

Lexical Cohesion

Dimensions and Linguistic Properties of Chains in English and German

*Ekaterina Lapshinova-Koltunski, José Martínez-Martínez, Kerstin
Kunz, Katrin Menzel, Erich Steiner*

FR4.6, UdS, Saarbrücken

May 29, 2016



Background



Research Project
**GECCo: German-English Contrasts
in Cohesion**

supported by the DFG

Project Team:

- Kerstin Kunz
- Ekaterina Lapshinova-Koltunski
- Erich Steiner
- Jose Manuel Martinez
- Katrin Menzel

Acknowledgement: Stefania Degaetano-Ortlieb, Marilisa Amoia

FR 4.6 Applied Linguistics, Interpreting and Translation Studies

www.gecco.uni-saarland.de

Overview

- 1 Defining Concept
- 2 Goals and Motivation
- 3 Methodology
 - Corpus data
 - Annotation Scheme
 - Annotation Procedures
- 4 Analyses

Defining Concept

Several studies have shown that two of the factors affecting regret are how much one feels personal responsibility for the result and how easy it is to imagine a better alternative. The availability of choice obviously exacerbates both these factors . When you have no options, what can you do? You will feel disappointment, maybe; regret, no. With no options, you just do the best you can. But with many options, the chances increase that a really good one is out there, and you may well feel that you ought to have been able to find it.

Mehreren Studien zufolge wird das Gefühl der Reue zum einen stärker, je mehr man sich für das Resultat persönlich verantwortlich fühlt, und zum anderen, je leichter man sich eine bessere Alternative vorstellen kann. Ein Auswahlangebot verschlimmert offensichtlich beide Faktoren. Was kann man schon groß anstellen, wenn man keine Wahl hat? Vielleicht ist man enttäuscht, aber Reue empfindet man nicht. Wenn es hingegen viele Optionen gibt, wächst das Risiko, dass man meint, eine besonders gute übersehen zu haben, und dies nun bereut.

Several studies have shown that two of the factors affecting regret are how much one feels personal responsibility for the result and how easy it is to imagine a better alternative. The availability of choice obviously exacerbates both these factors. When you have no options, what can you do? You will feel disappointment, maybe; regret, no. With no options, you just do the best you can. But with many options, the chances increase that a really good one is out there, and you may well feel that you ought to have been able to find it.

Lexical Cohesion

Mehreren Studien zufolge wird das Gefühl der Reue zum einen durch die Tatsache, dass man sich für das Ergebnis mehr persönlich verantwortlich fühlt, und zum anderen, je leichter man sich eine bessere Alternative vorstellen kann. Ein Auswahlangebot verschlimmert offensichtlich beide Faktoren. Was kann man schon groß anstellen, wenn man keine Wahl hat? Vielleicht ist man enttäuscht, aber Reue empfindet man nicht. Wenn es hingegen viele Optionen gibt, wächst das Risiko, dass man meint, eine besonders gute übersehen zu haben, und dies nun bereut.



Cohesion

Types of Cohesion (cf. Halliday & Hasan 1976)	Meaning relations
Coreference <i>An option ... it/this option</i>	identity
Substitution <i>Many options ... a good one</i>	Type reference/ comparison
Ellipsis <i>You will feel disappointment. []</i> <i>Maybe.</i> <i>Many options ... a good [].</i>	
Comparative Reference <i>One option ... another/ better option</i>	
Cohesive conjunction <i>X. But/ And/ However Y</i>	Logico-semantic relations (addition, contrast, cause, ...)
Lexical cohesion <i>factors ... factors</i>	Similarity (repetition, general nouns, meronymy, Synonymy ...)

Cohesive Chains

Differences across languages

EO-FICTION-002

*There was not one **way** but many. He followed first **one** and then **the other**.*

GTRANS-FICTION-002

*Nicht ein **Weg**, sondern eine Vielzahl von **Wegen** lag vor ihm. Er folgte bald dem einen **Pfad**, bald dem anderen.*

- DE: lexical chain with repetition and hyponym:
Weg – Wegen – Pfad
- EN: cohesive chain with various cohesive devices:
substitution and comparative reference with an ellipsis

Goals and Motivation

Goals

- analyse general properties of lexical cohesion
- special attention to
 - semantic variability in chains
 - cohesive strength
 - number and length of chains
 - degree of semantic specification of lexis
 - degree of variation along all these dimensions
- cross-linguistically: contrasts between English and German,
see House (1997, 2015), Hansen-Schirra et al. (2012) and
Neumann (2013)
- across modes: between spoken and written registers,
see Hundt & Mair (1999), Mair (2006) or Leech et al. (2009)

Background for Hypotheses

Kunz, K., E. Lapshinova-Koltunski, J. M. Martínez Martínez, K. Menzel and E. Steiner (*forthcoming*). Shallow features as indicators of English-German contrasts in lexical cohesion.

