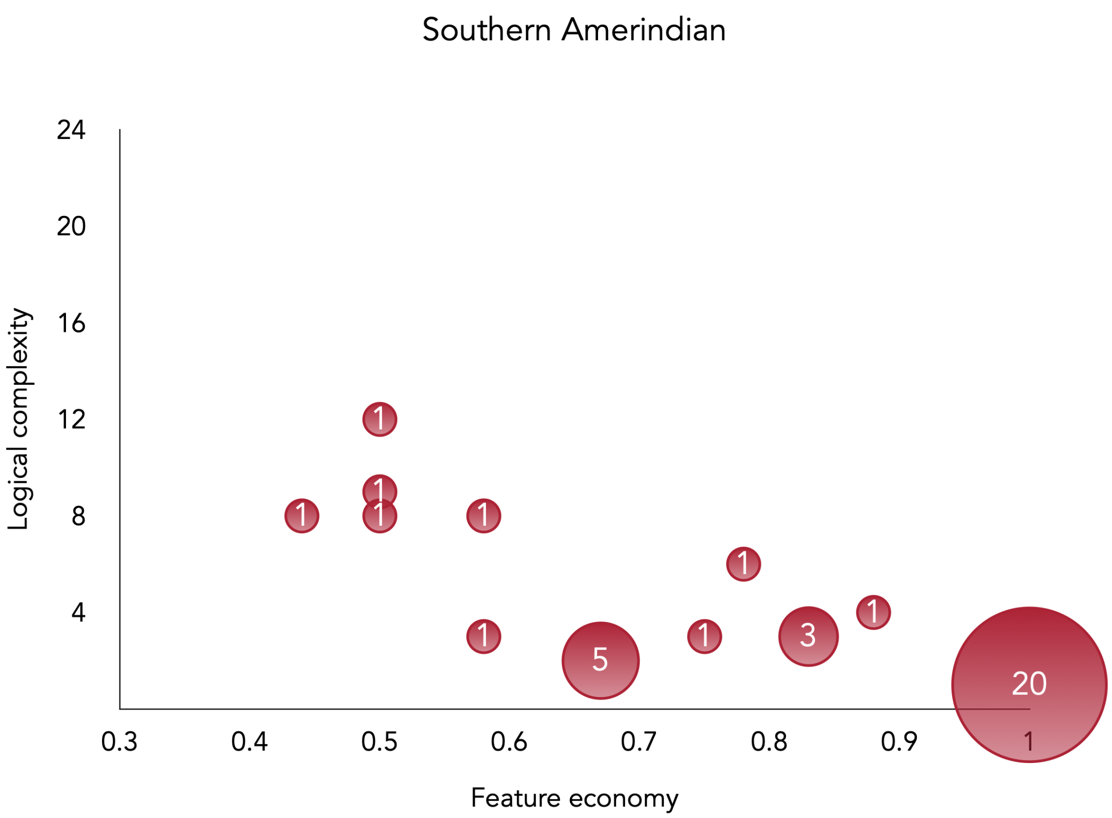


Figure 4.8: The distribution of feature economy and logical complexity in Southern Amerindian languages.

CHAPTER 5

Conclusion and discussion

This thesis has sought to answer the question of which complexity measure predicts the distribution of plosive inventories around the world best. We have analyzed the plosive inventories of 317 languages in order to determine their feature economy and logical complexity. From the results of this analysis, we have learned that a small majority of the languages in our sample has a regular plosive inventory. Since plosives only make up a small portion of the phoneme inventories of most languages, it would be interesting to conduct a similar analysis of the complete phoneme space of the languages in our sample. This would, of course, bring about a whole new set of challenges, such as affricates and double articulations, how to treat gaps in places of unpronounceable phonemes, and the large vowel space that cannot be easily split into clearly defined rows and columns.

From Figures 4.5 and 4.6, we have drawn the preliminary conclusion that logical complexity seems a better fit to our sample than feature economy. Looking at these figures, we must ask ourselves the question whether this comparison is fair. Both graphs are plotted in the same way on the same scale. But there is an important difference between the two: while feature economy can take on infinitely many values between zero and one, logical complexity can only have integer values. And even though there is technically no maximum to the possible values of logical complexity, in practice the highest possible logical complexity value in a spoken language is confined by the limited number of articulatory features humans have at their disposal. Thus, while we have found 29 different feature economy values, there are only 15 different values of logical complexity in our sample. Most importantly, two feature economy values that lie very close together (e.g. 0.50 and 0.51) would be represented in our graph by two different dots. And since 0.50 is a way more likely result of the division between a random inventory size and space size, we see a sort of "spiky" behavior in the graph, where high and low values of feature economy alternate.

To eliminate the feature economy spikes, we could group values that lie close together and plot them in a bar chart. This still leaves the questions of which values to group and how broad to make the bins. Figure 5.1 groups the values in bins of 0.05. In Figure 5.2 the bins are twice as broad. While the spikes are gone, there is still no ascending line to show that higher feature economies would be more frequent in our sample. We therefore stick to our earlier observation that logical complexity seems to fit the data better.

The last question we must answer is whether this conclusion holds when we consider the relatedness between languages within families. The figures in Appendix B show us that most families contain a great variety of different feature economies and logical complexities. Seinhorst (2021) confirms this presumption by calculating the Akaike information criterion (AIC) for both the feature economy and the logical complexity model. The AIC of the feature economy model is 252.21, whereas the AIC of the logical complexity model is 230.94. Since the AIC of logical complexity is so much lower than that of feature economy, the logical complexity model definitively seems to fit the data from our model best. We therefore conclude that logical complexity better predicts the distribution of plosive inventories in our sample than feature economy does.

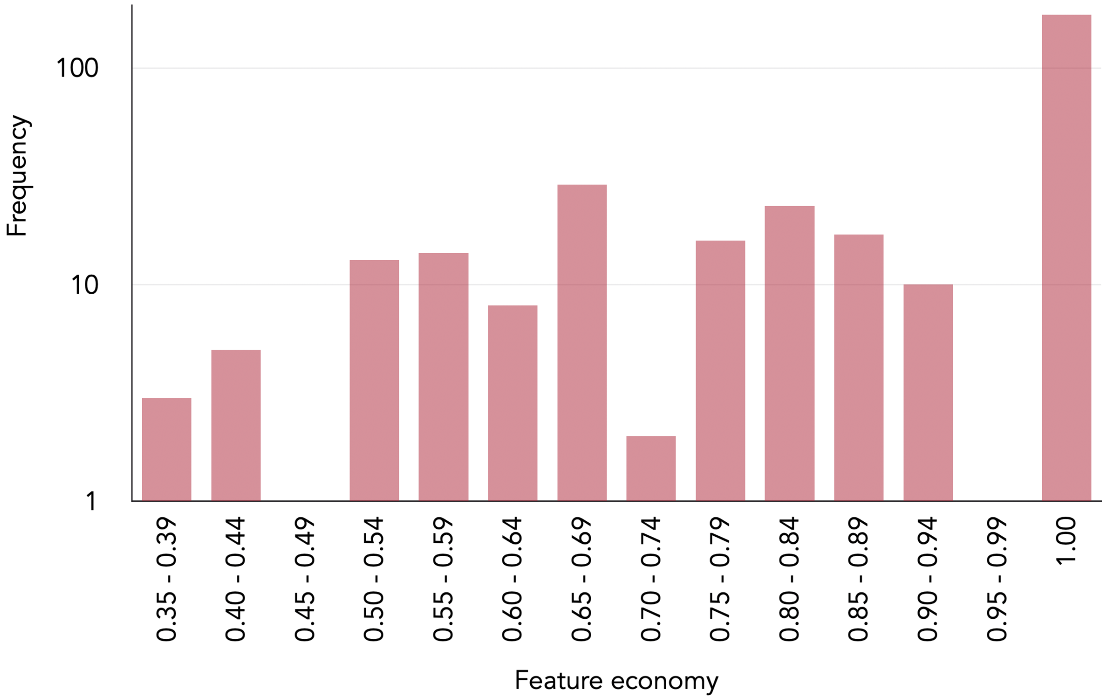


Figure 5.1: The frequency of all feature economies found in the sample, plotted in a bar graph with bins of 0.05.

