

Shakespeare's language: New perspectives from corpus linguistics

*Jonathan Culpeper,
Lancaster University, UK*

@ShakespeareLang



Arts & Humanities
Research Council



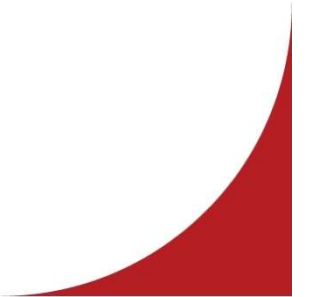
THE QUEEN'S
ANNIVERSARY PRIZES
FOR HIGHER AND FURTHER EDUCATION
2015





Shakespeare might not be a pleasurable experience ... for linguistic reasons

Comments from British undergraduates (TheStudentRoom):

- “Shakespeare uses overcomplicated sentence structure to say simple things which makes him seem pretentious and hard to understand, with old and outdated vocabulary.”
 - “I've never really seen the relevance of Shakespeare when it comes to teaching English. I mean he might as well have written his plays in a foreign language.”
 - “If I could go back in time, I'd kill Shakespeare so that today's schoolchildren wouldn't have to study his crap.”
- 

What will the project produce?

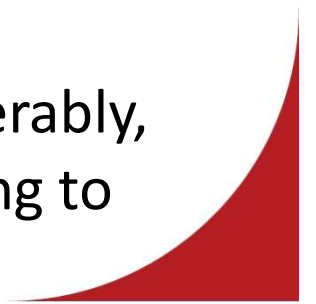
The key output: The encyclopedia

Two volumes:

- (1) a kind of dictionary, and
 - (2) a compendium of word patterns relating to characters, character groups, plays, play-genres, themes, etc..
- It will be the first systematic and comprehensive account of Shakespeare's language using methods derived from corpus linguistics – an approach that uses computers in large-scale language analysis.
 - It will be comparative, looking at language use in Shakespeare *and* his contemporaries.

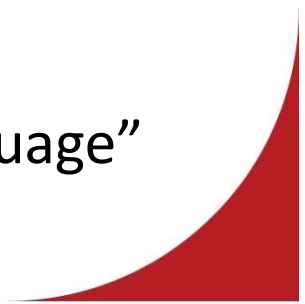
Why an approach from linguistics? And why now?

Jack Cade: "It will be proved to thy face that thou hast men about thee that usually talk of a noun and a verb, and such abominable words as no Christian ear can endure to hear." (2H6.IV.vii)

- Linguistics has changed – greater focus on actual language use, social issues, interdisciplinarity, etc.
 - Computer techniques have developed sufficiently to handle corpora of historical texts, provide evidence patterns of meanings, contribute to the description of style, etc.
 - Comparative historical resources have developed considerably, notably, *Early English Books Online* (EEBO-TCP), amounting to some 1.2 billion words.
- 

“Shakespeare coined more words than any other writer”

What can we ‘learn’ from the internet?


- “Shakespeare coined more words than other writers, around 1700 words ...”
 - “The English language owes a great debt to Shakespeare. He invented over 1700 of our common words ...”
 - “Shakespeare introduced nearly 3,000 words ...”
 - “Shakespeare is credited by the Oxford English Dictionary with the introduction of nearly 3,000 words into the language”
 - “Shakespeare invented a quarter of our language”
 - “Shakespeare invented half the words in the English language”
 - “Shakespeare *is* our language”
- 

“Shakespeare coined more words than any other writer” (contd.)

Work on neologisms:

- 1,502 words recorded in the Oxford English Dictionary as first citations in Shakespeare
- We are checking these in *Early English Books Online* (EEBO-TCP): 125,000 titles of printed material, 1473-1700. About 1.2 billion words.

Preliminary findings:

- If the current pattern continues, less than a quarter of those 1,502 words can reasonably be attributed to Shakespeare.
- 

“Shakespeare coined more words than any other writer” (contd.)

