



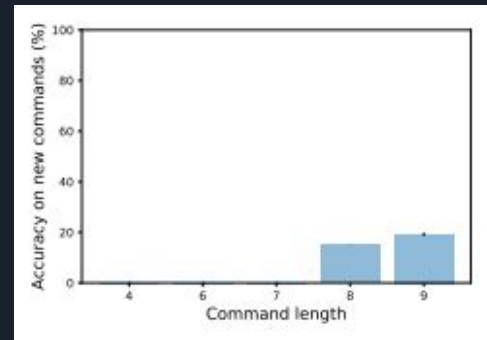
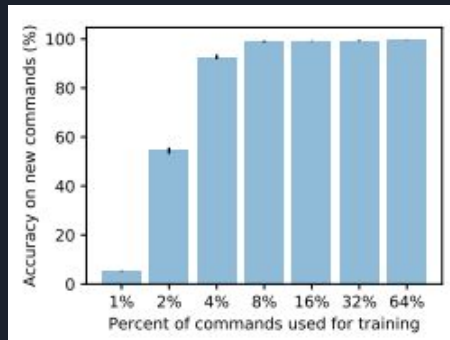
Compositionality in the era of neural networks

Dieuwke Hupkes

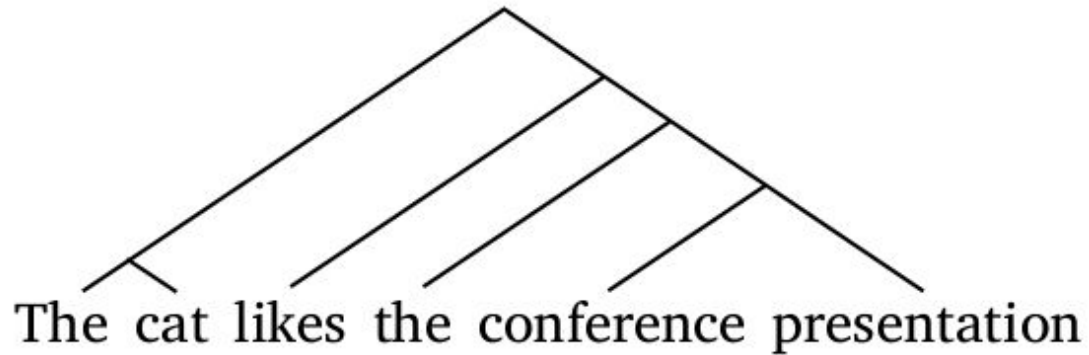
Compositionality

jump	⇒	JUMP
jump left	⇒	LTURN JUMP
jump around right	⇒	RTURN JUMP RTURN JUMP RTURN JUMP RTURN JUMP
turn left twice	⇒	LTURN LTURN
jump thrice	⇒	JUMP JUMP JUMP
jump opposite left and walk thrice	⇒	LTURN LTURN JUMP WALK WALK WALK
jump opposite left after walk around left	⇒	LTURN WALK LTURN WALK LTURN WALK LTURN WALK LTURN LTURN JUMP

Figure 1. Examples of SCAN commands (left) and the corresponding action sequences (right).

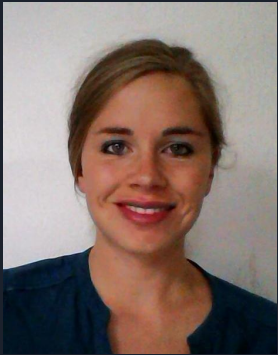


When is a network compositional?





Compositionality decomposed



Dieuwke Hupkes



Verna Dankers



Mathijs Mul



Elia Bruni

Compositionality Decomposed: How do Neural Networks Generalise?