Figure 5.2: The frequency of all feature economies found in the sample, plotted in a bar graph with bins of 0.1.

Bibliography

Abrams, Daniel and Steven Strogatz (2003). “Modelling the dynamics of language death”. In: *Nature* 424, p. 900. DOI: 10.1038/424900a.

Axtell, Robert (Oct. 2001). “Zipf Distribution of U.S. Firm sizes”. In: *Science (New York, N.Y.)* 293, pp. 1818–20. DOI: 10.1126/science.1062081.

Cederman, Lars-Erik (2003). “Modeling the Size of Wars: From Billiard Balls to Sandpiles”. In: *The American Political Science Review* 97.1, pp. 135–150. ISSN: 00030554, 15375943. DOI: 10.1017/S0003055403000571.

Clements, George (2003). “Feature economy in sound systems”. In: *Phonology* 20.3, pp. 287–333. DOI: 10.1017/s095267570400003x.

Cook, William, Paul Ormerod, and Ellie Cooper (Feb. 2004). “Scaling Behaviour in the Number of Criminal Acts Committed by Individuals”. In: *Journal of Statistical Mechanics: Theory and Experiment* 2004. DOI: 10.1088/1742-5468/2004/07/P07003.

Feldman, Jacob (2000). “Minimization of Boolean complexity in human concept learning”. In: *Nature* 407.6804, pp. 630–633. DOI: 10.1038/35036586.

Ladefoged, Peter and Sandra Ferrari Disner (2012). *Vowels and Consonants*. Wiley-Blackwell. ISBN: 978-1-4443-3429-6.

Levy, Moshe and Sorin Solomon (1997). “New evidence for the power-law distribution of wealth”. In: *Physica A: Statistical Mechanics and its Applications* 242.1, pp. 90–94. ISSN: 0378-4371. DOI: 10.1016/S0378-4371(97)00217-3.

Maddieson, Ian (1984). *Patterns of Sounds*. Cambridge University Press. ISBN: 978-0-521-11326-7. DOI: 10.1017/cbo9780511753459.

Seinhorst, Klaas (2016). “Mind the Gap: Inductive Biases in Phonological Feature Learning”. In: *The evolution of language: Proceedings of the 11th International Conference (EVOLANG11)*. Ed. by S. G. Roberts et al. EVOLANG.

— (2017). “Feature economy vs. logical complexity in phonological pattern learning”. In: *Language Sciences* 60, pp. 69–79. DOI: 10.1016/j.langsci.2016.10.002.

— (2021). *The complexity and learnability of phonological patterns: Simulations, experiments, typology*. LOT.

Shepard, Roger, Carl Hovland, and Herbert Jenkins (1961). “Learning and Memorization of Classifications”. In: *Psychological Monographs: General and Applied* 75.13, pp. 1–42. DOI: 10.1037/h0093825.

Sornette, Didier (2002). “Predictability of catastrophic events: Material rupture, earthquakes, turbulence, financial crashes, and human birth”. In: *Proceedings of the National Academy of Sciences* 99, pp. 2522–2529. DOI: 10.1073/pnas.022581999.

Zipf, George (1949). *Human Behavior and the Principle of Least Effort*. Addison-Wesley Press.

APPENDIX A

Calculations

The following twenty pages contain the feature economy and logical complexity calculations for the 317 languages in Maddieson (1984), in alphabetical order. The tables are structured as follows:

- Column 1: the name of the language;
- Column 2: the size of the language's plosive space;
- Column 3: the size of the language's plosive inventory;
- Column 4: the language's feature economy;
- Column 5: the disjunctive normal form of the plosive inventory;
- Column 6: the minimal formula that can be derived from the disjunctive normal form;
- Column 7: the language's logical complexity.

Language	Space size	Inventory size	Feature economy	Disjunctive normal form	Minimal formula	Logical complexity
!Xũ	24	12	0.50	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + ab''c + a'b''c + a''b''c + a''b'''c + a'bc' + a'b''c'	a'b + a'b'' + bc + b'c + b''c + a''c	12
Abipon	4	4	1.00	a + a' + a'' + a'''	*A*	1
Achumawi	4	4	1.00	a + a' + a'' + a'''	*A*	1
Acoma	8	7	0.88	ab + a'b + a''b + a'''b + ab' + a'b' + a''b'	a + a' + a'' + b	4
Adzera	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Ainu	3	3	1.00	a + a' + a''	*A*	1
Akan	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Alabama	6	4	0.67	ab + a'b + a''b + ab'	a + b	2
Alawa	10	10	1.00	ab + a'b + a''b + a'''b + a''''b + ab' + a'b' + a''b' + a'''b' + a''''b'	*A*	1
Albanian	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Aleut	3	3	1.00	a + a' + a''	*A*	1
Amahuaca	3	3	1.00	a + a' + a''	*A*	1
Amharic	12	8	0.67	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + a''bc' + a''b'c'	a'' + c	2
Amo	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Amoy	9	9	1.00	ab + a'b + a''b + ab' + a'b' + a''b' + ab'' + a'b'' + a''b''	*A*	1
Amuesha	6	4	0.67	ab + a'b + a''b + ab'	a + b	2
Andamanese	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1

Language	Space size	Inventory size	Feature economy	Disjunctive normal form	Minimal formula	Logical complexity
Angas	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Ao	3	3	1.00	a + a' + a''	*A*	1
Apinaye	8	7	0.88	ab + a'b + a''b + a'''b + ab' + a'b' + a''b'	a + a' + a'' + b	4
Arabana-Wanganura	6	6	1.00	ab + a'b + a''b + a'''b + a''''b + a'''''b	*A*	1
Arabela	3	3	1.00	a + a' + a''	*A*	1
Arabic	32	16	0.50	a'bcd + a''bcd + a'''bcd + ab'cd + a'b'cd + a''b'cd + a'bc'd + a''bc'd + a'''bc'd + ab'c'd + a'b'c'd + a''b'c'd + a'bcd' + a'b'cd' + a'bc'd' + a'b'c'd'	ab'd + a' + a''d + a'''bd	9
Aranda	12	12	1.00	ab + a'b + a''b + a'''b + a''''b + a'''''b + ab' + a'b' + a''b' + a'''b' + a''''b' + a'''''b'	*A*	1
Araucanian	4	4	1.00	a + a' + a'' + a'''	*A*	1
Armenian. Eastern	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Ashuslay	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Asmat	3	3	1.00	a + a' + a''	*A*	1
Atayal	4	4	1.00	a + a' + a'' + a'''	*A*	1
Auca	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Awiya	20	12	0.60	abc + a'bc + a'''bc + a''''bc + ab'c + a''b'c + a'''b'c + a''''b'c + a'''bc' + a''''bc' + a'''b'c' + a''''b'c'	a''' + a'''' + ac + a'bc + a''b'c	10
Azerbaijani	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Bambara	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1

