Experimental evidence on the individuality of lexicogrammar

Dr Andrea Nini   Michael Cameron   Colin Murphy

andrea.nini@manchester.ac.uk
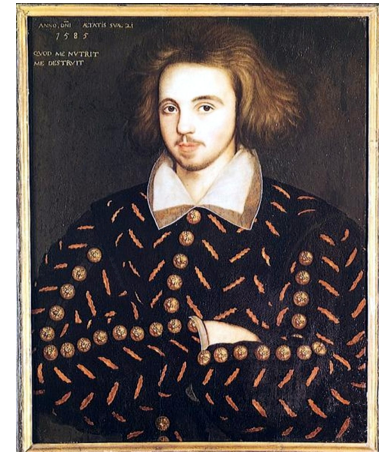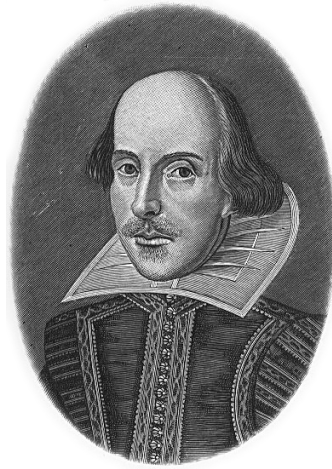
www.andreanini.com

11th International Conference on Construction Grammar

University of Antwerp

20th August 2021

Authorship Identification

Jenny Nicholl

Last seen 30th June 2005.

After her disappearance a series of **text messages** were sent from her phone.

Linguistic analysis showed that the later messages were more likely to have been sent by her ex-lover, **David Hodgson**.

Hodgson was **convicted of** Jenny's **murder** 19th February 2008 in part because of the linguistic evidence.

JENNY NICHOLL HISTORIC MESSAGES

Sum black+pink k swiss shoes and all the other **shit** like socks.We r **goin2the** indian.Only16quid.What u doin x

Yeah shud b gud.i just **have2get my** finga out and do anotha tape.wil do it on sun.will seems **keen2x**

Shit is it.fuck i**cant2day ive** allready **booked2go** bowling.cant realy pull out.wil **go2shop** and get her sumet soon.thanx4tdlin me x

No reason just seing what ur **up2.want2go** shopping on fri **and2will`s** on sun if ur **up2it**

Sorry **im not out2nite** havnt seen u 4a while aswel.ru **free2moro** at all x

No **im** out wiv jak sorry it took me so long **ive** had **fone off** coz havnt got much battery

Only just turned **my fone**.havnt lied bout anything.no it doesnt look good but ur obviously jst as judgmental than the rest.**cu** wen i **cu**&i hope its not soon

I havnt **lied2u**.anyway **im off back2sleep**

I know i waved at her we wer **suppose2go** at4but was a buffet on later on so waited.anyway he had a threesome it was great **cu** around

Im tierd of defending **myself** theres no point.bye

Happy bday!wil b round wiv ur **pressent2moz** sorry i cant make **it2day.cu2moz** xxx

SUMMARY OF STYLE DIFFERENCES

IM / **I AM**

AM NOT/I`M NOT / **AINT**

IVE / **AVE**

MY/MYSELF **/ME/MESELF**

OFF / **OF**

TO/2 - SPACE / **TO/2 + SPACE**

CU /**CYA**

FONE / **PHONE**

SHIT / **SHITE**

SUSPECT TEXT MESSAGES

Thought u wer grassing me up.mite b in trub wiv **me** dad told mum i was lving didnt giv a **shit**.**been2 kessick** camping was great.**ave2 go cya**

Hi jen tell jak **i am** ok know ever 1s gona b mad tell them **i am** sorry.living in scotland wiv **my** boyfriend.shitting **meself** dads gona kill **me** mum dont give a **shite**.hope nik didnt grass me up.keeping **phone of**.tell dad car jumps out of gear and stalls put it back in auction.tell him **i am** sorry

Y do u h8 me i know mum does.told her i was goin.i **aint** cumin back and the pigs wont find me.**i am** happy living up here.every1 h8s me in rich only m8 i got is jak.txt u couple wks tell pigs **i am** nearly 20 **aint** cumin back they can **shite off**

She got me in this **shit** its her fault not mine get blame 4evrything.**i am** sorry ok just had **2 lve** shes a bitch no food in and always searching **me** room eating **me** sweets.**ave2 go** ok **i am** very sorry x

in-group variation that influences perception (personality, cognitive abilities, …)