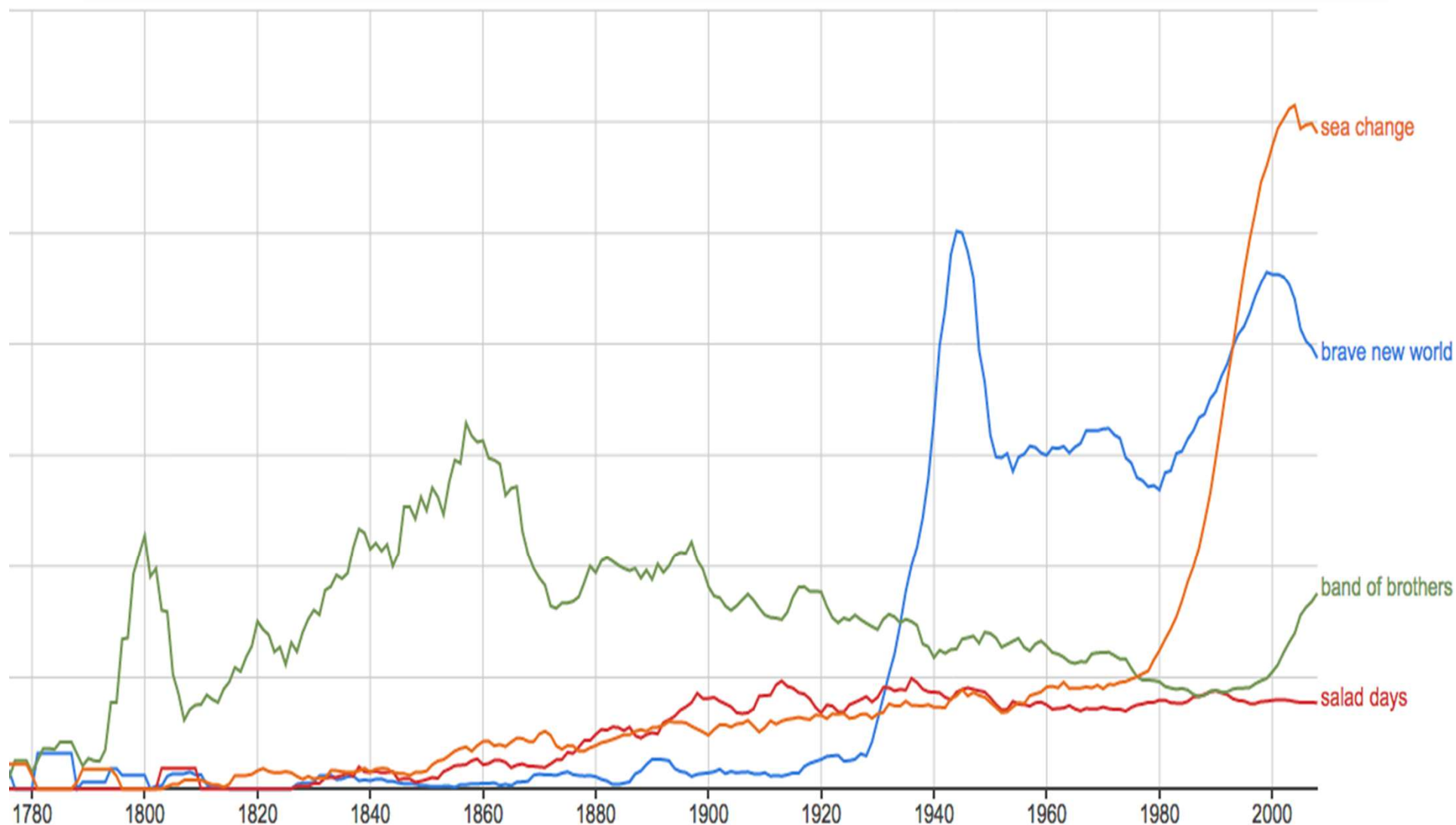
Issues

- How do we know that Shakespeare coined it as opposed to recorded it? Cf. *down staires* vs. *incarna[r]dine* (v.)
- Is it actually just a nonce word rather than neologism? Cf. *dropsied* vs. *domineering*
- People assume that Shakespeare’s coinages survive into today’s English – do they?

A glimpse of phrases first recorded in Shakespeare and their more recent life.

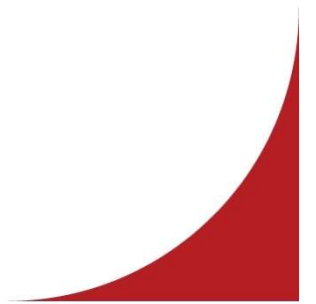


Four phrases first recorded in Shakespeare and their use in printed material over the last 200 years (Google's N-Gram Viewer)



A brief digression on the key problems in counting words

Word-forms and lexemes

- Dictionary headword:
do
 - Modern (morphological) word-forms:
do, does, doing, did, done
 - Early modern (morphological) word-forms:
do, does, do(e)st, doth, doing, did, didst, done
- 

A brief digression on the key problems in counting words (contd.)

Word-forms and lexemes

Dictionary headword:

do = 1

Modern (morphological) word-forms:

do, does, doing, did, done = 5

Early modern (morphological) word-forms:

do, does, do(e)st, doth, doing, did, didst, done = 8



A brief digression on the key problems in counting words (contd.)