- investigated the role of lexical cohesion in language variation
 - used various quantitative results for shallow features
 - standardised type-token-ratio (sttr)
 - lexical density (ld)
 - most frequent words (mfw)
 - top content words (tcw)
 - words with latinate affixes (lat)
- ⇒ insights about general lexical features
- ⇒ preliminary indicators for assumed properties in chains

Background for Hypotheses

Kunz, K., E. Lapshinova-Koltunski, J. M. Martínez Martínez, K. Menzel and E. Steiner (*forthcoming*). Shallow features as indicators of English-German contrasts in lexical cohesion.

	Language	Mode
Semantic variability in chains	E < G	S < W
Cohesive strength	E ≠ G	S > W
Number & length of (nominal) chains	E > G	S < W
Degree of specification	E ≠ G	S < W

Degree of variation	E < G?	S ≠ W
---------------------	--------	-------

Methodology

Methodology

we test our hypotheses by

- analysing lexical chains (their number, length and distance between members) annotated manually in our corpus data
- comparing distributions of various sense relations that chain members have to their antecedents
 - across languages and registers available

EO/GO-GECCOCHAIN

in tokens

register	EO	GO
ESSAY	27.171	31.407
FICTION	36.996	36.778
INTERVIEW	30.057	35.036
POPSCI	27.055	32.639
TOTAL	121.279	135.860

- subset of GECCo
(Lapshinova et al., 2012 and Hansen-Schirra et al., 2012)
- ESSAY and POPSCI: written discourse
- INTERVIEW: spoken discourse
- FICTION: spoken elements in form of dialogues

Sense Relations

- repetition
- antonymy
- synonymy

- hyperonymy
- hyponymy
- co-hyponymy

- holonymy
- meronymy
- co-meronymy

- type-instance
- instance-type
- co-instance

MMAX2 (Müller & Strube, 2006)



UNIVERSITÄT
DES
SAARLANDES

1.15.005 /home/katja/saar/gecco/docu/annotationsrichtlinien/LexicalCohesion/LexicalCohesion-GO/GO_POPSCI_005.mmaz

File Settings Display Tools Plugins Info Show MI Page

Hunderten von **Proteinen**, es sind die **Engme**, beginnen wir darum, daß wir den **da** dieser für ganz bestimme chemische **Reaktionen** warum und in welchem **Sinn** sie mit **Recht** als **Schlüssel** anzusehen sind. **Enzyme** sind **Eiweißkörper** und zusammengesetzt ist, braucht uns hier nicht zu interessieren. Wichtig ist dagegen, daß von den Hunderten oder mehr verschiedenen **Proteinen** in allen indischen **Lebewesen** nur 20 kommen. Alle **Eiweißarten**, die es in **Individuen** auf der Erde existierenden **Zeile** geben mag - ungefähr gleichen 20 **Aminosäuren** als **Elementen** oder **Baustein** zusammengesetzt. Passender noch als der Vergleich mit einem **Baustein** ist **Aneinanderlegung** der erwähnten immer wieder gleichen 20 **Aminosäuren** in **bauen**, wechselnder **Aufenthaltsorte**. Die Länge der **Kette** kann, das **Enzym Cytochrome c**, aus **diesem** Bau, wir näher eingehen wollen, ist ein aus 104 **Aminosäuren** zusammengesetztes **Ketten**, denen es besteht, ist darin durch ein bestimmtes graphisches **Symbol** gekennzeichnet, konkret dieser, und zwar gegen der hier **biochemischer Cytochrome** getan haben. Es ist, wie schon erwähnt, ein **Stoffwechsel-Schluß**, im Unterschied zu einer **reinen** ganz bestimmt und für eine chemische **Wirkung** auf konkreten Stoffen, angegeben, die der **zur** rein zufällig an ganz anderer Stelle existieren könnte. Das **charakteristische** **Wohlbefinden** einer zufälligen **Wiederholung** entspricht hier der Kontrast **zusammen** mit der das **Cytochrome** - **charakteristische Aminosäure-Sequenz** - durch **Zufall** entstehen könnte. Einfacher ausgedrückt: Wie oft müßte man 104 **Perlen** der 20 möglichen Farben innerhalb einer aus 104 **Perlen** bestehenden **Kette** gießen, um genau 2014 verschiedene **Möglichkeiten**, von diesen ist die der **Cytochrome** reproduzieren, beträgt mit 1 zu 201404 oder, auf die gewohnte **Basis** 10 umgerechnet, 1 zu 10310. Damit steht fest, daß der **Enzym** reinen **Zufall** entstanden sein kann, seit **Entstehung der Welt**, seit dem **Ur-Kanal**, sind erst 1017 **verschiedene Varianten** der aus 104 **Gliedern** jedes einzelne aller im ganzen **Kosmos** existierenden **Moleküle**, eine andere **Wirkung** der 104gliedrige **Kette** repräsentiert, selbst dann geben, wenn im ganzen **Universum** sind nur rund 1080 **Atoome** vorhanden! Man wird darüber müssen, daß die **Möglichkeit** einer rein zufällig