Language	Space size	Inventory size	Feature economy	Disjunctive normal form	Minimal formula	Logical complexity
Bandjalang	3	3	1.00	a + a' + a''	*A*	1
Barasano	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Bardi	5	5	1.00	a + a' + a'' + a''' + a''''	*A*	1
Bariba	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Bashkir	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Basque	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Batak	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Beembe	6	5	0.83	ab + a'b + a''b + ab' +a'b'	a + a' + b	3
Beja	16	9	0.56	a'bc + a''bc + a'''bc + ab'c + a'b'c + a''b'c + a'''b'c + a'''bc' + a'''b'c'	a''' + b'c + a'c + a''c	7
Bengali	16	16	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b' + ab'' + a'b'' + a''b'' + a'''b'' + ab''' + a'b''' + a''b''' + a'''b'''	*A*	1
Berta	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Birom	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Bisa	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Boro	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Brahui	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Breton	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Bribri	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1

Language	Space size	Inventory size	Feature economy	Disjunctive normal form	Minimal formula	Logical complexity
Bulgarian	12	12	1.00	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + abc' + a'bc' + a''bc' + ab'c' + a'b'c' + a''b'c'	*A*	1
Burera	4	4	1.00	a + a' + a'' + a'''	*A*	1
Burmese	9	9	1.00	ab + a'b + a''b + ab' + a'b' + a''b' + ab'' + a'b'' + a''b''	*A*	1
Burushaski	15	14	0.93	ab + a'b + a''b + a'''b + a''''b + ab' + a'b' + a''b' + a'''b' + a''''b' + ab'' + a'b'' + a''b'' + a'''b''	a + a' + a'' + a''' + b + b'	6
Campa	4	4	1.00	a + a' + a'' + a'''	*A*	1
Carib	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Cashinahua	6	4	0.67	ab + a'b + a''b + ab'	a + b	2
Cayapa	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Chacobo	3	3	1.00	a + a' + a''	*A*	1
Cham	15	12	0.80	ab + a'b + a''b + a'''b + a''''b + ab' + a'b' + a''b' + a'''b' + a''''b' + ab'' + a'b''	a + a' + b + b'	4
Chamorro	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Changchow	9	9	1.00	ab + a'b + a''b + ab' + a'b' + a''b' + ab'' + a'b'' + a''b''	*A*	1
Chatino	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Cheremis	6	5	0.83	ab + a'b + a''b + ab' +a'b'	a + a' + b	3
Chipewyan	9	5	0.56	ab + a'b + a''b + a''b' + a'b''	b + a''b' + a'b''	5
Chontal	6	5	0.83	ab + a'b + a''b + ab' +a'b'	a + a' + b	3

Language	Space size	Inventory size	Feature economy	Disjunctive normal form	Minimal formula	Logical complexity
Chuave	6	5	0.83	ab + a'b + a''b + ab' +a'b'	a + a' + b	3
Chukchi	4	4	1.00	a + a' + a'' + a'''	*A*	1
Chuvash	12	7	0.58	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + a'bc'	a'b + c	3
Cofan	12	12	1.00	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + abc' + a'bc' + a''bc' + ab'c' + a'b'c' + a''b'c'	*A*	1
Dafla	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Dagbani	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Dakota	9	7	0.78	ab + a'b + a''b + ab' + a'b' + a''b' + ab''	a + b + b'	3
Dan	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Dani	6	4	0.67	ab + a'b + a''b + ab'	a + b	2
Daribi	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Delaware	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Dera	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Diegueño	10	6	0.60	ab + a'b + a''b + a'''b + a''''b + ab'	a + b	2
Diola	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Diyari	12	8	0.67	ab + a'b + a''b + a'''b + a''''b + a'''''b + a''b' + a'''b'	a'' + a''' + b	3
Dizi	6	5	0.83	ab + a'b + a''b + ab' +a'b'	a + a' + b	3
Doayo	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Efik	6	4	0.67	a'b + a''b + ab' + a'b'	a' + ab' + a''b	5

Language	Space size	Inventory size	Feature economy	Disjunctive normal form	Minimal formula	Logical complexity
Evenki	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Ewe	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Farsi	8	7	0.88	ab + a'b + a''b + a'''b + ab' + a'b' + a''b'	a + a' + a'' + b	4
Fasu	3	3	1.00	a + a' + a''	*A*	1
Finnish	12	9	0.75	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + abc' + a'bc' + a''bc'	b + c	2
French	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Fuchow	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Fur	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Gã	8	6	0.75	ab + a'b + a'''b + ab' + a''b' + a'''b'	a + a''' + a'b + a''b'	6
Gadsup	6	4	0.67	ab + a'b + a''b + ab'	a + b	2
Garo	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Gbeya	9	9	1.00	ab + a'b + a''b + ab' + a'b' + a''b' + ab'' + a'b'' + a''b''	*A*	1
Georgian	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
German	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Gilyak	10	9	0.90	ab + a'b + a''b + a'''b + a''''b + a'b' + a''b' + a'''b' + a''''b'	a' + a'' + a''' + a'''' + b	5
Goldi	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Greek	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1

Language	Space size	Inventory size	Feature economy	Disjunctive normal form	Minimal formula	Logical complexity
Greenlandic	4	4	1.00	a + a' + a'' + a'''	*A*	1
Guahibo	12	6	0.50	ab + a''b + a'''b + a'b' + ab'' + a''b''	ab + a''b + a'''b + a'b' + ab'' + a''b''	12
Guajiro	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Guarani	6	4	0.67	ab + a'b + a''b + ab'	a + b	2
Gugu-Yalanji	4	4	1.00	a + a' + a'' + a'''	*A*	1
Gununa-Kena	8	6	0.75	ab + a'b + a''b + a'''b + ab' + a'b'	a + a' + b	3
Haida	20	14	0.70	abc + a'bc + a''bc + a'''bc + a''''bc + ab'c + a'b'c + a''b'c + a'''b'c + a''''b'c + a'''bc' + a''''bc' + a'''b'c' + a''''b'c'	a''' + a'''' + c	3
Hakka	9	9	1.00	ab + a'b + a''b + ab' + a'b' + a''b' + ab'' + a'b'' + a''b''	*A*	1
Hamer	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Hausa	18	9	0.50	abc + a'bc + a''bc + a''b'c + a''b''c + a'bc' + a''bc' + a''b'c' + a''b''c'	a'b + bc + a''	5
Hawaiian	2	2	1.00	a + a'	*A*	1
Hebrew	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Hindi-Urdu	20	17	0.85	ab + a'b + a''b + a'''b + a''''b + ab' + a'b' + a''b' + a'''b' + ab'' + a'b'' + a''b'' + a'''b'' + ab''' + a'b''' + a''b''' + a'''b'''	a + a' + a'' + a''' + b	5
Hopi	8	5	0.63	ab + a'b + a''b + a'''b + ab'	a + b	2
Hungarian	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1