Spelling variation:

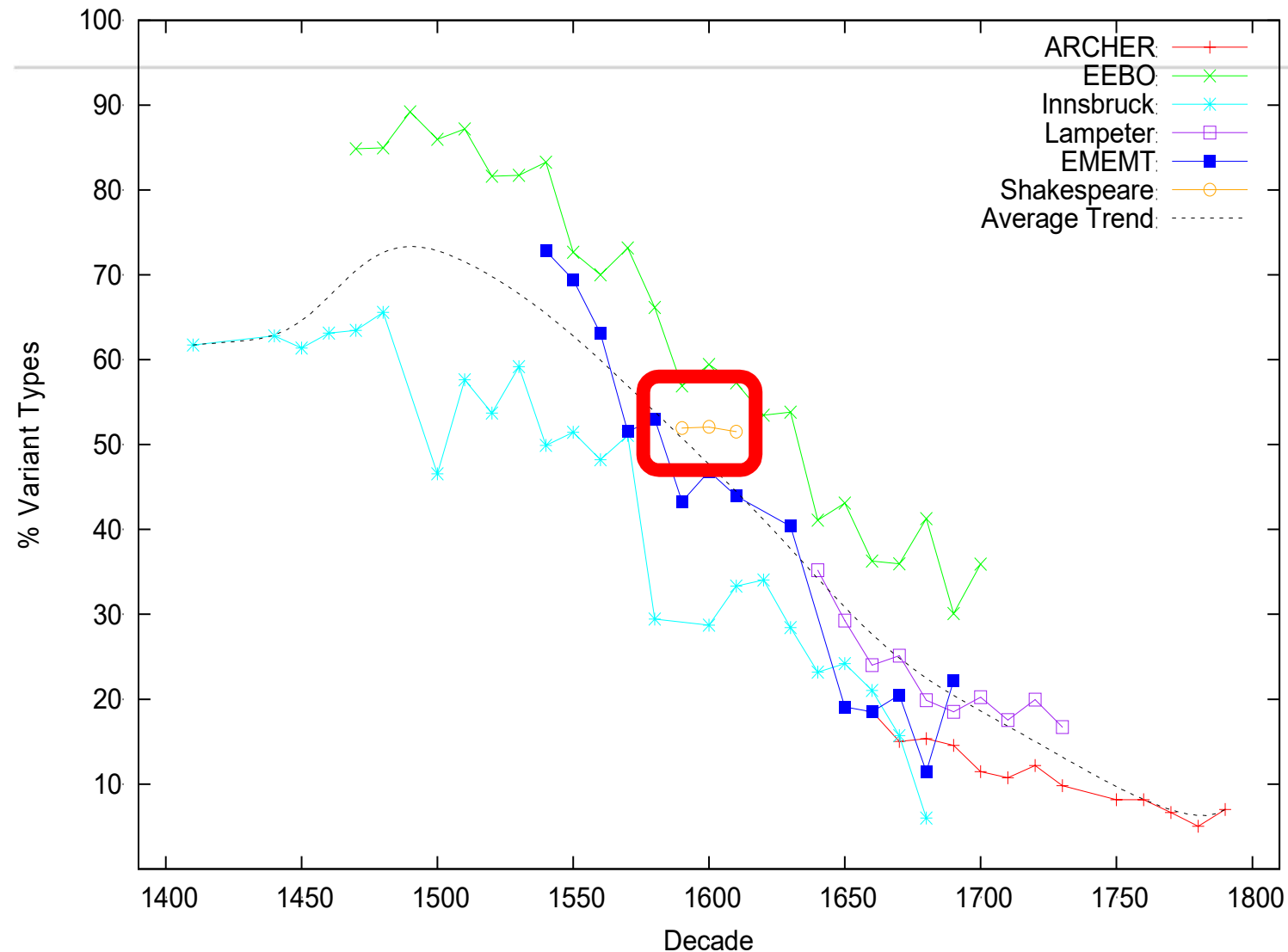
Problem: You decide to study the use of the word *would* in a corpus. You type it into your search program ... and look at the result.

But you miss: *wold*, *wolde*, *woolde*, *wuld*, *wulde*, *wud*, *wald*, *vvould*, *vvold*, etc., etc.

One orthographic word today; many in EModE.



A glance at the First Folio and spelling variation in English (Baron, Rayson & Archer 2009)

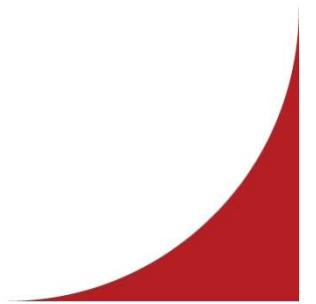


A brief digression on the key problems in counting words (contd.)

Solution: *Variant Detector* (VARD2) program, primarily devised by generations of scholars at Lancaster, but most recently given a significant boost by Alistair Baron.

Further problem: What do you regularize the spelling to? There is no standardised regular form in the way that there is today.