School   Law School   Lawyer

different life experiences → different texts → different exemplars
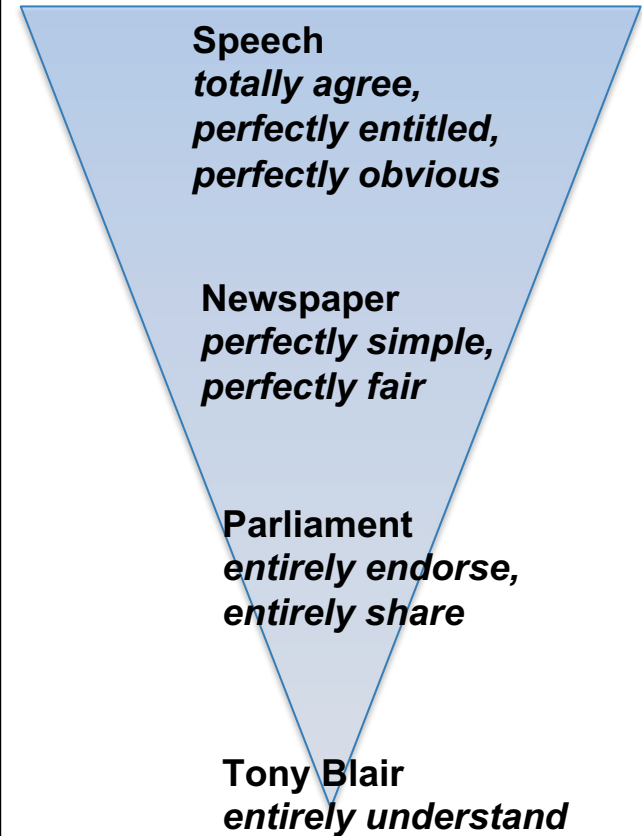
School   Electrician   Manager

Maximisers

Tony Blair

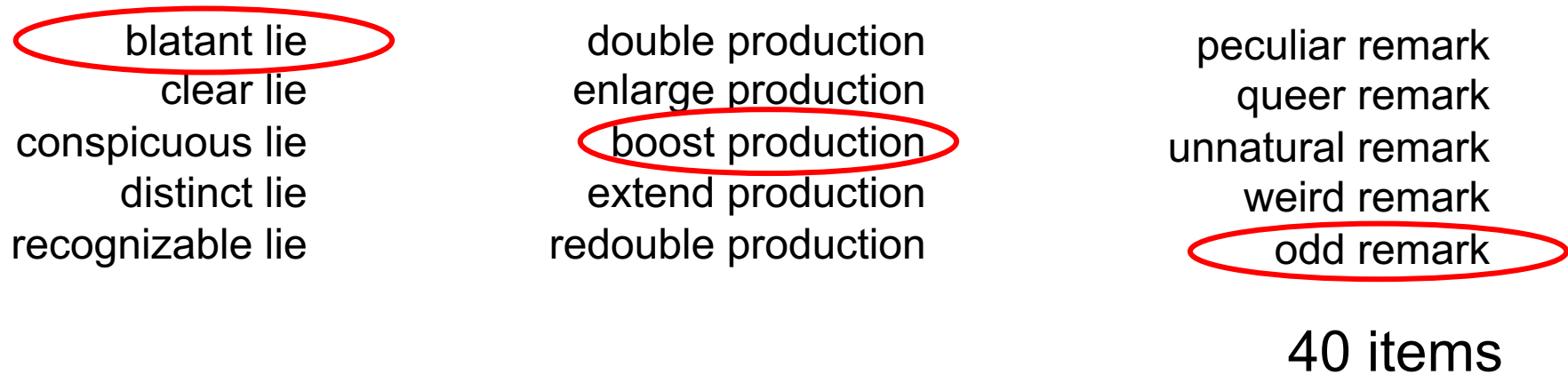3M words of interviews, speeches, PM question time

entirely absolutely altogether completely entirely extremely fully perfectly thoroughly totally utterly	critical agree central acceptable understand respect …

Speech
totally agree, perfectly entitled, perfectly obvious

Newspaper
perfectly simple, perfectly fair

Parliament
entirely endorse, entirely share

Tony Blair
entirely understand

TBC: max.+*understand* <*entirely* 54%, *totally* 38%, *fully* 7%, *completely* 1%, *perfectly* 1%, *wholly* 1%> (191 tokens)

BNC: max.+*understand* <*fully* 79%, *completely* 5%, *perfectly* 5%, *thoroughly* 4%, *entirely* 3%, *totally* 3%, *wholly* 2%, *absolutely* 1%> (199 tokens)

Mollin, Sandra. 2009. "I entirely understand" is a Blairism: The methodology of identifying idiolectal collocations. International Journal of Corpus Linguistics 14(3). 367–392. doi:10.1075/ijcl.14.3.04mol.

