International Journal of Language and Applied Linguistics
2016; 2(3):105-126
Published online December 15, 2016 (http://www.ijlal.ir)
ISSN: 2383-0514 (Online)

Analogy and Causation in Natural Science English: A Cross-Register View

Christoph Haase
Purkyne University, Usti nad Labem (UJEP), Czech Republic
Email: Christoph.Haase@ujep.cz

Abstract

This contribution examines two important devices in academic writing – analogy and causation - in a number of constructions of the X CAUSE Y type. The analogical relationship of the causer and the caused is investigated in a second step. The data basis is provided by a custom-made corpus of academic texts from the natural sciences which due to its dual nature – specialized texts are juxtaposed with popular texts – can answer questions of the distribution and features of these constructions. The findings may shed light on the underlying principles of analogy-making and causal chaining for authors but also for readers of these texts. Further, the idea of the adequate reflection of scientific results in language is addressed and the question is raised whether a more adequate language would enable "better" science.

Keywords: Causation, Analogy, Structure-mapping, Corpus linguistics, English for Academic Purposes (EAP)

Received September 28, 2016; Accepted December 14, 2016.

International Journal of Language and Applied Linguistics
2016; 2(3): 105-126
Published online December 15, 2016 (http://www.ijlal.ir)
ISSN: 2383-0514 (Online)

Introduction

This contribution intends to address a number of elements in current research and teaching of academic writing. It concerns different text-types and draws on theories largely from a corpus-cognitive linguistic perspective. For this end, a specialized and stratified corpus has been assembled and a variety of results – qualitative and quantitative – will be presented.

In teaching science writing or English for Academic Purposes (EAP) it often proves to be helpful to start with a difficult, almost impenetrably dense science text. Usually such evidence from original research shows students the enormous complications involved in understanding real science (cf. for example Canagarajah 2002, Hyland, 2006). But it is obvious that a perspective across text registers, aspects of research (or other academic content) can be much more comprehensible when we are looking at texts from popularized sciences. The obvious correspondence of these texts – original and its popularized variant – shows that the same or similar issues can be discussed at different levels so that science writing does actually occur at these levels and students can familiarize themselves with the strategies employed in these texts.

Text 1: 0065PN We show that the rates of diversification of the marine fauna and the levels of atmospheric CO2 have been closely correlated for the past 545 million years. These results, using two of the fundamental databases of the Earth's biota and the Earth's atmospheric composition, respectively, are highly statistically significant ($P<0.001$). The strength of the correlation suggests that one or more environmental variables controlling CO2 levels have had a profound impact on evolution throughout the history of metazoan life.

Text 2: 0065NS Rising levels of carbon dioxide could speed up evolution. It seems that the higher the levels of CO2 in the atmosphere, the faster new species appear. University of Kansas researchers modeled how the amount of CO2 in the atmosphere has changed over the past 545 million years. Levels of this greenhouse gas used to be up to 20 times as high as today.

International Journal of Language and Applied Linguistics
2016; 2(3): 105-126
Published online December 15, 2016 (http://www.ijlal.ir)
ISSN: 2383-0514 (Online)

This can be shown in the above examples: Text 1 and text 2 (both from the corpus described in §3) represent the same content but the linguistic question is what makes one text harder to penetrate and the other palatable. The answers to this question could help to understand EAP better and perhaps facilitate production of these texts at learner level.

Causation and Analogy

The Cause-Effect Relationship and Linguistic Causation

Interest in the issues of academic writing is relatively recent with the parallel development of the fields of EAP and English for Special Purposes (ESP), which is more workplace-oriented, but the true beginnings may date back to 1974 (Hyland 2006: 2). EAP is often seen as a subfield of second language education, in some educational contexts it is a main focus, for example in North America (Grabe & Kaplan 2010: 62). It developed in the mid- to late 1960s from a subfield of ELT, called EST (English for Science and Technology and had initially a strong "preoccupation with syllabus design, materials development, and pedagogy" (Benesch 2001: 4) but soon also investigated the linguistic side beyond the teaching of lexical bundles to inject some native-speaker competence into texts of non-native practitioners and has in-between embraced the 'centrality of context' (ibid).

For this contribution two integral elements have been selected, not only for the topic of academic writing itself but also for the science behind it. These elements are on the one hand side causation and causal relationships and on the other hand analogy-making. This is motivated by the observation that in science we need causation because making cause-effect relationships is the core of hard science and when those relationships hold when experimental results are analyzed but also when we look at the math behind it. Therefore, uncovering a natural law means that we are establishing a cause-effect relationship. A cause-effect relationship persists when one phenomenon is present and another will happen that would not have happened with the same repeatability if the first one would be absent. This causation permeates the philosophy of science, it is for example called the "Cement of the Universe" in the classic text by Mackie (1980). Causation gives the universe coherence and it gives us temporal order (Flaherty 2011).