This program also enabled us to tag, and thus easily track, any dialectal or 'foreign' words we encountered



“Shakespeare had ‘small Latin’”

(with Caterina Guardamagna)

Shakespeare corpus: 38 plays; 1,038,509 words

Comparative plays corpus: 46 plays by 24 other playwrights; 1,091,729 words

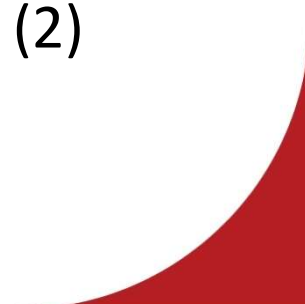
From our Latin counts, we excluded: mock Latin expressions, ambiguous cases (i.e. ambiguous between Latin and another language, e.g. Italian, French), proper nouns, stage directions



“Shakespeare had ‘small Latin’”

A glance at the top-50 items from the Latin list

hic (13), ergo (8), ad (6), pauca (6), quondam (6), videlicet (6), cum (5), inprimis (5), benedictus (4), extempore (4), aliis (3), bone (3), deum (3), horum (3), ibat (3), lapis (3), mater (3), nec (3), quis (3), quod (3), quo (3), senis (3), sigeia (3), suis (3), tellus (3), accusativo (2), aer (2), armiger (2), benedicite (2), bene (2), caret (2), cucullus (2), dii (2), domine (2), facit (2), fatuus (2), hac (2), haud (2), hoc (2), ignis (2), ipse (2), lentus (2), manes (2), mollis (2), monachum (2), mulier (2), nominativo (2), omne (2), perge (2), primus (2)



“Shakespeare had ‘small Latin’” (contd.)

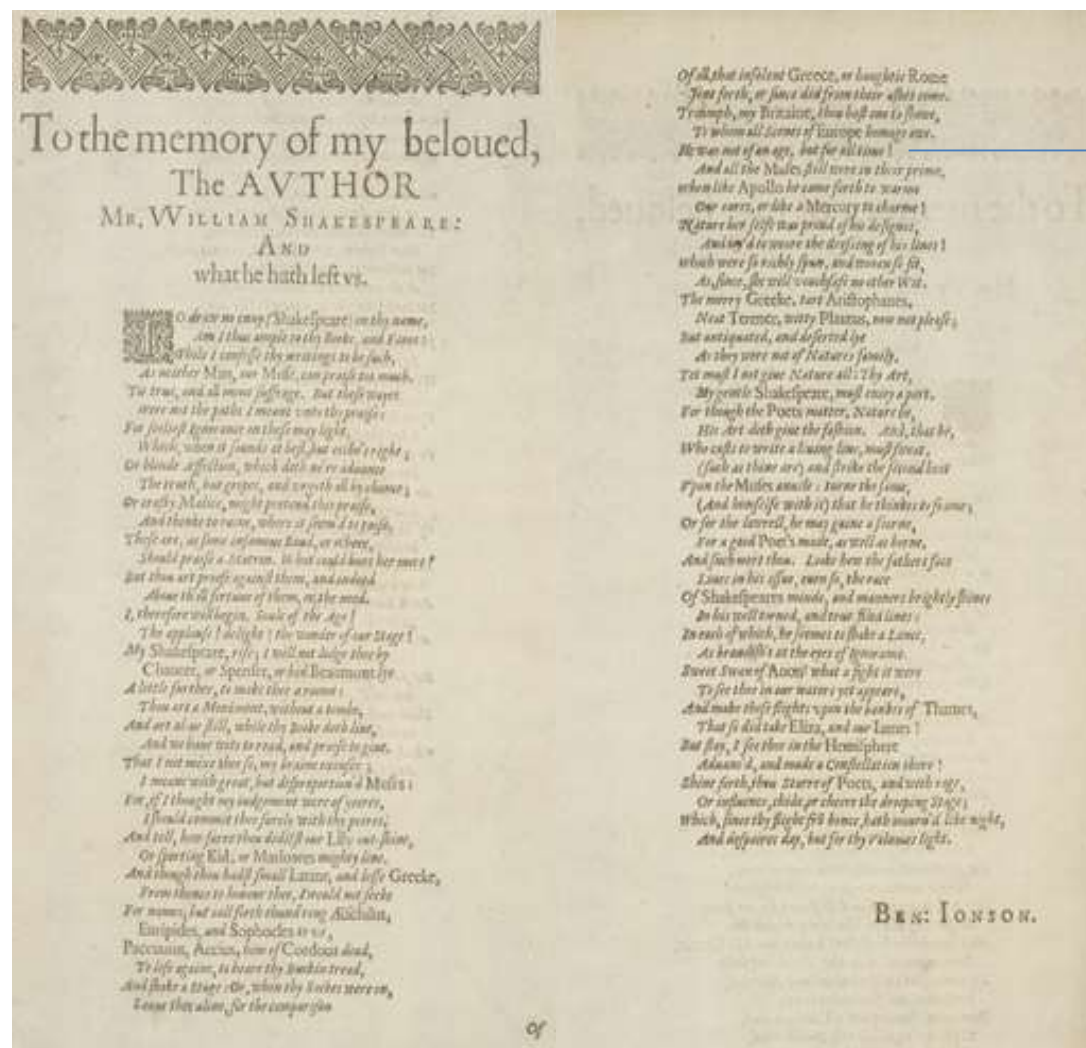
The number of Latin words in the Shakespeare Corpus and the occurrence of those words in the Comparative Corpus of contemporary Playwrights

	<i>Types</i>	<i>Tokens</i>	<i>Type-Token Ratio</i>	High = More lexical variation Low = Less lexical variation
<i>Shakespeare corpus (focussed list)</i>	245	362	67%	
<i>Comparative corpus of contemporary playwrights</i>	28	74	38%	



Shakespeare’s Latin vocabulary is more varied than his contemporaries put together

“Shakespeare’s language transcends time and space”




He was not of an age, but for all time!