Words that go together. Measuring individual differences in native speakers' knowledge of collocations

blatant lie
clear lie
conspicuous lie
distinct lie
recognizable lie

double production
enlarge production
boost production
extend production
redouble production

peculiar remark
queer remark
unnatural remark
weird remark
odd remark

40 items

% of correct responses   ~   linguistic experience   linguistic proficiency   IQ

Words that go together. Measuring individual differences in native speakers' knowledge of collocations
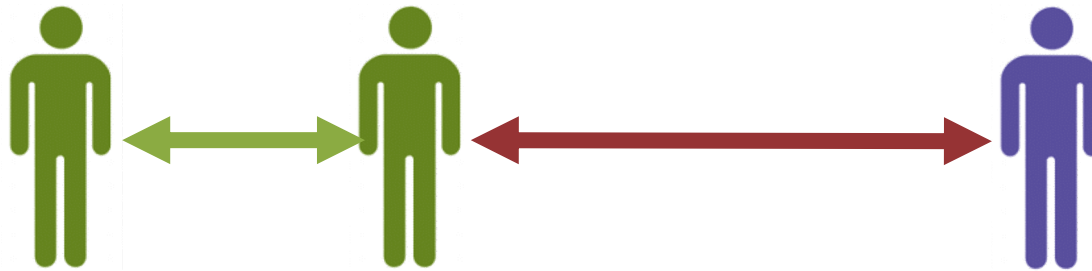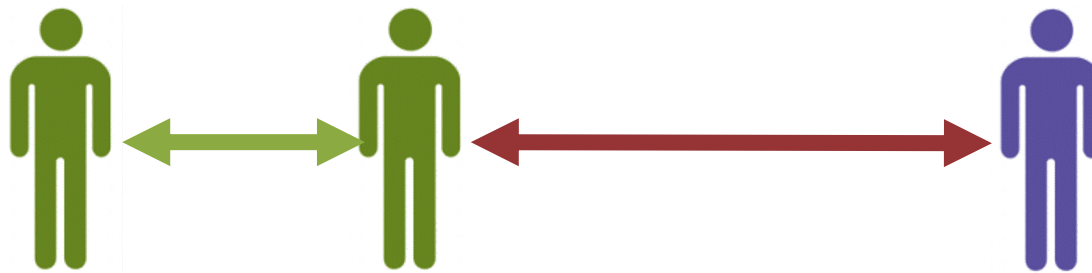
% of correct responses ~

linguistic experience   linguistic proficiency   IQ

blatant lie	double production	peculiar remark
clear lie	enlarge production	queer remark
conspicuous lie	boost production	unnatural remark
distinct lie	extend production	weird remark
recognizable lie	redouble production	odd remark

individual differences

How much **intra-speaker variation** is there within the same speaker?
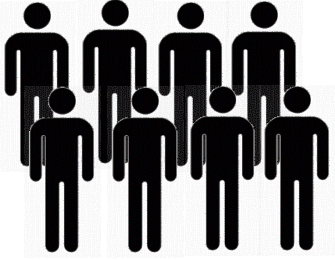
How much **inter-speaker variation** is there between two random speakers?

intra-speaker variation   **<<**   **inter-speaker variation**   **?**

8 participants

All academics in the humanities with PhD
All native speakers of British English

Questionnaire

Grammatical alternations test

Test on grammatical alternations

27 sentences divided in 12 categories

The author relies on speculation to make his point, _____ providing hard evidence to back up his argument.

      a. rather than

      b. instead of

British Academic Written Corpus (BAWE):

rather than V	544	*rather than providing*	5
instead of V	171	*instead of providing*	4

Test on grammatical alternations

27 sentences divided in 12 categories

Controlling for factors that influence variation

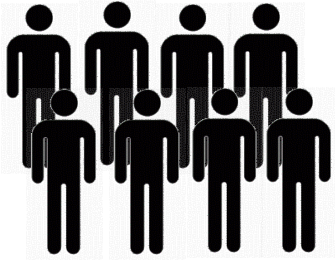| Preposition (7)	• rather than / instead of
Particle-verb alternation (5)	• take N off / take off N
Relative pronouns (3)	• which / that / ∅ / who / whom
Dative alternation (2)	• give N N / give N to N
Genitive alternation (2)	• s-genitive / of-genitive
Determiner (2)	• half / half of
Proper name (1)	• e.g. Ikea's website / the Ikea website
Help (1)	• help V / help to V
Try (1)	• try to V / try and V
Locative alternation (1)	• e.g. spray water on the grass / spray the grass with water
SKT (1)	• sort of / kind of / type of
Negation (1)	• e.g. not affected / unaffected