International Journal of Language and Applied Linguistics
2016; 2(3): 105-126
Published online December 15, 2016 (http://www.ijlal.ir)
ISSN: 2383-0514 (Online)

O the interface between scientific discovery and the language used to transport it, we can therefore state:

a) cause-effect relationship in research holds on basis of
   - experimental results
   - proof of a mathematical theorem
   - “plausibility” / “elegance”

When we further investigate the ways how causation and cause-effect relationships are grammaticalized in English, we find an entire spectrum of mechanisms from morphological and syntactic means to lexical means (e.g. the classes of causative and resultative verbs). In language therefore

b) cause-effect relationship in language holds on basis of
   - morpho-syntactic demands of the language used
   - the level of academic writing (e.g. specialized vs. popular)

The most unambiguous way to establish this relationship is the verb *to cause* itself. In this study, the X CAUSE Y construction will be discussed as can be seen in the following examples:

0086PN Cyclops is a crustacean copepod common in Africa and carrier of the guinea worm larvae that **CAUSES** dracunculiasis (31, 32).

0052PN mitochondrial decay is a significant factor in aging, **CAUSED** by the release of reactive oxygen species (ROS)

The reason for the choice is that *to cause* provides an unambiguous way to connect X and Y with X being the causer and Y being the caused, the result the effect, and sometimes the cause. Temporally the Xs precede the Ys but in academic writing the passivic structure often places the effects before the causes thus creating a focus on the phenomenon of interest and description.

International Journal of Language and Applied Linguistics
2016; 2(3): 105-126
Published online December 15, 2016 (http://www.ijlal.ir)
ISSN: 2383-0514 (Online)

Analogy and Language

The second integral element of academic writing is analogy-making for facilitated recognition and processing of largely abstract phenomena for subsequent analogical reasoning. Analogical reasoning is the traditional way to make scientific results comprehensible, especially for the layperson. Any theory that is laid out in mathematical terms is based on natural language analogs. We can say that analogy is a structural similarity (like metaphor) but it is a similarity in which we do not compare two isolated items but two domains. Two phenomena in one domain are linked by a particular relationship and this relationship is transferred onto a second domain where it is assumed that an analogical relationship exists. Therefore analogy is not the same as conceptual metaphor (where a cross-domain mapping is identified) although it is very similar. With Itkonen we can define analogy: “Analogy is generally defined as ‘structural similarity’” (2005) in which one aspect of one phenomenon is similar to a second aspect in a second phenomenon). To describe a secondary relationship, if the primary relationship holds, then other aspects of the one phenomenon may be similar to the second phenomenon.

Analogical mapping is the traditional method in the popularization of science because it systematically negotiates a structural connection between two domains on the basis of one or more relationships in between. The mapping occurs when an analogous relationship is hypothesized (or: observed, estimated) in the other domain. This is the basis for the classical structure-mapping hypothesis by Gentner (Gentner 1983, 2002). To make this plausible, a number of corpus examples show the process:

> 0034AX There are three main **mechanisms** and respective sites to accelerate particles in the Galaxy: supernova explosions either in the interstellar medium, in young and hot **starbubbles**, or in massive **star winds**.

In this, *mechanisms* refer to the underlying cause-effect machinery while *star bubbles* provide a visual analog to soap bubbles. *Star winds* combine the structural similarity of a logical phenomenon (directed movement) with a visual phenomenon (striations in astrophotographic images). Primarily, these expressions refer to visual analogues because the processes observed are not mechanical processes nor are they observable in any Newtonian way. There are no wind and bubbles in space but they are combined on the source domain side as visual analogues of observed objects which appear in astrophotography.

As mentioned, analogy-making is the major tool for understanding science, for example in science education:

Table 1
Analogical Thinking in Science (adapted from Harrison & Treagust 2006: 15)

Maxwell used **water pressure** in tubes to mathematically describe Faraday's **electric lines of force**
Robert Boyle imagined **elastic gas particles** as moving **coiled springs**
Huygens used **water waves** to theorise that **light** was wavelike
Konrad Lorenz used analogy to explain streamlined motion both in **birds** and **fish**
Kekulé derived his idea for a **benzene ring** from an image of a **snake biting its tail**

Table 1 shows a number of famous analogues in science and they structure how science is taught in schools today, for example we cannot see electricity but we can visualize the effects of water pressure vividly. Probably the most powerful and far-reaching analogue in the history of science is the wave an analogue for mechanical as well as electromagnetic waves. Thus, scientific discovery and analogical thinking are intertwined. The following text shows a variety of analogs for abstract phenomena:

> 0015AX Gamma Ray **Bursts** from the First Stars: Neutrino Signals
>
> If the first (PopIII) stars were very massive, their final **fate** is to collapse into very massive black holes. Once a proto-black hole has formed into the stellar core, accretion continues through a disk. It is widely accepted, although not confirmed, that magnetic fields **drive** an energetic jet which produces a burst of TeV neutrinos by photon-meson interaction, and eventually **breaks out** of the stellar **envelope** appearing as a Gamma Ray Burst (GRB). Based on recent numerical simulations and neutrino emission models, we **predict** the **expected** neutrino diffuse **flux** from these PopIII GRBs and compare it with the capabilities of present and planned detectors as AMANDA and IceCube. If **beamed** into 1% of the sky, we **find** that the rate of PopIII GRBs is = 4 × 106 yr-1. High energy

International Journal of Language and Applied Linguistics
2016; 2(3): 105-126
Published online December 15, 2016 (http://www.ijlal.ir)
ISSN: 2383-0514 (Online)

neutrinos from PopIII GRBs could dominate the overall **flux** in two energy **bands** [104-105] GeV and [105 - 106] GeV of neutrino telescopes. The enhanced sensitivities of forthcoming detectors in the high-energy band (AMANDA-II, IceCube) will provide a fundamental **insight** on the characteristic explosion energies of PopIII GRBs

The *flux* analogs for the subject matter are interspersed with words of perception for the methodology (*find, insight*), cf. Haase 2014 and 2016fc. and this is not even a popular text but a specialized astrophysics publication.

In an application of the structure mapping approach where the source domains provides the relationships between elements, out of these usually the easily perceived relationships emerge. We then assume that the same or a similar (analogue) relationship persists on the target side. In the structure-mapping hypothesis we find two levels of analogical description:

The computational level: Here, the analogy is due to direct perception and involves the systematic consistency of the relations

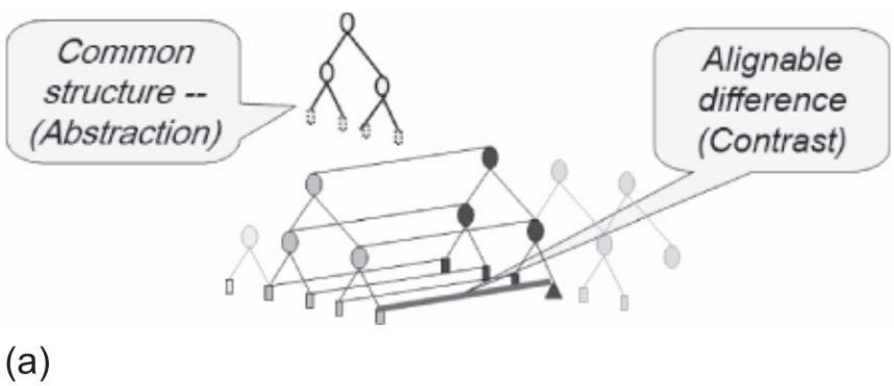

Figure 1. Level one (from Gentner & Christie 2010: 265)

Gentner observes that “People prefer to map connected systems of relations governed by higher-order relations with inferential import, rather than isolated predicates.” (Gentner 1989: 201). The inferential import provided by the visual or computational level can be likened to a “quick-and-dirty” approach, often in simple, vivid analogies as in the following example:

International Journal of Language and Applied Linguistics
2016; 2(3): 105-126
Published online December 15, 2016 (http://www.ijlal.ir)
ISSN: 2383-0514 (Online)

0055NS Most of the time, though, **chromosomes exist as long thin strands of more loosely wound DNA**. Until recently, their arrangement in the cell nucleus was seen as the random result of **their jumbling around together like lengths of string in a bag**. But recent work staining individual chromosomes with fluorescent tags has suggested they take up a highly complex but ordered architecture.

Chromosomes are compared to long, thin strands of DNA which jumble around etc. which makes for a very systematic analogue, not a one-off mental image. Still, these analogues do not hold on closer inspection and are often superseded by better, more powerful analogues as can be seen in the second example: One step higher we find that analogues hold better if they persist on a higher order of relationship.

