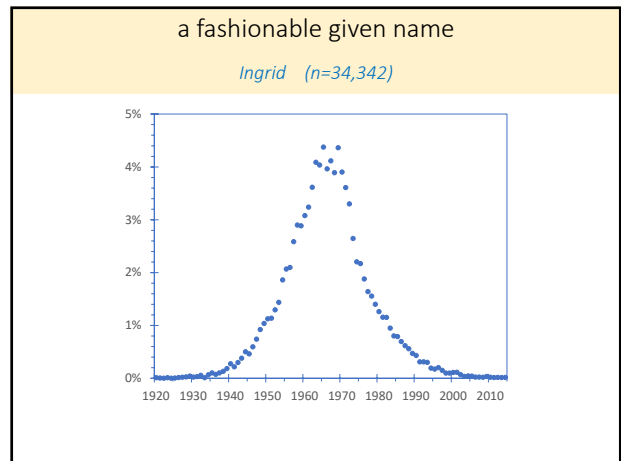


The emergence of Zipf's law in fashionable given names

Gerrit Bloothoof  
 Utrecht institute of Linguistics,  
 the Netherlands  
 g.bloothoof@uu.nl

1



2

understanding diffusion of fashion

given names are ideal

- no external influences (no advertisement)
- **all** namings are documented

3

work flow

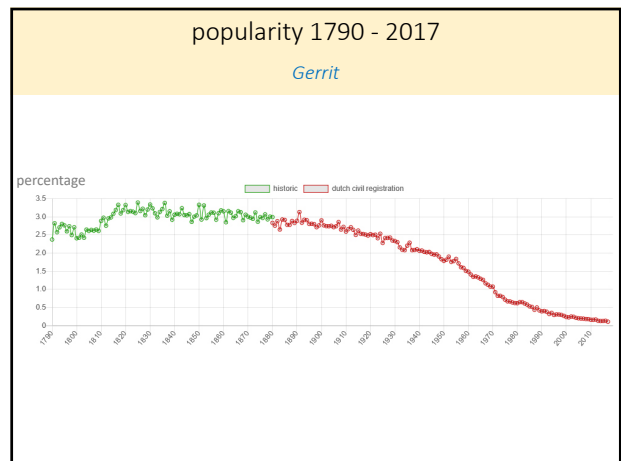
1. collecting birth dates of **all** children with a fashionable name
2. developing a **model of diffusion** of fashionable names (on the basis of the previous data)
3. studying the emergence of **Zipf's law** in fashionable names (predicted on the basis of the previous model)

4

a corpus of given names in the Netherlands

- 1790 - 2017
- given names of 25,9 million persons (total 35,2 million)
- their date of birth

5

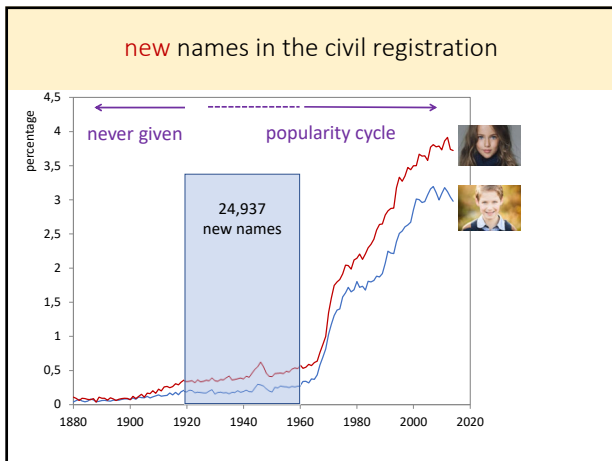


6

**Dia 1**

---

**BG(1)** Bloothoof, G. (Gerrit); 1-5-2019



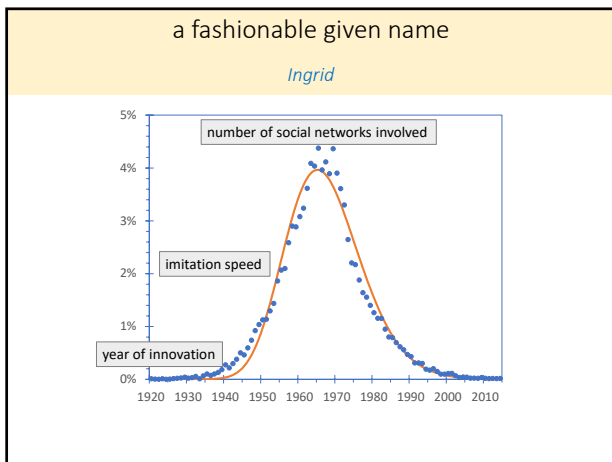
7

### a model of diffusion of new names

requires:



1. year of the innovation
2. imitation probability (imitation speed)
3. number of social networks for propagation

8



9

### (1) year of innovation

	1920 Robin		1920 Mariska
	1925 Luuk		1920 Marjolein
	1925 Roy		1925 Jessica
	1927 Remco		1927 Fleur
	1944 Milan		1930 Tamara
	1945 Danny		1934 Merel
	1946 Wesley		1936 Kim
	1948 Kevin		1937 Chantal
	1949 Jeffrey		1941 Wendy
			1945 Anouk
			1946 Amber
			1948 Britt
			1953 Melissa
			1955 Mandy
			1955 Cindy
			1958 Romy

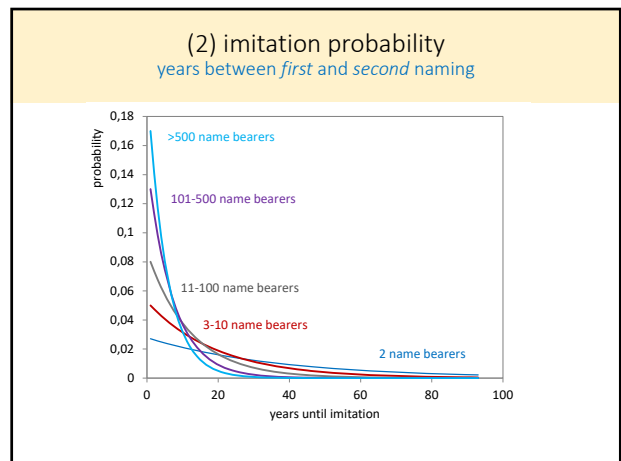
10

### (2) imitation probability

the probability of the time until *imitation* of the innovation

depends on future popularity

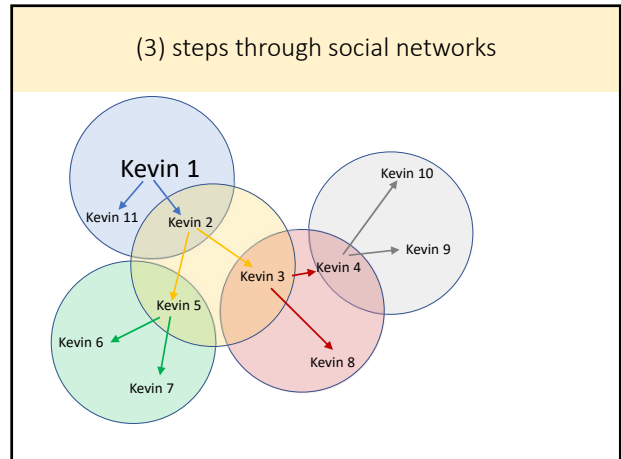
11



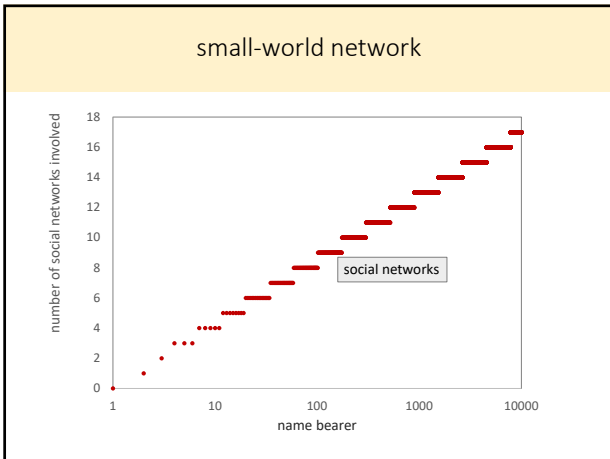
12

future popularity is built in