Dieuwke Hupkes

*Institute for Logic, Language and Computation
University of Amsterdam
Science Park 107, 1098 XG Amsterdam*

DIEUWKEHUPKES@GMAIL.COM

Verna Dankers

Mathijs Mul

*University of Amsterdam
Science Park, 1098 XH Amsterdam*

VERNADANKERS@GMAIL.COM

MATHIJSMUL@GMAIL.COM

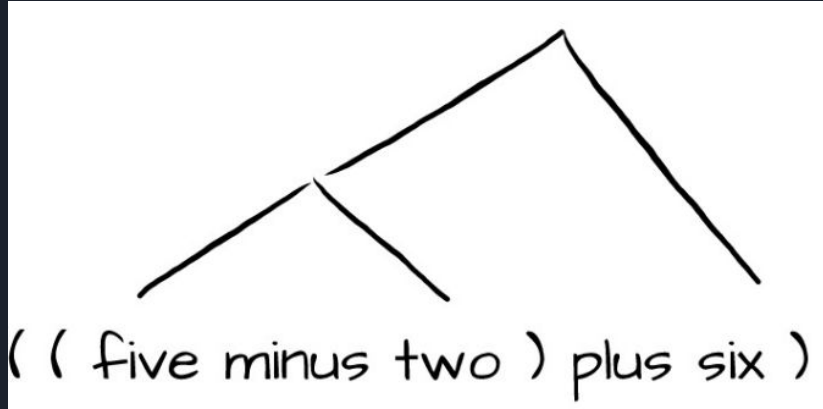
Elia Bruni

*University of Pompeu Fabra
Roc Boronat, 138 08018 Barcelona*

ELIA.BRUNI@GMAIL.COM

Compositionality in natural language

5 - 2 = 3



cherry on the cake

homonyms
Scope ambiguities
polysemous words
idioms
attachment ambiguities
SO MANY OTHER THINGS!



Levels of compositionality

The principle of compositionality

The meaning of a whole is a function of the meanings of the parts and of the way they are syntactically combined

(Partee, 1995)

The meaning of a complex expression is determined by the meanings of its constituents and by its structure

(Szabo, 2000)



Natural language

How do we evaluate compositionality “in the wild”?

A decorative graphic in the top-left corner consisting of a blue parallelogram and a light green parallelogram, both tilted at an angle. The background is a dark navy blue with faint, lighter blue diagonal stripes.

Hurdle 1: the task

The task

form

jump
jump left
jump around right
turn left twice
jump thrice
jump opposite left and walk thrice
jump opposite left after walk around left

⇒
⇒
⇒
⇒
⇒
⇒
⇒

JUMP
LTURN JUMP
RTURN JUMP RTURN JUMP RTURN JUMP RTURN JUMP
LTURN LTURN
JUMP JUMP JUMP
LTURN LTURN JUMP WALK WALK WALK
LTURN WALK LTURN WALK LTURN WALK LTURN WALK
LTURN LTURN JUMP

meaning

Figure 1. Examples of SCAN commands (left) and the corresponding action sequences (right).

form

4 The pilot was identified as Squadron Leader Dilokrit Pattavee.
5 Local media reports an airport fire vehicle rolled over while
responding.
6 28-year-old Vidal had joined Barça three seasons ago, from
Sevilla.
7 Since moving to the Catalan-capital, Vidal had played 49
games for the club.

proxy to meaning

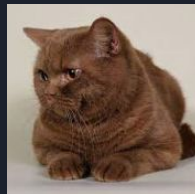
4 De piloot werd geïdentificeerd als majoor Dilokrit Pattavee.
5 De lokale media meldt dat er tijdens een actie op de
luchthaven een brandweerwagen is gekanteld.
6 De 28-jaar oude Vidal is drie seizoenen geleden van Sevilla
naar Barça overgestapt.
7 Sinds hij verhuisde naar de Catalaanse hoofdstad heeft hij 49
wedstrijden gespeeld voor de club.

Hurdle 2: what does the model know?

