

Ordposer

Hvordan kan en computer forstå en tekst – og hvordan kan du?

En pose ord er den simpleste form for en sprogmodel. Du tæller hvor mange gange hvert ord forekommer, og de mest brugte ord fortæller dig hvad teksten handler om. Google har brugt metoden i årevis.

Aktivitet:

Prøv selv at finde ud af, hvad disse ordposer handler om.

“Ordposen” er den simpleste version af en sprogmodel og derfor et godt sted at starte.

Forberedelse:

- Print ordposerne nederst på siden og klip ordene ud.
- Brug eventuelt eleverne til dele af klipningen. Det tager lang tid at klippe alle ordposerne ud. Som lærer kan du nøjes med at udklippe de små ordposer, så de er klar til undervisningen, og så kan du lade eleverne udklippe de store ordposer i fællesskab.
- Se ressource-slides som forberedelse til dialogen med eleverne.

Organisering:

Eleverne arbejder i par og afkoder indholdet af en ordpose. Når de har afkodet én ordpose, så udleverer du en ny. Brug ordposer med varierende sværhedsgrader for at motivere eleverne:

- Nogle har brug for en “quick win” (fx *Fødselsdagssangen* med “Hurra, hurra, hurra”)
- Andre har behov for en større udfordring (fx et stort datasæt fra H.C. Andersen)

Selvkørende aktiviteter og løbende samtale:

Opgaven er åben og u-instrueret. Den tilbyder deltagelsesmuligheder for alle, fordi mange strategier er mulige. Ordposen vækker nysgerrighed, og mens eleverne arbejder, er læreren friset til at observere deres strategier og tale med dem om forskellige sorteringsmetoder. Da aktiviteten også er en fysisk sorteringsopgave, så synliggøres elevernes tanker, valg og strategier til en vis grad. Ord-grupperinger og elev-dialog bliver vidnesbyrd om begyndende refleksion. Spørg ved hjælp af de faglige udtryk, som du gerne vil introducere eleverne for:

- Stammer I ordene, sådan at de ord der ligner hinanden, placeres sammen?
- Er der nogle ord som I ikke synes er vigtige? Renser I teksten for ubetydelige ord?
- Er der ord i teksten som stikker ud. - ord som *trigger* jer?

Brug disse indsigter i en fælles opsamling, hvor menneskelige strategier sammenlignes med computerens metoder til rensning og forståelse af tekst. Det analoge arbejde med ordposer suppleres med digitale aktiviteter fra maskinrummet.github.io (fx *Tekstrensning* og *ordskyer*).

Opsamling og kobling fra elevstrategier til sprogmodeller

(Benyt evt. slides 28–36 som visuel støtte)

Lærerfokus:

- Hvilken strategi brugte I i jeres gruppe?

Du har som lærer lyttet til elevernes samtaler og du har talt med dem om deres strategier og valg. Nu spørger du udvalgte grupper, som har brugt relevante strategier, og kobler dem til Maskinrummets arbejde med tekstforståelse – eller til eksempler fra virkeligheden, fx phishing-mails (Se beskrivelsen af Trigger-ord nederst)

Trin 1: Klassesamtale om typiske ord og stop-ord (Benyt evt. slide 28 som visuel støtte)

Vælg et par elever, som har arbejdet med fx *Wiki-artiklen* eller fødselsdagssangen og som har samlet “de typiske ord”.

Start samtalen med et simpelt spørgsmål, og spørg derefter ind til valg, strategier og udfordringer:

- Hvad var det første I kiggede efter, da I fik ordposen?
- Har I et gæt på, hvad jeres ordpose handler om?
- Hvordan kom I på sporet af det?
- Hvilken strategi brugte I i jeres gruppe?
- Samlede I de ord som lignede hinanden, så de typiske ord blev endnu mere tydelige: *elsk, elske, elsker, elskværdig*?
- Fjernede I ubetydelige ord (Stop ord)?

Trin 2: Klassesamtale om trigger-ord (Benyt evt slide 29-31 som visuel støtte)

Vælg et par som har undersøgt bageopskriften eller Den lille pige med svovlstikkerne og har lagt mærke til “trigger-ord”

Hvilke ord stikker ud i bunken?

- Gær: “Det må være en gærdej og der er brunfarin i, så måske er det kanelsnegle”
- Svovlstikkerne: Ordet er specielt og jeg kender historien, og så staves København på en gammeldags måde
- Havkongen: Ordet er specielt og jeg kender historien om Den lille havfrue.

