CONNECTING FROM CORPUS RESULTS TO EXTERNAL RESOURCES (see video)

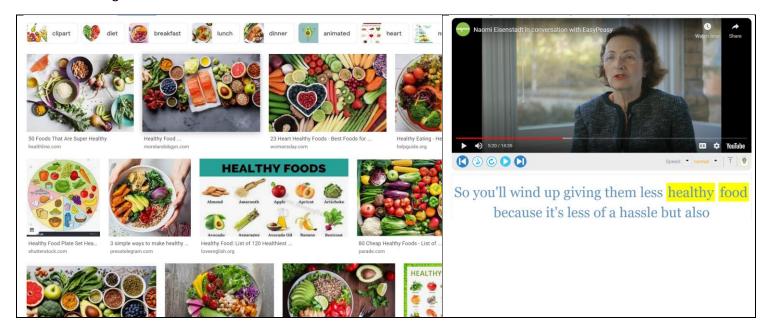
A unique feature of the corpora from English-Corpora.org is the ability to click on the corpus results to find and use external resources related to a word or phrase. In <u>all seventeen of the corpora</u>, the "results" page for a search contains a number of links to external resources. For example, the following are the results for <u>healthy NOUN</u> in the one billion word <u>COCA</u> corpus (the Corpus of Contemporary American English):



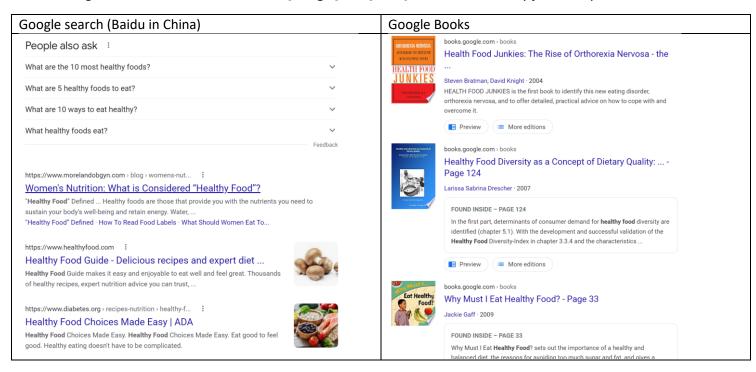
By default, clicking on any of the phrases will show the standard "concordance" or "keyword in context (KWIC)" display (as in the red link above). For example, the following are a few of the concordance lines for *healthy food*.



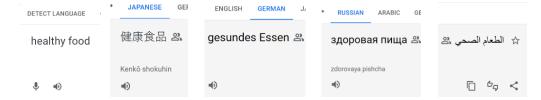
But at English-Corpora.org, you can also see translations for any of these words or phrases, as well as see links to a number of other external resources. For example, clicking on the [Image] or [Pron/Video] (pronunciation) link will show resources like the following:



The following are external resources from the [Google] and [Book] links for the healthy food entry in the results:



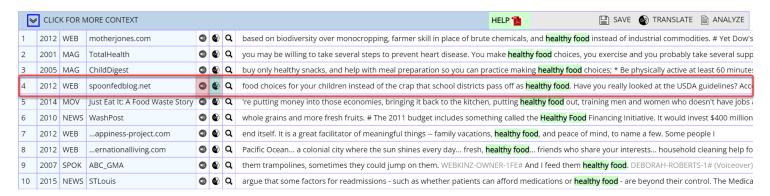
You can also select [**Translation**] and then click on any word for phrase in the results page to see the translation from Google Translate, as in the results from Japanese, German, Russian, or Arabic below (or any of 100+ additional languages):



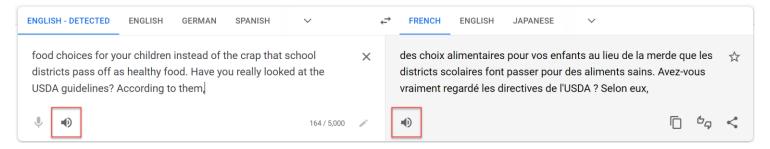
You can also click on [Entire page] in the corpus results to see the **translation of** *all of the entries*, into any of the 100+ languages, as with the results for the corpus search <u>soft NOUN</u>:

1120 soft tissue	Tejido suave	1120 soft tissue	软组织	1120 soft tissue	연조직	1120 soft tissue	мягких тканей
1109 soft drinks	bebidas sin alcohol	1109 soft drinks	软饮料	1109 soft drinks	청량 음료	1109 soft drinks	безалкогольные напитки
867 soft spot	Punto débil	867 soft spot	情有独钟	867 soft spot	약점	867 soft spot	Мягкое место
790 soft money	dinero blando	790 soft money	软钱	790 soft money	연화	790 soft money	мягкие деньги
721 soft drink	refresco	721 soft drink	汽水	721 soft drink	청량 음료	721 soft drink	безалкогольный напиток
546 soft voice	voz suave	546 soft voice	轻声细语	546 soft voice	부드러운 목소리	546 soft voice	мягкий голос
421 soft power	Poder suave	421 soft power	软实力	421 soft power	소프트 파워	421 soft power	мягкая сила
328 soft tissues	tejidos blandos	328 soft tissues	软组织	328 soft tissues	연조직	328 soft tissues	мягкие ткани
274 soft landing	Aterrizaje suave	274 soft landing	软着陆	274 soft landing	연착륙	274 soft landing	мягкая посадка
247 soft light	luz tenue	247 soft light	柔光	247 soft light	부드러운 빛	247 soft light	мягкий свет
219 soft skills	habilidades blandas	219 soft skills	软技能	219 soft skills	소프트 스킬	219 soft skills	мягкие навыки
216 soft peaks	picos suaves	216 soft peaks	软峰	216 soft peaks	소프트 피크	216 soft peaks	мягкие пики
213 soft touch	tacto suave	213 soft touch	柔软的触感	213 soft touch	소프트 터치	213 soft touch	мягкое прикосновение
207 soft skin	piel suave	207 soft skin	软皮	207 soft skin	부드러운 피부	207 soft skin	мягкая кожа
176 soft music	música suave	176 soft music	轻柔的音乐	176 soft music	부드러운 음악	176 soft music	мягкая музыка
167 soft hands	manos suaves	167 soft hands	柔软的手	167 soft hands	부드러운 손	167 soft hands	мягкие руки
158 soft leather	piel suave	158 soft leather	软质皮革	158 soft leather	부드러운 가죽	158 soft leather	мягкая кожа
148 soft spots	puntos débiles	148 soft spots	软点	148 soft spots	부드러운 반점	148 soft spots	мягкие пятна
146 soft targets	objetivos blandos	146 soft targets	软目标	146 soft targets	소프트 타겟	146 soft targets	легкие цели
143 soft glow	brillo suave	143 soft glow	柔和的光芒	143 soft glow	부드러운 빛	143 soft glow	мягкое свечение
9		1		1		1	

One of the problems of KWIC entries for language learners is simply knowing what the other words in the KWIC line mean. At English-Corpora.org, there are also links to external resources for each entry in the **concordance / Keyword in Context (KWIC)** page, as with the following KWIC entries for *healthy food*:



The link highlighted above in green will take you to a **translation of the KWIC entry**. You can also hear the KWIC line via the red link (and then via the "speaker" icon at Google Translate). (Baidu Translate is used in China.)

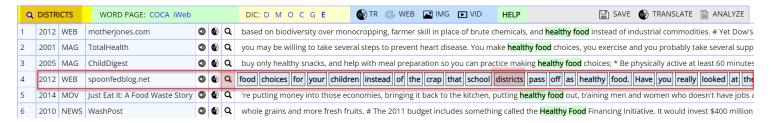


If you want, you can select multiple KWIC lines (up to 100 lines), and see a translation of all of the selected lines:

- 5 : the author would have pointed out the inconsistency with Michelle's garden and advocating for healthy food and Barack's appointment of Monsanto's Michael Taylor to the head of Food 5: el autor habría señalado la inconsistencia con el jardín de Michelle y la defensa de los alimentos saludables y el nombramiento de Barack de Michael Taylor de Monsanto a la cabeza de Alimentos
- 8: organizations feed the hungry 24/7/365. It is my sincere belief that everyone deserves a healthy food supply, regardless of his or her ability to pay for it. Your
 8: las organizaciones alimentan a los hambrientos 24/7/365. Creo sinceramente que todos merecen un suministro de alimentos saludables, independientemente de su capacidad para pagarlos. Su
- 10 : a minute or so longer. We're talking about how to feed your child healthy food, hopefully healthy food that he or she will like for a lifetime.