Knows (brown dog, black cat)



→ Knows (brown cat)



consistent with (brown dog, black cat)



Interpretation (brown cat)

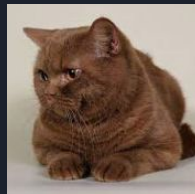


Hurdle 2: what does the model know?

Knows (brown dog, black cat)



→ Knows (brown cat)



Systematicity
(Szabo)

consistent with (brown dog, black cat)



Translation (brown cat)



How we measure
this in NMT models

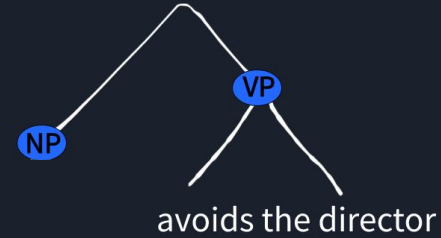


The tests

Systematicity

Can models understand novel combinations of known components?

$S \rightarrow NP VP$



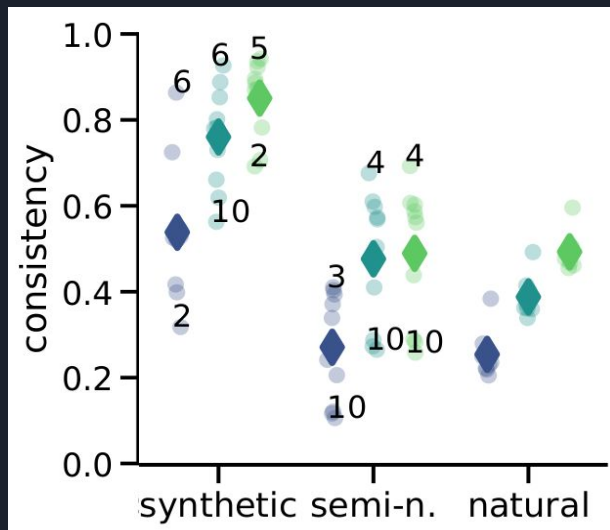
$S \rightarrow S CONJ S$



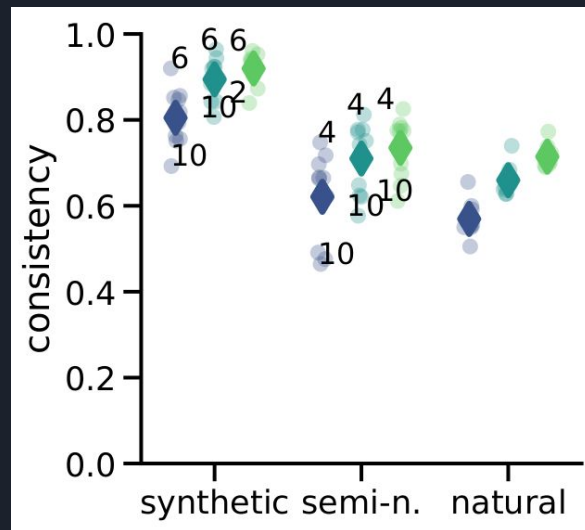
Systematicity

The father avoids the director, and the boys criticise the leader.

The mother avoids the director, and the boys criticise the leader.