Trin 2: Klassesamtale om trigger-ord (Benyt evt slide 29-31 som visuel støtte)

Vælg et par som har undersøgt bageopskriften eller Den lille havfrue og har lagt mærke til “trigger-ord”

Start samtalen med et simpelt spørgsmål, og spørg derefter ind til valg, strategier og udfordringer:

Særlig for disse ordposer er: Hvilke ord stikker ud i bunken?

- Gær: “Det må være en gærdej og der er brunfarin i, så måske er det kanelsnegle”
- Havkongen: Ordet er specielt og jeg kender historien om Den lille havfrue.
- Parceremoni: Det kender jeg kun fra Paradis Hotel
- Svovlstikkerne: Ordet er specielt og jeg kender historien, og så staves København på en gammeldags måde

Trin 3: Klassesamtale om negationer (Benyt evt slide 32 som visuel støtte)

Vælg et par som har undersøgt teksten “Does Luke like movies?”

Start samtalen med et simpelt spørgsmål, og spørg derefter ind til valg, strategier og udfordringer:

Særlig for denne ordposter er brugen af nægtende sætninger: Betyder det noget, hvor *ikke* står i sætningen?

- Tror I at Luke kan lide film?
- Hvilken strategi brugte I i jeres gruppe?
- Hvad gør man med en sætning som: “Luke kan (ikke) lide film?”
- Kan I være sikre?
- Betyder det noget om “ikke” forekommer ofte i ordposten? (Pointen og problematikken kan tages op senere i forbindelse med samtalen om n-gram.

Denne samtale om nægtende sætninger og placeringen af *ikke* kan senere kobles til muligheden for at benytte n-gram i forbindelse med generering af ny tekst.

N-gram. Med bi-gram låses to ord sammen: [ikke lide]

Begrebet **n-gram** forklares på slide 27. Du kan præsentere begrebet nu eller vente til I arbejder med lykkehjulene (slide 18)

Opsamling:

I har brugt tre strategier, som en sprogmodel også bruger, når den skal forstå en tekst. Det skal vi prøve i maskinrummet om lidt, og så har I oplevet betydningen af trigger-ord og af nægtende ord.

Trigger-ord: Ord som afslører eller aktiverer noget: gær (gær-dej), Havmanden (H.C. Andersen), København (Gammel retskrivning)

I din egen mailboks kender du måske ord som vækker din mistro: Mails med ord som *haster* og *penge* kan være vigtige, men de kan også være tegn på svindel. Uautoriserede måde at stave kendte afsendere på kan være et andet faretegn.

1. **Typiske ord:** Flest sedler med ordet “Danmark” fortæller, at det er en Wiki-artikel om Danmark
2. **Stamme:** Når du samler ord der ligner hinanden, så bliver de typiske ord endnu mere tydelige: elsk, elske, elsker, elskværdig, elskelig ...
3. **Stop-ord:** Se bort fra stop-ord. Stop ord er ord som ikke er vigtige for indholdet fx det, der, i, og

(20 min)

Hvilken wikipedia-artikel er det?

	og	det		op	t i l	Tyskland	mod	er	
af	de		nationer	sydvest	for	Sverige	og	syd	for
indgår	sammen	til	de	to	øvrige	rigsdele	grænser	Færøerne	
Danmark	Grønland	og	Færøerne	den	rigsdele	Grønland	har	selvstyre	og
	har	hjemmestyrer	Danmarks	forhold		et	land	i	Færøerne
kaldes	skandinaviske	med	Grønland	og	er	i	Kongeriget	Danmark	hvor
Danmark	Skandinavien	syd	Danmark	Det	sydligste		er	Norge	Rigsfællesskabet

Hvad er det her for noget?

smiler	Hurra	er	det	Kaspers	s o m	Hurra	venter	Hurra	ham
sikker	Hurra	Hvor	gave	får	Så	han	har	ønsket	sig
i	år	med	dejlig	chokolade	og	kager	til	vi	far
denne	Hurra	er	han	glad	Hurra	slut	Hurra	Men	Hurra
dag	er	også	rar	for	hjemme	og	med	Hurra	Hurra
med	fra	chokolade	Og	når	til	kager	kommer	han	hjem
råber	fødselsdag	går	Hurra	dem	Hurra	slut	skal	han	hjem
og	holde	skole	fest	og	som	dag	sig	som	gæst
		chokolade	til	I	til	Til	og	får	mor
højt	i	kor	og	sine	længe	Gid	Kasper	en	leve
må	og	kager	ønsker	opfyldt	få	han	dejlig	chokolade	og
kager	dejlig								dejlig

Kan Luke godt lide film?