beschränkten **Area B** **überdeckt**, vor dieser **Area A** mit hinreichendem **Abstand** ausgeschlossen werden kann. Dagegen aber beginnen wir Ihnen, davon waren wir ja ausgegangen auf dieser Seite wieder und wieder hincnach auch schon gewesen, die bisher darauf untersucht worden sind, nicht nur bei uns selbst, sondern bei **hunderten** und **tausenden** und ebenso bei **schwimpflichen** und den **ordinären Menschen**. Wie ist das möglich? Wenn der **Zugang** zu einer solchen **Überinflammation** mit so großer **Sicherheit** ausgeschlossen werden kann? Wie ist die **Überinflammation** unter diesen Umständen anders zu erklären als durch einen **Beispielzettel**, die alle die **genetischen Interaktionen** verbindet, bei denen sie das **Problem** ist? Und wie wäre diese Beziehung zu verstehen, wenn es sich um eine „durch“ **genetische Veränderungen** erfasste **Entzündung** handelt, die sich von **wiederholigen** ihren bestehenden Unterschieden unterscheidet? Und was kann man damit erklären, wenn **genetisch** derselben **Artzelle** angegesehen werden müssen? **Für** Fundamente, **entzündlich**, die **Reaktion** der **zellen** zu bestreiten, wel er (**immunologisch**) davon erwacht, daß es sich dabei **sehr** wirklich nur noch um eine **verschwundene Minderheit** handelt und daß die **Tatsache** der **Reaktion** heute auch von der **Mehrzahl** der **anderer** anerkannt wird, die ihre (**darmatische**) **Erklärung** noch immer ablehnen? **Für** Sie lautet: Es war offenbar kein **Zufall**. Aber ein genetischer **Zusammenhang** ist auch damit keineswegs bewiesen; denn **Gott** hat es bei **Schöpfung** eben so eingerichtet, daß also von ihm geschaffenes **Leben** mit diesen und unzähligen anderen identischen **Formen** ausgestattet worden sind. Aber auch auf diesen **Blickgrund** läßt sich noch etwas erwidern.

kennen die minimalen "Abweichungen" ins Spiel, von den Fischen her, die *die Röte* war. Wirklich abweichend sind die bei den verschiedenen Arten festgestellten Aminosäure-Sequenzen des Cytochrome c nämlich doch r. In dem Schema auf Seite 47 sind sie für den Menschen und 10 andere biologische Objekte im Vergleich untereinander gestellt. Die *Hierarchie der Objekte* unterliegt dabei abnehmenden Grade der Ähnlichkeit. Es geht sich bei diesem Schema nun ein weiteres, außerordentlich interessantes Dokument. Die Zahl der Unterschiede nimmt in dem Maße von oben nach unten zu, zwischen der Aminosäure-Sequenz des Menschen und der eines anderen Organismus.

Unterschieden, und so geht es fort, je entfernter die jeweilige Abstammungslinie ist. Die größten Unterschiede zwischen den Zellen eines Organismus sind daher zu erwarten, wenn man sich auf die Zellen im Bereich der Keimzellen konzentriert. Auf der Position 591, verglichen mit einer „eigenen“ Sequenz mit der entsprechenden, dann kann man schon Genaugkeiten erzielen, die bei der Sequenzierung der DNA mit dem Ergebnis der Sequenzierung der DNA aus dem Blut des Patienten 2093 übereinstimmen. Bei der Sequenzierung mit dem Ergebnis der Sequenzierung der DNA aus dem Blut des Patienten 2093 übereinstimmen. Diese Korrekturen ändern aber, vor allem dann, wenn man sich nicht allein auf den Bereich konzentriert, sondern die verschiedenen Arten untereinander vergleicht, an der **Beweisführung** grundsätzlich nichts. Die resultierenden Wahrscheinlichkeiten bleiben auch dann immer noch so extrem gering, daß die Möglichkeit, die Übereinstimmungen durch bloßen Zufall zu erklären, weiterhin ausscheint. Welche Ursache aber ist diese individuellen Abweichungen in denen uns dieser seine Abstammung nach ursprünglich identische **Allele** auf den Patienten übertragen hat? Der Antwort nähern wir uns, wenn wir uns einmal vor Augen halten, daß es sich um eine **Wiederholungssequenz** handelt, die es ermöglicht, daß von einem symmetrischen Bereich aus mehrere verschiedene Sequenzen abgelesen werden. In jeder neuen Zelle müßte es anhand dieser mitgebrachten **Präzision** dann von neuem symmetrisiert werden. Nur funktioniert der molekulargenetischen Apparatur, der für diesen Prozeß zuständig zwar mit einer geradezu unglaublicher **Präzision**. Absolut perfekter, fehlerloses, ohne Flickarbeit an der Zahl der Atome, die in jeder Sekunde in einer bestimmten **Anordnung** angeordnet muß, und ohne Rücksicht auf die unvorstellbaren **Zellzyklen**, über die hinweg das von **Augenblick zu Augenblick** zu geschehen hat, gelingt das aber auch ihm nicht. Auf molekularer Ebene verhindern das schließlich die **Genetischen Fehler**.

