

# Danish Anglicisms:

Corpora, dictionaries and news archives  
as finders and keepers

Henrik Gottlieb, University of Copenhagen

GLAD Symposium  
Cáceres, Spain, May 10, 2018



## Delimiting Anglicism candidates: Time

### **No pre-Anglomanie borrowings:**

\**Kirke* (from Old English *cyrice* > church)

\**Blæk* (from Old English *blæc* > ink)

### **Only items frequent after c. 1900:**

*Shagtobak* (frequent in the the 1900s; now obsolete)

*App* (frequent in the 2000s; unknown in the 1900s)

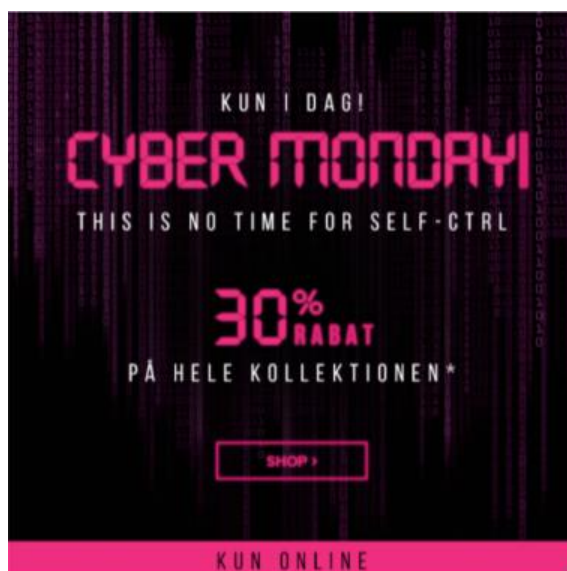
## Delimiting candidates: Text type & genre

### Only general-language items:

- No specialist technical expressions (e.g. chemical terms)
- No ephemeral (subculture-specific) slang terms
- No brand names (e.g. \*Google<sup>®</sup>) – only derivatives
- No personal and place names

EFL > ESL-related  
delimitation problem  
(a la South Africa):

**Anglicisms  
or code-switches?**





**You drive it.  
You define it.**

God stil handler om at turde være dig selv. Præcis som den helt nye KONA. Fuld af karakter og individualitet, og med et unikt, selvsikkert look, der adskiller den fra samtlige SUV'er i klassen. For det er ikke sjovt med mere af dét, man allerede kender.

Du har styringen over dit liv.

© Hyundai / denne brochure viser bilen med ekstraudrustning eller valgt, der er afhængig af variant. Der kan opkræves særlige, som ikke tilbydes i Danmark.

Finishing ninth grade at 53: "Yes, I did it!"



**GITTE OLSEN**  
MONTRICEPIGE

**Jeg har faktisk kun ottende klasse.**

02:35

## Structure

### Not to be included

#### Sentence-shaped code-switches:

*You drive it. You define it.*

(Heading in Danish Hyundai Kona brochure, 2017)

#### Loan creations

*færdigret* (synonym of *fast-food*)

#### Frequency-boosted items:

*ekspert* (ousting *fagmand*)

#### Transparent spin-offs:

*anti-amerikansk*; *online* + NP

#### Transparent hybrids:

*brytkræftscreening*; *screeningprogram*

### To be included

#### Imbedded stock phrases:

*(...), you name it!*

(Standard synonym of *etc.*)

#### Loan translations:

*kernefamilie* (< nuclear family)

**Semantic borrowings:** *morgen* (= forenoon)

#### Productive morphemes:

*anti-* ; *mega-* ; *post-* ; *online*

#### Non-transparent spinoffs:

*antiballistisk*; *antipsykiatri*

#### Hybrids including an original RL element:

*bijob* (= extra job)

## Provenance

### Not to be included

#### Relayed non-Anglosphere items:

*blitz*; *sushi*

#### Interference in expat discourse:

(all types of English influence)

### To be included

#### Relayed non-English Anglosphere items:

*Jungle*; *burrito*

#### Pseudo-Anglicisms

*Stationcar*, *happy end*, *bake-off*

## Delimiting candidates: Frequency

### **Cognitive criterion:**

Known to many educated speakers  
at any time between 1900 and 2020

### **Lexicographic criterion:**

**Included in standard or foreign-word dictionaries**

at any time between 1900 and 2020

### **Statistical criterion:**

Normalized frequency above a motivated threshold in general-language corpora  
at any time between 1900 and 2020

## Finding Danish candidates

### **Cognitive criterion** (do it yourself):

21st-century Danish discourse:

Personal newspaper-based and oral collection Nov. 2017 – April 2018:

c. **4,000** supposedly post-1997 candidates.