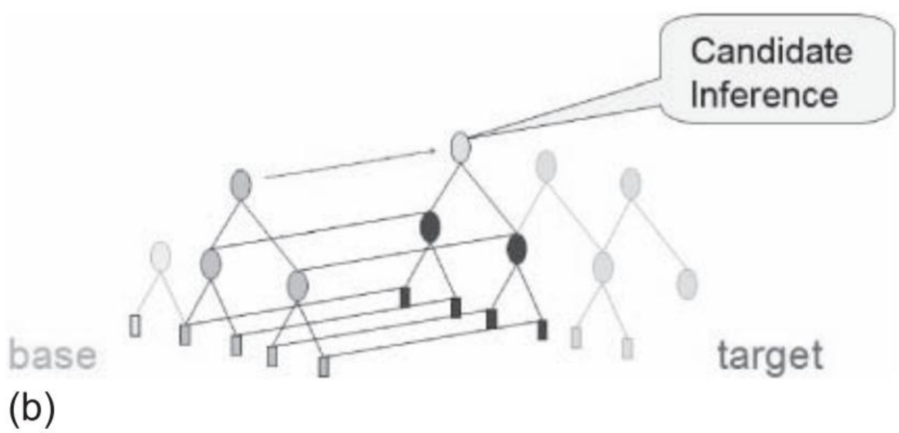

Figure 2. Gentner & Christie 2010: 265

b) The functional level: Here, the structural match is more highly resolved and the quality of the analogies depends on the knowledge of the reader, in what way it reflects actual experience

International Journal of Language and Applied Linguistics
2016; 2(3): 105-126
Published online December 15, 2016 (http://www.ijlal.ir)
ISSN: 2383-0514 (Online)

In the corpus example, a functional analogue (the cold virus) is what is likened to the studied virus in terms of behavior, not appearance:

> 0080NS To **insert** the gene, the team first **incorporated it into an adenovirus similar to the one that causes colds**. They also **plucked out** genes that enable the virus to replicate, so that it would **load its genetic cargo** into mouse cells without infecting the rest of the animal.

The above text talks about *inserting, incorporating, plucking out* and *loading the genetic cargo*. All these analogues are not visual but functional analogues and they are very systematic. Therefore, when looking at real science texts, the impressionistic analogues can be backed by solid data.

Data Discussion

Expectations

On studying analogy in academic writing we find that the mapping discussed above is constrained by a number of parameters, for example:

- by the knowledge of the observer/author
- by the knowledge of the reader
- by the knowledge of the author about the knowledge of the reader (the most complicated).

The emergence of genre types or registers is a consequence of shared and mutual knowledge between author and reader, thus establishing analogical mappings is revealing not only about the knowledge of the author but at the same time of the genre conventions (Swales 1993).

International Journal of Language and Applied Linguistics
2016; 2(3): 105-126
Published online December 15, 2016 (http://www.ijlal.ir)
ISSN: 2383-0514 (Online)

As mentioned above. *to cause* was selected as the most unambiguous way to map causes onto effects. Further, it provides a well-entrenched lexical field which enables quantification and comparison in different science text types.

We therefore have the following expectations:

- *cause* is dominant in the physical sciences compared with the biological sciences (discipline dominance)
- *cause* is dominant in academic science writing compared with popular science writing (genre/register dominance)

In the physical sciences, causation is the underlying assumption of mathematical formulæ therefore in the biological sciences, causation is expected to be established out of empirical observation (due to the diminished role of math).

Further, as academic science writing is based on logic and popular science writing is based more on temporal sequence (Haase 2010a), the cause dominance is expected in the specialized register with a secondary role in the popular register.

The Corpus

The corpus is called SPACE for corpus of Specialized and Popular ACademic English) which at the present moment has 1.5 million words with the two science domains of physical and biosciences as a stratified background. The sources are on the one hand side two preprint servers for academic publications, *arXiv.org* and *Proceedings of the National Academy of Sciences*. The corpus has a doubly binary structure in two registers: One is the original natural science register extracted from research papers. The unique second component is the popularized variant. Original research papers are routinely processed by science journalists who summarize and simplify them for popular science magazines like New Scientist in order to enable non-specialists with general academic skills to keep informed about research in other fields. Example titles of corpus texts from the two parallel variants can be found in the following table:

International Journal of Language and Applied Linguistics
2016; 2(3): 105-126
Published online December 15, 2016 (http://www.ijlal.ir)
ISSN: 2383-0514 (Online)

Table 2

SPACE Structure with Parallel Corpus Samples

Subcorpus	words	Title
arXiv	5,768	*Indeterminate-length quantum coding*
New Scientist	468	*The ultimate computer*
arXiv	3,852	*Quantum phase transitions and the breakdown of classical General Relativity*
New Scientist	2,134	*What lies beneath*
arXiv	2,226	*The disruption of stellar clusters containing massive Black Holes near the galactic center*
New Scientist	162	*Star shepherds*

In this study we looked for the distribution of the verb to cause and this was based on a subsample of 625,288 tokens which for the natural-science core of the corpus (under exclusion of the sections on medicine and psychology:

The Distribution of CAUSE

The distribution of all occurrences of the lemma *CAUSE* has been discussed against the background distribution in a number of standard corpora like BNC (queried via BNCWeb) and COCA (the Contemporary Corpus of American English). Both standard corpora have academic subsections which were queried only. Further, an academic standard corpus like MICUSP (Michigan Corpus of Upper-Level Student Papers) was used; however it was possible to specify the sections on physical and biosciences to match the queried sample from the SPACE corpus.

