Computational Linguistic Assessment of Textbooks and Online Texts by Means of Threshold Concepts in Economics

Andy Lücking [1,2*], Sebastian Brückner [3], Giuseppe Abrami [1], Tolga Uslu [1] and Alexander Mehler [1]

[1] Text Technology Lab, Faculty of Computer Science and Mathematics, Institute of Computer Science, Goethe University Frankfurt, Frankfurt, Germany, [2] Laboratoire de Linguistique Formelle, Laboratory of Excellence "Empirical Foundations of Linguistics", Université de Paris, Paris, France, [3] Department of Business and Economics Education, Johannes Gutenberg University Mainz, Mainz, Germany

The ongoing digitalization of educational resources and the use of the internet lead to a steady increase of potentially available learning media. However, many of the media which are used for educational purposes have not been designed specifically for teaching and learning. Usually, linguistic criteria of readability and comprehensibility as well as content-related criteria are used independently to assess and compare the quality of educational media. This also holds true for educational media used in economics. This article aims to improve the analysis of textual learning media used in economic education by drawing on threshold concepts. Threshold concepts are key terms in knowledge acquisition within a domain. From a linguistic perspective, however, threshold concepts are instances of specialized vocabularies, exhibiting particular linguistic features. In three kinds of (German) resources, namely in textbooks, in newspapers, and on Wikipedia, we investigate the distributive profiles of 63 threshold concepts identified in economics education (which have been collected from threshold concept research). We looked at the threshold concepts' frequency distribution, their compound distribution, and their network structure within the three kinds of resources. The two main findings of our analysis show that firstly, the three kinds of resources can indeed be distinguished in terms of their threshold concepts' profiles. Secondly, Wikipedia definitely shows stronger associative connections between economic threshold concepts than the other sources. We discuss the findings in relation to adequate media use for teaching and learning—not only in economic education.

Keywords: threshold concepts, corpus study, wikipedia, newspaper, specialized vocabulary, network model, economics, textbooks

1. INTRODUCTION

In recent years, research on how to facilitate teaching, curriculum development, and the diagnostic of competences acquired during higher education studies has intensified significantly in many disciplines, not only in Germany but also worldwide (Nicola-Richmond et al., 2018; Zlatkin-Troitschanskaia et al., 2018). As shown in various instructional models (e.g., the *offer-use model*

by Helmke, 2009), the quality of learning media is of crucial importance for the learning success of students. A central challenge for higher education lecturers from all disciplines is to select high-quality learning media for their teaching, for which learning media research provides corresponding findings. The investigation of the quality of learning media can be investigated on the basis of a variety of criteria. Expert ratings are often used for the evaluation of learning media using criteria, such as *Accuracy, Clarity, Comprehensiveness, Consistency, Grammar, Readability, Modularity,* and *Cultural Relevance* (Fischer et al., 2017). An important quality criterion and benchmark besides didactic, pictorial, and further media-structural characteristics is that the learning media address central concepts of a subject area and their interconnectedness, because the extent to which a digital medium supports learning success depends largely on the quality of the content presented in it (Devetak and Vogrinc, 2013).

In economic education in higher education in particular, a large amount of different media sources are frequently used because economic phenomena are the subject of everyday encounters and historical events (Simkins, 1999; Davies and Mangan, 2007; Meier, 2008; Hoyt and McGoldrick, 2012; Schuhen and Kunde, 2016). Traditionally, of course, the major learning resources are textbooks (Jadin and Zöserl, 2009; Maurer et al., 2019; Dalimunte and Pramoolsook, 2020), whose didactic purposes include, among others, the introduction of technical vocabulary. Currently, many textbooks are available to students as *Open Educational Resources* (OER), but the predominant use of textbooks as a learning resource has emerged over the years. This development has been attributed to the professional quality and the connection to lectures and courses (Devetak and Vogrinc, 2013; Fischer et al., 2017; Maurer et al., 2019; Dalimunte and Pramoolsook, 2020). Furthermore, in economics education, textbooks are central for teaching and learning in formal teaching-learning environments (Leet and Lopus, 2003; Richardson, 2004; Tinkler and Woods, 2013). In connection with the increasing digitization of university teaching, digital learning platforms, forums, and online encyclopedias are increasingly used by students as a complementary source of learning alongside textbooks because of their easy and often free access (Brooks, 2016; Kilgour et al., 2019; Maurer et al., 2019). According to several studies (Knight and Pryke, 2012; Steffens et al., 2017; Johinke and Di Lauro, 2020), Wikipedia is one the topmost used internet services. Online encyclopedias, such as Wikipedia, are often used to quickly access summaries and definitions or as a first encounter with subject-specific concepts (Jadin and Zöserl, 2009; Lim, 2009; Knight and Pryke, 2012; Maurer et al., 2019; Johinke and Di Lauro, 2020). Sources for learning-related purposes mainly are used by learners who explore the core contents and concepts of their respective fields (Knight and Pryke, 2012; Steffens et al., 2017; Maurer et al., 2019). In college-level economic education, an increase of the usage of digital learning tools in formal education has been acknowledged for many years (Simkins, 1999). Besides textbooks (Hu and Gao, 2019), Wikipedia is considered for initial orientation and for dealing with economic content, too (Meier, 2008; Haab et al., 2012; Freire and Li, 2016).

However, economic education is a differentiated field of study, the content may gradually change, for example, in the light of current news or changes in legislation. For current events, newspapers offer a way to stay up-to-date on the economic situation in businesses and countries (Croushore, 2012). In addition, many newspapers are easier to understand, especially for novice learners (Dalimunte and Pramoolsook, 2020), and are therefore sometimes read as frequently as online economic blogs (Haab et al., 2012). For a long time, newspapers have been one of the main resources used in economics education. Especially lecturers in introductory courses often use the variable prior knowledge of students regarding current issues in business and economics to encourage more active engagement with the subject. As current research suggests, students frequently come into contact with economic content in their everyday life by reading newspapers (Hoyt and McGoldrick, 2012). Especially in Germany, unlike in other industrial nations, business or economics has not yet been established as a school subject in Germany (Schuhen and Kunde, 2016). The majority of first-year students at German universities usually have previous knowledge that was acquired in an informal[1] context (cf. Schumann et al., 2010). The first-year students' knowledge of economics often comes from various media that are not directly related to a learning-intended purpose (e.g., online magazines, news magazines, videos) (Maurer et al., 2019), social interactions on financial topics (e.g., as a consumer in a supermarket or buying a mobile phone) (Davies and Mangan, 2007; Schuhen and Kunde, 2016), or other behavior of economic relevance (e.g., retirement planning). Consequently, students may also use textbooks, Wikipedia and newspapers as central learning media in economic education. In order to ensure that learning media with the highest possible quality of content are used in a way that is appropriate for the target group, lecturers are therefore inevitably faced with the question of which media to select for a given topic or concept to be taught. However, a comparative analysis of digital learning media in economics education with regard to concrete professional concepts is still pending. Since the core of these teaching-learning media in economics is always central focal content (Leet and Lopus, 2003), we will compare these media using domain specific economic concepts. In economics education the so-called *thresholds concepts* are a current and frequently discussed approach that seeks to identify the most important concepts for learning economics (Meyer and Land, 2006; Davies and Mangan, 2007). Therefore, the three media types can be compared using the linguistic features of these concepts. In order to provide teachers with a certain information basis for the selection of media based on the comparison of threshold concepts in learning media, we will address the following research question in this paper: To what extent do textbooks, Wikipedia, and newspapers used (by students) for learning about economic concepts differ in terms of

---

[1]Informal learning can—taking into account the variety of definitions—essentially be understood as learning *en passant*; i.e., learning that takes place implicitly when carrying out other activities (e.g., learn about costs when reading a newspaper article), is usually not consciously controlled by the learner (Neuweg, 2000; Hofhues, 2016).

the structure and linguistic characteristics of threshold concepts that are important for learning?

In section 2, threshold concepts are introduced in more detail and discussed in relation to domain-specificity, conceptual change, and specialized vocabularies. A theoretical linguistic perspective is outlined in section 3. In section 3, a theoretical linguistic perspective is outlined that shows how learning how learning can be construed in terms of a dynamic update semantics and how linked *mental files* represent relations between threshold concept terms in texts. Some terminological and conceptual distinctions that arise in this context are drawn in section 4. Section 5 then introduces a computational linguistic approach for deriving networks of linked threshold concepts on a large scale. The method is applied to three types of (online) resources, namely newspaper articles, textbooks, and Wikipedia article. The results are finally discussed in section 6.