### **Lexicographic criterion** (use the work of others):

20th-century discourse: *A Dictionary of Anglicisms in Danish* (Sørensen 1997): **6,180** items

Pre-1955 discourse: *Ordbog over det Danske Sprog* (standard dictionary of Danish)

Post-1955 discourse: *Nye Ord i Dansk* (including **4,573** Anglicisms; <https://dsn.dk/noid>)

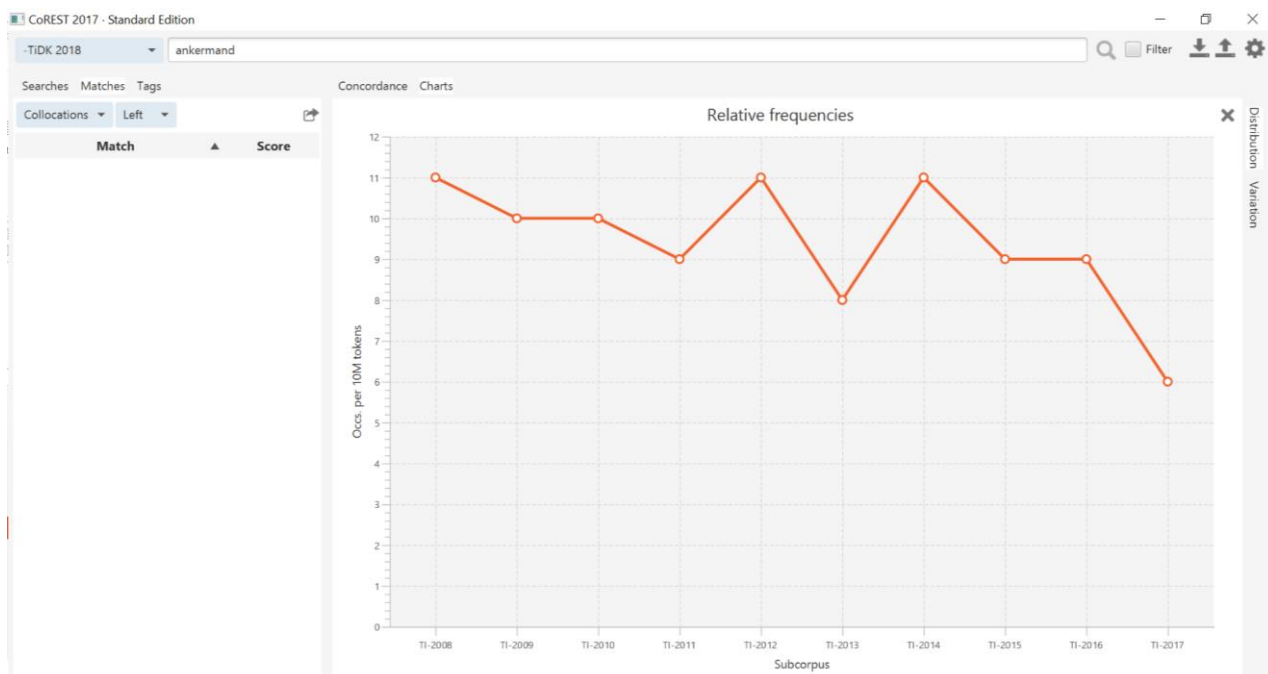
## Testing Danish candidates

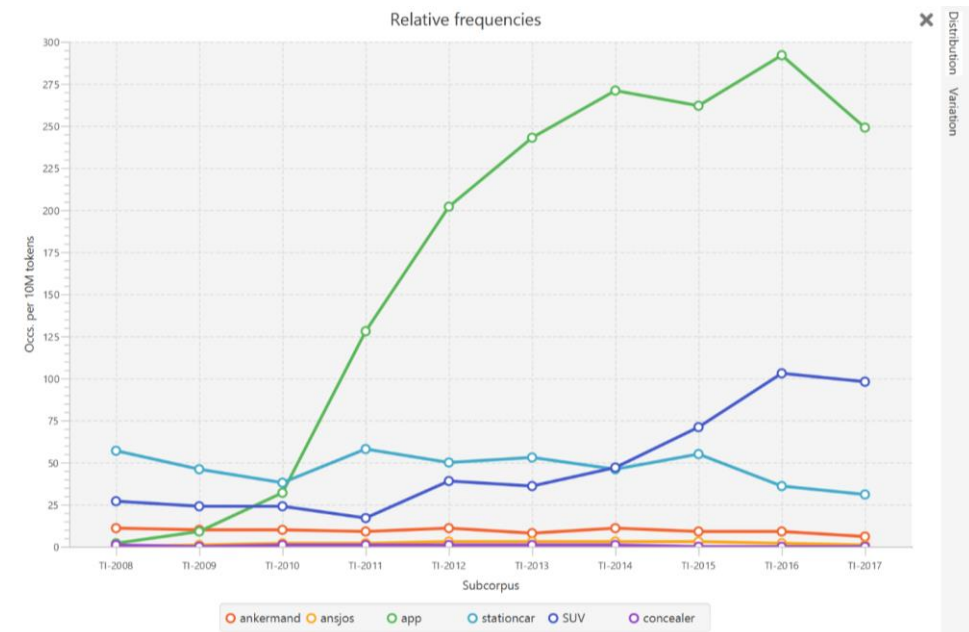
**Statistical criterion** for contemporary items (1990-):