“Shakespeare’s language transcends time and space” (contd.)

- Universal characters, emotions, themes language??

“Shakespeare has given us a universal language medium in which are crystalized the battle hymns, the intellectual conceptions and the spiritual aspirations of the Anglo-Saxons.”

(Rutherford, N.J. and Bennett, E.H., 1918-1922, *English Speaking World*, Vol.2
(8): 14)

- No Shakespearean dictionary has treated Shakespeare’s language as relative, i.e. put Shakespeare’s usage in the context of that of his contemporaries.
- 

Good

Crystal & Crystal (2004:201-202):

(1) [intensifying use] real, genuine ('love no man in good earnest'). (2) kind, benevolent, generous. (3) kind, friendly, sympathetic. (4) amenable, tractable, manageable. (5) honest, virtuous, honourable. (6) seasonable, appropriate, proper. (7) just, right, commendable. (8) intended, right, proper. (9) high-ranking, highborn, distinguished. (10) rich, wealthy, substantial.

+ phrases and compounds

good (n.) 1 goodwill, goodness *HF* IV.ii.50 [Griffith to Katherine, of the fall of Wolsey's foundation at Ipswich] *Drowning to further the good that did it*

2 good fellow *Per* II. Chorus:9 [Gower alone, of Pericles] *The good in conversation ... / Is still at Jansan* *Tim* I.ii.3, 15; *Tim* II.ii.75

3 interest, advantage, benefit *Cor* II.ii.234 [Sicinius to Brutus, of Coriolanus] *It shall be to him then as our good will, / As sure destruction*

good, do prosper, succeed, triumph *2H6* V.ii.14 [Dick to Cade] *If we mean to thrive and do good, break open the good*

good, do one 1 make prosper, enable to succeed *MA* I.ii.269 [Claudio to Don Pedro] *Your highness now may do me good* *2H6* I.ii.77; *2H6* III.ii.38

2 be of use to, provide assistance to *2H6* III.ii.119 [King to all, of Gloucester] *myself ... / Look after him, and cannot do him good*

good (adj.) 1 [intensifying use] real, genuine *AF* I.ii.25 [Celia to Rosalind] *love no man in good earnest* *Tim* II.ii.232

2 kind, benevolent, generous *Tim* I.ii.242 [Timon to Apemantus] *if thou wert not sudden, I would be good to thee* *Tim* III.ii.24

3 kind, friendly, sympathetic *2H6* V.ii.274 [Bonifacio to all] *The King was sometime ripe for a good word*

4 amenable, tractable, manageable *TS* I.ii.127 [Hortensio to Gremio, of a husband for Katherine] *there be good fellows in the world, as a man could fight on them*

5 honest, virtuous, honourable *Cor* II.ii.58 [Menenius to Brutus and Sicinius, of their supporters] *they be deadly that tell you have good faces*

6 seasonable, appropriate, proper *CE* II.ii.70 [Antipholus of Syracuse to Dromio of Syracuse] *learn to just in good time*

7 just, right, commendable *2H6* II.ii.28 [Suffolk to Gloucester, of malice] *no worse than well because / So good a quarrel*

8 intended, right, proper *RJ* I.ii.45 [Mercutio to Romeo] *Take our good meaning*

9 high-ranking, highborn, distinguished *RJ* I.ii.40 [Bolingbroke to Mowbray] *Thou art ... a miscreant, / Too good to be so* *2H6* better (adj.)

10 rich, wealthy, substantial *Cor* I.ii.15 [First Citizen to all] *We are accounted poor citizens, the patricians good* *2H4* IV.ii.15; *2H6* I.ii.12

good, 'tis very well *Tit* I.ii.412 [Saturninus to Bassianus] *'Tis good, sir. You are very short with us*

good (adv.) [intensifying use] really, genuinely *2H4* III.ii.44 [Folstaf to Bardolph] *the sack that thou hast drunk we would have bought we lights us good cheap at the dearest chandler's in Europe*

good-conceited (adj.) cleverly devised, ingeniously composed *Cym* II.ii.16 [Cloten to musicians, of their playing to Innogen] *a very excellent good-conceited thing*

