



Challenge *Sprachenerkennung* (Abgabe bis 13. Dezember)

Schreibe ein Programm, das erkennen kann, in welcher (von mindestens vier) Sprachen ein Text geschrieben ist!



Häufigkeitsprofil einer Sprache

- **Beobachtung:** Jede Sprache hat ein spezifisches **Buchstabenhäufigkeitsprofil**
- Beispiel Deutsch (nur A-Z): (Wikipedia)
 - relative Häufigkeit von A: 6.51%
 - relative Häufigkeit von B: 1.89%
 - relative Häufigkeit von C: 3.06%
 -
 - relative Häufigkeit von Z: 1.13%



Häufigkeitsprofil einer Sprache

- **Beobachtung:** Jede Sprache hat ein spezifisches **Buchstabenhäufigkeitsprofil**
- Beispiel Deutsch (nur A-Z):

	Englisch
■ relative Häufigkeit von A: 6.51%	8.17%
■ relative Häufigkeit von B: 1.89%	1.49%
■ relative Häufigkeit von C: 3.06%	2.78%
■	
■ relative Häufigkeit von Z: 1.13%	0.07%



In welcher Sprache ist ein Text geschrieben?

- Idee eines Erkennungsprogramms
 - Lies einen Text ein und bestimme die relativen Häufigkeiten der Buchstaben A-Z im Text
 - Vergleiche das so erhaltene Profil des Textes mit den Profilen verschiedener Sprachen
 - Gib die Sprache aus, deren Sprachprofil dem Textprofil am nächsten ist.



Ähnlichkeit von Profilen

- Sprachprofil (p_A, p_B, \dots, p_Z)
- Textprofil (t_A, t_B, \dots, t_Z)
- Mögliches Abstandsmass:

$$\sum_{\rho=A..Z} (t_{\rho} - p_{\rho})^2$$



Testdaten (Deutsch)

Und wenn ein Lied meine Lippen verlässt
(Söhne Mannheims)

und wenn ein lied meine lippen verlässt,
dann nur damit du liebe empfängst.
durch die nacht und das dichteste geäst
damit du keine ängste mehr kennst

sag ein kleines stückchen wahrheit
sieh wie die wüste lebt
schaff ein kleines bisschen klarheit
und schau wie sich der schleier hebt
eine wüste aus beton und asphalt
doch sie lebt und öffnet einen spalt
der dir neues zeigt
zeigt das alte weicht
auch wenn dein schmerz bis an
den himmel reicht

und wenn ein lied, meine lippen verlässt,
dann nur damit du liebe empfängst
durch die nacht und das dichteste geäst
damit du keine ängste mehr kennst

dieses lied ist nur für dich
schön, wenn es dir gefällt
denn es kam so über mich
wie die nacht über die welt
schlanke faust der dunkelheit
bin ich zum ersten schlag bereit

ich bin der erste der dich befreit
und einer der letzten der um dich weint

in unserer sanduhr fällt das letzte korn
ich hab gewonnen und hab ebenso verlorn
jedoch missen möchte ich nichts
alles bleibt unser gedanklicher besitz
und eine bleibende erinnerung
zwischen tag und nacht legt sich die
dämmerung



Testdaten (Englisch)

Yesterday
(The Beatles)

Yesterday, all my troubles seemed so far away
Now it looks as though they're here to stay
Oh, I believe in yesterday

Suddenly, I'm not half the man I used to be
There's a shadow hanging over me.
Oh, I yesterday came suddenly

Why she had to go I don't know she wouldn't say
I said something wrong, now I long for yesterday

Yesterday, love was such an easy game to play
Now I need a place to hide away
Oh, I believe in yesterday

Why she had to go I don't know she wouldn't say
I said something wrong, now I long for yesterday

Yesterday, love was such an easy game to play
Now I need a place to hide away
Oh, I believe in yesterday



Testdaten (Italienisch)

Felicità
(Al Bano & Romina Power)

Felicità e tenersi per mano
andare lontano la felicità.
E il tuo sguardo innocente
in mezzo alla gente la felicità.
E restare vicini come bambini
la felicità, felicità.

Felicità e un cuscino di piume
l'acqua del fiume che passa e va.
E' la pioggia che scende
dietro le tende la felicità.
E abbassare la luce per fare pace
la felicità, felicità.

Felicità e un bicchiere di vino
con un panino la felicità.
E lasciarti un biglietto dentro al cassetto la felicità
E cantare a due voci
quanto mi piaci la felicità, felicità.

Senti nell'aria c'è già
la nostra canzone d'amore che va.
Come un pensiero che sa di felicità.

Senti nell'aria c'è già
un raggio di sole più caldo che va
come un sorriso che sa di felicità.

Felicità e una sera a sorpresa
la luce accesa e la radio che va
e' un biglietto d'auguri pieno di cuori la felicità.
E una telefonata non aspettata
la felicità, felicità.

Felicità e una spiaggia di notte
l'onda che batte la felicità.
E una mano sul cuore piena d'amore la felicità.
E aspettare l'aurora per farci ancora
la felicità, felicità.

Senti nell'aria c'è già . . .

Felicità



Testdaten (Französisch)

Non, je ne regrette rien
(Edith Piaf)

Non, rien de rien
Non, je ne regrette rien
Ni le bien qu'on m'a fait, ni le mal
Tout ça m'est bien égal
Non, rien de rien
Non, je ne regrette rien
C'est payé, balayé, oublié
Je me fous du passé

Avec mes souvenirs
J'ai allumé le feu
Mes chagrins, mes plaisirs
Je n'ai plus besoin d'eux
Balayés mes amours
Avec leurs trémolos
Balayés pour toujours
Je repars à zéro

Non, rien de rien
Non, je ne regrette rien
Ni le bien qu'on m'a fait, ni le mal
Tout ça m'est bien égal
Non, rien de rien
Non, je ne regrette rien
Car ma vie
Car mes joies
Aujourd'hui
Ça commence avec toi...



