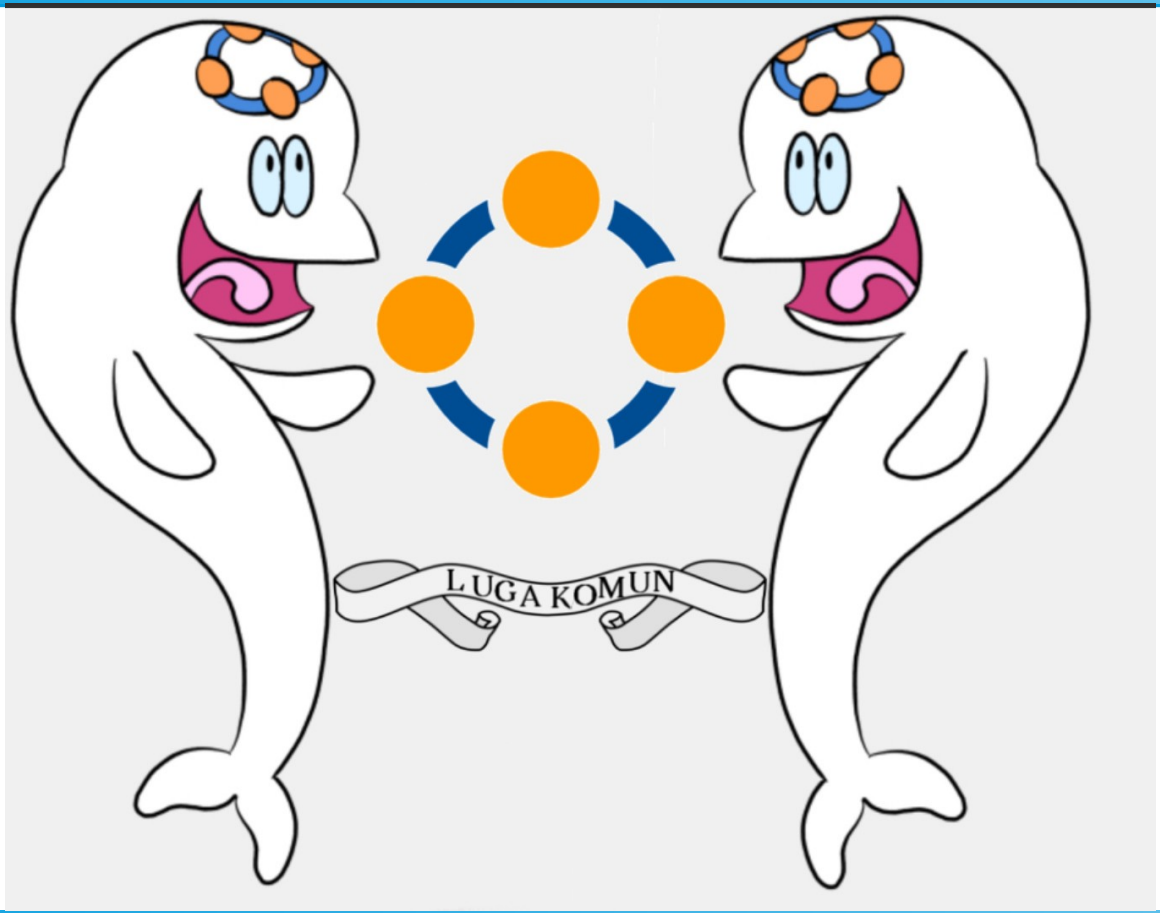
Lugamun

Eine mit Computerunterstützung erstellte „Worldlang“

Lugamun (luga komun), a worldlang



Text sample: Mi ha sonya

Mi xvo a tum si den, mis rafiki, malgre to nas konfronta yo difisil de si den va den tali, mi ha yexo sonya. It xi sonya gen jar profon ni sonya emeriki. Mi ha sonya to aru den si nasion ga kaixu janli va fa real o manan sahi de iti imanin: “Nas opin to yo si sahi eviden: to ol jen bi krea kom igval.” Mi ha sonya to aru den ni yo komonte lal de Georgia, yo bet de esklave prei va bet de haja esklave prei ga bisa side gengen ni mesa de kikanes. — Martin Luther King Jr.