natürliche radioaktive Strahlung der Umgebung und die von der unvermeidbaren Wärmeabstrahlung ausgehenden Störungen. Es schleichen sich also Fehler ein. Erstthalten seien, wenn man die Subtilität der molekularen Reparationsabläufe bedenkt, aber im Laufe der Zeit sammeln sie sich eben an. Da stößt dann plötzlich, nach der Neubildung des Moleküls in einer neu entstandenen Zelle, sozusagen eine Perle „falscher Farbe“ an einem bestimmten Punkt der Kette: Eine Amibiosine ist durch eine andere ausgetauscht worden. Die Biologie bezeichneten eine solche Änderung des erblichen Materials bekanntlich als „Mutation“. Ob der Austausch Folgen hat, ist abhängig von dem Punkt, an dem die Mutation erfolgt ist. Ist sie bei einem der ersten, dann noch nicht vollständig ausgereiften Bausteine, so ist sie ohne Ausnahmen absterblich. Das aktive Zentrum ist so etwas wie der „Bart“ der „Dreiecks“-Molekül, auf dem seine Stoßwellenfunktion entweder fehlt oder unzureichend ist. Beim Schneiden C besteht diese Fähigkeit schon erwähnt, in der Sauerstoffübertragung innerhalb der Zelle. Eine Mutation im Bereich des aktiven Zentrums von Cytochrome c führt daher in der Regel zum sofortigen Absterben der neu entstandenen Zelle als Folge ihrer Entstörung. Die Mutation hat sich als „letal“ erwiesen. Die verschwindende Ausnahme bedeutete in diesem Fall, daß die neue Amibiosine, die da plötzlich innerhalb der aktiven Zentren aufgetaucht ist, zu einer Verbesserung der funkzionierenden Moleküle führte. Daß ihr Auftreten den „Bart“ des Moleküls allein zufällig in einer Weise veränderte, welche die Sauerstoffübertragung in irgendeiner Hinsicht verbessert, indem sie sie etwas schneller und effektiver werden läßt. Es läßt sich leicht denken, daß das nur außerordentlich selten vorkommt. Andererseits steht ebenso fest, daß jeder einzelne derartige „letale“ Erfolg darstellt, an dem die Evolution von da an brachialisch festhalten wird. Das geschieht einfach in der Form, daß ein weiterer „Bart“, durch einen solchen Glückstreffer in der Mutationslotterie eine Verbesserung seiner Ausstattung erhält, bessere Überlebenschancen hat. Er wird durch mit großer Sicherheit die größere Anzahl von Nachkommen hinterlassen, was wiederum die Wahrscheinlichkeit erhöht, daß er auch wieder glücklich gegenüber den wenigen mit ausgesparten Konkurrenz seinen eigenen Anteil innerhalb weniger Generationen erhält. Denkbare Typ setzt sich daher in der endgültigen Konkurrenz gegen die anderen fort. Ein Beispiel: Wenn das Ergebnis eines langen Zeiträume auf diesem Wege entstanden ist, führt er Sovjet für Sovjet; denn „glücklicher“ werden festgehalten, z.B. Zellen mit einer negativ wirksamen Mutation innerhalb des aktiven Zentrums (und die war ohne Zweife die überwiegende Mehrzahl) starben abgehen. Sie schieden aus, all ihre Spuren sind längst getilgt. Daneben aber gab es nun noch die Mutationen, die nicht das aktive Zentrum des Moleküls betrafen, nicht den „sonden sozusagen den“ Griff, den bildet stattdessen die chemischen Molekulgerüste. Sie sind ganz offensichtlich die Ursache der Verschiedenheiten, die wir heute zwischen den Cytochrome-c-Molekülen unterschiedlicher Herkunft - ob vom Menschen, einem Nagetier oder einer Blattwespe - finden. Von den „letalen“ Mutationen gibt es keine Zeugnisse mehr. Die Träger dieser Mutationen konnten nicht überleben (und ihre negativen

[Organismen]

One-click annotation Panel Settings

LexicalCohesion

lexical_type: synonym
 distance_type: 0-immediate
 lexical_chain: set_405

Suppress check Warn on extra attributes

Apply **Undo changes**

Auto-apply is OFF

Auto-apply is OFF

Analyses

Hypotheses

	Language	Mode
Semantic variability in chains	E < G	S < W
Cohesive strength	E ≠ G	S > W
Number & length of (nominal) chains	E > G	S < W
Degree of specification	E ≠ G	S < W

Degree of variation

E < G?