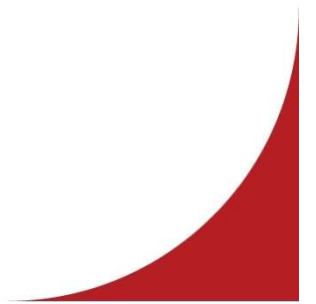
good deed (adv.) in truth, in reality *WT* I.ii.42 [Hermione to Leontes] *yet, good deed, Leontes, / I love thee not a jar o' th' clock behind / What lady she her lord* *2H6* **DISCOURSE MARKERS**, p.127

good-den / e'en / even / morrow *2H6* **GREETINGS**, p.206; **FAREWELLS**, p.170

good-faced (adj.) smooth-faced, pretty *WT* IV.ii.112 [Autolycus to Clown] *good-faced sir ... sweet sir*

good life comfortable position, respectable way of life *2H6* III.ii.112 [Mistress Page to Mistress Ford] *defend your reputation, or bid farewell to your good life for ever* *2H6* **SONG OF GOOD LIFE**

Crystal & Crystal (2004:201-202):


- (1) [intensifying use] real, genuine ('love no man in good earnest').
 - (2) **kind**, benevolent, generous.
 - (3) **kind**, friendly, sympathetic.
 - (4) amenable, tractable, manageable.
 - (5) honest, virtuous, honourable.
 - (6) seasonable, appropriate, **proper**.
 - (7) just, **right**, commendable.
 - (8) intended, **right**, **proper**.
 - (9) high-ranking, highborn, distinguished.
 - (10) rich, wealthy, substantial.
- 

Good

No	Text	Solution 1 to 50	Page 1 / 71
1	WT WT_1_2	's some comfort . What ? Camillo there ? Aye , my	good Lord . Go play (Mamillius) thou art an honest man
2	H5 H5_3_6	It is well . The Fig of Spain . Exit . Very	good . Why , this is an arrant counterfeit Rascal , I remember
3	TC TC_2_1	if he knock out either of your brains , he were as	good crack a fusty nut with no kernel . What with me too
4	TN TN_4_2	Who I sir , not I sir . God be with you	good sir Topas : Marry Amen . I will sir , I will
5	1H4 1H4_1_3	my Love , and your high Majesty . The circumstance considered ,	good my Lord , Whatever Harry Percy then had said , To such
6	R3 R3_2_4	Because sweet Flowers are slow , and Weeds make haste .	Good faith , good faith , the saying did not hold In him
7	MND MND_1_1	and Demetrius . Happy be Theseus , our renowned Duke . Thanks	good Egeus : what 's the news with thee ? Full of vexation
8	Oth Oth_3_4	this is from some Mistress , some remembrance ; No , in	good troth Bianca . Why , whose is it ? I know not
9	Cor Cor_1_1	Country ? Very well , and could be content to give him	good report for it , but that he pays himself with being proud
10	CE CE_3_2	hold you still : I 'll fetch my sister to get her	good will . Exit . Enter Dromio , Syracuse . Why how now
11	R2 R2_1_1	heaven . Thou art a Traitor , and a Miscreant ; Too	good to be so , and too bad to live , Since the
12	WT WT_2_3	commanded None should come at him . Not so hot (good Sir) I come to bring him sleep . T is such
13	1H6 1H6_3_1	than I do ? except I be provoked . No , my	good Lords , it is not that offends , It is not that
14	1H4 1H4_4_4	how much they do import , You would make haste . My	good Lord , I guess their tenor . Like enough you do .
15	WT WT_2_3	and frame of Hand , Nail , Finger .) And thou	good Goddess Nature , which hast made it So like to him that
16	Cym Cym_4_2	In this place we left them ; I wish my Brother make	good time with him , You say he is so fell . Being
17	Oth Oth_3_3	To let you know my thoughts . What dost thou mean ?	Good name in Man , & woman (dear my Lord) Is
18	KJ KJ_4_3	he may inquire us out . Exeunt Lords . Here 's a	good world : knew you of this fair work ? Beyond the infinite
19	WT WT_3_2	(cracking it) Break too . What fit is this ?	good Lady ? What studied torments (Tyrant) hast for me ?
20	2H4 2H4_2_1	I warrant he is an infinitive thing upon my score .	Good Master Fang hold him sure : good Master Snare let him not
21	Cor Cor_4_6	very many of us , that we did we did for the	best , and though we willingly consented to his Banishment , yet it
22	2H6 2H6_4_9	you love your Prince & Country : Continue still in this so	good a mind , And Henry though he be unfortunate , Assure yourselves

Good

Pretend some alteration in **good** will? What's heere? I haue vpon
My selfe, and my **good** Cousin Buckingham, Will to your Mother,
she is low voic'd. Cleo. That's not so **good**: he cannot like her long.
Goodmorrow (**good** Lieutenant) I am sorrie For your displeasure:
Father Frier. Duk. And you **good** Brother Father; what offence
an enuious emulator of euery mans **good** parts, a secret & villanous
she shall be there. Ro. And stay thou **good** Nurse behind the Abbey wall,
Mar. Patience deere Neece, **good** Titus drie thine eyes. Ti. Ah Marcus,
Anthonio; that I had a title **good** enough to keepe his name company!
the singlenesse. Mer. Come betweene vs **good** Benuolio, my wits faints.
Enter Count Rossillion. Par. **Good**, very good, it is so then: good, very



'good' (contd.)