I say to you today, my friends, so even though we face the difficulties of today and tomorrow, I still have a dream. It is a dream deeply rooted in the American dream. I have a dream that one day this nation will rise up and live out the true meaning of its creed: “We hold these truths to be self-evident; that all men are created equal.” I have a dream that one day on the red hills of Georgia the sons of former slaves and the sons of former slave owners will be able to sit down together at the table of brotherhood.

Phonology: Vowels

	front	central	back
close	i		u
mid	e		o
open	a		

Name	Representation	Description	class
Search	Search	Search	vowel
i	2779 (92%)	LATIN SMALL LETTER I	vowel
u	2646 (88%)	LATIN SMALL LETTER U	vowel
a	2600 (86%)	LATIN SMALL LETTER A	vowel
e	1841 (61%)	LATIN SMALL LETTER E	vowel
o	1826 (60%)	LATIN SMALL LETTER O	vowel
ɛ	1129 (37%)	LATIN SMALL LETTER OPEN E	vowel
ɔ	1070 (35%)	LATIN SMALL LETTER OPEN O	vowel

Three diphthongs:
ai /aɪ̯/ · **au** /aʊ̯/ · **oi** /oɪ̯/

Based on PHOIBLE, WALS, and (diphthong frequencies) LAPSyD.

Phonology: Consonants

	Labial	Alveolar	Post-alv.	Palatal	Velar	Glottal
Nasal	m	n				
Plosive	p b	t d			k g	
Fricative	f	s	ʃ (x)			h
Affricate			t͡ʃ (c) d͡ʒ (j)			
Rhotic		r (r)				
Approximant		l		j (y)	w (v)	

Based on PHOIBLE and WALS.

Phonology: Consonants

	Labial	Alveolar	Pos				
Nasal	m	n					
Plosive	p b	t d				k g	
Fricative	f	s	ʃ (x)				h
Affricate			ʧ (c				
Rhotic		r (r)					
Approximant		l					

THE WORLD ATLAS OF LANGUAGE STRUCTURES ONLINE

Value	Representation
No voicing contrast	182
Voicing contrast in plosives alone	189
Voicing contrast in fricatives alone	38
Voicing contrast in both plosives and fricatives	158
Total:	567

Based on PHOIBLE and WALS.

Phonology: Syllable structure

Consonant pairs allowed to start a syllable:

- **bl, fl, gl, kl, pl, sl**
- **br, dr, fr, gr, kr, pr, tr**
- **by, cy, fy, ky, my, ny, py, xy**
- **cv, dv, gv, hv, kv, sv, tv, xv**

THE WORLD ATLAS OF LANGUAGE STRUCTURES ONLINE

	Value	Representation
○	Simple syllable structure	61
●	Moderately complex syllable structure	274
●	Complex syllable structure	151
	Total:	486

Consonants allowed to end a syllable:
l, m, n, r, s, t.

Based on WALS and own study of the 10 source languages.