$S1 \rightarrow S3$



$S1 \rightarrow S1'$

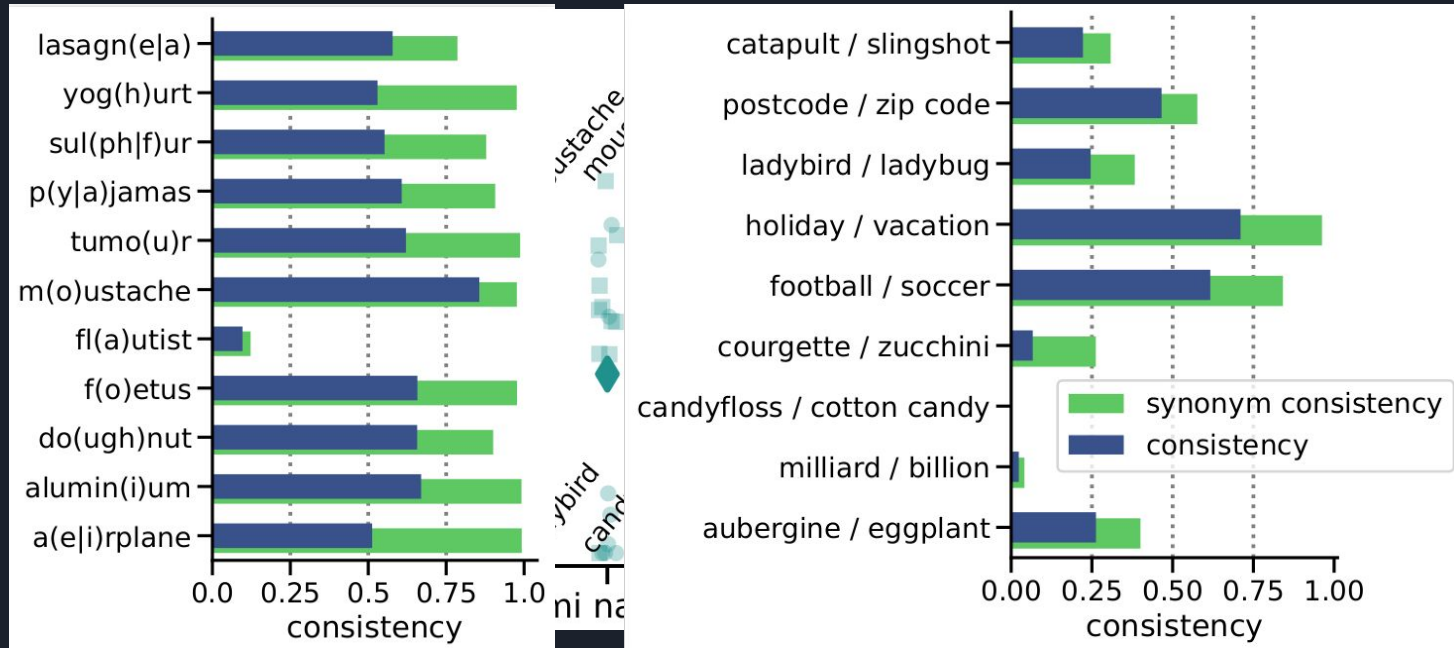
training size


- small
- medium
- full

Substitutivity

Okay, this is not as sensational as a radio-control **aeroplane**, but it's a lot quieter.

Okay, this is not as sensational as a radio-control **airplane**, but it's a lot quieter.





An example of our substitutivity test

The mother that travels by (aeroplane|airplane) wants to facilitate the development of innovative practices in education and training at tertiary level, and their transfer, including from one participating country to others.

De moeder die per vliegtuig reist wil de ontwikkeling van innovatieve praktijken op de gebied van onderwijs en opleiding op tertiair niveau, en de overdracht daarvan, ook van de ene deelnemende land naar de andere, vergemakkelijken.