In the following table, the entire verbal paradigm of *CAUSE* is counted in different corpora, with the absolute frequency f_v and the normalized frequency per million words.

International Journal of Language and Applied Linguistics
2016; 2(3): 105-126
Published online December 15, 2016 (http://www.ijlal.ir)
ISSN: 2383-0514 (Online)

Table 3

Distribution of CAUSE in Different Corpora

Corpus	f_V	f_V per 10^6	Ac f_V	Ac f_V per 10^6
BNC	5,667	58.87	1,260	82.18
COCA	24,282	52.29	5,574	61.21
MICUSP (phys & bio)	1,061	204.04		
SPACE	215	343.84		

The first figures for the standard corpora (BNC and COCA) give the overall frequency, the second columns display the frequencies in the academic components of the standard corpora. Thus, for comparison, the last two columns for BNC and COCA and the first two columns for MICUSP and SPACE shall be compared. As an overall estimation, divergent figures for *CAUSE* in all corpora as well as in the academic subcorpora can be found. The British National Corpus shows the overall lowest figures, with its academic components it still shows only the second lowest counts. The considerable gap between SPACE and BNC (344 per million words of verbal CAUSE vs. ca. 60 per million words) can be explained by the fact that the academic texts collected in the BNC are not science texts per se as they are mainly law texts and political science where clear causation may be not as relevant. In the BNC we find an overall rather low frequency of 5,667 verbal occurrences absolute (out of 12,889 for CAUSE as a lexical item, including nominal and informal 'cause from the spoken component)

In COCA, the academic frequency is even lower (61.21).

Only MICUSP and SPACE are roughly in the same dimension but the counts for SPACE are higher by a factor of ca. 1.5. A reason could be that MICUSP collects intermediate to upper intermediate student papers (Römer & Swales 2010) which resemble popular science writing more. Unhedged, unambiguous causation is a rhetoric device students are apparently cautious to use. When hedged, the application, however, is twofold: The dual nature of hedges – one interpersonal, the other epistemic – are in essence Hyland's content-based and discourse-based hedges and fulfill his research-oriented vs. text-oriented function that he attributes to lexical use, cf. Pan, Reppen & Biber 2016: 67). The usually are connected in the application of politeness

International Journal of Language and Applied Linguistics
2016; 2(3): 105-126
Published online December 15, 2016 (http://www.ijlal.ir)
ISSN: 2383-0514 (Online)

when the force of a statement is mitigated, mainly hedges are used to negotiate the commitment to the truth of a statement.

The data can be interpreted in different ways. MICUSP shows considerable spread in the use of *CAUSE* in different science domains. So for example in the biological sciences we find roughly 260 occurrences per one million words with physics coming in lower than 200 occurrences (generated with http://micase.elicorpora.info/).

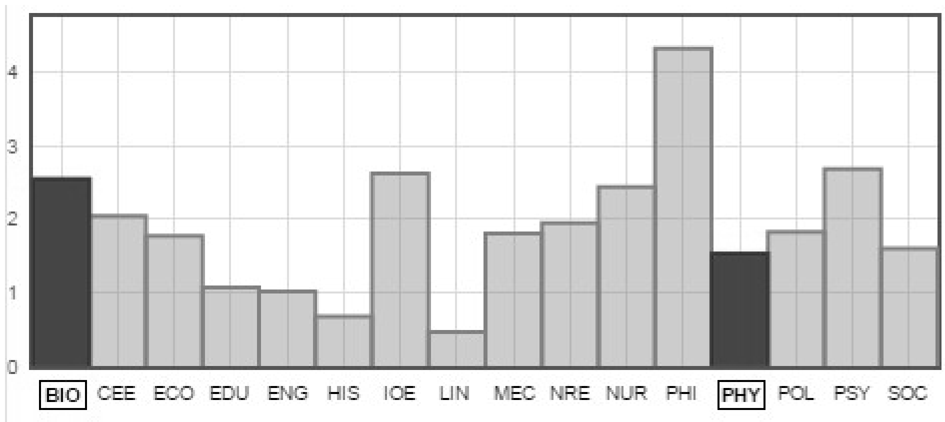

Figure 3. MICUSP proportional distribution of CAUSE in biology and physics (generated with http://micase.elicorpora.info/), (adapted from Haase 2015fc.)

Compared with the different domains in the SPACE corpus, the picture diversifies further. This is primarily due to the component of popular science texts which provide the counterparts to the original research papers, thus making these figures comparable.