the	Not	they	way	look	or	sound	My	doesnt	hair
just	raise	there	because	is	some	cool	thing	screen	on
in	passion	to	art	be	able	to	enjoy		it
an	and	I	explosion	need	to	see	more	and	story
name	My	Luke	is	and	I	definitely	dont	movies	love

Er det fra en artikel i DR eller BT?

ske	forstår	godt	Jeg	Nikolaj	Jacosben	at	rasende	er	Alt
ikke	nok	Niklas	godt	er	gået	Landin	og	forrest	har

markeret	med	flere	sig	redninger	men	vigtige	er	det	som
om	fejlene	står	at	kø	Fejl	i	IKKE	der	må
skal	i	anden	ændres	Vågn	op	halvleg			
Flere	profilerne	har	af	indtil	videre	svigtet	er	det	bare

Hvilken bog er det?

eneste	være	ville	stads	stor	en	i	krone	dronning	
glas	det	men	meget	er	dybere	dybt	noget	end	ankertov
på	dejligste	d e n	og	kornblomst	klart	så	det	som	reneste
imellem	ligesom	grenene	fuglene	heroppe	luften	i	det	På	allerdybeste
når	kirketårne	mange	stilles	måtte	på	oven	for	hinanden	at
række	bunden	fra	over	op	Dernede	vandet	havfolkene	bor	Nu
må	slet	man	tro	ikke	der	at	er	kun	den
nøgne	sandbunden	hvide	der	nej	de	vokser	træer	forunderligste	og
de	den	ved	bevægelse	mindste	vandet	af	sig	rører	ligesom
Langt	i	ude	e r	havet	så	vandet	som	blåt	bladene
om	var	de	Alle	levende	små	fiskene	store	og	smutter
sted	havkongens	ligger	murene	slot	af	er	og	koraller	de
lange	vinduer	spidse	det	a f	rav	allerklareste	taget	men	er
muslingskaller	åbner	der	lukker	og	eftersom	sig	går	vandet	det
ser	ud	dejligt	i	thi	ligger	hver	perler	strålende	én
planter	er	som	smidige	så	stilk	i	blade	og	at

Hvad er det for en opskrift?

Mælk	Smør	Gær	kakaopulver	Stødt	farin	2	Æg		Hvedemel
ca	8½dl	2	tsk	Blødt	kardemomme	1	tsk	Groft	salt
200g		smør	200g	Brun		1	spsk	Stødt	kanel
2dl	Flormelis	1½	spsk	Vand	Evt	600g	100g	2½dl	25g

Hvad er det her for noget?

spillet	i	som	foregår	at	handler	være	der	sidste	står
som	køn	er	i	på	det	luksuriøst	hotel	et	i

af	Antallet	spiller	er	næsten	Mexico	ligeligt	melle	delt	drenge
og	piger	disse	skal	så	og	i	par	forene	Der
altid	dog	én	eller	flere	er	af	det	tilover	ene
bor	som	på	solo	singleværel	køn	gang	om	En	ugen
en	minds	person	af	og	sendes	til	at	pigern	være
af	ud	Paradis	Hotel	det	overtal	ved	den	sker	ugentli
parceremonier	Ved	skiftes	drenge	der	parceremo	drengen	og	vælger	hver
således	skud	at	det	hver	for	uge	er	anden	pigerne
igen	som	skaber	et	anden	ind	omvend	Ugen	uge	efter
et	der	antal	nye		kommer	den	tilbag	deltag	om
drenge	antal	og	piger	Det	ulige				

Hvad er denne sætning?

Du	se	i	være	den	a	t	som			
	du	verden	må	selv	ønsker	forandring				

Hvad er denne sætning?

komme	skal	ikke	og	I	kong	spille	her	gulerod
-------	------	------	----	---	------	--------	-----	---------

Hvad er svaret på denne gåde?

det	er	ingen	der	Hvad	vegne				
kommer	går	går	og	og					

Hvad er denne sætning?

elsker	elskes	at	Du	drøm	En	et	om	menneske
--------	--------	----	----	------	----	----	----	----------

Hvad er denne sætning?

Du	se	i	være	den	at	som			
	du	verden	må	selv	ønsker	forandring			

Hvad er denne sætning?

komme	skal	ikke	og	I	kong	spille	her	gulerod
-------	------	------	----	---	------	--------	-----	---------

Hvad er svaret på denne gåde?

det	er	ingen	der	Hvad	vegne				
kommer	går	går	og	og					

Hvad er denne sætning?

elsker	elskes	at	Du	drøm	En	et	om	menneske
--------	--------	----	----	------	----	----	----	----------

Hvad er denne sætning?

Du	se	i	være	den	at	som			
	du	verden	må	selv	ønsker	forandring			