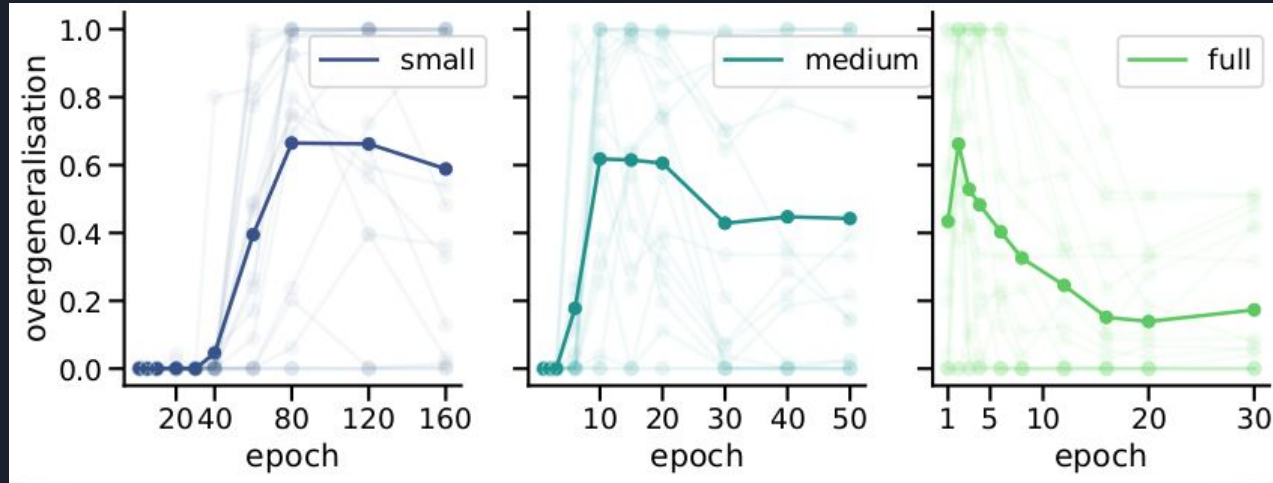
De moeder die per vliegtuig reist wil de ontwikkeling van innovatieve praktijken op de gebied van onderwijs en opleiding op tertiair niveau, en hun overdracht, waaronder van de ene deelnemende land naar de andere, vergemakkelijken.

Overgeneralisation

You guys are out of your minds!

Jullie zijn niet goed wijs.

Jullie zijn uit je hoofd.





Discussion

- Models are often **not maximally compositional**, but sometimes **too compositional**
- The question we should ask: what is their **level of compositionality**?
 - ◆ What is the right level?
 - ◆ We need measures for that!
- Does compositionality actually matter for performance?
 - ◆ Do compositional models generalise better?
 - ◆ Do they learn better in low-resource scenarios?



Thank you



Verna Dankers



Elia Bruni

dieuwkehupkes.nl
dieuwkehupkes@fb.com

Synthetic data

<i>n</i>	Template
1	The N_{people} V the $N_{\text{elite}}^{\text{sl}}$. <i>The poet criticises the king .</i>
2	The N_{people} Adv V the $N_{\text{elite}}^{\text{sl}}$. <i>The victim carefully observes the queen .</i>
3	The N_{people} P the $N_{\text{vehicle}}^{\text{sl}}$ V the $N_{\text{elite}}^{\text{sl}}$. <i>The athlete near the bike observes the leader .</i>
4	The N_{people} and the N_{people} V the $N_{\text{elite}}^{\text{sl}}$. <i>The poet and the child understand the mayor .</i>
5	The $N_{\text{quantity}}^{\text{sl}}$ of $N_{\text{people}}^{\text{pl}}$ P the $N_{\text{vehicle}}^{\text{sl}}$ V the $N_{\text{elite}}^{\text{sl}}$. <i>The group of men near the bike forgets the queen .</i>
6	The N_{people} V that the $N_{\text{people}}^{\text{pl}}$ V . <i>The farmer sees that the lawyers cry .</i>
7	The N_{people} Adv V that the $N_{\text{people}}^{\text{pl}}$ V . <i>The mother probably thinks that the men scream .</i>
8	The N_{people} V that the $N_{\text{people}}^{\text{psl}}$ V Adv . <i>The mother thinks that the men scream carefully .</i>
9	The N_{people} that $V V$ $N_{\text{elite}}^{\text{sl}}$. <i>The poets that sleep understand the queen .</i>
10	The N_{people} that $V \text{Pro}$ V the $N_{\text{elite}}^{\text{sl}}$. <i>The man that criticises him recognises the queen .</i>