International Journal of Language and Applied Linguistics
2016; 2(3): 105-126
Published online December 15, 2016 (http://www.ijlal.ir)
ISSN: 2383-0514 (Online)

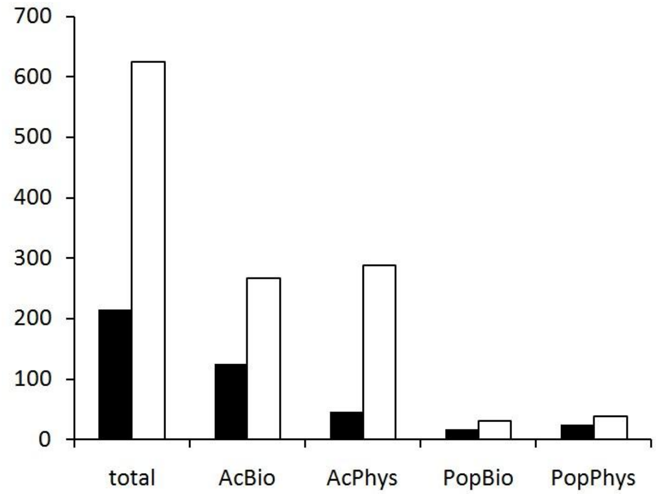

Figure 4. CAUSE in SPACE breakdown (counts in absolute numbers) (ordinate not to scale)

In the graph, the first set represents the total counts; the other sets are the subcorpora with the academic biosciences (AcBio), academic physics (AcPhys), popular biosciences (PopBio) and popular physics (PopPhys). The black bars represent the total frequency and are juxtaposed by the total word number in the subcorpora (which was subdivided by 1000 for comparison, thus the ordinate axis is not to scale). As can be seen, both specialized registers are larger subcorpora which is due to the fact that the popular variants are only summaries of the original research papers so they are necessarily shorter.

Table 4

Different Verbal Uses of CAUSE in the Four SPACE Components

Tag	AcBio	AcPhys	PopBio	PopPhys	total
f_V per 10^6	471.94	159.25	578.31	652.42	343.84

In SPACE, the normalized occurrences appear highly heterogeneous with popular physics leading the field with 652 occurrences per one million words. This is a substantial gap to the domain that was expected to provide the highest count as the "hardest" science, academic physics. However, this domain has the overall lowest score with 159 hits per million. Closer together are academic and popular bioscience. Although again the popular text type has a higher score of 578 (compared to academic bioscience with 472), the difference is not as pronounced.

International Journal of Language and Applied Linguistics
2016; 2(3): 105-126
Published online December 15, 2016 (http://www.ijlal.ir)
ISSN: 2383-0514 (Online)

But overall, both popular registers outnumber both academic registers. We can therefore assume a higher in-corpus variability than intercorpus variability but the same ranking in academic biology and physics persists as in MICUSP.

The explanation for this effect is partly speculation but it ties in with similar observations on the use of hedge expressions with equally surprising results.

In an earlier study (Haase 2010b) the expected use of hedge expressions and the subsequent negotiation of the commitment of the author/s to their result was found to be stronger with the popular-science authors. While they are removed from the actual research, the distance also allows for a lack of cautiousness and for a desire for simplification. Therefore, in tendency, what is gray becomes black and white. Thus, two phenomena that might be causally linked in the original research become partners in an unambiguous cause-effect relationship in the popularized version.

If the parameter of research discipline is considered, the difference within the academic sections in SPACE between academic biosciences and academic physics is a striking 471.94 vs. 159.25. The following sample illustrated the usage:

Academic biosciences

0065PN Increasing plant diversity beginning in the Silurian (425 million years ago) led to increasing weathering of rocks that had two effects: atmospheric CO_2 levels decreased, **causing** a decrease in carbon isotope fractionation in marine deposits;

Academic physics

0004AX the number of systems that change from sA = +1 to sA = -1 is unequal to the number that change from sA = -1 to sA = +1, **causing** an imbalance that changes the outcome ratios at A. In other words, in general the statistical distribution of outcomes at A is altered by the distant shift

International Journal of Language and Applied Linguistics
2016; 2(3): 105-126
Published online December 15, 2016 (http://www.ijlal.ir)
ISSN: 2383-0514 (Online)

The trend continues if all bioscience and physics subcorpra are compared: In the overall discipline breakdown (aggregated from the academic and popular sections) we find the relation to be 525.12 per million in biosciences vs. 392.33 in physics.

While 578.31 per million in popular biosciences stands against 652.42 in popular physics, the samples below show all hallmarks of popular science writing:

Popular biosciences

0080NS To insert the gene, the team first incorporated it into an adenovirus similar to the one that **causes** colds. They also plucked out genes that enable the virus to replicate, so that it would load its genetic cargo into mouse cells without infecting the rest of the animal.

Popular physics

0003NS The unpredictable accelerations **caused** by these spurts are at least 10 times as big as the Pioneer acceleration, and make it impossible to measure the effect.