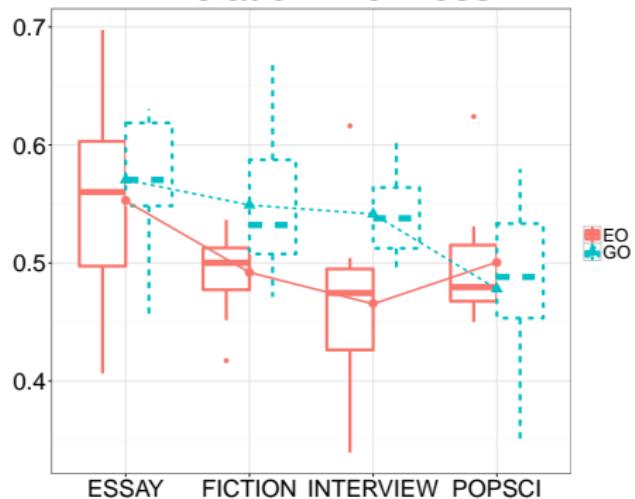
S ≠ W

Features under Analysis

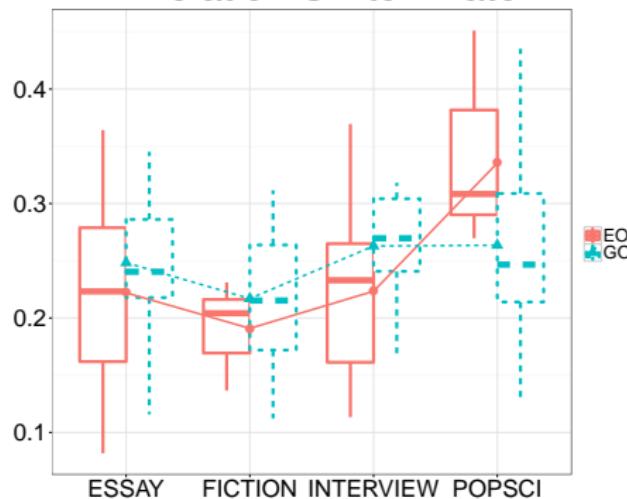
- Semantic variability in chains: number of different relations per chain
- Cohesive strength: distance between elements in a chain
- Number and length of nominal chains: chain number and chain length
- Degree of specification: amount of more specific and more general sense relations
- Degree of variation: variation in the observed features