Language	Space size	Inventory size	Feature economy	Disjunctive normal form	Minimal formula	Logical complexity
Hupa	6	5	0.83	ab + a'b + a''b + ab' +a'b'	a + a' + b	3
Iai	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Igbo	36	20	0.56	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + ab''c + a'b''c + a''b''c + ab'''c + a'b'''c + a''b'''c + abc' + ab'c' + ab''c' + ab'''c' + a''bc'' + a''b'c'' + a''b''c'' + a''b'''c''	c + ac' + a''c''	5
Ik	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Iraqw	16	10	0.63	abc + a'bc + a''bc + a'''bc + ab'c + a'b'c + a''b'c + a''bc' + a''b'c' + a'''b'c'	ac + a'c + a'' + a'''b	7
Irish	24	12	0.50	ab'c + ab''c + a'bc + a''bc + a'''bc + a'''b'c + ab'c' + ab''c' + a'bc' + a''bc' + a'''bc' + a'''b'c'	ab' + ab'' + a'b + a''b + a'''b + a'''b'	12
Island Carib	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Itonama	12	6	0.50	abc + a'bc + a''bc + ab'c + a'b'c + a'bc'	a'b + bc + ac + a'c	8
Iwam	3	3	1.00	a + a' + a''	*A*	1
Japanese	12	9	0.75	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + abc' + a'bc' + a''bc'	b + c	2
Jaqaru	10	10	1.00	ab + a'b + a''b + a'''b + a''''b + ab' + a'b' + a''b' + a'''b' + a''''b'	*A*	1
Javanese	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Jingpho	9	9	1.00	ab + a'b + a''b + ab' + a'b' + a''b' + ab'' + a'b'' + a''b''	*A*	1
Jivaro	3	3	1.00	a + a' + a''	*A*	1

Language	Space size	Inventory size	Feature economy	Disjunctive normal form	Minimal formula	Logical complexity
K'ekchi	8	5	0.63	ab + a'b + a''b + a'''b + ab'	a + b	2
Kabardian	12	8	0.67	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + abc' + ab'c'	a + c	2
Kadugli	10	9	0.90	ab + a'b + a''b + a'''b + a''''b + a'b' + a''b' + a'''b' + a''''b'	a' + a'' + a''' + a'''' + b	5
Kaliai	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Kan	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Kanakuru	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Kanuri	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Karen	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Kariera-Ngarluma	6	6	1.00	ab + a'b + a''b + a'''b + a''''b + a'''''b	*A*	1
Karok	3	3	1.00	a + a' + a''	*A*	1
Kashmiri	12	12	1.00	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + abc' + a'bc' + a''bc' + ab'c' + a'b'c' + a''b'c'	*A*	1
Katcha	10	9	0.90	ab + a'b + a''b + a'''b + a''''b + a'b' + a''b' + a'''b' + a''''b'	a' + a'' + a''' + a'''' + b	5
Kefa	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Ket	16	7	0.44	a'bc + a''bc + a'''bc + ab'c + a'b'c + a'bc' + a'b'c'	a' + a''bc + a'''bc + ab'c	10
Kewa	12	5	0.42	a'b + a''b + a'''b' + ab'' + a'b''	a'b + a''b + a'''b' + ab'' + a'b''	10
Khalaj	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1

Language	Space size	Inventory size	Feature economy	Disjunctive normal form	Minimal formula	Logical complexity
Kharia	16	16	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b' + ab'' + a'b'' + a''b'' + a'''b'' + ab''' + a'b''' + a''b''' + a'''b'''	*A*	1
Khasi	9	8	0.89	ab + a'b + a''b + ab' + a'b' + a''b' + ab'' + a'b''	a + a' + b + b'	4
Khmer	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Kirghiz	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Klamath	10	10	1.00	ab + a'b + a''b + a'''b + a''''b + ab' + a'b' + a''b' + a'''b' + a''''b'	*A*	1
Koiari	6	5	0.83	ab + a'b + a''b + ab' +a'b'	a + a' + b	3
Koma	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Komi	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Korean	9	9	1.00	ab + a'b + a''b + ab' + a'b' + a''b' + ab'' + a'b'' + a''b''	*A*	1
Kota	10	10	1.00	ab + a'b + a''b + a'''b + a''''b + ab' + a'b' + a''b' + a'''b' + a''''b'	*A*	1
Kpelle	12	8	0.67	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + abc' + ab'c'	a + c	2
Kullo	6	5	0.83	ab + a'b + a''b + ab' +a'b'	a + a' + b	3
Kunama	6	5	0.83	ab + a'b + a''b + ab' +a'b'	a + a' + b	3
Kunimaipa	8	7	0.88	ab + a'b + a''b + a'''b + ab' + a'b' + a''b'	a + a' + a'' + b	4
Kunjen	10	10	1.00	ab + a'b + a''b + a'''b + a''''b + ab' + a'b' + a''b' + a'''b' + a''''b'	*A*	1

Language	Space size	Inventory size	Feature economy	Disjunctive normal form	Minimal formula	Logical complexity
Kurdish	8	7	0.88	ab + a'b + a''b + a'''b + ab' + a'b' + a''b'	a + a' + a'' + b	4
Kurukh	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Kwakw'ala	20	12	0.60	abc + a'bc + a''bc + a''''bc + ab'c + a'b'c + a''b'c + a''''b'c + a'''bc' + a''''bc' + a'''b'c' + a''''b'c'	a'''c + a'''' + ac + a'c + a''c	9
Lahu	12	11	0.92	ab + a'b + a''b + ab' + a'b' + a''b' + ab'' + a'b'' + a''b'' + ab''' + a'b'''	a + a' + b + b' + b''	5
Lak	48	17	0.35	ab'cd + a'b'cd + a''b'cd + a'''b'cd + ab''cd + a'b''cd + a''b''cd + a'''b''cd + a''b'c'd + a'''b'c'd + a'''b''c'd + abcd' + a'bcd' + a''bcd' + a'''bcd' + a''bc'd' + a'''bc'd'	a''b'd + a'''b'd + a'''b''d + a'''bd' + a''bd' + bcd' + b'cd + b''cd	24
Lakkia	27	10	0.37	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + ab''c + a''bc' + a''b'c' + a''b'c''	a''b' + a''bc' + b'c + bc + ac	11
Lappish	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Lelemi	8	5	0.63	ab + a'b + a'''b + ab' + a''b'	a + a'b + a''b' + a'''b	7
Lithuanian	12	12	1.00	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + abc' + a'bc' + a''bc' + ab'c' + a'b'c' + a''b'c'	*A*	1
Logbara	9	8	0.89	ab + a'b + a''b + ab' + a'b' + a''b' + ab'' + a'b''	a + a' + b + b'	4
Luiseño	8	6	0.75	ab + a'b + a''b + a'''b + ab' + a'b'	a + a' + b	3
Lungchow	9	8	0.89	ab + a'b + a''b + ab' + a'b' + a''b' + ab'' + a'b''	a + a' + b + b'	4
Luo	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Luvale	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Maasai	3	3	1.00	a + a' + a''	*A*	1