Source Languages

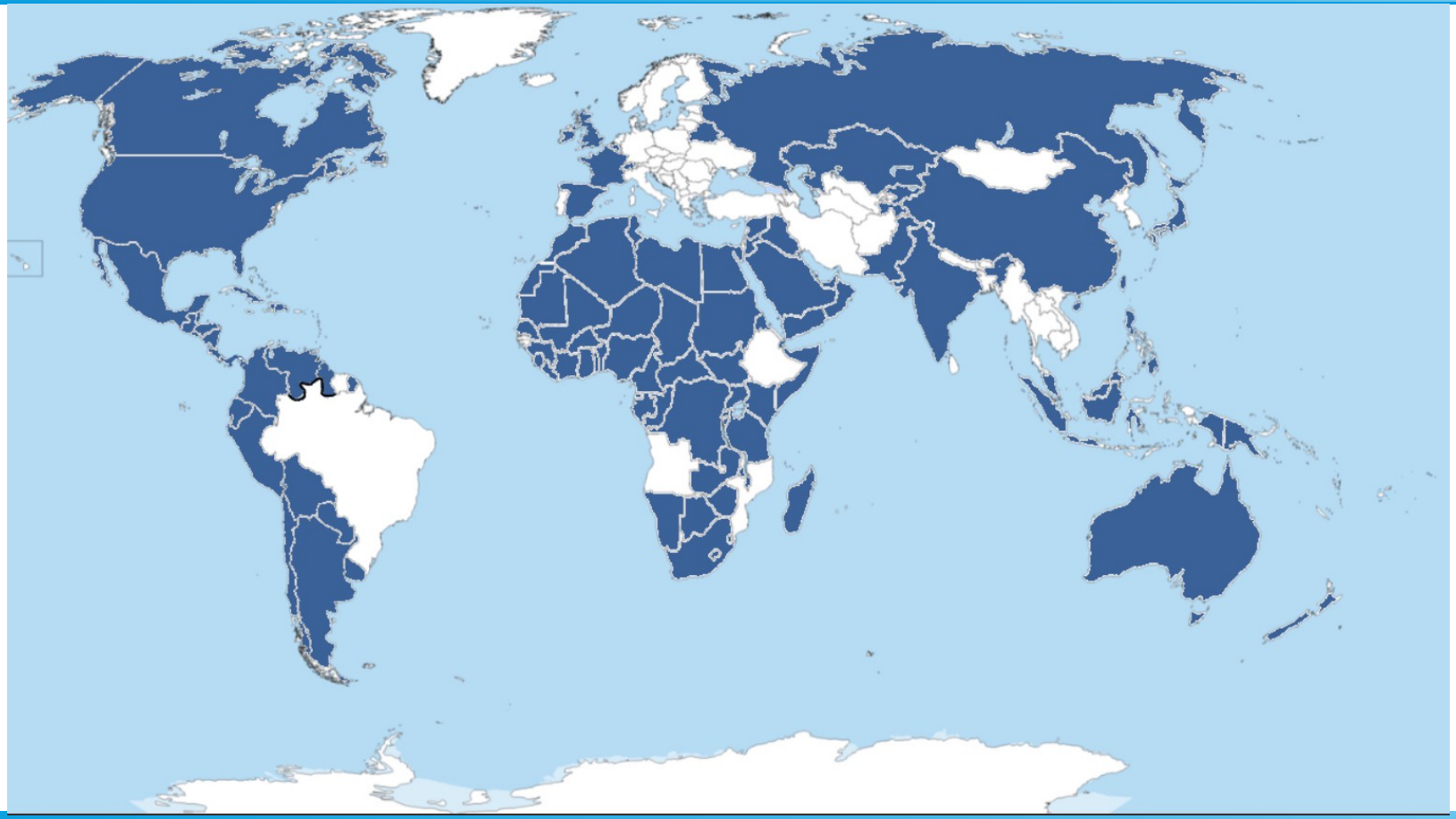
5x Indo-European:

- English (Germanic)
- Hindi/Urdu (Indo-Iranian)
- Spanish (Romance)
- Russian (Balto-Slavic)
- French (Romance again)

5 other language families:

- Mandarin (Sino-Tibetan)
- Swahili (Niger-Congo)
- Standard Arabic (Afroasiatic)
- Indonesian/Malay (Austronesian)
- Japanese (Japonic)

Source languages as official languages



The top 5

5x Indo-European:

- **English (Germanic)**
- **Hindi/Urdu (Indo-Iranian)**
- **Spanish (Romance)**
- Russian (Balto-Slavic)
- French (Romance again)

5 other language families:

- **Mandarin (Sino-Tibetan)**
- Swahili (Niger-Congo)
- **Standard Arabic (Afroasiatic)**
- Indonesian/Malay (Austronesian)
- Japanese (Japonic)

If a word doesn’t have a related candidate in another source language, it must come from one of the top 5 languages.

Word selection example: “peace”

```
++ Candidates with 4 related candidates:
[1]  hi:aman (P:0.667 – C:0.000+S:0.000+I:0.667, related candidates in 4
languages: ar:amne, id:damai, ja:hean, sw:amani)
++ Candidates with 2 related candidates:
[2]* ja:hean (P:0.143 – C:0.000+S:0.143+I:0.000, related candidates in 2
languages: cmn:hepin, hi:aman, hi:cen)
[3]  ja:pisu (P:0.147 – C:0.000+S:0.147+I:0.000, related candidates in 2
languages: en:pis, es:pas)
[4]  sw:amani (P:0.403 – C:0.000+S:0.180+I:0.222, related candidates in 2
languages: ar:amne, hi:aman)
[5]  en:pis (P:1.026 – C:0.000+S:0.137+I:0.889, related candidates in 2
languages: es:pas, ja:pisu)
[6]  es:pas (P:1.190 – C:0.000+S:0.190+I:1.000, related candidates in 2
languages: en:pis, ja:pisu)
[SKIPPED (too distorted)]  ar:amne (P:1.268 – C:0.400+S:0.312+I:0.556, related
candidates in 2 languages: hi:aman, sw:amani)
...
Candidate #1 "aman" added to the dictionary on 2022-09-08 23:27:29.
```