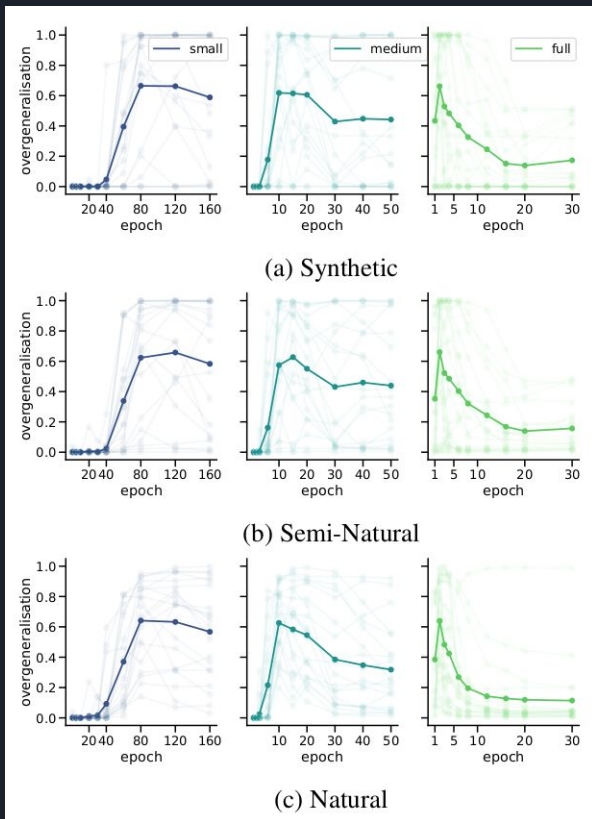
Semi-natural data

<i>n</i>	Template
1	The N_{people} (VP (TO) (VP (VB) (NP (NP) (PP (IN) (NP (NP) (PP (IN) (NP)))))))) E.g. The woman wants to use the Internet as a means of communication .
2	The N_{people} (VP (VBP) (VP (VBG) (S (VP (TO) (VP (VB) (S (VP (TO) (VP))))))))) E.g. The men are gon na have to move off-camera .
3	The N_{people} (VP (VB) (NP (NP) (PP (IN) (NP))) (PP (IN) (NP (NP) (PP (IN) (NP))))) E.g. The doctors retain 10 % of these amounts by way of collection costs .
4	The N_{people} reads an article about (NP (NP) (PP (IN) (NP (NP) (PP (IN) (NP (NP) (PP (IN) (NP)))))))) E.g. The friend reads an article about the development of ascites in rats with liver cirrhosis .
5	The N_{people} reads an article about (NP (NP (DT) (NN)) (PP (IN) (NP (NP) (SBAR (S (WHNP (WDT)) (VP)))))) . E.g. The teachers read an article about the degree of progress that can be achieved by the industry .
6	An article about (NP (NP) (PP (IN) (NP (NP) (PP (IN) (NP (NP) (PP (IN) (NP)))))))) is read by the N_{people} . E.g. An article about the inland transport of dangerous goods from a variety of Member States is read by the lawyer .
7	An article about (NP (NP) (PP (IN) (NP (NP) (, .) (SBAR (S (WHNP (WDT)) (VP)))))) , is read by the N_{people} . E.g. An article about the criterion on price stability , which was 27 % , is read by the child .
8	Did the N_{people} hear about (NP (NP) (PP (IN) (NP (NP) (PP (IN) (NP (NP) (PP (IN) (NP)))))))) . E.g. Did the friend hear about an inhospitable fringe of land on the shores of the Dead Sea ?
9	Did the N_{people} hear about (NP (NP (DT) (NN)) (PP (IN) (NP (NP) (SBAR (S (WHNP (WDT)) (VP)))))) ? E.g. Did the teacher hear about the march on Employment which happened here on Sunday ?
10	Did the N_{people} hear about (NP (NP) (SBAR (S (VP (TO) (VP (VB) (NP (NP) (PP (IN) (NP)))))))) ? E.g. Did the lawyers hear about a qualification procedure to examine the suitability of the applicants ?