8 participants

All academics in the humanities with PhD
All native speakers of British English

Questionnaire

Grammatical alternations test   Grammatical alternations test

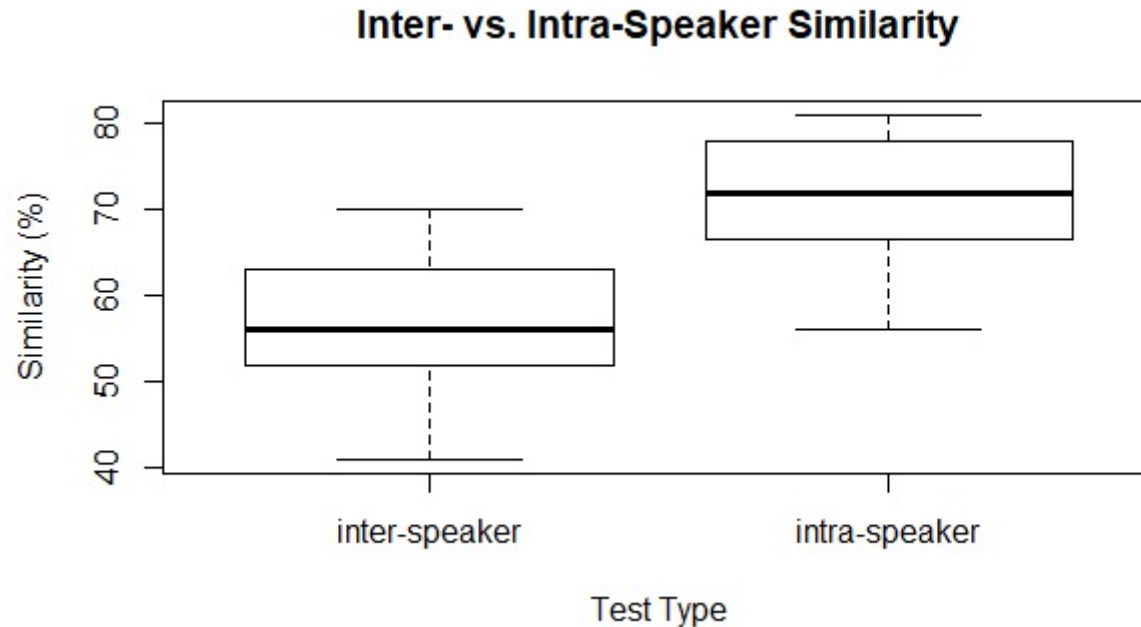**Changed the order and as much lexical material as possible**

at least 2 weeks

"another questionnaire"

Results

Inter- vs. Intra-Speaker Similarity

Wilcoxon test *p* = 0.0007

Higher levels of similarity for grammar test than collocations test

Cohen's *d* = 1.8

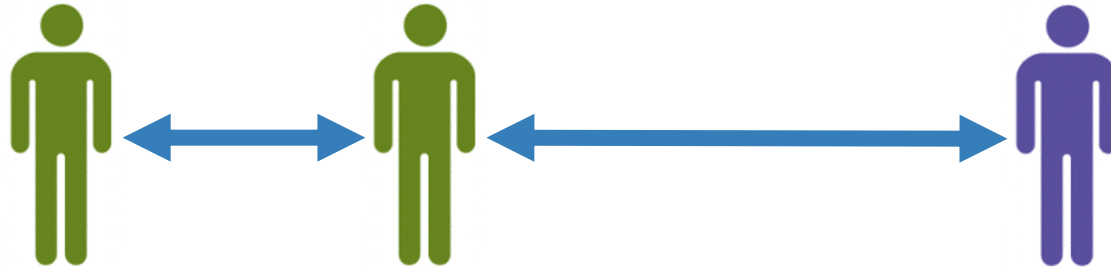with *p* < 0.05 power is 0.99 even with this small sample size!

Results
Sanity check

Inter-Speaker Similarity: Test 1 vs Test 2

Conclusions

No two profiles of alternations are identical

Same person tests more similar to themselves than random pairs

Results are consistent with usage-based approaches to lexicogrammar

Results have massive implications for forensic linguistics

Is there really a probability of selection of a variant?

Could it be a memory test?

Anecdotal evidence from debriefing points
to weak memory effect

Larger replication in the works

Experimental evidence on the individuality of lexicogrammar

Dr Andrea Nini   Michael Cameron   Colin Murphy

andrea.nini@manchester.ac.uk

www.andreanini.com

11th International Conference on Construction Grammar

University of Antwerp

20th August 2021