Semantic Variability

Relation Richness

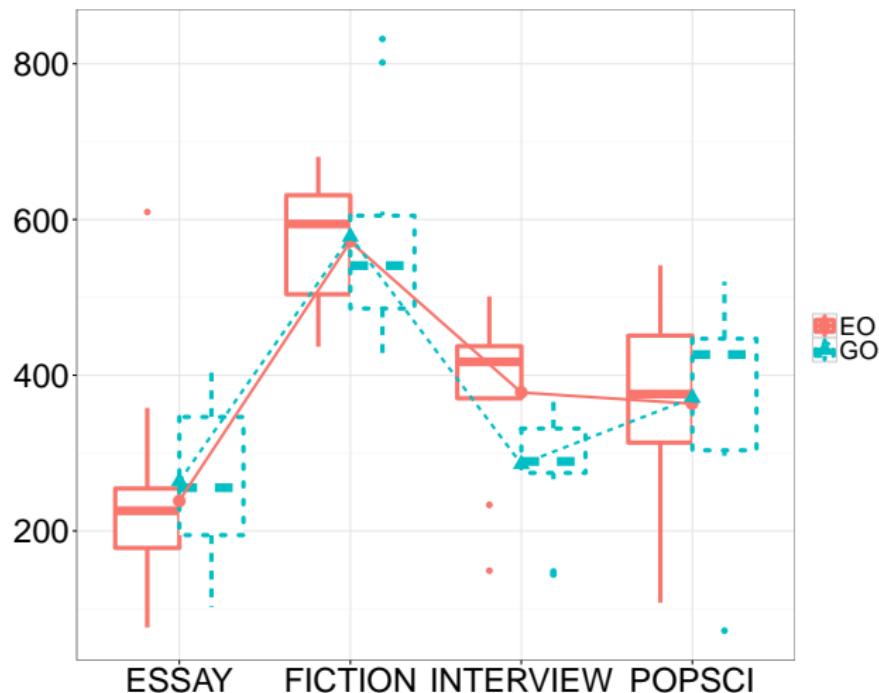


Relation Switch Rate



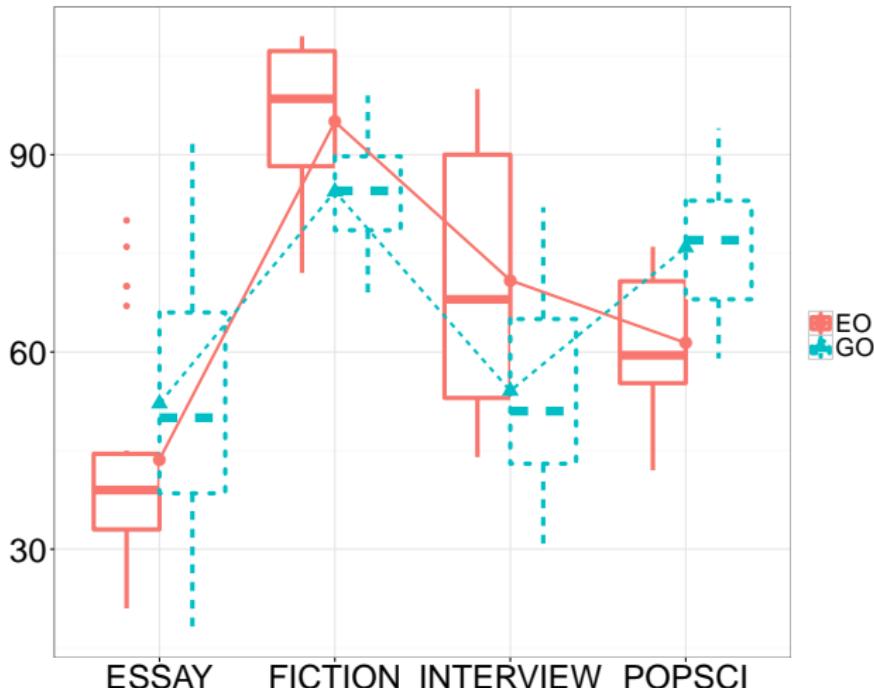
Cohesive Strength

Average distance among chain members per text



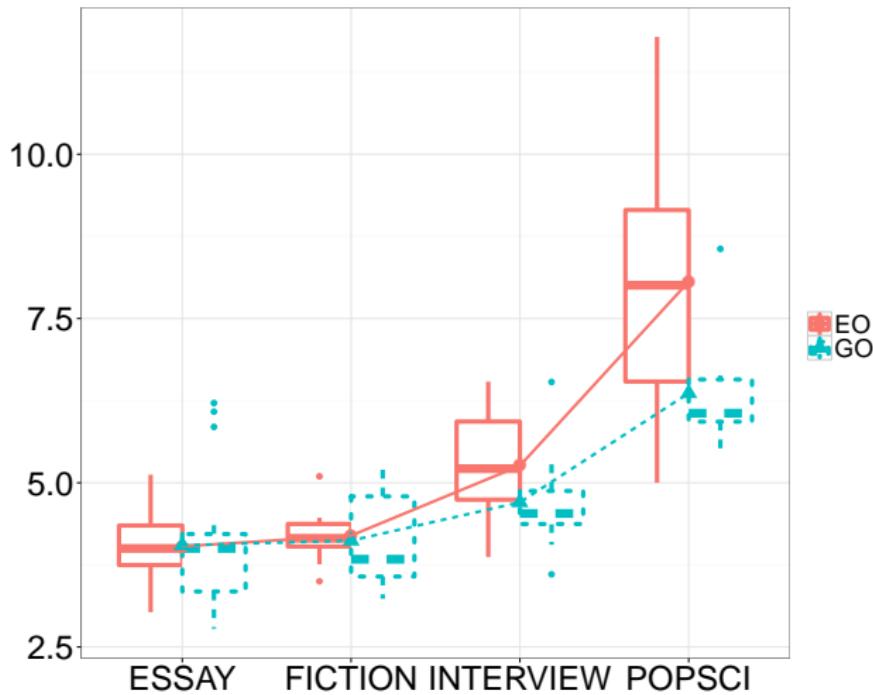
Chain Number

Average number of chains per text



Chain Length

Average length of chains per text



Degree of Specification

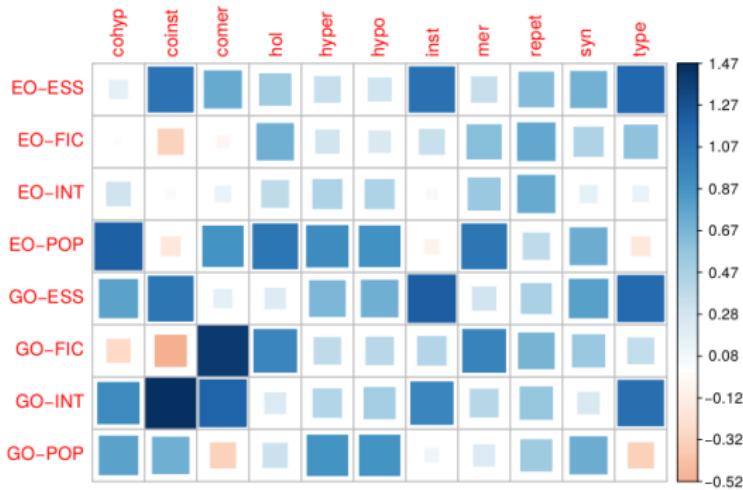
same type, more general, more specific

- repetition
- antonym
- synonym
- hyperonym
- hyponym
- cohyponym
- holonym
- meronym
- comeronym
- type
- instance
- coinstance