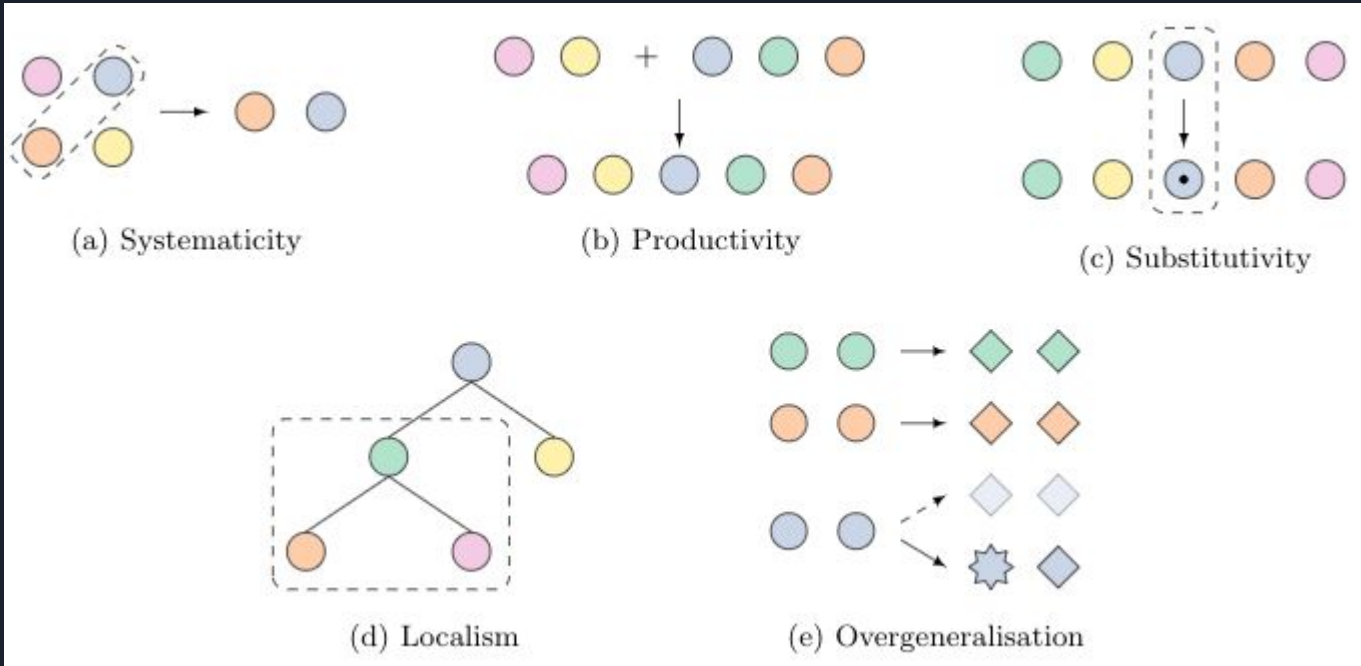
All our idioms

Idiom	Subordinate clause	Local translation
<u>once in a while</u>	that said " I will play it once in a while "	eens in een tijdje
do the <u>right thing</u>	that said " Just do the right thing "	doen het juiste ding
out of your <u>mind</u>	that said " Have you gone out of your mind "	uit je hoofd
<u>state of the art</u>	that said " This is a state of the art, official facility "	stand van de kunst
from <u>scratch</u>	that said " We are cooking from scratch every day "	van kras
take <u>stock</u>	that said " Take stock of the lessons to be drawn "	nemen voorraad
across the <u>board</u>	that said " I got red lights all across the board "	aan boord
in the final <u>analysis</u>	that said " In the final analysis, this is what matters "	in de laatste analyse
out of the <u>blue</u>	that said " It just came out of the blue "	uit het blauwe
in <u>tandem</u>	that said " We will work with them in tandem "	in tandem
by <u>heart</u>	that said " I knew the formula by heart "	door hart