2. THRESHOLD CONCEPTS APPROACH AND CONCEPTUAL CHANGE

An approach using threshold concepts rather than simplified content categories (Kricks et al., 2013) has been introduced into didactic discussions that focuses on highest potentials for developing a professional disciplinary understanding for both novice and experienced learners (Meyer and Land, 2013). The authors describe threshold concepts as "akin to a portal, opening up a new and previously inaccessible way of thinking about something" (Meyer and Land, 2006, p. 3). Due to their special character within a discipline, they thus represent a threshold that needs to be crossed and that fundamentally changes the learner's understanding of the discipline. Concepts can thus describe principles and rules, objects, theories, modeling methods on an abstract level, which contribute to the development of a comprehensive understanding of the learner within an individual discipline (Sender, 2017).

Often the threshold concepts approach refers to learning in the sense of conceptual change (Davies and Mangan, 2007): it is assumed that knowledge gain is not just an accumulative process of mere addition of new knowledge, but that the learner's existing knowledge structures are (possibly fundamentally) transformed (Davies and Mangan, 2007). If the learner develops a new understanding of a concept, the conceptual change can be very sudden and unexpected, namely when the learner experiences the new concept as expanding his or her previous field of imagination. This initial change of concepts can be demonstrated didactically by a change of perspective for the learner, e.g., by looking at a purchase decision from the roles of buyer and supplier and thus better understanding the formation of prices (Sender, 2017, p. 56). This illustrates a short-term event in the learning process. If the learner is able to adapt and transfer his new concept to other contexts and examples, or if he experiences the limits of his newly developed conceptions, the knowledge structures are gradually changed and consolidated, so that a threshold concept also has a long-term effect (Sender, 2017). Thus, the more short- and long-term support the understanding of a concept has, the more irreversible the understanding is

(Cousin, 2008). Accordingly, *irreversibility* is one characteristic of threshold concepts, alongside *transformativity*, *integrativity*, *limitedness*, and *difficulty* (Meyer and Land, 2005, 2006). The constant transformation and application of the acquired knowledge to a variety of known phenomena promotes the intertwining of knowledge. Integrativity leads to the fact that different knowledge structures, which previously could not be put into context for the learner, are increasingly brought into a semantic relation. Threshold concepts are also limited, since the new conceptual spaces created by linking content-related ideas simultaneously create new boundaries that distinguish the discipline from other academic disciplines (Meyer and Land, 2005).

2.1. Threshold Concepts in Business and Economics

A large number of studies focus on the identification of threshold concepts (Sender, 2017; Brückner and Zlatkin-Troitschanskaia, 2018; Hatt, 2018; Lamb et al., 2019; van Mourik and Wilkin, 2019; Ivan Montiel and Antolin-Lopez, 2020). *Opportunity costs* was the initial threshold concept that has been identified for the discipline of economics (Meyer and Shanahan, 2003) and has since been taken up in several studies (Shanahan et al., 2006; Davies and Mangan, 2007). The critical discourse and empirical examination as to which concepts can be considered threshold concepts and which are important for the curriculum but not mandatory is ongoing and has since been discussed in a number of papers (Davies and Mangan, 2007; Lucas and Mladenovic, 2009; Ivan Montiel and Antolin-Lopez, 2020). Over the years, in addition to opportunity costs, a large number of concepts have been proposed and empirically tested in economics, e.g., on depreciation (Lucas and Mladenovic, 2009), elasticity (Reimann and Jackson, 2006), information asymmetry (Hoadley et al., 2015), and many more, on the basis of multiple research methods, e.g., using interviews with teachers or learners, videographies, curriculum analyses or standardized tests. For example, in a Delphi study, Hatt (2018) use interviews with entrepreneurs to investigate which concepts they regard as threshold concepts. Ivan Montiel and Antolin-Lopez (2020) conduct a literature analysis and develops 33 threshold concepts for corporate sustainable management. Davies and Mangan (2007) identify threshold concepts in economics on the basis of literature analysis and Hoadley et al. (2015) use expert interviews to find out whether or not a pre-selected sample of threshold concepts actually consists of threshold concepts. Some studies also examine facets of conceptual change on this basis. For example, Sender (2017) analyzes how affective and cognitive states develop in liminal phases of understanding when confronted with threshold concepts in economics courses. Brückner and Zlatkin-Troitschanskaia (2018) examine how confident students are in their ability to assess their solution behavior in tests when the complexity of threshold concepts increases. A number of studies also describe that the relationships established between the threshold concepts by the learner are of great importance for generating a deeper understanding (Davies

and Mangan, 2007; Vidal et al., 2015; Ivan Montiel and Antolin-Lopez, 2020). A central area of research also lies in various types of conceptual change. Davies and Mangan (2007) distinguish three forms, i.e., the *basic*, *discipline*, and *procedural* form of conceptual change. This three-part categorization has been taken up frequently, especially in recent years, by integrating further concepts from the economic sciences and further developing existing concept attributions (Lucas and Mladenovic, 2009; Kricks et al., 2013; Hoadley et al., 2015; Sender, 2017; Brückner and Zlatkin-Troitschanskaia, 2018; van Mourik and Wilkin, 2019). A basic conceptual change is defined as "Understanding of everyday experience transformed through integration of personal experience with ideas from discipline" (Davies and Mangan, 2007, p. 715). This is a conceptual change, which is fundamental and which a learner experiences as soon as he develops a first disciplinary understanding, e.g., of the concept of *cost*. Concepts documented along the basic threshold are accessible to most learners, as they are confronted with their everyday life (e.g., in their behavior as consumers) (Davies and Mangan, 2007). At the level of the disciplinary threshold, the learner succeeds in developing and linking conceptual understandings based on a theoretically elaborated perspective, which is hardly accessible from everyday life. This concerns concepts that are mainly accessible within the economic sciences (e.g., the concept of opportunity costs, hedging; depreciation; see Davies and Mangan, 2007; Lucas and Mladenovic, 2009; Hoadley et al., 2015). Some of the concepts require that a first encounter with a subject has already taken place and that the learner has a basic level of knowledge (Davies and Mangan, 2007), for example, the concept of costs should be understood before the opportunity cost principle is understood. The procedural threshold comprises concepts that are deeply integrated in the subject structures and require an understanding of modeling in economics. These are abstract modeling methods, procedures or argumentations that are used to analyze economic phenomena, but also to further develop economic theories (e.g., comparative statics, intertemporality; Davies and Mangan, 2007; Sender, 2017; Brückner and Zlatkin-Troitschanskaia, 2018). However, it can be seen that the studies mainly focus on learning processes and learning success as well as on personal prerequisites. According to the *offer-use-model* (Helmke and Schrader, 2008), it is important to investigate whether learning media also offer learners the possibility to go through this conceptual change and to connect concepts with each other. It is therefore important to investigate to what extent the threshold concepts are represented in the learning media used by the learners. The frequency of occurrence in learning media and the cross-linking of threshold concepts (Davies and Mangan, 2007) are thus a central aspect of the investigation of the potential of learning media.

Due to their fundamental character for the genesis of a disciplinary economic understanding, threshold concepts are often a central content in textbooks and are sometimes referred to as "building blocks" (Davies and Mangan, 2007, p. 724). A number of studies also start in their investigations in textbooks, often analyzing the variable views and differences in their understanding by learners (Lucas and Mladenovic, 2009). Less frequently, linguistic characteristics and representations of

threshold concepts are considered, although these have been shown to be of great importance for learning and understanding processes (Mayer, 2005). For example, Shanahan et al. (2006, p. 105) explicate: "Many first-year economics students report, that they find 'economic jargon' the most difficult barrier to their understanding. For economists 'learning the language' is one of the necessary elements to 'think like an economist'."
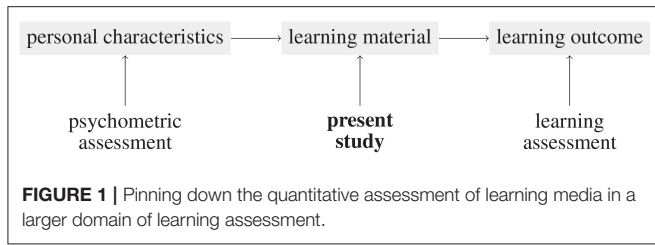
2.2. Threshold Concepts and Specialized Vocabularies

Since threshold concepts in business and economics are addressed by words it comes as no surprise that there is a connection to investigations from linguistics, in particular in studies of a certain kind of a manner of speaking (a socio-, functo-, or technolect) known as *specialized languages*, or the "language of science." A specialized language is more than just a specialized vocabulary since it involves grammatical aspects as well (Crystal, 1997, p. 384)—however, the vocabulary is the most salient part of a scientific sociolect and threshold concepts are no exception to this impression. Accordingly, there is a branch of linguistics specialized on specialized languages (see Roelcke, 2010 for an introduction), in particular in lexicography (Hoffmann et al., 1998). Interestingly, lexicographic work on specialized vocabularies distinguishes three classes of scientific expressions: "technical terms, semi-technical terms, and general vocabulary frequently used in a specialized domain" (Motos, 2011, p. 9, quoted from Nagy, 2014, p. 267). Obviously, there is a coincidence with the 3-fold distinction of threshold concepts into basic, discipline, and procedural, which could be worth pursuing. The present study, however, investigates textual features with regard to threshold concepts, based on linguistic considerations concerning specialized languages.