Degree of Specification

association between lang/register and sem.relations (likelihood ratio)

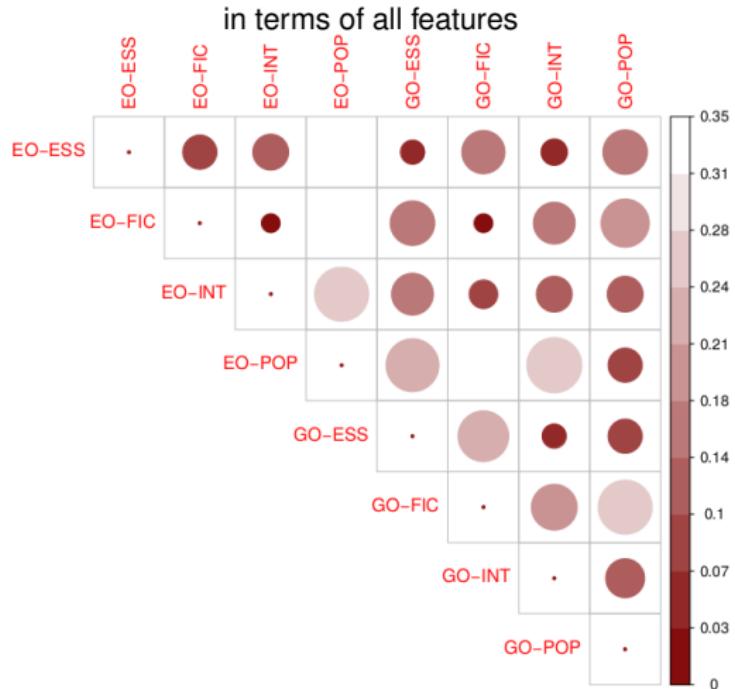
- If ratio < 1 => $\log(\text{ratio}) < 0$ (negative values) => red color
- If ratio > 1 => $\log(\text{ratio}) > 0$ (positive values) => blue color



INTERPRETATION: essays and popsci in both languages have a similar degree of specification (type-inst rel.); fiction have partly similar relations, English are more general (hol and type); interviews are different: English are much more general.

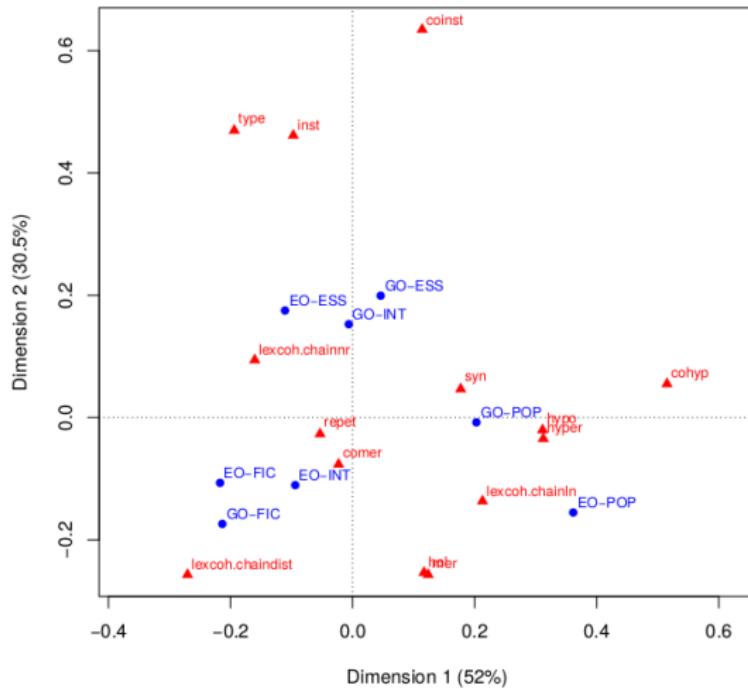
Variation

squared distances between lang/register profiles



intralingually: FIC and INT in English, ESS and INT in German (bigger distance);
 cross-lingually: register pairs, esp. for ESS and FIC

Variation: CA with all features



three groups of multilingual registers; interviews behave different cross-lingually



Summary

	Language	Mode
Semantic variability in chains	E < G	S < W
Cohesive strength	E ≠ G	S > W
Number & length of (nominal) chains	E > G	S < W
Degree of specification	E ≠ G	S < W
Degree of variation	E < G	S ≠ W

Thank you!

Questions?