Word selection example: “peace”

```
++ Candidates with 4 related candidates:
[1]  hi:aman (P:0.667 – C:0.000+S:0.000+I:0.667, related candidates in 4
languages: ar:amne, id:damai, ja:hean, sw:amani)
++ Candidates with 2 related candidates:
[2]* ja:hean (P:0.143 – C:0.000+S:0.143+I:0.000, related candidates in 2
languages: cmn:hepin, hi:aman, hi:cen)
[3]  ja:pisu (P:0.147 – C:0.000+S:0.147+I:0.000, related candidates in 2
languages: en:pis, es:pas)
[4]  sw:amani
languages: ar
[5]  en:pis (P:1.026 – C:0.000+S:0.137+I:0.889, related candidates in 2
languages: es:pas, ja:pisu)
[6]  es:pas (P:1.190 – C:0.000+S:0.190+I:1.000, related candidates in 2
languages: en:pis, ja:pisu)
[SKIPPED (too distorted)]  ar:amne (P:1.268 – C:0.400+S:0.312+I:0.556, related
candidates in 2 languages: hi:aman, sw:amani)
```

Criterion 1: number of related candidates and similarity

aman – peace, from Hindi अमन (aman), Swahili *amani*, Arabic أَمْن (ʾamn), and loosely related to Indonesian *damai* and Japanese 平安 (heian).

Word selection example: “peace”

```
++ Candidates with 4 related candidates:
[1]  hi:aman (P:0.667 – C:0.000+S:0.000+I:0.667, related candidates in 4
languages: ar:amne, id:damai, ja:hean, sw:amani)
++ Candidates with 2 related candidates:
[2]* ja:hean (P:0.143 – C:0.000+S:0.143+I:0.000, related candidates in 2
languages: cmn:hepin, hi:aman, hi:cen)
[3]  ja:pisu (P:0.147 – C:0.000+S:0.147+I:0.000, related candidates in 2
languages: en:pis, es:pas)
[4]  sw:amani (P:0.403 – C:0.000+S:0.180+I:0.222, re
languages: ar:amne, hi:aman)
[5]  en:pis (P:1.026 – C:0.000+S:0.137+I:0.889, related candidates in 2
languages: es:pas, ja:pisu)
[6]  es:pas (P:1.190 – C:0.000+S:0.190+I:1.000, related candidates in 2
languages: en:pis, ja:pisu)
[SKIPPED (too distorted)]  ar:amne (P:1.268 – C:0.400+S:0.312+I:0.556, related
candidates in 2 languages: hi:aman, sw:amani)
```

Criterion 2: distortion

aman – peace, from Hindi अमन (aman), Swahili *amani*, Arabic أَمْن (ʾamn), and loosely related to Indonesian *damai* and Japanese 平安 (heian).

Word selection example: “peace”

```
++ Candidates with 4 related candidates:
[1]  hi:aman (P:0.667 – C:0.000+S:0.000+I:0.667, related candidates in 4
languages: ar:amne, id:damai, ja:hean, sw:amani)
++ Candidates with 2 related candidates:
[2]* ja:hean (P:0.143 – C:0.000+S:0.143+I:0.000, related candidates in 2
languages: cmn:hepin, hi:aman, hi:cen)
[3]  ja:pisu (P:0.147 – C:0.000+S:0.147+I:0.000, related candidates in 2
languages: en:pis, es:pas)
[4]  sw:amani (P:0.403 –
languages: ar:amne, hi:am
[5]  en:pis (P:1.026 – C:0.000+S:0.137+I:0.889, related candidates in 2
languages: es:pas, ja:pisu)
[6]  es:pas (P:1.190 – C:0.000+S:0.190+I:1.000, related candidates in 2
languages: en:pis, ja:pisu)
[SKIPPED (too distorted)]  ar:amne (P:1.268 – C:0.400+S:0.312+I:0.556, related
candidates in 2 languages: hi:aman, sw:amani)
```

Criterion 3: influence of the source language

aman – peace, from Hindi अमन (aman), Swahili *amani*, Arabic أَمْن (ʾamn), and loosely related to Indonesian *damai* and Japanese 平安 (heian).