3. THEORETICAL LINGUISTICS PERSPECTIVE: THRESHOLD CONCEPTS IN DISCOURSE REPRESENTATION STRUCTURES

Three general factors from the complex network of factors that influence learning introduced in sections 1 and 2 can be extracted: personal characteristics, learning material, and learning outcome (cf. **Figure 1**). There are statistical assessments for both the personal characteristics and the learning outcome (Lodico et al., 2006). However, assessments regarding the learning material (e.g., quantification of texts) are rare. The aim of the present study is to develop a methodological proposal in this respect. Threshold concepts seem to be particular suited for obtaining a reference frame that is needed for frequentist analyses and comparisons. Threshold concepts are especially suited for this task since the corresponding word forms are easily identifiable in texts (see section 4 on words and concepts) and they are related to conceptual change (cf. section 2).
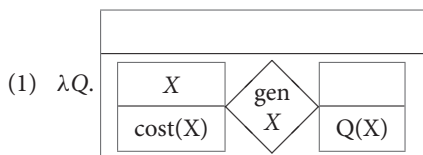
Let us illustrate this with a very simple example, namely *Kosten* "cost." The everyday sense of *cost* is derived from buying events. This is encoded in natural language grammar where the *lexical frame* (Fillmore et al., 2012) for the noun *cost* has four core

FIGURE 1 | Pinning down the quantitative assessment of learning media in a larger domain of learning assessment.

elements: *asset*, *goods*, *intended_event*, and *payer*[2]. Accordingly, we can think of the psycholinguistic, everyday concept of *cost* as a *sensorimotor simulator* of such buying events (Barsalou, 1999). Now this does not square easily with the economic sense of *cost*. The associated German Wikipedia page, for instance, starts as follows [translated by AL][3]:

> Costs are the negative consequences of the use of production factors with an impact on profits. The exact definitions differ depending on the subject area. In the economic sense of cost accounting, costs are usually understood to be the consumption of production factors valued in monetary units.

The economic definition of *cost* is at most indirectly related to buying events (each "production factor" has eventual to be paid in the everyday sense, though, hence providing evidence that *cost* is to be classified as a *basic* threshold concept). The subject noun *costs* which starts the Wikipedia article compiles a new *mental file* (Heim, 2002; Murez and Recanati, 2016) or *discourse referent* (Karttunen, 1969; Kamp and Reyle, 1993) which becomes the information structural *topic* (Cohen and Erteschik-Shir, 2002). Since *costs* is a bare plural noun, it introduces a plurality (represented by capital X) and receives a generic interpretation (Link, 1983; Krifka, 2003). Using the graphical discourse representation format of Asher (1993) and Kamp and Reyle (1993), the semantic representation of *costs* at this point is as follows:


The file or discourse universe is in the following populated with the predication, regimented by a syntax-driven construction algorithm (cf. subsection 4.1.3). Plural *be*, "are," triggers the parsing hypothesis (Demberg et al., 2013) that *are* is a copula

which initiates a predication on its subject (this parsing prediction turns out to be correct). The copula is interpreted in terms of the identity function (Russell, 1919) and introduces the corresponding condition (the predicate variable Q provides the interface for composition). Having processed the remainder of the sentence in this fashion, the semantic representation given in (2) is obtained (note that the deverbal noun *use* receives an eventive interpretation, as does the identity relation; processing present tense introduces the condition that the main event e_1 holds at the indexical time point n, "now").
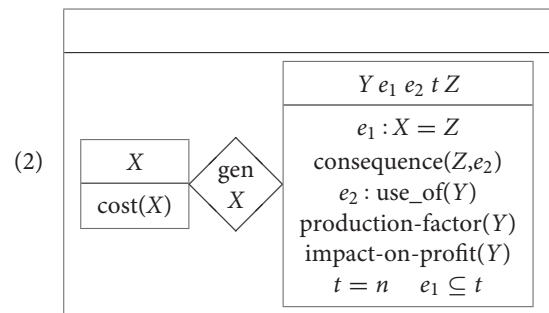
Part of the predicative content in (2) is the information that costs follow from production factors. Since this sounds different from what the learner knows from his or her everyday language competence, the new mental file *costs* is not merged with the eponymous pre-theoretic one (though both remain related at least due to phonological identity). Furthermore, if the learner already has a (rich) mental file for the noun compound *production factors* (Y), integration of both files will happen at this point. This integration obviously depends on the learner's prior knowledge[4].

The Wikipedia article continues with mentioning opportunity costs alongside costs. This mention again compiles a mental file. Since *costs* and *opportunity costs* share a great deal of surface form (namely the head noun *costs*) they will be connected, but their precise connection at this point is still unspecified (given that the learner has no prior knowledge in this regard). Thus, already after a few sentences, two threshold concepts will be initialized and connected—in terms of operations on a knowledge base: the knowledge base is expanded (by introducing mental files) and denser connected (by a *Consequence* relation) (Chi and Ohlsson, 2005, p. 376 f.).

Textbooks often refrain from an initial definition of costs in favor of a distinction of different *types* of costs (namely elaborating on the production factors mentioned in the above-given quotation)[5]. Instead they list examples for costs, such as delivery costs, holding costs, production costs, retooling costs, etc. Accordingly, the mental file is populated with sub-types of costs. These sub-types are

---

[2]https://framenet2.icsi.berkeley.edu/fnReports/data/lu/lu9191.xml?mode= lexentry (accessed October 29, 2020).

[3]https://de.wikipedia.org/wiki/Kosten (accessed October 29, 2020). "Kosten sind die negativen Konsequenzen einer erfolgswirksamen Nutzung von Produktionsfaktoren. Die genauen Definitionen unterscheiden sich je nach Fachgebiet. Im betriebswirtschaftlichen Sinn der Kostenrechnung wird darunter meist der in Geldeinheiten bewertete Verbrauch an Produktionsfaktoren verstanden."

[4]Additionally the learner may associate, for instance, personal experiences with any of the mental files, but this is not part of text meaning, see subsection 4.1.2.
[5]This enumerative way is taken, for instance, in Weber et al. (2014), Mumm (2015), and Blum (2017).

connected to the file's header by means of *Elaboration* relations, which involve a *part-of* condition by default (Asher and Lascarides, 2003, p. 160). Hence, the knowledge base is expanded and this expansion receives a more fine-grained representation (Chi and Ohlsson, 2005, p. 376, p. 382).

Since "opportunity costs" follows the same compound structure as the other just mentioned sub-types of costs, the initial hypothesis is to add it *via Elaboration* to *costs*'s mental file, too. However, in this case linguistic structure is deceptive: while all sub-type of costs are related to the everyday *buying* concept, opportunity costs are not. Hence, they eventually have to be compiled in a file of their own.

This sketch of a linguistic analysis shows that different texts present what can be assumed to be the same topic in different ways. These different ways can be made precise in terms of a dynamic update semantics (a closely related, cognitive model of text meaning has been developed by Asher, 1993), which then can be used as a model of learning (Lücking, 2019)[6]. Semantic updates are equivalent to changes in a knowledge base, which characterizes (declarative) learning. We have seen three types of changes or updates. In general the following types of changes can be distinguished (see Chi and Ohlsson, 2005 for details): *larger size*, *denser connectedness*, *increased consistency*, *finer grain of representation*, *greater complexity*, *higher level of abstraction*, and *shifted vantage point*. Acquiring threshold concepts from the *discipline* category essentially involves *denser connectedness* changes, where acquiring those of category *procedural* rest on a *higher level of abstraction* or even a *shifted vantage point*. Now there is no large-scale implementation of construction algorithms leading to semantic representations as studied in theoretical linguistics, nor is there a construction algorithm for further operations on mental files. For that reason, current computational linguistics employs shallow processing methods that aim at approximating such representations (cf. subsection 4.1.3). An approach to applying computational linguistics methods to texts in order to derive networks of threshold concept expressions is developed in section 5.