Language	Space size	Inventory size	Feature economy	Disjunctive normal form	Minimal formula	Logical complexity
Maba	10	8	0.80	a'b + a''b + a''''b + ab' + a'b' + a''b' + a'''b' + a''''b'	a' + a'' + a'''' + b'	4
Mabuiag	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Maidu	4	4	1.00	a + a' + a'' + a'''	*A*	1
Malagasy	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Malakmalak	4	4	1.00	a + a' + a'' + a'''	*A*	1
Malay	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Malayalam	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Manchu	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Mandarin	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Maori	3	3	1.00	a + a' + a''	*A*	1
Maranungku	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Margi	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Maung	4	4	1.00	a + a' + a'' + a'''	*A*	1
Mazahua	18	10	0.56	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + a''b''c + a''bc' + a''b'c' + a''b''c'	a'' + bc + b'c	5
Mazatec	9	5	0.56	ab + a'b + a''b + ab' + a''b''	b + ab' + a''b''	5
Mixe	6	5	0.83	ab + a'b + a''b + ab' +a'b'	a + a' + b	3
Mixtec	6	4	0.67	ab + a'b + a''b + ab'	a + b	2
Mongolian	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1

Language	Space size	Inventory size	Feature economy	Disjunctive normal form	Minimal formula	Logical complexity
Moro	8	6	0.75	ab + a'b + a''b + a'''b + ab' + a'''b'	a + a''' + b	3
Moxo	6	5	0.83	ab + a'b + a''b + ab' +a'b'	a + a' + b	3
Muinane	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Mundari	16	16	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b' + ab'' + a'b'' + a''b'' + a'''b'' + ab''' + a'b''' + a''b''' + a'''b'''	*A*	1
Mura	6	5	0.83	ab + a'b + a''b + ab' +a'b'	a + a' + b	3
Mursi	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Nama	3	3	1.00	a + a' + a''	*A*	1
Nambakaengo	27	18	0.67	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + ab''c + a'b''c + a''b''c + abc' + ab'c' + ab''c' + a''bc' + a''b'c' + a''b''c' + abc'' + ab'c'' + ab''c''	a + c + a''c'	4
Nambiquara. Southern	12	8	0.67	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + abc' + ab'c'	a + c	2
Nasioi	6	4	0.67	ab + a'b + a''b + ab'	a + b	2
Navaho	12	6	0.50	abc + a'bc + a''bc + a'b'c + a''b'c + a''b'c'	a''b' + a'c + bc	6
Neo-Aramaic	8	7	0.88	ab + a'b + a''b + a'''b + ab' + a'b' + a''b'	a + a' + a'' + b	4
Nera	6	5	0.83	ab + a'b + a''b + ab' +a'b'	a + a' + b	3
Nez Perce	4	4	1.00	a + a' + a'' + a'''	*A*	1
Ngizim	32	18	0.56	abc + a'bc + a''bc + a'''bc + ab'c + a'b'c + a''b'c + a'''b'c + ab''c + a'b''c + a''b''c + a'''b''c + ab'''c + a'b'''c + a''b'''c + a'''bc' + a'''b'c' + a'''b''c'	a'''b + a'''b' + a'''b'' + ac + a'c + a''c	12

Language	Space size	Inventory size	Feature economy	Disjunctive normal form	Minimal formula	Logical complexity
Nimboran	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Nootka	8	6	0.75	ab + a'b + a''b + a'''b + ab' + a'b'	a + a' + b	3
Norwegian	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Nubian	6	5	0.83	ab + a'b + a''b + ab' +a'b'	a + a' + b	3
Nunggubuyu	6	6	1.00	a + a' + a'' + a''' + a'''' + a'''''	*A*	1
Nyangi	4	4	1.00	a + a' + a'' + a'''	*A*	1
Nyangumata	16	6	0.38	abc + a'bc + a''bc + a'''bc + a'''b'c + a'bc'	a'b + bc + a'''c	6
Ocaina	12	7	0.58	abc + a'bc + a''bc + ab'c + a''b'c + a'bc' + a'b'c'	ac + a'b + a'c' + a''c	8
Ojibwa	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Osmanli	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Ostyak	4	4	1.00	a + a' + a'' + a'''	*A*	1
Otomi	24	13	0.54	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + ab''c + a'b''c + a''b''c + ab'''c + a'b'''c + a''bc' + a''b''c'	a''b + a''b'' + ac + a'c + b'c	10
Paez	12	8	0.67	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + abc' + ab'c'	a + c	2
Papago	8	7	0.88	ab + a'b + a''b + a'''b + ab' + a'b' + a''b'	a + a' + a'' + b	4
Pashto	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Pawaian	3	3	1.00	a + a' + a''	*A*	1
Po-Ai	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Pomo. South Eastern	10	7	0.70	ab + a'b + a''b + a'''b + a''''b + ab' + a'b'	a + a' + b	3

Language	Space size	Inventory size	Feature economy	Disjunctive normal form	Minimal formula	Logical complexity
Puget Sound	16	10	0.63	abc + a'bc + a''bc + a'''bc + ab'c + a'b'c + a''b'c + a''bc' + a'''bc' + a''b'c'	a'' + ac + a'c + a'''b	7
Punjabi	24	24	1.00	abc + a'bc + a''bc + a'''bc + ab'c + a'b'c + a''b'c + a'''b'c + ab''c + a'b''c + a''b''c + a'''b''c + abc' + a'bc' + a''bc' + a'''bc' + ab'c' + a'b'c' + a''b'c' + a'''b'c' + ab''c' + a'b''c' + a''b''c' + a'''b''c'	*A*	1
Quechua	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Quileute	16	9	0.56	abc + a'bc + a''bc + a'''bc + ab'c + a'b'c + a''b'c + a''bc' + a'''bc'	ac + a'c + a''c + a''b + a'''b	10
Romanian	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Roro	6	4	0.67	ab + a'b + a''b + ab'	a + b	2
Rotokas	6	4	0.67	ab + a'b + a''b + ab'	a + b	2
Rukai	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Russian	12	11	0.92	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + abc' + a'bc' + a''bc' + ab'c' + a'b'c'	a + a' + b + c	4
Sa'ban	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Saek	9	8	0.89	ab + a'b + a''b + ab' + a'b' + a''b' + ab'' + a'b''	a + a' + b + b'	4
Sara	12	10	0.83	ab + a'b + a'''b + ab' + a'b' + a'''b' + ab'' + a'b'' + a''b'' + a'''b''	a + a' + a''' + b''	4
Sebei	4	4	1.00	a + a' + a'' + a'''	*A*	1
Sedang	12	11	0.92	ab + a'b + a''b + ab' + a'b' + a''b' + ab'' + a'b'' + a''b'' + ab''' + a'b'''	a + a' + b + b' + b''	5