Word selection example: “peace”

```
++ Candidates with 4 related candidates:
[1]  hi:aman (P:0.667 – C:0.000+S:0.000+I:0.667, related candidates in 4
languages: ar:amne, id:damai, ja:hean, sw:amani)
++ Candidates with 2 related candidates:
[2]* ja:hean (P:0.143 – C:0.000+S:0.143+I:0.000, related candidates in 2
languages: cmn:hepin, hi:aman, hi:cen)
[3]  ja:pisu (P:0.147 – C:0.000+S:0.147+I:0.000, related candidates in 2
languages: en:pis, es:pas)
[4]  sw:amani (P:0.403 – C:0.000+S:0.1
languages: ar:amne, hi:aman)
[5]  en:pis (P:1.026 – C:0.000+S:0.137+I:0.889, related candidates in 2
languages: es:pas, ja:pisu)
[6]  es:pas (P:1.190 – C:0.000+S:0.190+I:1.000, related candidates in 2
languages: en:pis, ja:pisu)
[SKIPPED (too distorted)]  ar:amne (P:1.268 – C:0.400+S:0.312+I:0.556, related
candidates in 2 languages: hi:aman, sw:amani)
...
Candidate #1 "aman" added to the dictionary on 2022-09-08 23:27:29.
```

Summed penalty (lower is better)

The final choice is always human

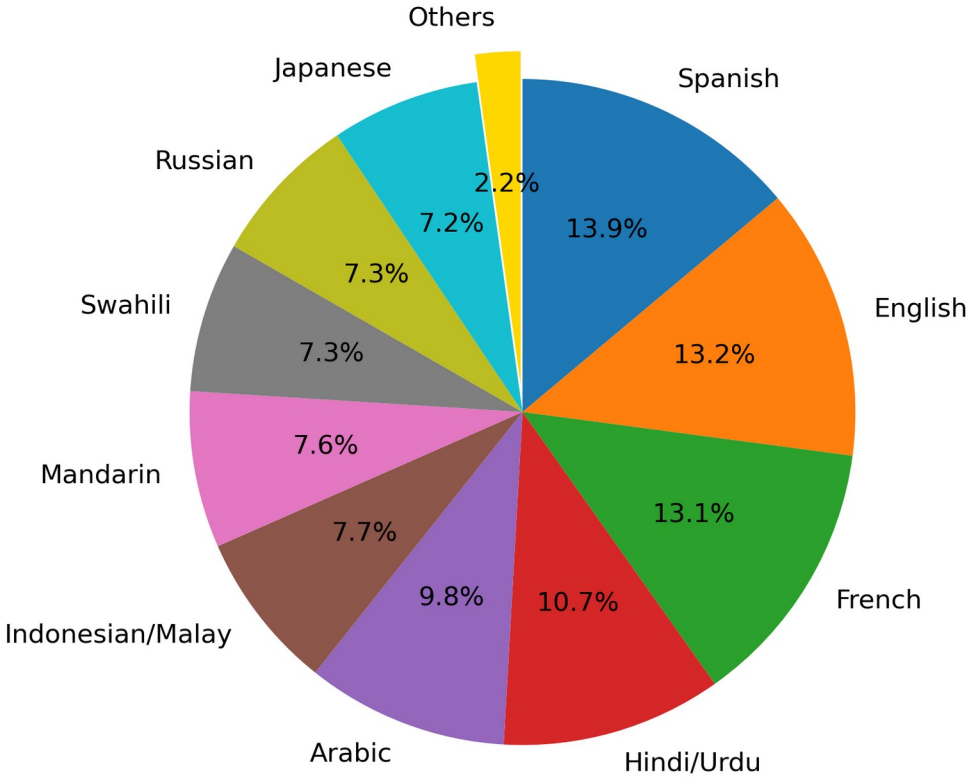
```
++ Candidates with 3 related candidates:
[1]* ru:etiketka (P:0.254 – C:0.000+S:0.142+I:0.111, related candidates in 3
languages: es:etiketa, fr:etiket, id:etiket)
[2]  id:etiket (P:0.433 – C:0.000+S:0.099+I:0.333, 1 identical candidate, related
candidates in 3 languages: es:etiketa, fr:etiket, ru:etiketka)
[3]  id:label (P:0.473 – C:0.000+S:0.140+I:0.333, related candidates in 3
languages: en:lebel, hi:lebal, ja:raberu)
[4]  hi:lebal (P:0.838 – C:0.000+S:0.171+I:0.667, related candidates in 3
languages: en:lebel, id:label, sw:lebo)
[5]  fr:etiket (P:0.877 – C:0.000+S:0.099+I:0.778, 1 identical candidate, related
candidates in 3 languages: es:etiketa, id:etiket, ru:etiketka)
[6]  es:etiketa (P:1.000 – C:0.000+S:0.000+I:1.000, related candidates in 3
languages: fr:etiket, id:etiket, ru:etiketka)
[7]  en:lebel (P:1.169 – C:0.200+S:0.080+I:0.889, related candidates in 3 ...
...
Candidate #2 "etiket" added to the dictionary on 2022-09-10 15:34:47.
Selection rationale: We prefer this candidate since it's shared by two languages
and one syllable shorter than #1, while being otherwise identical to it.
```