Linguistic semantics (and pragmatics, for that matter) studies the *normative* dimension of meaning: the interpretation of words and sentences of a language that any speaker should get if he or she is a speaker of that language. This does not guarantee that the speaker *actually* or *de facto* gets the normative interpretation; nor does it follow that the normative interpretation exhausts the speaker's understanding. So let us first elaborate on this and related issues to avoid any possibility of confusion.

---

[6]It should be emphasized that we just informally sketched how to derive conceptual discourse *representations*, which are just the first step of semantic interpretation. The second step consists in interpreting these representations in models. The propositional meaning of a discourse representation is the set of input-output assignments that provide a successful embedding in a model: its *context change potential*. Construing a learner as a model (knowledge base), it is suggestive to define conceptual change in terms of context change potential change.

4. THRESHOLD CONCEPTS: MENTAL, REFERENTIAL, AND DIFFERENTIAL MEANING

As outlined in section 2, threshold concepts from the disciplines of business and economics can be approached from various perspectives: they are defined as specialized terms, they are building blocks of students' learning development and they are expressed by words. Each of these perspectives corresponds to different scientific (sub-)disciplines (namely business and economics, learning psychology and education, and linguistics and lexicography, in that order; for a related view see Lenci, 2008). But how are they related?

4.1. Different Concepts of "Threshold Concepts"

According to a widely accepted sign-based conception, a word is a couple of a *form* (hereafter also called *expression*) and a *meaning*. The form side can be a token, an inflected morpho-syntactic expression of a type (lemma), or it can be the lemma itself. With respect to the meaning side, any scholar dealing with meaning faces a dilemma: she has to use meaningful words in order to describe the meaning of words (cf. Neurath, 1932). In order to avoid vicious circles, a distinction between *metalanguage* (the language used to describe meanings) and *object language* (the language whose meanings are described) is to adhered to (cf. subsubsection 4.1.1). The basic idea is that the metalanguage provides an interpreted descriptive framework according to which meanings (of the object language) can be specified. In fact, there are (good) reasons to assume that such an approach cannot be circumvented—the irreducibility of language principle (cf. either Wittgenstein, 1984 for a usage-based view or Hjelmslev, 1961 for a structuralist view of this argument).

Now one can think that the meanings of words *are* concepts. However, the concept a speaker associates with a word includes private episodes. Such private episodes do not belong to the shared (i.e., *normative*) lexical meanings of words. Accordingly, we also distinguish between the (idealized) lexical meaning of a threshold concept expression and (a student's) concept of it (subsubsection 4.1.2).

But one can just look up the meaning of a word in a dictionary, can't one? Although there is a kernel of truth in it, dictionaries completely avail themselves on the meanings of the object language of the dictionary; in other words, dictionaries contain *paraphrases* of meanings (subsection 4.1.3).

4.1.1. Lexical Meanings

The term *meaning* applies to various relations, as pointed out by means of the examples (3a–c) by Murphy (2010, p. 30):

(3) a. *Happiness* means "the state of being happy."

    b. Happiness means never having to frown.

    c. *Glädje* means *happiness* in Swedish.

    d. By *happiness* Peter means *ecstasy*.

In (3) only the first example (3a) involves lexical meaning. In (3b) a consequence relation is expressed and in (3c) a translation relation. (3d) finally is a about speaker meaning (Linsky, 1971). Speaker meaning is usually conceived as pragmatic while lexical meaning is semantic ("Speaker's Reference and Semantic Reference," re-published in Kripke, 2011).

Besides lexical meaning there is *compositional meaning* (which for instance accounts for the ambiguity within a simple sentence, such as *every dog chased a cat*, which as a relational (a single cat is chased) and a dependent (there are as many cats as dogs, that is, a plural interpretation of the singular noun phrase *a cat*) reading; see e.g., Zeevat, 2018).

Lexical meaning has to be distinguished into *sense* and *denotation* (this distinction goes back to Frege's, 1892)[7]. The denotation relation gives rise to the phenomenon that natural language expressions *are about* something in the first place. The denotation of a word is the set of things (potentially) "picked out" by that word. In lexical semantics, senses are directly represented in terms of semantic components (see Jackendoff, 1983, 1991, 2002; Pustejovsky, 1995; Wierzbicka, 1996). We know, however, of no lexical semantic analysis of threshold concept. Thus, describing the meaning of threshold concept expressions in terms of a (existing or specifically developed) metalanguage and their interactions w.r.t. to compositionality and inference could be a desideratum for further studies.

So far, meanings have been ascribed to both words and thoughts. The tension is resolved when considering that senses are *types*, that is, abstract properties which have a normative (and therefore also coordinative) dimension (this issue will be briefly taken up in subsubsection 4.1.2). These sense types are tokened in thoughts of individuals. Accordingly, in cognitive sciences concepts are construed as "temporary constructions in working memory" (Barsalou, 1993, p. 34). Each speaker instantiating a lexical sense instantiates his or her *perspective* or *understanding* of the lexical sense, or indexed concept.
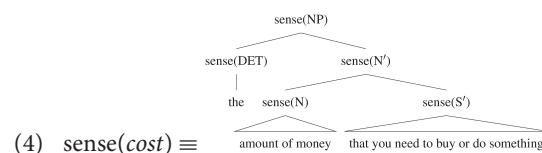
4.1.2. Indexed Concepts

A *concept* is a psychological entity, namely a mental representation and therefore a property of an individual. A concept in the sense of the threshold concept approach (cf. section 2) integrates a disciplinary perspective—an normative description of an economic fact or a principle identified by experts—with the individual perspective—the individual mental representations that the learner associates with a fact—within learning, the individual perspective matches the disciplinary one (Sender, 2017). This means that (i) concepts are not directly observable (they can be evinced by learning assessments or (neuro-)psychological testing, however); (ii) concepts are charged with individual-specific content (which partly accounts for individual-specific understanding); (iii) that concepts are the place where learning takes place.

Now speakers have knowledge about the meaning of lexical items; that is, part of speakers' lexicalized concepts is *their understanding* of the sense of an expression—this is also one of

the hallmarks of Cognitive Grammar (Langacker, 2013, p. 29)[8]. Hence, the senses identified and modeled in lexical semantics are idealizations; these senses are only realized in meaning-making minds[9]. Thus, when we talk about *the* meaning or *the* concept of an expression, we rely on an idealization, namely the assumption that we share meanings and have a common understanding. Of course, this issue has not gone unnoticed. In fact, there are several genealogical reasons that prevent a "conceptual solipsism." These include: *coordination* (Lewis, 1969; meanings get coordinated between communities of language users *via* situation of language use), and *evolution* (Millikan, 1984; meanings have a historic yet normative force acquired as biological functions in evolutionary processes). Following a semiotic variant of the principle of methodological individualism (Keller, 1995), socially accepted concepts have to be explained in terms of individual concepts (further examples are known from social ontologies; Searle, 2006). Following the advice of Klein and Kracht (2014, p. 304), namely "the more we talk to each other, the easier it gets, and the more we can come to understand each other," natural language dialog is the best way for securing mutual understanding. Such an approach is actually pursued in learning studies, where, e.g., classroom interactions are observed. In particular non-verbal behavior of the learners provide evidence on their conceptualizations (Cook and Goldin-Meadow, 2006), in line with the dictum that, for instance, manual gestures are "postcards from the mind" (de Ruiter, 2007).

4.1.3. Dictionary Concepts

While lexical semantics is a useful tool for linguistic analyses of word meanings (cf. subsubsection 4.1.1), it is less useful for everyday use and computational applications. After all, when one wants to know what a word means, one looks it up in a dictionary. According to the British English Online Dictionary[10], the meaning of *cost* is "the amount of money that you need to buy or do something." In contrast to lexical semantics, a dictionary describes object language terms in terms of object language terms[11]. The sketch of meanings from subsubsection 4.1.1 suffices in order to make more precise what claim a dictionary entry makes.

(4) $sense(cost) \equiv$

---

[7]This pair of kinds of meanings are often translated as *sense* and *reference*. However, since most semanticists would agree that reference is a pragmatic notion (Searle, 1969; Roberts, 2019), we reserve it for that purpose.

