

Det Centrale Ordregister (COR)

Peter Juel Henriksen, *Dansk Sprognævn*

Bolette S. Pedersen, *Center for Sprogteknologi, KU*

Sanni Nimb, *Det Danske Sprog- og Litteraturselskab*



DET DANSKE SPROG- OG
LITTERATURSELSKAB

Dansk
Sprognævn



Det Centrale Ordregister (COR)



- COR er et samarbejde mellem DSN, DSL og CST
- COR-projektet løber 2021-2023
- Projektet skaber en grundresurse for dansk sprogteknologi
 - *Rammeværk (hvert dansk ord og lemma får et indeks)*
 - *Orddatabaser (de centrale lemmaer forsynes med betydning)*

Det Centrale Ordregister – rammeværket

Det Centrale Ordregister – rammeværket

CPR

141294-1234 ←



020587-4321 ←



170134-1122 ←



Det Centrale Ordregister – rammeværket

CPR

COR

141294-1234 ←



banan → COR.23456

020587-4321 ←



stærk → COR.34567

170134-1122 ←



køre → COR.45678

jeg → COR.00123

fordi → COR.00234

Det Centrale Ordregister – rammeværket

"banan"

COR.23456

Det Centrale Ordregister – rammeværket

"banan"	COR.23456
banan	COR.23456.101
bananen	COR.23456.102
bananer	COR.23456.103
bananerne	COR.23456.104



banan sb., -en, -er, -erne, i sms. banan-, fx banankage, bananpalme

Det Centrale Ordregister – rammeværket

"banan"	COR.23456
banan	COR.23456.101
bananen	COR.23456.102
bananer	COR.23456.103
bananerne	COR.23456.104



banan sb., -en, -er, -erne, i sms. banan-, fx banankage, bananpalme




Den Danske Ordbog



Den Danske Ordbog • DET DANSKE SPROG- OG LITTERATURELSKAB

↳ Gå til ordbogens hovedside: ordnet.dk/ddo

banan  substantiv, fælleskøn (opslaget er forkortet – læs hele artiklen på ordnet.dk)

-en, -er, -erne

FON [ba'næ'n]

OPRINDELSE fra spansk eller portugisisk *banana*, oprindelig vestafrikansk ord

- 1 krum, aflang frugt fra bananpalmen, med gul skræl og blødt, sødt frugtkød • vokser på palmen *Musa paradisiaca*
- 2 (*slang*) krone eller anden form for penge

SYN [dask](#) | [bobs](#) | [gryn](#)

- 3 (*slang*) penis

[hård banan](#)

- person der virker hård og ufølsom – *kendt fra 1967, uformelt*

SYN [hård negl](#)

[sløv banan](#)

- langsom, ugidelig og doven person – *uformelt*

SYN [sløv padde](#)

[vride bananen \(regnormen, ...\)](#)

- (*slang*) tisse • om mand

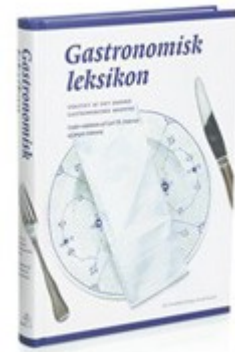
Den Store Dansk Encyklopædi



INGEFÆRORDENEN (ZINGIBERALES)

banan

Musa, slægt i bananfamilien med ca. 45 vildtvoksende arter i tropisk Afrika og Asien. Den er blandt verdens største urter og er ofte træliggende, da bladskederne danner en falsk stamme. Bladpladen kan være op til 3-4 m lang