No.	Word	Total no. in whole corpus	Expected collocate frequency	Observed collocate frequency	In no. of texts	Z- score
1	morrow	332	8.45	113	28	35.793
2	Lord	2,591	65.948	287	33	27.159
3	.	66,858	1701.727	2,703	36	24.26
4	night	785	19.981	88	24	15.105
5	good	2,756	70.148	184	30	13.534
6	cheere	56	1.425	17	14	12.627
7	my	12,103	308.056	526	36	12.389
8	Sooth	66	1.68	18	13	12.206
9	bad	157	3.996	28	18	11.758
10	newes	278	7.076	38	16	11.437

Good (= 3507)

Encyclopaedia of Shakespeare's Language

Definition preview: definition 104 for entry [good_ADJ](#)

good**** *adj.* (good, better, best):

1. A polite address: '(my) good Lord/friend/Sir/Master/Lady/Madam/etc.'. Typically used when meeting or parting, thanking or making suggestions. *But (good my Lord) do it so cunningly* TGV, III. 1.
2. Honest, truthful, principled; of high moral standards. (This sense also shapes the discourse markers '(in) good faith/sooth/troth', which mean truly or honestly). *a man of good repute, carriage, bearing, & estimation* LLL, I. 1.
3. Positive rather than negative. Typically, contrasted with 'bad'. *Is thy news good or bad?* ROM, II. 5.
4. In one's favour, especially favourable wishes or blessings. *The Gods be good to us* COR, V. 4.
5. A welcoming, cheerful manner. *Therefore for Gods sake entertain good comfort, And cheer his Grace with quick and merry eyes* R3, I. 3.

good will As now.

good morrow Good morning.

good night As now.



Do the word-networks of *good*, *bad*, *ill* and *evil* overlap? Are there strong links amongst them?

(*ill* = 120)
(*evil* = 21)

Does Shakespeare's usage reflect that of his contemporaries? A glance at collocates in EEBO 1560-1640

<i>Good</i> (= 577,864)	<i>Bad</i> (= 28,137)	<i>Evil</i> (= 61,561)	<i>Ill</i> (= 46,366)
Works A Evil Thought Success Bad Conscience For Do Will	Debtor Witch Bad Worse Good Dealing Success Humours Usage News	Dooers Willers Concupiscences Conditioned Speakers Livers Adulteries Cogitations Favoured Eschew	Favouredly Willers Hap Favoured Conditioned Beseeming Speed Presage Husbandry Entreated

Character profiles: *Romeo and Juliet*



Lily James
and Richard
Madden.

(Photo: Johan Perrson)

- What language characterizes Romeo and what language, Juliet?
- What are their linguistic styles, their style markers, their keywords?

Character profiles: *Romeo and Juliet*

Rank-ordered keywords for Romeo and Juliet (raw frequencies in brackets)

Romeo	Juliet
<p>beauty (10), love (46), blessed (5), eyes (14), more (26), mine (14), dear (13), rich (7), me (73), yonder (5), farewell (11), sick (6), lips (9), stars (5), fair (15), hand (11), thine (7), banished (9), goose (5), that (84)</p>	<p>if (31), be (59), or (25), I (138), sweet (16), my (92), news (9), thou (71), night (27), would (20), yet (18), that (82), nurse (20), name (11), words (5), Tybalt's (6), send (7), husband (7), swear (5), where (16), again (10)</p>

Character profiles: *Romeo and Juliet*

Romeo:

- She hath, and in that sparing makes huge waste; For beauty, starv'd with her severity, Cuts beauty off from all posterity. She is too fair, too wise, wisely too fair, To merit bliss by making me despair: She hath forsworn to love, and in that vow Do I live dead that live to tell it now. (I.i)
- If I profane with our unworthiest hand This holy shrine, the gentle sin is this; Our lips, two blushing pilgrims, ready stand To smooth that rough touch with a tender kiss. (I.v)

Character profiles: *Romeo and Juliet*

Juliet:

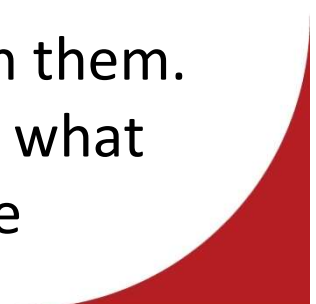
- If he be married, / Our grave is like to **be our** wedding-bed (I.v.)
- If they do see thee, they will murder thee (II.ii.)
- But if thou meanest not well (II.ii.)
- Is thy news good, or bad? answer to that; Say either, and I'll stay the circumstance: Let me be satisfied, is 't good or bad? (II.ii)
- 'Tis almost morning; I would have thee gone; And yet no further than a wanton's bird [...] (II.ii.)