Language	Space size	Inventory size	Feature economy	Disjunctive normal form	Minimal formula	Logical complexity
Selepet	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Senadi	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Seneca	6	3	0.50	ab' + a'b + a''b	ab' + a'b + a''b	6
Sentani	6	4	0.67	ab + a'b + a''b + ab'	a + b	2
Shasta	3	3	1.00	a + a' + a''	*A*	1
Shilha	24	13	0.54	a'bcd + a''bcd + ab'cd + a'b'cd + a''b'cd + a'bc'd + a''bc'd + ab'c'd + a'b'c'd + a''b'c'd + a'bcd' + a''bcd' + a'b'cd'	a'c + a'd + a''d + b'd + a''bc	11
Sinhalese	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Siona	16	8	0.50	abc + a'bc + a'''bc + ab'c + a''b'c + a'''b'c + a'''bc' + a'''b'c'	a''' + ac + a'bc + a''b'c	9
Siriono	18	8	0.44	abc + a'bc + a''bc + ab'c + ab''c + a'b''c + a''b''c + a''bc'	ac + bc + b''c + a''b	8
Socotri	6	5	0.83	ab + a'b + a''b + ab' +a'b'	a + a' + b	3
Somali	30	12	0.40	a'bc + a'''bc + ab'c + a'b'c + a'''b'c + a''''b'c + a''b''c + ab'c' + a'b'c' + a'''b'c' + a''''b'c' + a''b''c'	a'bc + a'''bc + ab' + a'b' + a'''b' + a''''b' + a''b''	16
Songhai	12	7	0.58	a'bc + a''bc + ab'c + a'b'c + a''b'c + a'bc' + a'b'c'	a' + a''c + b'c	5
Spanish	3	3	1.00	a + a' + a''	*A*	1
Squamish	8	6	0.75	ab + a'b + a''b + a'''b + ab' + a'b'	a + a' + b	3
Standard Thai	9	8	0.89	ab + a'b + a''b + ab' + a'b' + a''b' + ab'' + a'b''	a + a' + b + b'	4
Suena	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1

Language	Space size	Inventory size	Feature economy	Disjunctive normal form	Minimal formula	Logical complexity
Sui	16	12	0.75	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b' + ab'' + a'b'' + ab''' + a'b'''	a + a' + b + b'	4
Sundanese	8	6	0.75	ab + a'b + a'''b + ab' + a''b' + a'''b'	a + a''' + a'b + a''b'	6
Swahili	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Tabi	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Tacana	6	5	0.83	ab + a'b + a''b + ab' +a'b'	a + a' + b	3
Tagalog	6	5	0.83	ab + a'b + a''b + ab' +a'b'	a + a' + b	3
Taishan	12	8	0.67	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + abc' + ab'c'	a + c	2
Tama	8	6	0.75	a'b + a'''b + ab' + a'b' + a''b' + a'''b'	a' + a''' + b'	3
Tamang	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Tampulma	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Taoripi	3	3	1.00	a + a' + a''	*A*	1
Tarascan	18	11	0.61	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + ab''c + a'b''c + a''b''c + a''bc' + a''b'c'	c + a''b + a''b'	5
Tarok	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Tavgy	8	8	1.00	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b'	*A*	1
Teke	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Telefol	12	7	0.58	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + a''bc'	c + a''b	3

Language	Space size	Inventory size	Feature economy	Disjunctive normal form	Minimal formula	Logical complexity
Telugu	16	15	0.94	ab + a'b + a''b + a'''b + ab' + a''b' + a'''b' + ab'' + a'b'' + a''b'' + a'''b'' + ab''' + a'b''' + a''b''' + a'''b'''	a + a'' + a''' + b + b'' + b'''	6
Temein	10	9	0.90	ab + a'b + a''b + a''''b + ab' + a'b' + a''b' + a'''b' + a''''b'	a + a' + a'' + a'''' + b'	5
Temne	8	7	0.88	ab + a'b + a''b + a'''b + ab' + a'b' + a''b'	a + a' + a'' + b	4
Ticuna	12	7	0.58	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + abc'	ab + c	3
Tiddin Chim	9	8	0.89	ab + a'b + a''b + ab' + a'b' + a''b' + ab'' + a'b''	a + a' + b + b'	4
Tigre	6	5	0.83	ab + a'b + a''b + ab' +a'b'	a + a' + b	3
Tiwa	18	9	0.50	abc + a'bc + a''bc + ab'c + a'b'c + ab''c + a'b''c + a''b''c + a''bc'	ac + a'c + a''b + b''c	8
Tiwi	5	5	1.00	a + a' + a'' + a''' + a''''	*A*	1
Tlingit	16	12	0.75	abc + a'bc + a''bc + a'''bc + ab'c + a'b'c + a''b'c + a'''b'c + a''bc' + a'''bc' + a''b'c' + a'''b'c'	a'' + a''' + c	3
Tolowa	12	5	0.42	abc + a'bc + a''bc + a'b'c + a''bc	a''b + a'c + bc	6
Tonkawa	6	4	0.67	ab + a'b + a''b + ab'	a + b	2
Totonac	4	4	1.00	a + a' + a'' + a'''	*A*	1
Tsou	6	4	0.67	ab + a'b + a''b + ab'	a + b	2
Tuareg	20	9	0.45	a'bc + a'''bc + a''''bc + ab'c + a'b'c + a''b'c + a'''b'c + a'bc' + a'b'c'	a' + ab'c + a''b'c + a'''c + a''''bc	12
Tucano	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1

Language	Space size	Inventory size	Feature economy	Disjunctive normal form	Minimal formula	Logical complexity
Tunica	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Tuva	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Tzeltal	9	6	0.67	ab + a'b + a''b' + ab'' + a'b'' + a''b''	b'' + ab + a'b + a''b'	7
Vietnamese	6	4	0.67	ab + a'b + a''b + ab'	a + b	2
Wantoat	12	8	0.67	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + abc' + ab'c'	a + c	2
Wapishana	9	7	0.78	ab + a'b + a''b + a'b' + a''b' + ab'' + a'b''	b + a' + a''b' + ab''	6
Wappo	8	7	0.88	ab + a'b + a''b + a'''b + ab' + a'b' + a''b'	a + a' + a'' + b	4
Washkuk	12	8	0.67	a'bc + a''bc + ab'c + a'b'c + a''b'c + ab'c' + a''bc' + a''b'c'	a'' + ab' + a'c	5
Western Desert	5	5	1.00	a + a' + a'' + a''' + a''''	*A*	1
Wichita	8	6	0.75	abc + a'bc + ab'c + a'b'c + a'bc' + a'b'c'	a' + c	2
Wik-Munkan	4	4	1.00	a + a' + a'' + a'''	*A*	1
Wintu	12	8	0.67	ab + a'b + a''b + a'''b + ab' + a'b' + ab'' + a'b''	a + a' + b	3
Wiyot	12	8	0.67	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + abc' + ab'c'	a + c	2
Wolof	16	9	0.56	abc + a'bc + a''bc + a'''bc + ab'c + a'b'c + a''b'c + ab'c' + a''b'c'	ab' + a″b' + a'c + bc	8
Yagaria	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Yakut	12	9	0.75	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + abc' + a'bc' + a''bc'	b + c	2