come to <u>terms</u> with	that said " I have come to terms with my evil past "	komen overeen met
by the same <u>token</u>	that said " By the same token I will oppose what is evil "	bij dezelfde token
at your <u>fingertips</u>	that said " The answer is right at your fingertips "	binnen handbereik
look the other <u>way</u>	that said " We cannot look the other way either "	kijken de andere manier
follow <u>suit</u>	that said " And many others follow suit "	volgen pak
keep <u>tabs</u> on	that said " I keep tabs on you "	houden tabs
in the short <u>run</u>	that said " In the short run it clearly must be "	in de korte lopen
by <u>dint</u> of	that said " We are part of it by dint of our commitment "	door de int
set <u>eyes</u> on	that said " I wish I had never set eyes on him "	set ogen op

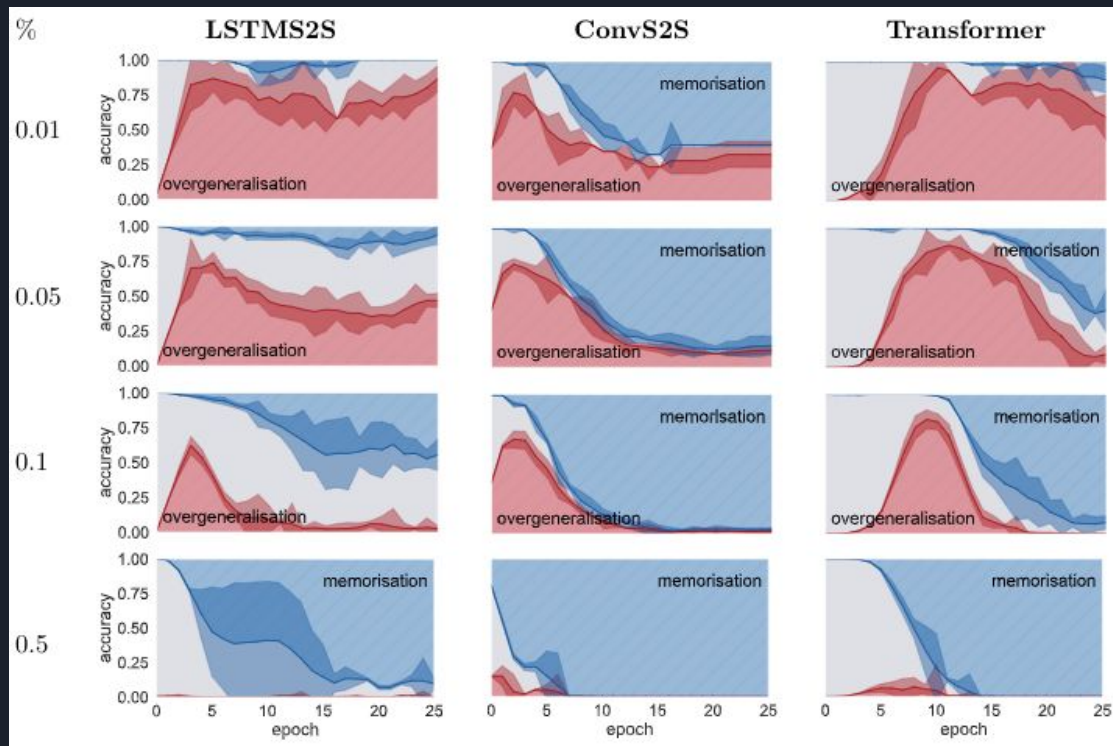
Overgeneralisation -- all data



Compositionality decomposed: tests



Compositionality decomposed: overgeneralisation



Compositionality decomposed: localism