 10: un minuto más o menos. Estamos hablando de cómo alimentar a su hijo con alimentos saludables, con suerte alimentos saludables que le gustarán toda la vida.
- 12: fatty foods. It has taken some time to reshape those and find pleasure in healthy food. But I have had a good influence on my father, he is
- 12: alimentos grasos. Ha llevado algún tiempo remodelarlos y encontrar placer en la comida saludable. Pero he tenido una buena influencia en mi padre, él es
- 17: of the human potential movement will refresh the spirit. Much of the hearty and healthy food is grown onsite at the Esalen Gardens, and the hot springs overlooking the
- 17: del movimiento del potencial humano refrescará el espíritu. Gran parte de la comida abundante y saludable se cultiva en los Jardines de Esalen, y las aguas termales con vistas al

You can also click on the |Q| icon (highlighted in red, below) to see a wide range of information for any word in the KWIC entry, such as the word *districts*. This information includes a detailed "word page" (discussed later in this document), links to definitions from external dictionaries, translations, and web, image, and video links. So with a couple of clicks, learners have quick and easy access to any words in the KWIC entry. Unlike the results at other corpus sites, at English-Corpora.org there is never any reason for non-native speakers to feel intimidated by words that they don't know in the KWIC entries.



In terms of KWIC results, you can also see a translation of the "expanded KWIC" entry, as is shown for line #4 below:

Google Translate English - detected → Spanish ✓

don't et from home. Stuff their parents should be teaching them! How to be respectful, caring, compasionate, truthful, trustworthy, and healthy. I better stop otherwise I'll be up on my soapbox and there's no telling when I'll come down. Because we all know (insert sarcastic tone here) that a cheese stick and a graham cracker is a top-notch breakfast that will awaken a young mind! **barf!! ***barf!! ***

no et desde casa. ¡Cosas que sus padres deberían enseñarles! Cómo ser respetuoso, cariñoso, compasivo, veraz, digno de confianza y saludable. Será mejor que me detenga, de lo contrario estaré arriba en mi tribuna y no se sabe cuándo bajaré. ¡Porque todos sabemos (inserte un tono sarcástico aqui) que un palito de queso y una galleta Graham es un desayuno de primer nivel que despertará una mente joven! ** vomitar!! ** # OK. Seguí el enlace y no me pude resistir. Envié un correo electrónico a estos padres locos. (ver más abajo) ¡Me pregunto si obtendré una respuesta! # " Como padre y maestro, estoy sorprendido, no, conmocionado de que no desee opciones de alimentos mejores y más saludables para sus hijos en lugar de la basura que los distritos escolares hacen pasar por <u>saludable</u> alimento. ¿Realmente has mirado las pautas del USDA? Según ellos, las patatas fritas cuentan como verdura. ¿¿¿En serio??? ¿Déjame preguntarte esto? ¿Cuántas veces se cocina la comida de sus hijos? En el distrito en el que trabajo, la comida se cocina dos veces. Para cuando los niños lo obtienen, ya no tiene ningún valor nutricional y el sabor también se ha ido. No hay absolutamente nada de malo en que los niños tengan opciones más saludables. ¿Por qué deberían estar continuamente sujetos a chatarra procesada? No es bueno para sus cuerpos. Y los cuerpos enfermos hacen que los estudiantes sean lentos. Los maestros enseñan. Alimentación del Servicio de Alimentos. Si el Servicio de Alimentos quiere mejorar el menú en Cobbles, entonces, como padres, deben apoyarlos. Sin interponerse en su camino. Lo que mis hijos (mis estudiantes y yo mismo) ponen en

Finally, in each of the seventeen corpora from English-Corpora.org, you can quickly and easily create "Virtual Corpora", based on words or phrases in the texts, or based on metadata (e.g. date, author, or source). For example, you could create a "Virtual Corpus" dealing with *telescopes* in about 2-3 seconds, and then find the keywords from this corpus in 1-2 seconds more. And then via the [Translate all Entries] link, you can see the **translation of all of the keywords** for your Virtual Corpus:

HELP	ENTRY	SAVE	WORD (CLICK FOR CONTEXT)	light-year	Lichtjahr	light-year	光年	light-year	سنة ضوئية	light-year	световой год
			TRANSLATE ALL ENTRIES	spectrograph	Spektrograph	spectrograph	スペクトログラフ	spectrograph	مطياف	spectrograph	спектрограф
1	\$	*	LIGHT-YEAR	astronomer	Astronom	astronomer	天文学者	astronomer	عالم الفلك	astronomer	астроном
2	\$	*	SPECTROGRAPH	sunspot	Sonnenfleck	sunspot	黒点	sunspot	بقعة شمسية	sunspot	солнечное пятно
3	\$	*	ASTRONOMER	quasar	Quasar	quasar	クエーサー	quasar	الكوازار	quasar	квазар
4	\$	*	SUNSPOT	telescope	Teleskop	telescope	望遠鏡	telescope	تلسكوب	telescope	телескоп
5	\$	*	QUASAR	eyepiece	Okular	eyepiece	接眼レンズ	eyepiece	العدسة	eyepiece	окуляр
6	\$	*	TELESCOPE	refractor	Refraktor	refractor	屈折望遠鏡	refractor	المنكسر	refractor	рефрактор
7	\$	*	EYEPIECE	12000	Wellenlänge	wavelength	波長	wavelength	The same and the same	wavelength	длина волны
8	\$	*	REFRACTOR	wavelength	J	reflector	リフレクター	reflector	الطول الموجي	reflector	отражатель
9	\$	*	WAVELENGTH	reflector	Reflektor			100 market and 100 miles	العاكس	optics	оптика
10	\$	*	REFLECTOR	optics	Optik	optics	光学	optics	بصريات	-00 8 -00-00-00-0	
11	\$	*	OPTICS	aperture	Öffnung	aperture	絞り	aperture	فتحة	aperture	отверстие
12	\$	*	APERTURE	kilometer	Kilometer	kilometer	キロメートル	kilometer	كيلومتر	kilometer	километр
13	\$	*	KILOMETER	supernova	Supernova	supernova	超新星	supernova	سوبرنوفا	supernova	сверхновая
14	\$	*	SUPERNOVA	observatory	Observatorium	observatory	天文台	observatory	المرصد	observatory	обсерватория
15	\$	*	OBSERVATORY	astronomy	Astronomie	astronomy	天文学	astronomy	الفلك	astronomy	астрономия
16	\$	*	ASTRONOMY	finder	Finder	finder	ファインダ	finder	مكتشف	finder	искатель
17	\$	*	FINDER	nebula	Nebel	nebula	星雲	nebula	سديم	nebula	туманность
18	\$	*	NEBULA	galaxy	Galaxis	galaxy	銀河	galaxy	المجرة	galaxy	галактика
19	\$	*	GALAXY	11	>-70006-0-4000000	90 00 000000000000000000000000000000000			DE40.00		

The COCA corpus and the iWeb corpus allow you to browse through the top 60,000 words in the corpus, and to see a wide range of information on each of these 60,000 words. And as before, you can find links to a wide range of external information for each of these words. The entries below are for words near #25,200 (i.e. the 25,200th most frequent word in the corpus). For each word, you can hear the word pronounced, see videos with that word highlighted, find images, or see translations. For example, the following external resources are for word #25,205: gasket.

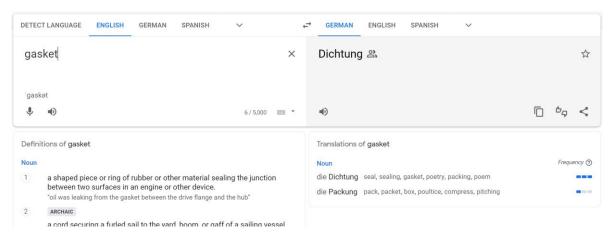
	RANK	FREQ	Word	PoS	Audio	Video	Image	DE	
1	25203	576	ergonomic	ADJ	•	(D)		©	
1 . of or relating to ergonomics									
2	25204	576	tailgate	VERB	•	D		•	
1 . follow at a dangerously close distance									
3	25205	576	gasket	NOUN	•	(D)		€ ₹	
1 . seal consisting of a ring for packing pistons or sealing a pipe joint									
4	25206	576	reopening	NOUN	4))	(D)		©	
1 . reopen + ing									
5	25207	576	impolite	ADJ	•	(D)	_	•	
1 . not polite									

Images and videos (external links)