Above-threshold frequency ( $> 0.008$  ppm) in at least two Danish corpora

**Lexicographic criterion** for pre-1990 items:

Inclusion in at least two major Danish dictionaries

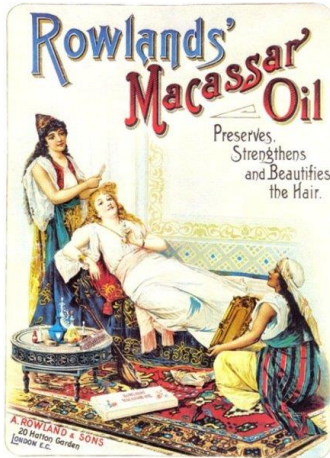




## Makassar – a city in Indonesia



Antimacassar: an obsolete term, perhaps ...

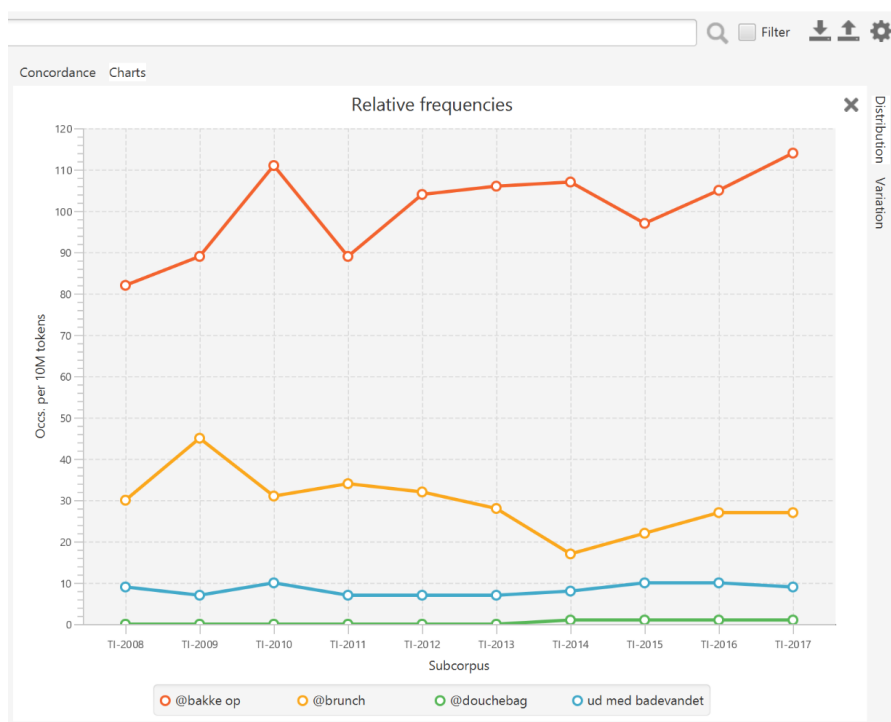


... but still useful (yet rarely used)





Danish Anglicism candidates	Dating	ADL -1940 14 million	Meyer -1924	ODS -1955	Korpus DK 1983-2002 56 million	DDO 1955-	TiDK corpus 2008-2017 384 million	Infomedia 1990-2018 20,000 million	Inclusion ?
antimakassar	1865	0	✓	✓	3 (0.054)	✓	0	20 (0.002)	*
[VP] ud med badevandet	1884	1 (0.07)	---	✓	17 (0.304)	✓	317 (0.826)	7,998 (0.800)	**
cykle	1892	6 (0.43)	✓	✓	1,538 (27.46)	✓	11,580 (30.16)	357,616 (35.76)	***
shag(tobak)	1919	3 (0.21)	✓	✓	0	✓	1 (0.003)	61 (0.006)	*
bakke op	1938	0	---	✓	1,135 (20.27)	✓	3,870 (10.08)	666,661 (66.67)	***
brunch	1974	0	∅	∅	73 (1.30)	✓	1,132 (2.95)	41,921 (4.92)	**
antiballistisk	1987	0	∅	∅	3 (0.054)	∅	3 (0.008)	88 (0.009)	*
and all that jazz	1998	0	∅	∅	1 (0.018)	∅	8 (0.021)	350 (0.035)	*
douchebag	2010	0	∅	∅	0	✓	6 (0.016)	200 (0.020)	*



- A-B items in DEA (Dictionary of European Anglicisms): c. 400 etymons
- All items in DEA: c. 3,800 etymons (Görlach 2003)

**A-B items in Danish database: c. 1,200 words = c. 1,000 etymons**

- All Danish items (following DEA's profile): c. 9,500 words = c. 7,900 etymons
- All Danish items (following DDO's profile): c. 9,450 words = c. 7,850 etymons

**Expected size of complete GLAD database: in excess of 10,000 etymons**

**Scope: 16 European languages + Japanese, Mandarin, Cantonese & Arabic**