[8]Despite claims that concepts and meanings are complementary contents (e.g., Barsalou et al., 1993). Note further that according to Cognitive Grammar "meanings are in the minds of the speakers who produce and understand the expressions" (Langacker, 2013, p. 27). Obviously this claim can only be made because Cognitive Grammar lacks a notion of denotation, leaving it with the identity problem of conceptual content.

[9]There are positions that postulate an objective existence of senses, though—Frege's (1892) "third realm" is a classic example.

[10]https://dictionary.cambridge.org/ (accessed May 14, 2020).

[11]Murphy (2010, p. 34) is very explicit: "Such paraphrases, also called glosses, are indicated in single quotation marks. One must keep in mind, however, that these glosses are not themselves the meanings of the words (as they are represented in our minds)—they are descriptions of the meanings of the words."

The lexical meaning of *cost* is the sense of the syntactic parse (compositional meaning) of the gloss. The reader learns the meaning of *cost*, if he or she knows sense(NP). Furthermore, in order to derive sense(NP) not only the lexical meanings but also the compositional meanings have to be computed (cf. also section 3). In order to avoid this, a further simplification can be made by abstracting away from compositional meanings. Now the lexical meaning of *cost* is related (but not equivalent any more) to the lexical meanings of the content words from the gloss, as in (5)

(5) sense(*cost*) is related to sense(*amount*), sense(*money*), sense(*need*), sense(*buy*), sense(*do*), and sense(*something*)

Interestingly, for the dictionary user (5) is nearly as helpful as (4). Most notably, however, dictionary concepts give rise to a notion of *context* of a learning media (cf. Braun et al., 2014): the context in (5) is just the collection of expressions of the dictionary gloss. But in general a context can be any stretch of text from a few words to entire corpora or online resources. Given a context of expressions (dictionary entry, corpus, …), the expressions are transferred into a claim about their senses, as is made precise in (5). What happens here is that a statement about meanings is given in purely relational manner in terms of the object language—just like in a dictionary paraphrase. That is, (5) exemplifies the scheme of a *differential* rather than *referential* approach to word meaning (Sahlgren, 2008)[12]. Ultimately based on word frequency measures within text corpora, the relata of an expression can also be assigned different strengths by means of vector-valued word representations (Spärck Jones, 1972; Mikolov et al., 2013; Levy et al., 2015)—reflecting their respective "importance." Now dictionary concepts have a further property which is useful for present purposes: for any two non-identical contexts c_1 and c_2, the dictionary concept of a random expression will differ with respect to c_1 and c_2. In other words, dictionary concepts are *text-bound*, and text-boundedness is a prerequisite for comparing different resources in the first place. From a learning perspective, an interpreter of a dictionary entry has to entertain an indexed concept for each of its elements—amounting to the transient nature of threshold concepts and the mental linkage emphasized in subsection 2.1.

4.1.4. Concept Expressions and the "Law of Denotation"

(Lexical) semantics discovered a couple of principles and generalizations. The most important one for current purposes is what Murphy (2010, p. 36) calls the *Law of Denotation* (LoD): the "bigger" a word's sense (i.e., the more conditions that it places on what counts as a referent for that word), the smaller its extension will be. There are several phenomena to which this principle applies. For instance, the hypernym–hyponym relation fulfills the law of denotation, as does compounding. A broader term like *dog* has less lexical meaning components than a narrower term like

dachshund[13]. Since the modifying noun of a nominal compound adds its meaning in some way or other to the head noun, the law of denotation is trivially fulfilled.

Since every expression is bound up with a sense[14], larger constituents are necessarily accumulative (in fact, compositional). Now assuming expressions, sentences or discourses to be coherent (a notion on which see Asher and Lascarides, 2003, p. 21, and various other places and Ginzburg, 2012, p. 208), this gives rise to the simple but useful generalization: *the more expressions, the more elaborate the combined sense* (where "combined" is intended to cover both compositional derivation as well as accumulation).

The relation between senses and denotations is regimented by LoD. It applies likewise to words, phrases and sentences. The more fine-grained the senses of these constituents, the more detailed are their denotations. The connection to sciences and the language of sciences is obvious: (natural) sciences aim at precise descriptions of the world. That is, scientific languages are about very detailed denotations. In order to achieve this level of detail, guided by LoD, the expressions of the specialized vocabularies need to have elaborate senses, which, by dint of compositional meanings, get even more specific in phrases and sentences. Since natural languages are devices of ontology construction, as has been pointed out by some versions of semantics (e.g., Barwise and Perry, 1983), it is also possible to "postulate new denotations," so to speak, as has famously been done in the history of physics several times, for instance. LoD and *making things precise* has repercussions to linguistic expressions. Against this backdrop, we discuss observable features of expressions of threshold concepts in the following.

4.2. Linguistic Features

Following the guideline that threshold concepts are instances of specialized vocabularies, we expect their expressions to exhibit features which will be described in more detail in the following: (1) compounding potential, (2) large nominal groups, and (3) web of threshold expressions.

1. **Compounding potential**. Of how many compounds is an expression a part? The compounding potential is a long-known feature of specialized vocabulary where specialized languages are characterized by a large number of compounds (Widdowson, 1974). It has also been highlighted by business and economics studies on threshold concepts (e.g., Meyer and Land, 2006). A large number of (compound) nouns is also confirmed in the textbook study by Hu and Gao (2019). In light of the above-mentioned specificity demand of languages of science, this feature is expected. But why are compounds semantically specific and distinguish themselves from *prima vista* synonymous syntactic realizations? Most

---

[12]This line of thought is rooted in structuralism (de Saussure, 1916; Hjelmslev, 1961).

[13]In this case one must of course know that *dachshund* is a hyponym of the hypernym *dog*. According to dictionary approaches, such knowledge is part of the speaker's mental lexicon, according to conceptual semantics it is computed based on semantic componential representations.

[14]This is less clear, however, for syncategorematic expressions, such as conjuncts. However, since they do not *remove* any sense components, they do no harm to the generalization.

nominal compounds [that are compounds whose head is a noun while the modifying component may be an adjective (*green tea*), a verb (*swimming pool*), or a further noun (*football*)] are determinative, meaning that the modifying expression determines the head noun. For instance, a football is not just a ball, but a ball meant to be moved along by one's feet. But there are more interesting properties of compounds. Most importantly, a compound induces a *kind reading* (Bücking, 2010). Given this feature, we expect compounding (as a form of name-giving) to be coupled to the dynamic ontological modifications within the sciences, as is evinced by findings for specialized vocabulary (Widdowson, 1974).

If we conceive the kind-reading of compounds in relation to LoD and the specificity demands of scientific languages, a few trends can be derived:

(a) For all compounds that share the same threshold concept expression head it holds that the more modifying constituents the compound has, the more specific it is. This follows trivially from sense accumulation. For instance, both *Grenzkosten* "terminal cost" and *Marginalkosten* "marginal cost" are more specific than *Kosten* "cost."

(b) The inverse formulation of the previous item is that the more specific a given threshold concept head is, the less compounds it will show. Note that this is a recursive notion: (more) complex compounds may consist of (less) complex heads.

(c) Going from expressions to the use of these expressions in sentences and texts it is very likely that the more compounds a sentence or text contains, the more specific the sentence or text is (see also the following linguistic feature, "large nominal groups").

These trends can directly be read off the concept expressions.

2. **Large nominal groups**. Related to the compounding potential is the elaborateness of the whole nominal group of which a concept expression (compound or not) is a part. Expressions of specialized vocabularies tend to occur in elaborate environments (Strevens, 1977). Contexts of elaborateness are constructed by adjectives and relative clauses (mainly restrictive ones). Obviously, nominal groups are more specific according to LoD. This feature is a further linguistic feature of threshold concept expressions to look for.

3. **"Web of threshold expressions."** Based on postulations of threshold concept research from subsection 2.1 and the linguistic perspective sketched in section 3, concept expressions are to be expected to be related to each other, i.e., forming a "web" of threshold expressions (Davies and Mangan, 2007). Thus, in terms of subsubsection 4.1.4 we can make the claim more precise in saying that the web of threshold concepts is a context of weighted expressions where the context consists exclusively of threshold concepts. Now the different contexts under consideration (textbooks, newspaper, Wikipedia) trivially give rise to different dictionary concepts. However, since the different contexts are an independent variable, differences can point at meaningful differences in the independent variable (i.e., contexts). Further support

for this claim comes from qualitative investigations of specialized vocabularies, where the context is accredited to be most important feature of special terms (Vaňková, 2018). From that we can derive the expectation that the web of threshold concepts is "stronger woven" in formal than in informal contexts.