Contact:

e.lapshinova@mx.uni-saarland.de

j.martinez@mx.uni-saarland.de

Information:

www.gecco.uni-saarland.de

Sense Relations

- REPETITION: orthographical repetition of nominal expressions such as *London* and *London*, or *place* and *place*.
 - *I live in a town called Reigate. It's between London and the countryside which is quite nice. It takes us about 25 minutes to get to London on the train. It's very nice actually, it's a nice place to live. And I grew up in a place called Banstead.*
- * In case of compounding, the second element is the determining factor: *stem cell* but not *stem cell behaviour* or *stem cell research*.
 - *For instance, just identifying a true stem cell can be tricky. For scientists to be able to share results and gauge the success of techniques for controlling stem cell behavior... Of course, the goal of stem cell research is to replace...*

- ANTONYMY: relation of contrast *halving* and *doubling*
 - Dazu gehören zum Beispiel *die Halbierung* der Energie- und Rohstoffintensität bis 2020 gegenüber 1990 (bzw. 1994) und *die Verdoppelung* des Anteils erneuerbarer Energien am Energieverbrauch bis 2010.
- SYNONYMY: total synonymy (*Lebewesen – Organismen*) but also near synonymy, such as between technical and common-language terms, e.g. *belly* and *abdomen*
 - Wie ist die Übereinstimmung unter diesen Umständen anders zu erklären als durch eine Beziehung, die alle die *Lebewesen* miteinander verbindet...? Und wie anders wäre diese Beziehung zu verstehen als die einer durch Vererbung, durch "genetische Überlieferung" entstandenen Gemeinschaft von *Organismen*

- HYPERONYMY: in case the superordinate term follows the more specific term: *village* and *place*
- HYPONYMY: in case the specific term follows the superordinate one: *place* and *town*
- COHYPONYMY: between two elements on the same level of specification: *town* and *village*
 - *I live in a town called Reigate. It's between London and the countryside which is quite nice. It takes us about 25 minutes to get to London on the train. I say it's a town, it's more of a village. It's quite small. It's very nice actually, it's a nice place to live. And I grew up in a town called Banstead.*

Sense Relations

- **HOLONMY**: relation, where the whole follows the part: *the EU* following *Britain*
- **MERONYMY**: part-whole relation, where the part follows the whole: *Bulgaria* following *the EU*
- **CO-MERONYMY**: succession of two parts that belong to a whole: *Romania, Croatia, Britain* following *Bulgaria*
 - *This means that this will not be the last enlargement of the EU. Bulgaria and Romania are waiting in the wings and likely to join in the next 2-3 years. The candidatures of Turkey and Croatia are well advanced. We in Britain welcome both. In particular we believe that Turkey, if she meets all the criteria, would add a new dimension to the EU, and represent a vital reaching out to the Islamic world at a time when such links are more than ever needed.*

Sense Relations

- **TYPE**: relation between a common noun and a named entity: *town* and *Reigate*
- **INSTANCE**: relation where the named entity follows the common noun: *Reigate and town*, *Banstead and town*
- **CO-INSTANCE**: relation between two named entities (*Reigate and London*)
 - *I live in a town called Reigate. It's between London and the countryside which is quite nice. It takes us about 25 minutes to get to London on the train. I say it's a town, it's more of a village. It's quite small. It's very nice actually, it's a nice place to live. And I grew up in a town called Banstead which is fairly close to Reigate.*

Cohesive Chains

Differences across modes

GO-INTERVIEW-014

Die sind sehr laut, die Burschis. Und wir haben auch dann verschiedene **Tiere** jeweils da, also einmal **Seelefanten**, die da vorkommen, **die nördlichen Seelefanten**, und in dem Fall hier **der Stellersche Seelöwe**, das ist so mit **der größte**. Dem gegenüber sitzt der **Kalifornische Seelöwe**, nur um den Größenunterschied mal zu sehen. Und da können sich die Leute eben hinstellen und sagen: "So groß ist **der** wirklich", ne, und so bissig ist **er** dann auch. Das größte **Tier**, wie gesagt, der **Seelefant**, die wohnen da im Grunde zusammen miteinander. Hier wird ein Flossenvergleich nachher stattfinden. Was kann der **Seelöwe**, was der **Seehund** nicht kann?

GO-POPSCI-005

Dennoch aber begegnen wir ihr nun, davon waren wir ja ausgegangen, auf dieser Erde wieder und wieder, nämlich bei allen **Lebewesen**, die bisher darauf untersucht worden sind: nicht nur bei uns selbst, sondern bei **Affen, Hunden und Ameisen**, bei **Fischen, Fröschen und Schmetterlingen** und ebenso bei **Schimmelpilzen, Weizen** und der ordinären **Bäckerhefe**. Wie ist das möglich, wenn der Zufall als Ursache einer solchen Übereinstimmung mit so großer Sicherheit ausgeschlossen werden kann? Wie ist die Übereinstimmung unter diesen Umständen anders zu erklären als durch eine Beziehung, die alle die **Lebewesen** miteinander verbindet, bei denen sich das gleiche, hochspezifische Enzym-Muster findet?

Background for Hypotheses

Krielke, M-P. (2015). Lexikalische Kohäsion und Koreferenz im Deutschen und Englischen. Eine corpusbasierte Studie populärwissenschaftlicher Texte.

- variability in semantic relations between nouns: DE > EN
- DE: various relations under analysis
- EN: preferences for repetitions and hypernyms, few antonyms