If both parameters (research discipline and register) are considered, the most surprising result of this study is obtained: the use in the popular text-types is considerably higher than in the academic text types. In the aggregated discipline/register breakdown, 315.59 occurrences of *CAUSE* per million words in academic science writing stand against a colossal 615.35 occurrences in popular science writing. These findings are summarized in the following table:

Table 5

CAUSE in all SPACE subcorpora

genre/discipline	Biosciences	Physics	total
Popular	578.31	652.42	615.35
Academic	471.94	159.25	315.59
f_V per 10^6	525.12	392.33	343.84

International Journal of Language and Applied Linguistics
2016; 2(3): 105-126
Published online December 15, 2016 (http://www.ijlal.ir)
ISSN: 2383-0514 (Online)

While bioscience causation with *CAUSE* outnumber physical sciences, the main hypothesis of its prevalence in the academic registers needs to be rejected. Both popular subcorpora outnumber the corresponding specialized academic subcorpora.

Analogical Mapping in the X CAUSE Y Construction

The proposed causal link of analogical domains in cause-effect relationships can be studied by assigning these domains different types that correspond with the nature of the domains. As the analogical mapping occurs usually from concrete domains to abstract domains (thus facilitating comprehension and recall), the concrete-abstract level is a parameter in this typology. The suggested typology can be seen in the following table:

Table 6

Types of Domain and Directionality Assignments

domain cause	domain effect	type
concrete-specific	concrete-specific	A
concrete-specific	concrete-general	B
concrete-general	concrete-general	C
concrete-general	concrete-specific	D
abstract-specific	abstract-specific	E
abstract-specific	abstract-general	F
abstract-general	abstract-general	G
abstract-general	abstract-specific	H

Applied to the different registers we can then map the conceptual elements of cause and effect and determine their direction. This is done as an example below for the register of popular biosciences:

International Journal of Language and Applied Linguistics
2016; 2(3): 105-126
Published online December 15, 2016 (http://www.ijlal.ir)
ISSN: 2383-0514 (Online)

Table 7

Conceptual Element CAUSED Collocates and their Mapping Types - PopBio

domain cause	domain effect	type
fragment release	*treatment effects*	C
oxygen release	*mitochondrial decay*	B
Blade	*amber split*	A
Pain	*activity decline*	H
ROS formation	*Finding*	F
ALCARLA addition	*oxidant decline*	B
Replenishment	*potential reversal*	C
oxidative stress	*ALCARLA levels*	D

Comparatively few purely abstract elements were chosen to stand in for causes and effects with mainly concrete phenomena discussed, occasionally concrete causes (like a *blow*) would precipitate a more general effect (*injury*).

In academic biosciences there are minor variations into the purely abstract when for example simply *findings* are caused. This is summarized in the following table:

Table 8

Conceptual Element CAUSED Collocates and their Mapping Types - AcBio

domain cause	**domain effect**	**type**
fragment release	*treatment effects*	C
oxygen release	*mitochondrial decay*	B
Blade	*amber split*	A
Pain	*activity decline*	H
ROS formation	*Finding*	F
ALCARLA addition	*oxidant decline*	B
Replenishment	*potential reversal*	C
oxidative stress	*ALCARLA levels*	D

The domain type attributions overall show a less heterogeneous profile although this largely depends on the classification. The classification itself is hard to formalize, therefore more research is needed for a comprehensive quantitative view. However, in the light of the discussion

International Journal of Language and Applied Linguistics
2016; 2(3): 105-126
Published online December 15, 2016 (http://www.ijlal.ir)
ISSN: 2383-0514 (Online)

on the efficiency of analogy-making, if the second half of the typology (E to H) prevails, comprehension of abstract-to-abstract may not be facilitated. It would be a question of modifying EAP courses to make students aware of the domains they use for clarity and precision.

Conclusion

The study presented related two eminent mechanisms of academic writing – the establishment of cause-effect relationships on the one side and analogy-making on the other to a survey of the verbal paradigm of CAUSE in different corpora. The results show remarkable spread in subcorpora dedicated to different academic ends (physical science texts use it differently than bioscience texts) but primarily a register difference was obtained as a striking result. While the standard corpora overall due to poor sampling on the side of their makers (exclusion of "real" science for example in the BNC) show insufficient results, the specialized SPACE corpus shows considerable spread when considered from the research discipline and the register (specialized and popular).

We can conclude that all registers of academic writing rely on causal relationships and on analogous mapping but to varying degrees. Further, it brings two strategies – both at the heart of academic writing – together in one analytical context.

International Journal of Language and Applied Linguistics
2016; 2(3): 105-126
Published online December 15, 2016 (http://www.ijlal.ir)
ISSN: 2383-0514 (Online)

References

Benesch, S. (2001). *Critical English for academic purposes: theory, politics, and practice*. Mahwah, NJ: Lawrence Erlbaum.