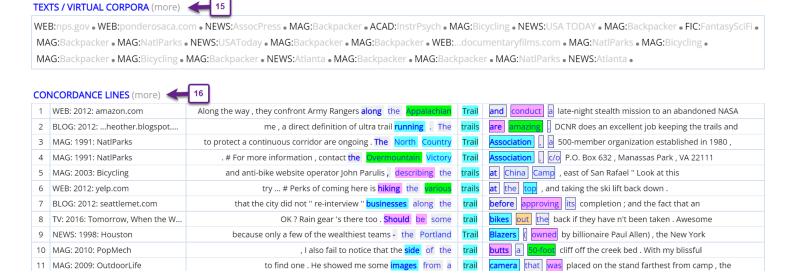


Translation (Google Translate: Baidu in China)



In COCA and iWeb, each of the top 60,000 words (lemmas) has a "word sketch" or "home page", with the following 16 pieces of information: 1) lemma and part of speech 2) add word to "favorites" list 3) see "history" of past words 4) frequency information 5) definitions, including links to external definitions and etymology 6) images 7) pronunciation and external videos 8) definitions from several external websites 9) synonyms 10) word used in definitions for other words, and more specific and more general words from WordNet 11) related topics 12) collocates 13) morphologically-related words 14) 2-4 word clusters 15) texts or websites where the word is very frequent 16) concordance (KWIC) lines.





Several of these pieces of information come from external resources. For example:

[5] Definitions from several external websites, include etymology

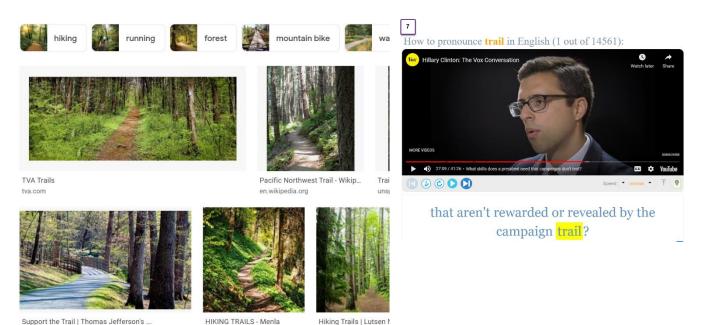


c. 1300, "to hang down loosely and flow behind" (of a gown, sleeve, etc.), from Old French trailler "tow; pick up the scent of a quarry," ultimately from Vulgar Latin "tragulare" to drag," from Latin tragula "dragnet, javelin thrown by a strap," probably related to trahere "to pull" (see **tract** (n.1)). Transitive sense of "to tow or pull along the ground" is from c. 1400. The meaning "follow the trail c (an animal, etc.) is first recorded late 14c. Meaning "to lag behind" is from 1957. Related: Trailed; trailing.

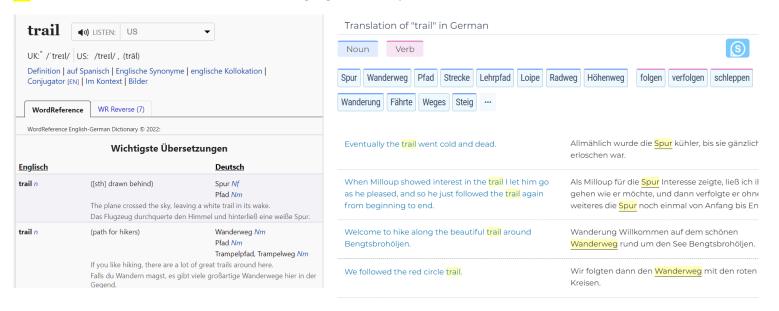
trail (n.)

early 14c., "trailing part of a robe, gown, etc.," from **trail** (v.). The meaning "track or smell left by a person or animal" is also from 1580s. Meaning "path or track worn in wilderness" is attested from

[6, 7] Images (from Google Images) and videos (from several external websites), which allows you to hear the word in context.



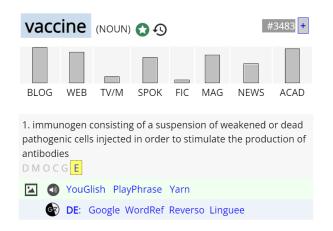
[8] Translations of the word into hundreds of languages, at multiple external sites



Finally, in COCA you can enter and analyze entire texts. You can find the keywords for the text:



You can also click on **any word** in either the text or the frequency/keyword lists (as with *vaccine* above) to see a **detailed** "word sketch" of that word (this is just a partial view of the page for *vaccine*).