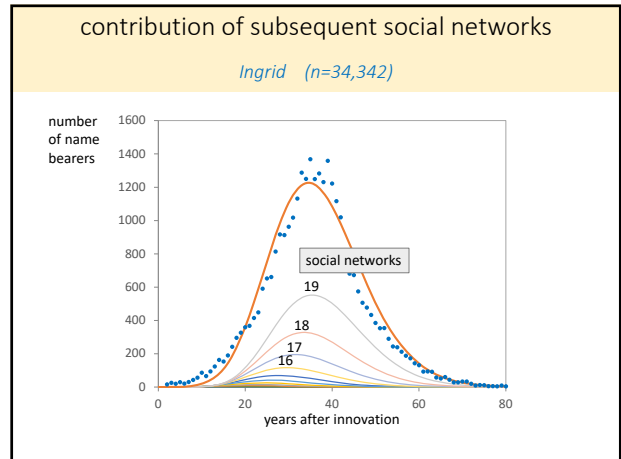
13



14



15



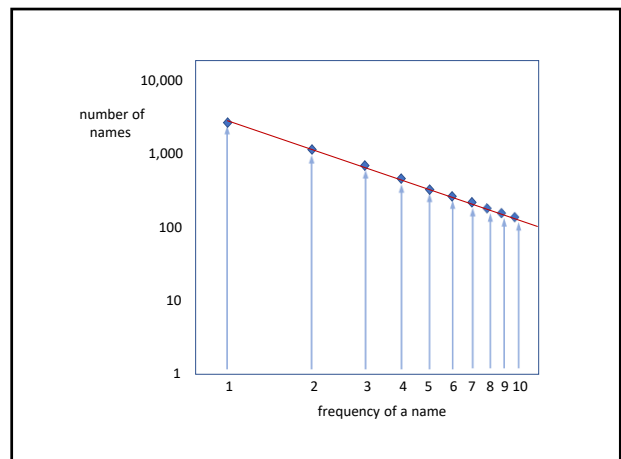
16

Zipf's law

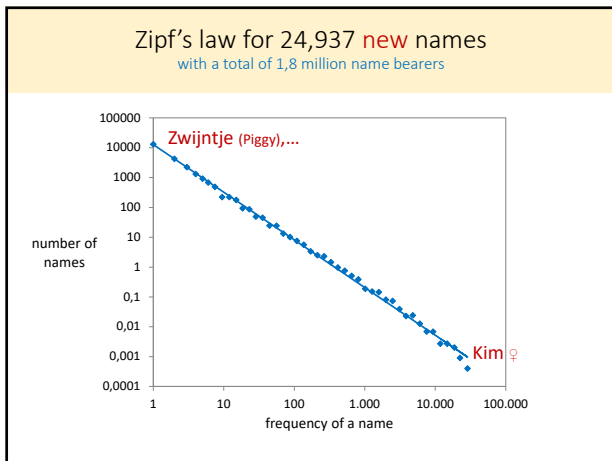
- many rare names
- a few very popular names

count number of names with frequency 1, 2, 3, 4,..... 100,..... 10,000

17



18

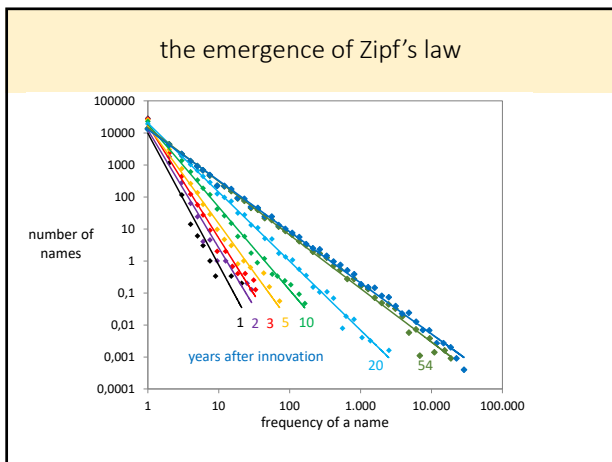


19

### Zipf's law *from time zero*

- synchronise names at year of innovation
- time zero: 24,937 names with frequency 1
- Zipf's law after 1 year, 2 year, ..., 54 years

20



21

### achievements so far

- given name data (1790-2017)
- a diffusion model for fashionable names
- Zipf's law shown from time zero onwards

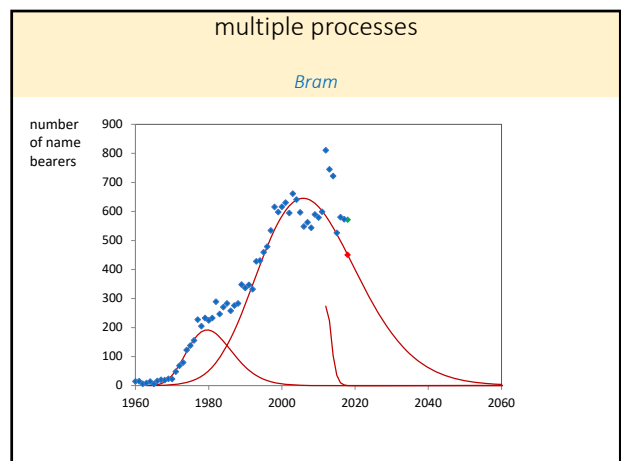
*but*

- how is future popularity built in a name?

22



23



24