5. METHODS

5.1. Guiding Questions

From subsection 2.2 we take the assumption that resources from formal learning environments are more specific than resources from informal learning environments, since formal environments are characterized by special vocabularies, among others. What tends to be more specific in its use, however, will also form more specific associations with similarly used units: threshold concepts should therefore be more strongly associated with each other if they tend to be used together in specific and equally rare contexts. In this way of thinking, specificity and associative strength seem to be two related concepts that help to compare the use of threshold concepts in different corpora. Thus, it is reasonable to operationalize the above introduced linguistics of threshold concepts by *quantifying* their specificity properties and association relations: the former will be carried out by means of a classical distribution analysis using appropriately quantified specificity values; the latter will be performed by means of a network analysis in which threshold concepts are the nodes whose association relations are interpreted as node connections or links, weighted by the strengths of these associations. In this way we gain access to two types of information: a node-related one (*specificity*) and a link-related one (*association strength*). This enables us to explore both sources of information independently as well as simultaneously using a unified, network-based representation format. However, let us first look at which guiding questions can be formulated either node- or link-related in more general terms[15]:

1. **Q1:** *Do formal corpora show "longer" compounds (i.e., words composed by two or more other words) than informal ones, that is, do formal corpora have more modifying constituents for a given threshold expression head?*
2. **Q2:** *Are there more compounds with threshold concepts (whether as head or not) in formal corpora than in informal ones?*
3. **Q3:** *Are threshold concepts within formal corpora part of larger nominal groups than in informal corpora?*
4. **Q4:** *Does the "web of threshold concepts" derived from formal corpora give rise to a stronger connected threshold concept context than the one derived from informal corpora?*

More formally speaking, the questions Q1–Q3 are all node-related: by operationalizing answers to these questions, we quantify the specificity of threshold concepts in the underlying

---

[15]Here we focus on threshold concepts within formal and informal learning contexts. For an assessment of the three classes of threshold concepts—basic, discipline, modeling—see the study of Brückner and Lücking (2019).

corpora. Question Q4 is link-related: this question addresses the association strengths in networks of threshold concepts. In any event, according to the current state of our explanations, questions Q1–Q4 are formulated too unspecifically: what does it mean to be connected, for example, and how should this be numerically weighted? In other words, Q1–Q4 cannot yet be tested by means of an exact measurement procedure. To ensure this, we must first translate them into a formal language: in our case this is network theory. This will also mean that we consider variants of selected hypotheses addressing these questions. Ultimately, this approach serves to precisely measure the two core hypotheses about the greater specificity and stronger associativity of threshold concepts in textbooks. To this end, 63 threshold concepts (see **Appendix A**) are compared across several corpora where the textbook corpus consists of the textbooks listed in **Appendix B**. This comparison is based on the measurement procedure described in the next section.

5.2. A Two-Part Procedure for Measuring the Use of Threshold Concepts

To tackle the guiding questions Q2 and Q4, we develop a two-part procedure to measure significant differences in the use of threshold concepts. Our first aim is to quantify the difference in the specificity of uses of threshold concepts. In order to operationalize this notion, we start from the following assumptions:

- The more often a threshold concept x manifests itself as a component in compounds and the higher the frequencies of these compounds in corpus C, the higher the *degree of specification* of x and thus its use in C. We call this sort of specificity *compounding-related specificity* or just *compounding-specificity* of x in C. Furthermore, the more frequently the concept occurs in C as a whole, the higher its polytextuality in the sense of Köhler (1986) (i.e., the higher the number of sentences by which it is semantically specified), the higher its degree of specification. We call this sort of specificity *sentence-related specificity* or just *sentence-specificity*. Finally, the higher the number of threshold concepts with the higher degrees of compounding- or sentence-specificity, the higher the overall specificity of this set of concepts in the underlying corpus.
- The more compounding- or sentence-specific the use of a threshold concept in a corpus, the more detailed and differentiated knowledge can be acquired about this concept by reading texts of this corpus (i.e., the larger the context of the dictionary concept of the threshold concept expression in question).

Starting from these considerations we arrive at the following hypothesis about the difference between formal and informal language corpora (manifesting formal and informal learning contexts) in terms of the compounding- and sentence-specificity with which they manifest threshold concepts:

H1: *The use of threshold concepts in formal language corpora is more compounding- or sentence-specific than in informal language corpora.*

Our second aim is to quantify the differences in the associative networks of threshold concepts as induced by corpora of three different genres, i.e., of press communication, encyclopedic communication, and technical communication. From subsection 2.2 we know that newspapers are an example for informal learning contexts, whereas textbooks make up formal contexts. Since to our knowledge there is no *linguistic* judgment of Wikipedia in this respect yet, we remain neutral and will see how Wikipedia compares to formal and informal resources used in the following. For this purpose, we start from the following consideration:

- The greater the differences in the ways threshold concepts are used in two corpora, the more different the associative relations that can be learned as a result of reading homogeneous subsets of texts of these corpora.

By a *homogeneous subset* we mean a set of texts sampled from the same corpus. It should be noted that we do not directly observe the acquisition of semantic associations between threshold concepts. Rather, this acquisition will be estimated by means of word embeddings (Mikolov et al., 2013). The embeddings are compared for the purpose of measuring the semantic associations of the embedded concepts, in the sense of the *Weak Contextual Hypothesis* (WCH) of Miller and Charles (1991): Words that tend to be used in similar contexts are then regarded as semantically similar and correspondingly more strongly associated. That is, if a corpus exhibits such contextual similarities, reading subsets of texts from that corpus makes the acquisition of corresponding syntagmatic or paradigmatic associations, as we assume, more likely. Thus, if the semantic associations of a corpus deviate significantly from those that can be expected, for example, from a thematically similar corpus of textbooks, this may have negative consequences for the acquisition of the concepts concerned. Even if we do not investigate this consequence ourselves, we at least measure the previously mentioned similarity or dissimilarity of association networks. These considerations are a prerequisite for operationalizing the falsification of the following hypothesis about the difference between formal and informal language corpora in terms of the semantic networking of threshold concepts:

H2: *Due to their usage contexts in formal language corpora, threshold concepts are more strongly associated than due to their usage in informal language corpora.*

By falsifying the alternative hypotheses of H1 and H2, we obtain evidence that the threshold concepts we are looking at are used significantly differently in the genres under consideration, insofar as their uses correspond to different degrees of specificity (a), while spanning different semantic networks (b). However, what differs in two ways, in that it induces the acquisition of concepts of different specificity (*node-related*) and different associations (*edge-related*), ultimately represents a different learning basis or learning context. From this point of view, it becomes clear that

we understand the structure induced by threshold concepts as a *network of concept nodes and their association relations*, whose "shape" depends on what is said about them in the underlying corpus or how they are specified by means of compounding. More precisely, let $T = \{a_1, \ldots, a_n\}$ be a set of threshold concepts and $C = \{x_1, \ldots, x_m\}$ a text corpus. Then, we denote by

$$C(T) = (V, E, \mu, \nu, \lambda) \qquad (1)$$

the *Threshold Concept Network* (TCN) induced by C over T where $E \subseteq V^2$, $\mu: V \rightarrow \mathbb{R}_0^+$ is a function measuring the specificity $\mu(v)$ of each $v \in V \subseteq T$ in C, $\nu: E \rightarrow \mathbb{R}$ is a function measuring the semantic association $\nu(\{v, w\})$ between v and w for each $\{v, w\} \in E$ and $\lambda: V \rightarrow T$ is an injective vertex labeling function. More specifically, $\nu(\{v, w\})$ equals the cosine similarity of the embedding vectors computed for v and w, respectively, by the operative embedding method that is used to explore C.