The final choice is always human

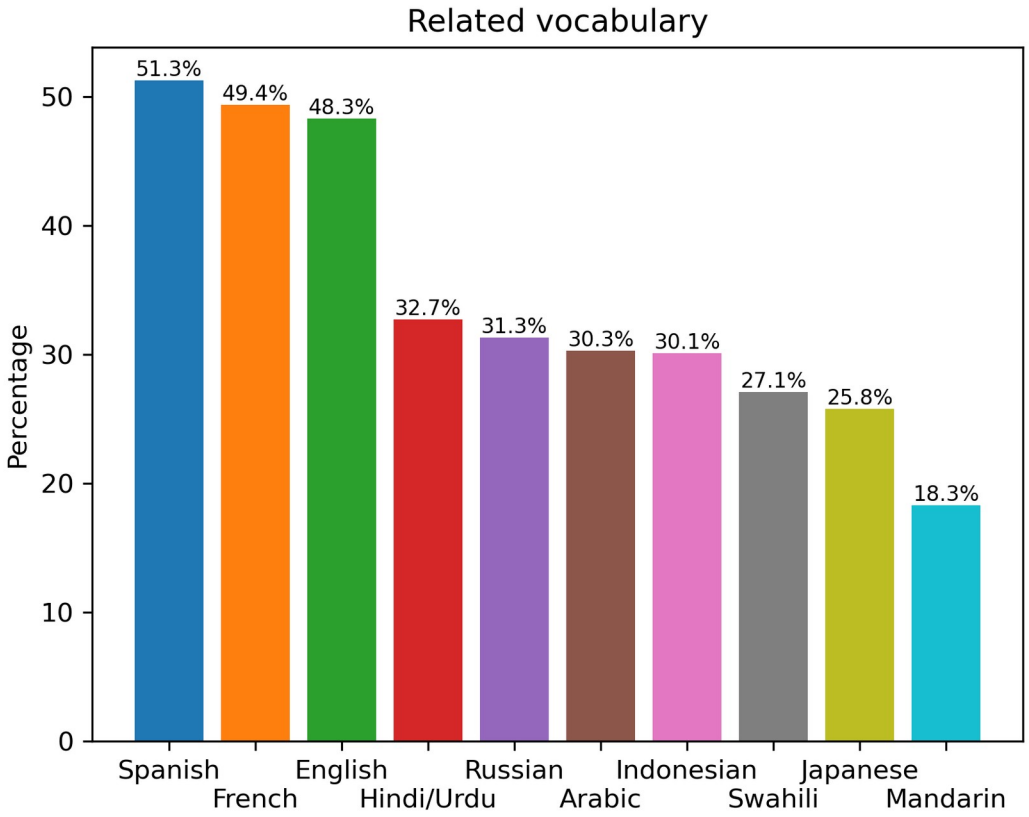
```
++ Candidates with 3 related candidates:
[1]* ru:etiketka (P:0.254 – C:0.000+S:0.142+I:0.111, related candidates in 3
languages: es:etiketa, fr:etiket, id:etiket)
[2]  id:etiket (P:0.433 – C:0.000+S:0.099+I:0.333, 1 identical candidate, related
candidates in 3 languages: es:etiketa, fr:etiket, ru:etiketka)
[3]  id:label (P:0.473 – C:0.000+S:0.140+I:0.333, related candidates in 3
languages: en:lebel, hi:lebal, ja:raberu)
[4]  hi:lebal (P:0.838 – C:0.000+S:0.171+I:0.667, related candidates in 3
languages: en:lebel, id:label, sw:lebo)
[5]  fr:etiket (P:0.877 – C:0.000+S:0.099+I:0.778, 1 identical candidate, related
candidates in 3 languages: es:etiketa, id:etiket, ru:etiketka)
[6]  es:etiketa (P:1.000 – C:0.000+S:0.000+I:1.000, related candidates in 3
languages: fr:etiket, id:etiket, ru:etiketka)
[7]  en:lebel (P:1.169 – C:0.200+S:0.080+I:0.889, related candidates in 3 ...
```