SYNONYMS (> CONCEPT) NEW: DEFIN +SPEC +GENL

inoculation inoculation, preparation, serum, shot, vaccine

TOPICS (more)

disease, virus, vaccination, infect, outbreak, flu, infection, epidemic, vaccinate, immunity, immune, inject, dose, infectious, injection, antibiotic, immunization, influenza, complication, transmit

COLLOCATES (more)

NOUN flu, influenza, dose, autism, virus, polio, disease, safety

VERB cause, receive, develop, prevent, protect, produce, contain, administer

ADJ effective, available, safe, pneumococcal, live, seasonal, oral, experimental

ADV currently, eg, genetically, routinely, safer, ie, commercially, ideally

RELATED WORDS

vaccinate

CLUSTERS (more)

vaccine •	vaccine for • vaccine in • vaccines for • vaccines in • vaccine against • vaccine safety • vaccine has • vaccines have
• vaccine	flu vaccine • influenza vaccine • polio vaccine • in vaccines • new vaccine • about vaccines • smallpox vaccine • b vaccine

TEXTS / VIRTUAL CORPORA (more)

WEB:scienceblogs.com • BLOG:...ncebasedmedicine.org • BLOG:scienceblogs.com • ACAD: Patient-Reported Outcomes • WEB:...ncofarm.blogspot.com • BLOG:scienceblogs.com • BLOG:scienceblogs.com • ACAD:Texas Law Review • WEB:scdc.gov • BLOG:scienceblogs.com • MOV:Vaxxed: From Cover-Up to Catastrophe • MAG:USNWR •

CONCORDANCE LINES (more)

1	SPOK: 1991: CNN_Health	interests include treatments for breast cancer and developing		
2	MAG: 2007: ScienceNews	weight , says Janda . He notes that a combination of	vaccines	against ghrelin and other weight-loss drugs currently on the
3	ACAD: 2017: Clinical Therapeutic	in HSCT recipients (at which time varicella	vaccine	and MMR vaccine may be administered to seronegative HSCT
4	ACAD: 2008: DrugIssues	. Eligible injectors were offered the three doses of	vaccine	and questioned about their past and current drug use , their
5	WEB: 2012: cdc.gov	the most preventable type of female cancer , with both HPV	vaccines	and regular screening tests . What is HPV ? # HPV
6	SPOK: 2002: CNN_YourHealth	need it , let me go . And with the flu	vaccine	and the nasa spray , then it 's much easier . FEIG

Again, from this "word sketch", you can access many **external resources**, such as definitions, etymology, images, videos, pronunciation, and translations.



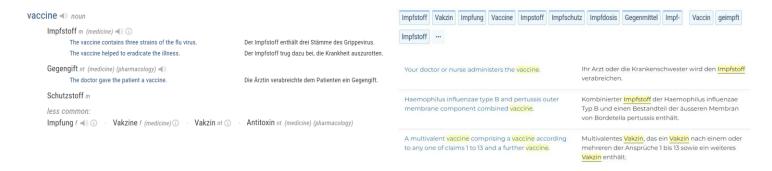


most effective vaccine roll-outs in the entire world.

vaccination (n.)

1800, used by British physician Edward Lenner (1749-1822) for the technique he publicized of preventing smallpox by injecting people with the similar but much milder cowpox virus (vario vaccinae), from vaccine (adj.) "pertaining to cows, from cows" (1798), from Latin vaccinus "frocows," from vacca" "cow," a word of uncertain origin. A mild case of cowpox rendered one immune thereafter to smallpox. "The use of the term for diseases other than smallpox is due to Pasteur" [DED].

The earlier 18c, method of smallpox protection in England was by a kind of inoculation called variolation (from variola, the medical Latin word for "smallpox"). There are two forms of smallpox: a minor one that killed 2% or less of the people who got it, and a virulent form that I about a 30% mortality rate and typically left survivors with severe searring and often blinded them. Those who got the minor form were noted to be immune thereafter to the worse. Doctor would deliberately infect healthy young patients with a local dose of the minor smallpox, usual resulting in a mild case of it at worst, to render them immune to the more deadly form. Jenner method was safer, as it involved no smallpox exposure.



Only at English-Corpora.org can you 1) enter an entire text and then 2) see detailed "word sketches" for each of the words in the text, including external information.

We believe that the corpora from English-Corpora.org provide the best data for English, especially for studies of <u>variation in English</u>. But English-Corpora.org also goes far beyond what other corpus sites offer, by linking you to an extremely wide range of external resources, which can be invaluable for language learners and teachers.