Let $C_i(T) = (V_i, E_i, \mu_i, \nu_i, \lambda_i)$ and $C_j(T) = (V_j, E_j, \mu_j, \nu_j, \lambda_j)$ be two TCNs induced by the corpora C_i and C_j. For any pair of vertices $v \in V_i, w \in V_j$, for which $\lambda_i(v) = \lambda_j(w)$, we will write $\dot{v} = \dot{w}$. To operationalize the falsification of H1 and H2, we now specify the functions μ and ν in more detail:

- *On μ and H1:* We consider a simple frequency-related definition of μ, according to which $\mu(v)$ corresponds to the number of tokens of the lemma v in C plus the number of occurrences of compounds in C that contain v as a component (*compounding- + sentence-specificity*). A first variant of μ, denoted by μ', considers only the former number (*compounding-specificity*), a second, denoted by μ'', only the latter number (*sentence-specificity*). Let μ be any of these variants, then we derive the following rank-frequency distribution

$$\mu(V) = ((v_{i_1}, \mu(v_{i_1})), \ldots, (v_{i_n}, \mu(v_{i_n})), \ \mu(v_{i_1}) \geq \ldots \geq$$
$$\mu(v_{i_n}), \ v_{i_1}, \ldots, v_{i_n} \in V \quad (2)$$

for which we compute the exponent α of the power law that best fits this rank distribution. In this way, we test the skewness of the distribution of the specificities of threshold concepts as induced by C: the higher the value of α, the faster the frequency-related transition from high-rank (frequent or highly specified) to low-rank (rare or rarely specified) concepts; note that we always consider small numbers of concepts for the distributions, so the slope cannot be the result of a larger number of rare concepts and especially *hapax legomena*. The alternative to H1 is now considered falsified if the corpus length-normalized rank specificity distribution of formal language corpora is above that of informal language ones, under the condition of a Zipfian, power law-like character of such distributions as normally observed for word frequency distributions (Zipf, 1949; Tuldava, 1998) and also assumed for threshold concepts. Beyond that, we assume that power laws better fit the use of threshold concepts in textbook corpora or in formal language corpora in general than in informal language corpora (e.g., of press communication). Furthermore, we assume that the rank specificity distributions

of formal language corpora differ significantly from those obtained for informal language corpora. Finally, we assume that the rank correlation between the rank specificity distributions of formal and informal language corpora is lower than in cases where the corpora manifest either both formal or informal language—provided that these corpora are all sufficiently similar thematically. If we succeed in falsifying the alternative to H1 in these senses, we get the information that formal language contributes to the development of more specific threshold concepts, the specificity distribution of which follows a Zipfian distribution in a more pronounced and significantly different way compared to corpora of informal language, that the specificity of the concepts in the latter corpora tends to be lower, and that, finally, thematically and formally similar corpora are more similar to each other than corpora of different formality.

- *On ν and H2:* The association strength of TCNs in relation to the degree of formality of the underlying corpus will be measured using methods of network theory (Newman, 2010) and especially of the theory of linguistic networks (Mehler et al., 2020a). More specifically, we test H2 by quantifying the densities of TCNs derived from different corpora using the approach of Mehler et al. (2020b). That is, we utilize the notion of α-cuts, as introduced in the description of fuzzy sets, and apply it to weighted graphs as follows: let $C(T) = (V, E, \mu, \nu, \lambda)$ be a TCN. Then we define:

$$a(C(T)) = (\alpha_1, \ldots, \alpha_l)^T \qquad (3)$$
$$\alpha_1 = \min\{s(\nu(e)) \mid e \in E\} \qquad (4)$$
$$\forall k \in \{2, \ldots, l\}: \alpha_k = \min\{s(\nu(e)) \mid s(\nu(e)) > \alpha_{k-1}\} \quad (5)$$
$$\forall e \in E: s(\nu(e)) = \frac{\nu(e) - \text{Min}}{\text{Max} - \text{Min}} \in [0, 1] \quad (6)$$

where Max (Min) is the theoretical maximum (minimum) that ν can assume. Then we define the α-cut of $C(T) = (V, E, \mu, \nu, \lambda)$, that is, the so-called *alα-cut graph* $C(T, \alpha) = (V, E|_\alpha, \mu|_\alpha, \nu|_\alpha, \lambda|_\alpha)$ where

$$E|_\alpha = \{e \in E \mid \nu(e) \geq \alpha\} \qquad (7)$$

and $\mu|_\alpha, \lambda|_\alpha$ are the restrictions of μ, λ to the vertex set induced by $E|_\alpha$ and where $\nu|_\alpha: E|_\alpha \, to [0, 1], \forall e \in E|_\alpha: \nu|_\alpha(e) = s(\nu(e))$. This allows us to define the graph series

$$cuts(a(C(T))) = (C(T, \alpha_1), \ldots, C(T, \alpha_l)) \qquad (8)$$

Finally, for any graph index $\iota: \mathbb{G} \rightarrow \mathbb{R}$, we get a series of index values:

$$\iota(cuts(a(C(T)))) = (\iota(C(T, \alpha_1)), \ldots, \iota(C(T, \alpha_l))) \quad (9)$$

In this paper, we experiment with graph cohesion and graph clustering (Newman, 2010). For each of these indices, we want to know how (1) early, (2) fast, and (3) differently its values for the different series of alα-cut graphs calculated for the targeted corpora are decreasing or increasing. Now Hypothesis H2 is considered falsified if the cohesion of the series of alpha-cut

TABLE 1 | Summary of corpora used in the study.

	No. of articles	No. of token	Period of publication
SZ-Eco	288,792.000,0	85,826,410.000,0	1992–2014
SZ-All	1,707,666.000,0	630,588,082.000,0	1992–2014
WP-Eco	653,397.000,0	265,063,077.000,0	2001–2016
WP-Top-1	37,895.000,0	20,090,166.000,0	2001–2016
WP-Top-3	71,013.000,0	28,145,793.000,0	2001–2016
WP-All	1,760,875.000,0	736,071,291.000,0	2001–2016
ZEIT	184,186.000,0	179,327,441.000,0	1994–2014
TB	14.000,0 books	2,326,374.000,0	2015–2020

See main text for a description. The acronyms are resolved in section 5.3.

graphs calculated for the textbook corpus decreases later than in the case of alpha-cut graphs calculated for non-textbook corpora, and in such a way that the behaviors of these series differ significantly from each other. Further, we expect the same behavior with regard to the corresponding series of graph clustering or transitivity values.

In a nutshell: H1 is considered falsified if the alternative hypotheses to H1 and H2 are falsified. If such a double falsification succeeds, we obtain evidence that formal language corpora support the development of more strongly specified threshold concepts that are at the same time more strongly associated with each other or semantically networked. According to our guiding idea, such an observation is linked to the assumption that reading formal language corpora facilitates the acquisition of threshold concepts according to the associated learning objective.

5.3. Data and Pre-processing

We consider corpora from press communication, encyclopedic communication and technical communication (see **Table 1**):
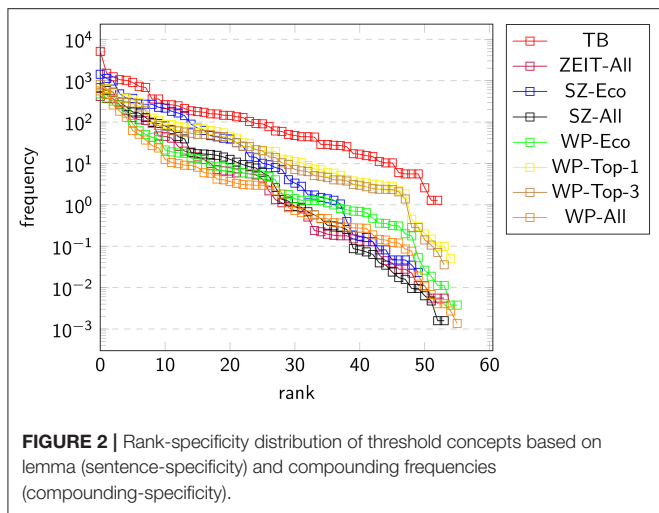
1. *Corpus SZ-Eco:* as an informal language corpus of texts about economics, we process 288,792.000,0 texts from the *Süddeutsche Zeitung* (SZ), one of the largest daily German newspapers, all of which belong to the register *Wirtschaft* [*economics*]—see **Table 1** for the corpus statistics.
2. *Corpus SZ-All:* SZ-Eco is contrasted with SZ-All, that is, the corpus of al 1,707,666.000,0 articles of SZ published in the years 1992 to 2014 (see **Table 1**). In this way we get access to the usage regularities of threshold concepts in arbitrary press articles of whatever topic.
3. *Corpus WP-Top-1:* as a formal language corpus of texts on economics, we determine the subset of all Wikipedia articles whose top-level topic category corresponds to the *Dewey Decimal Classification* (DDC) *Category 330* (*Economics*). In other words, we DDC-categorize all Wikipedia articles of the German Wikipedia using text2ddc (Uslu et al., 2019) and select those articles whose top-level topic category corresponds to DDC category 330. In this way, we obtain a subset of Wikipedia articles that can be very reliably assigned to our target topic of economics: anyone who reads articles of the Wikipedia article network, which is spanned by these articles, navigates, so to speak, in the thematically homogeneous area of economically relevant articles.