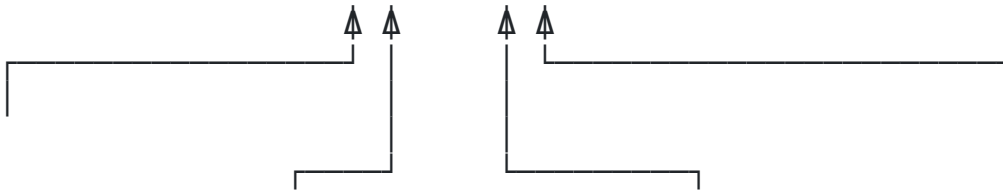
Søg

melbanan

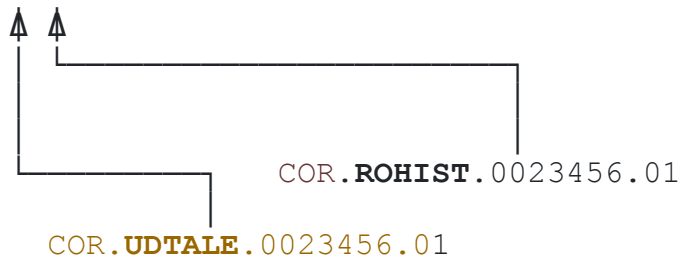
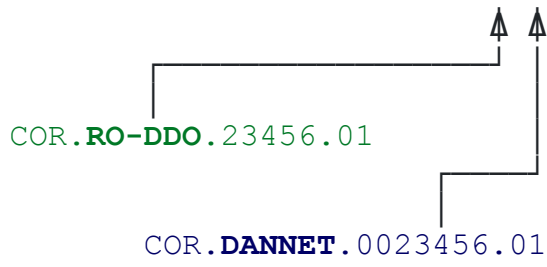
plantain, fælles betegnelse for en hel række sorter el. arter af banan, hvis stivelse ikke omdannes til sukker ved modning. De anvendes i tropiske lande kogte el. stegte som stivelseskilde i maden. Se også [banan](#).

COR. 23456

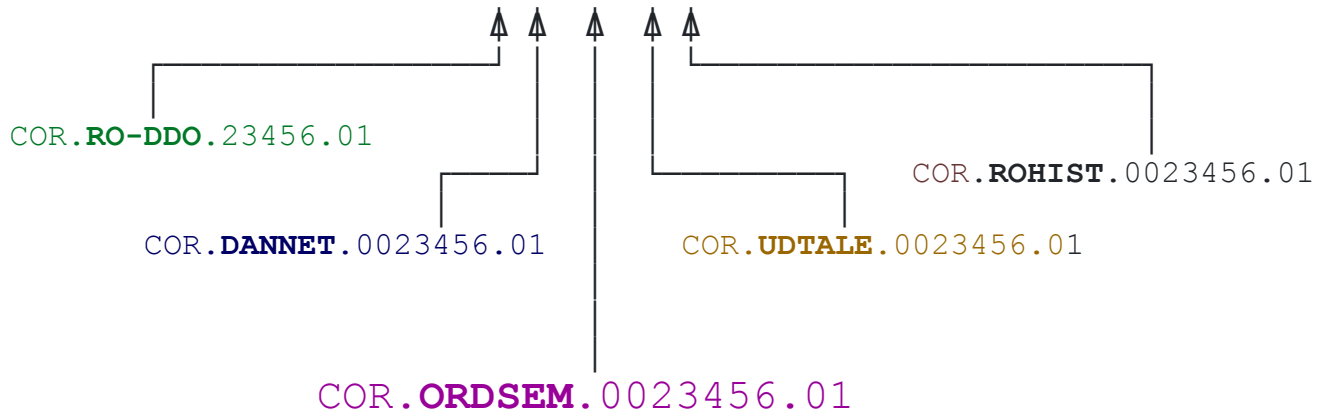
COR. 23456



COR.23456



COR.23456



Projektgruppe S og SX

COR-tagging

COR-tagging

"Bananen er en yndet spise, mens melbananen nok mest er for gourmeter."

COR-tagging

"Bananen er en yndet spise, mens melbananen nok mest er for gourmeter."

[bananen] [er] [en] [yndet] [spise] , [mens] [melbananen] ...
A B C D E F G ...

COR-tagging

"Bananen er en yndet spise, mens melbananen nok mest er for gourmeter."

[bananen] [er] [en] [yndet] [spise] , [mens] [melbananen] ...
A B C D E F G ...

A COR.23456.102

COR-tagging

"Bananen er en yndet spise, mens melbananen nok mest er for gourmeter."

[bananen] [er] [en] [yndet] [spise] , [mens] [melbananen] ...
A B C D E F G ...

A COR.23456.102
E COR.18181.101

COR-tagging

"Bananen er en yndet spise, mens melbananen nok mest er for gourmeter."

[bananen] [er] [en] [yndet] [spise] , [mens] [melbananen] ...
A B C D E F G ...

A COR.23456.102
E COR.18181.101
G COR.00000

COR-tagging

"Bananen er en yndet spise, mens melbananen nok mest er for gourmeter."

[bananen] [er] [en] [yndet] [spise] , [mens] [melbananen] ...
A B C D E F G ...

A COR.23456.102
E COR.18181.101
G COR.00000

Lov om dansk retskrivning

VI MARGRETHE DEN ANDEN, af Guds Nåde Danmarks Dronning, gør vitterligt: Folketinget har vedtaget og Vi ved Vort samtykke stadfæstet følgende lov:

§ 1. Dansk retskrivning fastlægges af Dansk Sprognævn og offentliggøres i Dansk Sprognævns retskrivningsordbog.

Stk. 2. Ved fastlæggelsen af dansk retskrivning følger Dansk Sprognævn de regler, der er fastsat i eller i medfør af lov om Dansk Sprognævn.