etiket – label, from Indonesian *etiket*, French *étiquette*, Spanish *etiqueta*, and Russian этикетка (etikétka).

Influence distribution



Related vocabulary percentages



Frequent Source Language Combinations

Combinations involving three languages:

- English/Spanish/French: 34.3%
- Spanish/French/Russian: 21.3%
- English/Spanish/Russian: 20.3%
- English/French/Russian: 20.0%
- Spanish/French/Indonesian: 15.6%
- English/French/Indonesian: 15.1%
- English/Spanish/Indonesian: 14.9%
- Spanish/French/Hindi: 14.1%
- Spanish/Indonesian/Russian: 13.7%
- English/Indonesian/Russian: 13.6%
- French/Indonesian/Russian: 13.5%
- English/French/Hindi: 13.1%

Combinations involving two languages:

- Spanish/French: 42.5%
- English/French: 37.3%
- English/Spanish: 37.0%
- Spanish/Russian: 24.1%
- English/Russian: 23.2%
- French/Russian: 23.2%
- Arabic/Swahili: 18.1%
- Spanish/Indonesian: 17.6%
- French/Indonesian: 17.6%
- English/Indonesian: 17.4%
- English/Japanese: 16.1%
- Spanish/Hindi: 16.0%
- French/Hindi: 16.0%

Sample words and their origin

reforme – reform (verb+noun)
From French *réformer, réforme*, English *reform*, Spanish *reformar, reforma*, Hindi रिफ़ार्म (rifārm), and Russian реформировать, реформа (reformírovat', refórma).

sali – pray
From Swahili *-sali* and Arabic صَلَّى (ṣallā).

sereal – cereal (type of grass and its grains), grain
From Indonesian *sereal*, English/Spanish *cereal*, and French *céréale*.

sore – evening
From Indonesian *sore* and loosely related to French *soir* and Japanese 夜 (yoru).
Derivative: **sore hau** – good evening (greeting)

sosa – suck
From Russian сосать (sosát') and loosely related to Arabic مَصَّ (maṣṣa), Chinese 咂 (za), French *sucer*, and Swahili *-fyonza*.

tenko – weather
From Japanese 天候 (tenkō) and loosely related to Chinese 天氣/天气 (tiānqì) and Spanish *tiempo*.

State and community

- Project started in Feb. 2021
- Publicly announced in Aug. 2021 (on Discord and Reddit)
- By now (Nov. 2022) there are 1400+ words in the dictionary, 870 of which are roots
- Community:
  - Website (open wiki): https://www.lugamun.org/

    Dictionary · Grammar · Background documents · Several translation projects · Statistics page …
  - Reddit: r/auxlangs

  - Discord Server: auxlangs / helplingvoj (#lugamun channel)

Thank you and … / Xukur va …

Questions?

Yo ven ka?

https://www.lugamun.org/