Testdaten (Deutsch?)

Baby
(Subzonic)

en tag wie jede anderi, nüüt isch andersch xii, i de city voll mit
lüüt, es laufed alli anenand verbii, eine zieht siis handy, rännt a de
fenschterwand verbii, chunnt uf miich zue, xeet mi nöd und rännt i
miich, ich flueche vor miich hii, da bemerk ich sie, sie lächled miich
aa und zieht flott a miir verbii, ich traue miine auge nöd, dräh mi
nomol um, aber sie isch e wäg und verschwunde

baby, säg wo chunnsch du her, bisch mer eifach devoo, wie wänn i luft
für di wär, baby, säg wo gaasch du hii, so schnäll wie du cho bisch,
bisch a miir verbii

sii isch undertaucht, inmitte vo de mängi, ich laufe sofort naa und
dräng mi dur d'ängi, ich chan sie nöd mal rüefe, ich känn nöd mal ihre
name, ich xee sie vorne laufe, doch ich chum nöd a sie ane, es isch
wie wänn das alles in ziitlupe passiert, wie am ferseh, wänn
d'schwiizer nati es goal kassiert, de traum nimmt es ändi, e tragischi
wändi, wo sie in zug gaat und de vor miir abfahrt

die begägnig hätt en schluss, aber nie en aafang gnoo, bis uf villi,
offni frage hätt sie nüüt übrig glaa

bevor i zrugglauf, dräh i mi nomol um, aber sie isch e wäg und
verschwunde

miin zug fahrt ab, ich beobachte ihn zum fenschter uus, han gar nöd
gschnallt, dass er miir g'folgt isch, ich schtaane uf und rüef so luut
i chan zum fenschter uus, aber er isch e wäg und verschwunde, er isch
wäg und verschwunde



Testdaten (Deutsch?)

I han es Zündhölzli azündt
(Mani Matter)

i han es zündhölzli azündt
und das het e flamme gäh
und i ha für d'zigarette
welle füür vom hölzli näh
aber ds hölzli isch dervo-
gspickt und uf e teppich cho
und es hätt no fasch es loch i teppich gäh dervo
ja me weis was cha passiere
we me nid ufpasst mit füür
und für d'gluet ar zigarette
isch e teppich doch de z'tüür
und vom teppich hätt o grus
chönne ds füür i ds ganze hus
und wär weis, was da nid alles no wär worde drus

s'hätt e brand gäh im quartier
und s'hätti d'füürwehr müesse cho
hätti ghornet i de strasse
und dr schluch vom wage gno
und sie hätte wasser gsprützt
und das hätt de glych nüt gnützt
und die ganzi stadt hätt brönnt, es hätt se nüt meh
geschützt

und d'lüt wären umegsprunge
i dr angscht um hab und guet
hätte gmeint s'heig eine füür gleit
hätte ds sturmgewehr gno ir wuet
alls hätt brüelet: wär isch tschuld?
ds ganze land i eim tumult
dass me gschosse hätt uf d'bundesrät am rednerpult

d'uno hätt interveniert
und d'uno-gägner sofort o
für ir schwyz dr fride z'rette
wäre beid mit panzer cho
s'hätt sech usdehnt nadisnah
uf europa, afrika
s'hätt e wältchrieg gäh und d'mönschheit wär jitz
nümme da

i han es zündhölzli azündt
und das het e flamme gäh
und i ha für d'zigarette
welle füür vom hölzli näh
aber ds hölzli isch dervo-
gspickt und uf e teppich cho -
gottseidank dass i's vom teppich wider furt ha gno



Testdaten (Deutsch??)

Fiesta Mexicana
(Rex Gildo)

Hossa, Hossa, Hossa, Hossa
Fiesta, Fiesta Mexicana, heut geb
ich zum Abschied für alle ein Fest.
Fieste, Fiesta Mexicana es gibt viel
Tequila der glücklich sein läßt.
Alle Freunde sind hier, feiern
noch einmal mit mir.
Wir machen Fiesta, Fiesta Mexicana,
weil ihr dann den Alltag, die Sorgen
schnell vergesst.
Adio, Adio Mexico,
ich komme wieder zu dir zurück
Adio, Adio Mexico
ich grüß mit meinem Sombrero, te
quiero, ich habe dich so lieb.

Fiesta, Fiesta Mexicana auf der
kleinen Plaza da lacht man und singt.
Fiesta, Fiesta Mexicana, wenn zum
letzten Tanz die Gitarre erklingt.
Juanita, Pepe ja die zwei sagen
noch einmal Goodbye, wir machen
Fiesta, Fiesta Mexicana, weil das
bunte Leben die Liebe zu uns
bringt.

Adio, Adio Mexico,
ich komme wieder zu dir zurück
Adio, Adio Mexico
ich grüß mit meinem Sombrero, te
quiero, ich habe dich so lieb.

Hossa, Hossa, Hossa, Hossa
Fiesta, Fiesta Mexicana und schon
wird es hell und der Morgen ist nah.
Und ich küsse Carmen Sita, denn der Abschied ist nah.
Weine nicht muss ich gehn, weil wir
uns ja wieder sehn, bei einer
Fiesta, Fiesta Mexicana bald wird
wieder alles so schön wie es mal war.

Hossa

Fiesta, Fiesta Mexicana bald wird
wieder alles so schön,
ja du wirst sehn, darum bin ich bald
wieder da.

Hossa, Hossa, Hossa, Ole



Testdaten

- Eigene Testdaten aus dem Internet