§ 2. Dansk retskrivning skal følges af alle dele af den offentlige forvaltning, af Folketinget og myndigheder med tilknytning til Folketinget samt af domstolene. Det samme gælder ikkeoffentlige uddannelsesinstitutioner, som modtager dækning af driftsomkostningerne på halvdelen eller mere, og private og selvejende skoler, hvor børn opfylder undervisningspligten.

Bekendtgørelse om Dansk Sprognævns virksomhed og sammensætning

§ 3. Sager om ændringer af retskrivningen skal behandles i **repræsentantskabet**. Repræsentantskabet tager stilling til, om en ændring er af principiel karakter.

§ 4. Repræsentantskabet består af indtil 43 medlemmer. De 40 indstilles af følgende:

1. Folketingets administration
2. Statsministeriet
3. Justitsministeriet
4. Undervisningsministeriet
5. Kulturministeriet
6. **Det Danske Sprog- og Litteraturselskab**
7. **Center for Sprogteknologi**
8. Stednavneudvalget
9. Modersmål-Selskabet
10. Forbundet Kommunikation og Sprog
11. Danmarks Radio
12. TV 2
13. Dansk Standard
14. Det Danske Akademi
15. It- og Telestyrelsen
16. Dansk Forfatterforening
17. Danske Skønlitterære Forfattere
18. Dansk Journalistforbund
19. Dansk Skuespiller Forbund
20. Dansklærerforeningen
21. Danmarks Lærerforening
22. Foreningen Norden
23. Danske Dagblades Forening
24. Danske Universiteter.
25. Professionshøjskolernes Rektorkollegium
26. Center for Læseforskning
27. EU's sprog tjenester
28. DANTERMcentret
29. Grænseforeningen
30. Lektoratsudvalget under Ministeriet for Videnskab, Teknologi og Udvikling
31. Landsforbundet af Voksen- og Ungdomsundervisere
32. Translatørforeningen

COR-indeksering

<i>COR-indeks</i>	<i>Resurse</i>	<i>"Myndighed"</i>
niv. 1	Ortografisk normerede	Lov om Dansk Retskrivning (DSN)
niv. 2	Kontrollerede ordkilder	Repræsentantskabet (DSL, CST, DSN, ...)
niv. 3	Frie ordkilder	Alle andre

COR som grundresurse for dansk sprogteknologi



COR som grundresurse for dansk sprogteknologi



- COR niv. 1 **synkroniseret** med dansk retskrivning

COR som grundresurse for dansk sprogteknologi



- COR niv. 1 **synkroniseret** med dansk retskrivning
- **Kontrollerede** COR-serier i niv. 2

COR som grundresurse for dansk sprogteknologi



- COR niv. 1 **synkroniseret** med dansk retskrivning
- **Kontrollerede** COR-serier i niv. 2
- **Frie** COR-serier i niv. 3

COR som grundresurse for dansk sprogteknologi



- COR niv. 1 **synkroniseret** med dansk retskrivning
- **Kontrollerede** COR-serier i niv. 2
- **Frie** COR-serier i niv. 3
- Også for resurser bag **Licensmure**

COR som grundresurse for dansk sprogteknologi



- COR niv. 1 **synkroniseret** med dansk retskrivning
- **Kontrollerede** COR-serier i niv. 2
- **Frie** COR-serier i niv. 3
- Også for resurser bag **Licensmure**
- **Globale** regler for indeksering

COR som grundresurse for dansk sprogteknologi



- COR niv. 1 **synkroniseret** med dansk retskrivning
- **Kontrollerede** COR-serier i niv. 2
- **Frie** COR-serier i niv. 3
- Også for resurser bag **Licensmure**
- **Globale** regler for indeksering
- Projektets afleveringer:

COR som grundresurse for dansk sprogteknologi



- COR niv. 1 **synkroniseret** med dansk retskrivning
- **Kontrollerede** COR-serier i niv. 2
- **Frie** COR-serier i niv. 3
- Også for resurser bag **Licensmure**
- **Globale** regler for indeksering
- Projektets afleveringer:
 - **2022**: COR-projektets **rammeværk**
 - **COR niv. 1** (ca. 60k lemmaer)
 - Manual til **COR-indeksering**
 - Værktøj til **COR-tagging**
 - Online-tjek af **COR-kompatibilitet**

COR som grundresurse for dansk sprogteknologi



- COR niv. 1 **synkroniseret** med dansk retskrivning
- **Kontrollerede** COR-serier i niv. 2
- **Frie** COR-serier i niv. 3
- Også for resurser bag **Licensmure**
- **Globale** regler for indeksering
- Projektets afleveringer:
 - **2022**: COR-projektets **rammeværk**
 - **COR niv. 1** (ca. 60k lemmaer)
 - Manual til **COR-indeksering**
 - Værktøj til **COR-tagging**
 - Online-tjek af **COR-kompatibilitet**
 - **2023**: COR-projektets **semantiske databaser**