Blevins, J. P. (2006). English Inflection and Derivation. In: *The Handbook of English Linguistics*, Edited by Aarts, B. & McMahon, A. Cambridge: Blackwell, 507-532.

Canagarajah, S.A. (2002). *Critical Academic Writing and Multilingual Students*. Ann Arbor, MI: The University of Michigan Press.

Flaherty, M.G. (2011). *The textures of time. Agency and temporal experience*. Philadelphia: Temple University Press.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science* 7: 155–170.

Gentner, D. (1989). The mechanisms of analogical learning, In Vodniadou, S. & Ortony, A. (Eds.). *Similarity and analogical reasoning.* Cambridge: CUP.

Gentner, D. (2002). Analogy in scientific discovery: The case of Johannes Kepler. In *Model-based reasoning: Science, technology, values.* ed. by L. Magnani & N.J. Nersessian. New York: Kluwer, pp. 21-39.

Gentner, D. & Christie, S. (2010). Mutual bootstrapping between language and analogical processing. *Language and Cognition*, 2(2). 261-283.

Grabe, W. & Kaplan, R.B. (2010). Regional Studies: North America. In Berns, M. & Brown, K. (Eds.), *Concise encyclopedia of applied linguistics*. Amsterdam: Elsevier, 59-65.

Haase, C. (2006). Contrastive causation: A corpus view. In: Walinski J., Kredens, K., & Gozdz-Roszkowski, S. (eds.), PALC 2005, proceedings. München: Lang.

Haase, C. (2010a). Verb classes and the grammaticalization of causativity in discourse. In: Witczak-Plisiecka, I. (ed.) *Pragmatic Perspectives on Language and Linguistics* 2009. Vol.1: Speech Actions in Theory and Applied Studies. Sheffield: Cambridge Scholars.

Haase, C. (2010b). Mediating between the “two cultures” in academia: The role of conceptual metaphor. *Discourse and Interaction* 1 (3) (2010), 5-18.

Haase, C. (2014). Registers of Analogy: Popular Science and Academic Discourse. In *Communication Across Genres and Discourses*, edited by R. Vogel. Brno: Muni Press.

International Journal of Language and Applied Linguistics
2016; 2(3): 105-126
Published online December 15, 2016 (http://www.ijlal.ir)
ISSN: 2383-0514 (Online)

Haase, C. (2015). Causal chaining and the active/passive ratio in science texts. *Discourse and Interaction* 5(2), 1-16.

Haase, C. (2017fc.). Verbs of perception as metaphorical devices in the natural sciences: A corpus study of specialized and popular text types. In *Metaphor in Science Writing*, edited by T. Smith and A. Beger. Amsterdam: John Benjamins.

Harrison, A.G. & D.F. Treagust (2006). Teaching and learning with analogies. In *Metaphor and analogy in science education*, ed. by Aubusson, P.J., A.G. Harrison & S.M. Ritchie. Berlin: Springer, pp. 11-24.

Hyland, K. (2006). *English for Academic Purposes. An Advanced Resource Book*. London & New York: Routledge.

Itkonen, E. (2005). Analogy as structure and process. Approaches in linguistics, cognitive psychology and philosophy of science. Amsterdam: Benjamins.

Mackie, J.L. (1980). *The Cement of the Universe: A Study of Causation*. Clarendon Library of Logic and Philosophy. Oxford: OUP.

Pan, F., Reppen, R. & Biber, D. (2016). Comparing patterns of L1 versus L2 English academic professionals: Lexical bundles in Telecommunications research journals. *Journal of English for Academic Purposes* 21, 60-71.

Römer, U. & Swales, J. (2010). The Michigan Corpus of Upper-level Student Papers (MICUSP). *Journal of English for Academic Purposes* 9, 249.

Swales, J. (1993). Genre and Engagement. *Revue belge de philologie et d'histoire*, 71, 3, 687-698.

International Journal of Language and Applied Linguistics
2016; 2(3): 105-126
Published online December 15, 2016 (http://www.ijlal.ir)
ISSN: 2383-0514 (Online)

Author

Christoph Haase, PhD., studied at first physics, later English and German with particular focus on linguistics at the Ernst-Moritz-Arndt University in Greifswald, Germany. His studies abroad involved Generative linguistics at the University of Oviedo in Oviedo, Spain and Cognitive Science as a visiting scholar at Carleton University in Ottawa, Canada. In 2000 Christoph came to Chemnitz University of Technology to work on the Internet Grammar. In 2001 he became Assistant Professor. He defended his PhD dissertation in 2002. In 2009 he came to Usti to work at KAJ. His research interests revolve around cognitive linguistics, lexical semantics, universals and typology, grammaticalization and the philosophy of language.