Concluding remarks

I hope we have shown:

- Shakespeare's language is relative, as is all language.
- Our approach to meanings is very different from the traditional approach, and especially effective in providing wholistic distant readings.
- It is evidence-based, and thus can be used to probe 'myths' about language.
- It can also contribute to styles ... styles that create characters, plays, etc.

Our approach uses computers, but it is never dependant on them. A human must decide on what goes into the computer and what the computer does with it, and, crucially, must interpret the results.



Project website

<http://wp.lancs.ac.uk/shakespearelang/>



Encyclopaedia of Shakespeare's Language Project
Contents of comparative corpus for Shakespeare's plays

Author	Title	First production*	First publication*	Edition in corpus**
Comedy				
John Lyly	<i>Alexander and Campaspe</i>	c.1583	1584	1584
John Lyly	<i>Gallathea</i>	1585	1592	1592
Robert Greene	<i>Friar Bacon and Friar Bungay</i>	1589	1594	1594
George Peele	<i>The Old Wives Tale</i>	1590	1595	1595
George Chapman	<i>The Blind Beggar of Alexandria</i>	1596	1598	1598
Thomas Heywood	<i>The Fair Maid of the West Part I</i>	1604	1631	1631
George Chapman	<i>An Humorous Days Mirth</i>	1597	1599	1599
Henry Porter	<i>The Two Angry Women of Abington</i>	c.1598	1599	1599
Anonymous	<i>Mucedorus</i>	1590	1598	1598
Thomas Dekker	<i>Old Fortunatus</i>	1599	1600	1600
Thomas Heywood	<i>How a Man May Chuse</i>	1602	1602	1602
Ben Jonson	<i>Volpone</i>	1606	1616	1616
Francis Beaumont and John Fletcher	<i>The Woman Hater</i>	1606	1607	1607
George Wilkins	<i>The Miseries of Inforced Marriage</i>	1606	1607	1607
Francis Beaumont	<i>The Knight of the Burning Pestle</i>	1607	1613	1613
John Fletcher	<i>The Faithful Shepherdess</i>	1608	c.1610	1610
Francis Beaumont and John Fletcher	<i>Philaster</i>	1609	1620	1620
Thomas Middleton	<i>The Roaring Girl</i>	1611	1611	1611
Ben Jonson	<i>Bartholomew Fayre</i>	1614	1631	1631
Philip Massinger	<i>The Bondman</i>	1623	1624	1624

*Dates of first production and first publication are from the Database of Early English Playbooks: <http://deep.sas.upenn.edu/>

**Play-texts sourced from Early English Books Online: <http://eebo.chadwyck.com/home>

Author	Title	First production*	First publication*	Edition in corpus**
History				
Robert Greene	<i>The Scottish History of James the Fourth</i>	c.1590	1598	1598
Christopher	<i>Tamburlaine Part I</i>	c. 1587	1590	1590
Christopher	<i>Edward II</i>	1592	1594	1594
George Peele	<i>The Famous Chronicle of Edward I</i>	1591	1593	1593
Christopher	<i>The Massacre at Paris</i>	1593	c.1594	1594
George Peele	<i>The Battle of Alcazar</i>	1589	1594	1594
Anthony Munday	<i>The Death of Robert Earl of Huntingdon</i>	1598	1601	1601
Thomas Heywood	<i>Edward IV Part I</i>	1599	1600	1600
Thomas Heywood	<i>Edward IV Part II</i>	1599	1600	1600
Anonymous	<i>The Life of Sir John Oldcastle</i>	1599	1600	1600
Thomas Heywood	<i>If You Know Not Me, You Know Nobody Part I</i>	1604	1605	1605
Thomas Dekker	<i>Sir Thomas Wyatt</i>	1602	1607	1607
Robert Armin	<i>The Valiant Welshman</i>	1612	1615	1615
Thomas Drue	<i>The Duchess of Suffolk</i>	1624	1631	1631
Tragedy				
Thomas Kyd	<i>The Spanish Tragedy</i>	1587	1592	1592
Christopher	<i>The Jew of Malta</i>	1589	1633	1633
Christopher	<i>Dr Faustus</i>	1592	1604	1604
Christopher	<i>Dido, Queen of Carthage</i>	1586	1594	1594
Thomas Heywood	<i>A Woman Killed With Kindness</i>	1603	1607	1607
John Marston	<i>The Malcontent</i>	1604	1604	1604
Ben Jonson	<i>Sejanus</i>	c.1604	1604	1604
Francis Beaumont and John Fletcher	<i>The Maid's Tragedy</i>	1610	1619	1619
John Webster	<i>The White Devil</i>	1612	1612	1612
John Webster	<i>The Duchess of Malfi</i>	1614	1623	1623
Thomas Middleton and William	<i>The Changeling</i>	1622	1653	1653
Thomas Middleton	<i>Women Beware Women</i>	1621	1657	1657