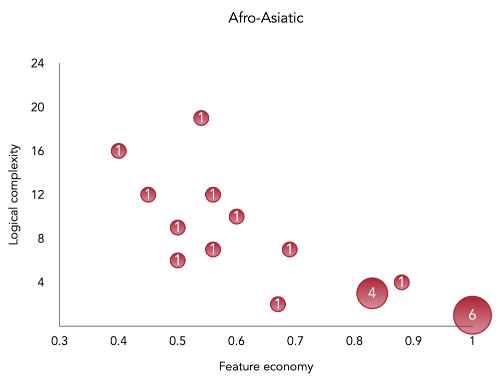
Language	Space size	Inventory size	Feature economy	Disjunctive normal form	Minimal formula	Logical complexity
Yana	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Yao	12	12	1.00	abc + a'bc + a''bc + ab'c + a'b'c + a''b'c + abc' + a'bc' + a''bc' + ab'c' + a'b'c' + a''b'c'	*A*	1
Yaqui	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Yareba	6	5	0.83	ab + a'b + a''b + ab' +a'b'	a + a' + b	3
Yay	12	10	0.83	ab + a'b + a''b + a'''b + ab' + a'b' + a''b' + a'''b' + ab'' + a'b''	a + a' + b + b'	4
Yuchi	9	9	1.00	ab + a'b + a''b + ab' + a'b' + a''b' + ab'' + a'b'' + a''b''	*A*	1
Yukaghir	8	7	0.88	ab + a'b + a''b + a'''b + ab' + a'b' + a''b'	a + a' + a'' + b	4
Yulu	12	11	0.92	ab + a'b + a''b + a'''b + ab' + a'b' + a'''b' + ab'' + a'b'' + a''b'' + a'''b''	a + a' + a''' + b + b''	5
Yurak	12	7	0.58	abc + a'bc + a''bc + ab'c + abc' + a'bc' + ab'c's	a + a'b + bc	5
Zande	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Zoque	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Zulu	6	6	1.00	ab + a'b + a''b + ab' + a'b' + a''b'	*A*	1
Zuni	12	4	0.33	abc + a'bc + a''b'c + a''b'c'	abc + a'bc + a''b'	8

APPENDIX B

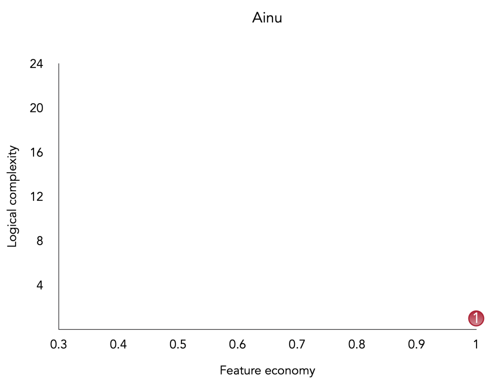
Language families

The graphs in this appendix show the distribution of feature economy and logical complexity per language family, following the division described in Maddieson (1984). Larger circles denote a higher number of languages with that specific combination of feature economy and logical complexity. The number between brackets indicates the number of languages our sample contains from that family.

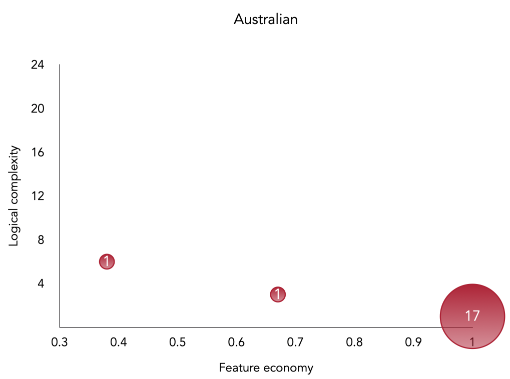
Afro-Asiatic (21)



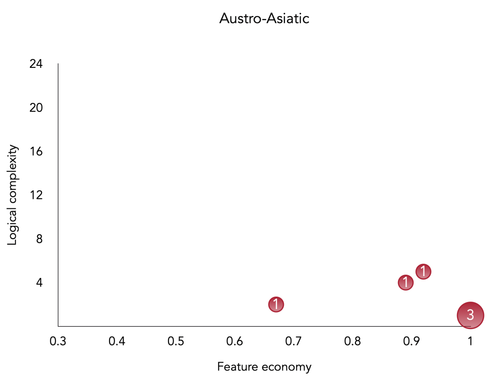
Ainu (1)



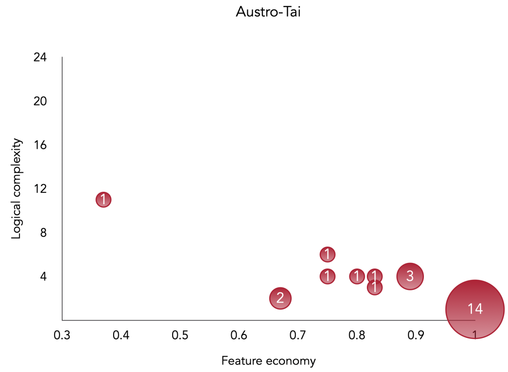
Australian (19)



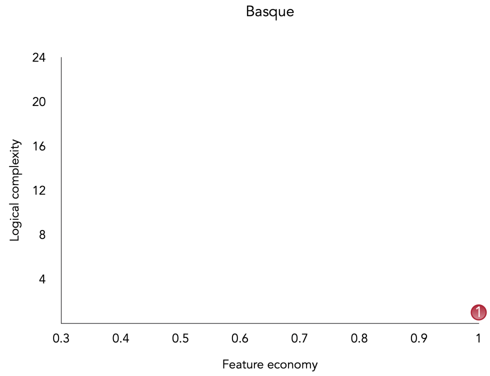
Austro-Asiatic (6)



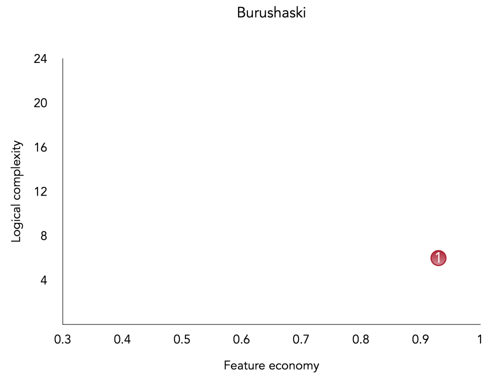
Austro-Tai (25)



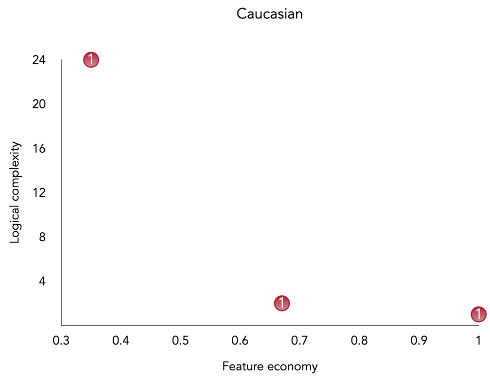
Basque (1)



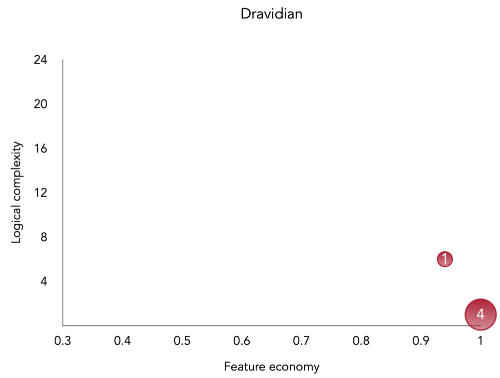
Burushaski (1)



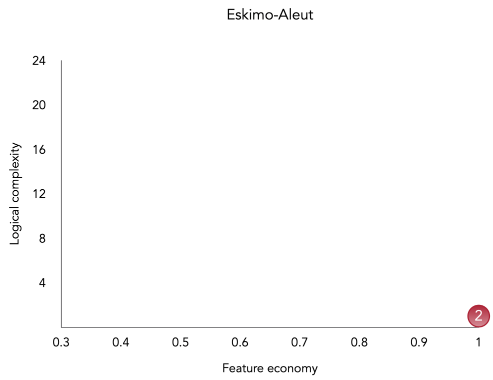
Caucasian (3)



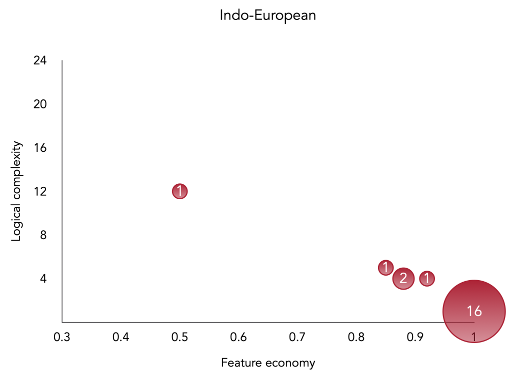
Dravidian (5)



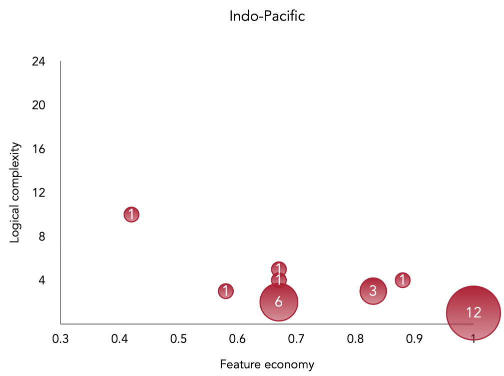
Eskimo-Aleut (2)



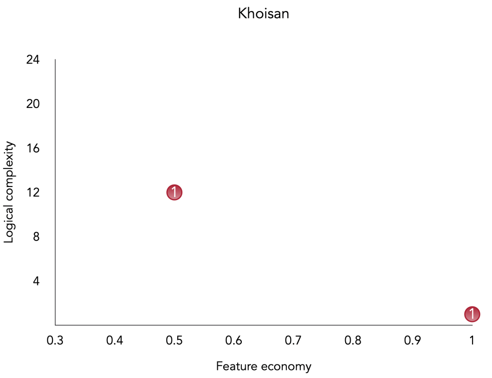
Indo-European (21)



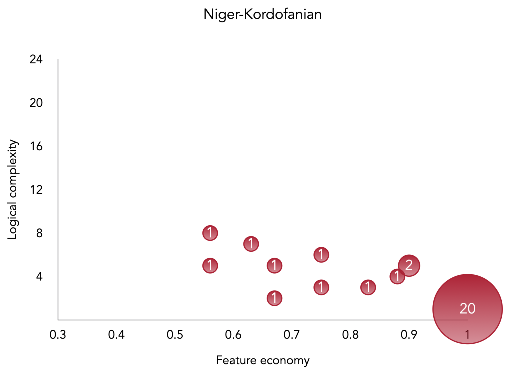
Indo-Pacific (26)



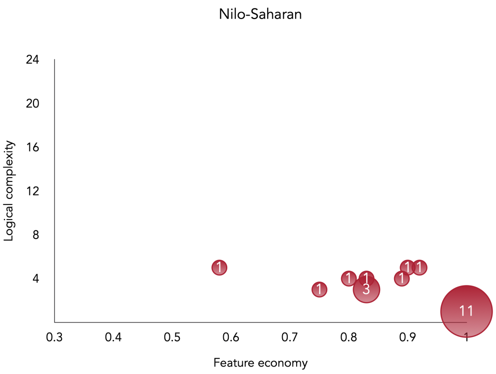
Khoisan (2)



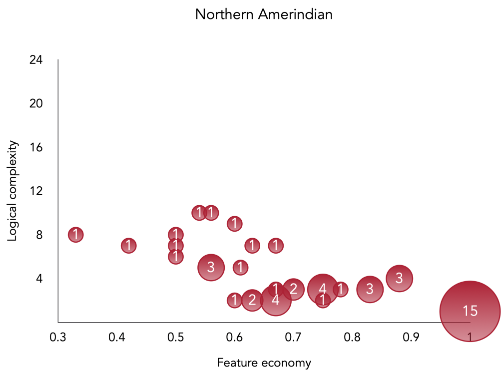
Niger-Kordofanian (31)



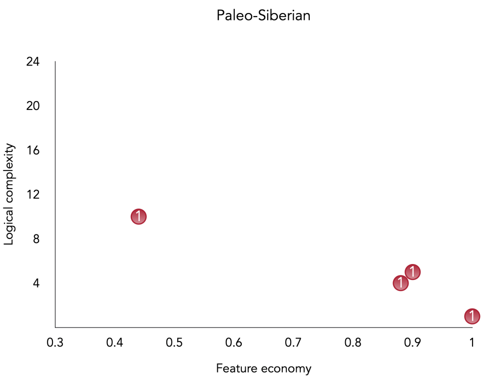
Nilo-Saharan (21)



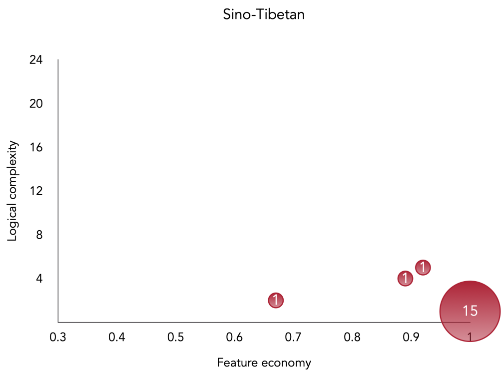
Northern Amerindian (51)



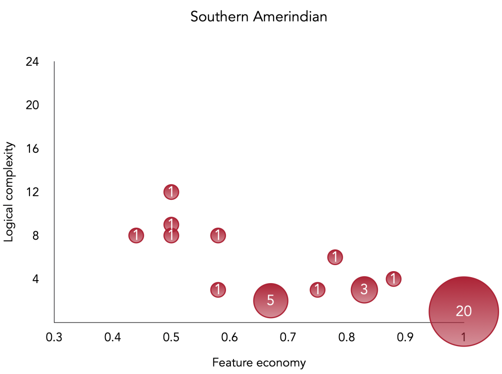
Paleo-Siberian (4)



Sino-Tibetan (18)



Southern Amerindian (37)



Ural-Altaic (22)