4. *Corpus WP-Top-3:* in analogy to WP-Top-1, WP-Top-3 is the set of all German Wikipedia articles where the DDC category 330 is among the first three DDC categories assigned to this article by text2ddc with a membership value of at least 10%. Obviously WP-Top-3 contains larger parts of WP-Top-1 (10% threshold) or even this corpus as a whole, but likely also articles whose relation to economics is less confirmed, even if they do not fall below the 10% threshold.
5. *Corpus WP-Eco:* WP-Eco is the corpus of all articles in Wikipedia that are directly or indirectly assigned to the category *Wirtschaft* [*economics*] from Wikipedia's category system. WP-Eco contains 653,397.000,0 articles and thus about a third of all 1,760,875.000,0 articles of German Wikipedia; WP-Eco also contains articles that are (possibly) only (very) indirectly related to the topic of economics. Whoever reads articles from the corresponding article network navigates, so to speak, in the wider area of economics-related articles, while possibly changing the topic (starting from economics), but in a frame that still has to do with economics.
6. *Corpus WP-All:* the largest corpus we look at includes the 1,760,875.000,0 articles from the German Wikipedia, most of which are not related to economics (see **Table 1**).
7. *Corpus ZEIT:* as a second corpus of informal language of press communication, we process the 184,186.000,0 texts of the German weekly newspaper *Die Zeit* published in the years 1994–2014.
8. *Corpus TB:* Last but not least we analyze a corpus of formal language, that is, a corpus of 14 textbooks all about economics in the narrow sense (see **Appendix B**).

In total, we consider eight corpora, three of which are informal language corpora of press communication (SZ-Eco, SZ-All, ZEIT), three of which mainly comprise texts that are not related to economics (SZ-All, WP-All, ZEIT) and five of which are formal language corpora (WP-All, WP-Eco, WP-Top-1, WP-Top-3, TB). Moreover, one of the informal language corpora (SZ-Eco) and four of the formal language corpora (WP-Eco, WP-Top-1, WP-Top-3, TB) focus more or less on economics. For preprocessing all these corpora, we use TextImager (Hemati et al., 2016). That is, the corpora are tokenized, part of speech-tagged and lemmatized. Furthermore, sentences are split and tokens are segmented to identify candidate compounds, their heads and modifiers. Text classification regarding the second level of the DDC is performed by means of text2ddc (Uslu et al., 2019). Embeddings are computed for all corpora separately using word2vec based on standard settings (i.e., word vector size = 100, window size = 5, with five training iterations) for skip-gram and cbow (see Mehler et al., 2020c for a related procedure). Finally, the embeddings are used to induce TCNs according to section 5.2, which are then processed with GraphMiner, a network analysis software under development at TTLab (www.texttechnologylab.org).

5.4. Results

In **Figure 2** we show the rank-specificity distribution of our set of threshold concepts based on the variant μ of vertex weights in TCNs. It is remarkable that the specificity values of threshold concepts in textbooks are above all distributions induced by

FIGURE 2 | Rank-specificity distribution of threshold concepts based on lemma (sentence-specificity) and compounding frequencies (compounding-specificity).
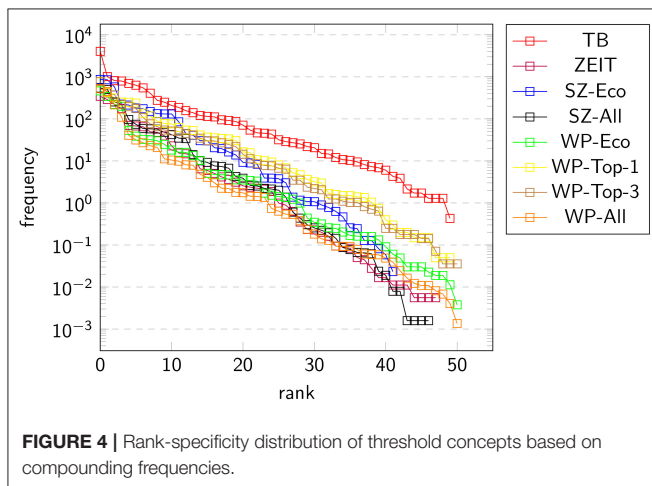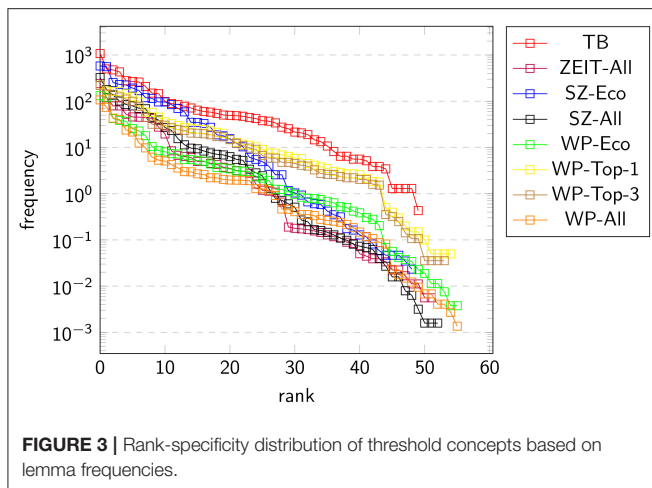
the comparison corpora. Furthermore, the specificity values for concepts from formal language corpora dedicated to economics, such as WP-Top-1 and WP-Top-3 are also higher. In contrast, specificity values from corpora of more general content (WP-All, SZ-All, ZEIT-All) do not achieve such high levels. In the middle of the spectrum of specificity distributions we observe SZ-Eco and WP-Eco, two corpora of medium size, which deal with economic issues in a larger thematic context. Note that we calculate relative frequencies in order to rule out size effects and scale the distributions (by multiplying with 1,000,000.000,0) in order to enhance readability.

In order to estimate whether the distributions actually differ from each other, we perform pairwise Kolmogorov-Smirnov goodness-of-fit tests. If the p-values of any such fit is high, then we cannot reject the hypothesis that the distributions of the two samples are the same. In other words: small p-values indicate a significant difference between two distributions. Results are collected in **Table 2**, where $p < 0.1$ is highlighted in green (likewise for **Tables 3–8** below): obviously, in most cases the distributions differ from each other. Remarkable exceptions are SZ-Eco in relation to SZ-All (the latter contains the former), WP-Top-1 and WP-Top-3 (also a matter of inclusion) and especially SZ-Eco in relation to WP-All.

The scenario observed in **Figure 2** is also displayed by **Figure 3** (sentence-specificity) and **Figure 4** (compounding-specificity): the specificity distributions are all topped by the distribution for textbooks. In this sense, it can be said that the threshold concepts considered here are most specifically described in the formal language textbook corpus, followed by the two formal language Wikipedia-based corpora WP-Top-1 and WP-Top-3 and least specifically in the informal newspaper corpora SZ-All and ZEIT-All, although in the case of compounding-specificity the situation is not so obvious. A borderline case is WP-Eco, a corpus that consists of Wikipedia articles that are directly or indirectly assigned to the thematic field of economics.

TABLE 2 | P-values of the Kolmogorov-Smirnov goodness-of-fit test applied to the pairwise combinations of the distributions in **Figure 2**.

	TB	SZ-Eco	SZ-All	WP-Eco	WP-Top-1	WP-Top-3	WP-All	Zeit-All
TB	—	0,000,5	$7.242{,}4 \times 10^{-06}$	0,001,8	0,057,8	0,041,6	0,000,3	0,006,7
SZ-Eco	—	—	0,441,1	0,583,4	0,015,4	0,018,1	0,823,3	0,441,1
SZ-All	—	—	—	0,053,1	0,001,7	0,002,3	0,219,2	0,087,4
WP-Eco	—	—	—	—	0,130,2	0,156,6	0,778,9	0,033,8
WP-Top-1	—	—	—	—	—	0,961,5	0,008,6	0,068,7
WP-Top-3	—	—	—	—	—	—	0,011,2	0,052,7
WP-All	—	—	—	—	—	—	—	0,033,8
Zeit-All	—	—	—	—	—	—	—	—

TABLE 3 | *P*-values of the Kolmogorov-Smirnov goodness-of-fit test applied to the pairwise combinations of the distributions in **Figure 3** (sentence-specificity).

	TB	SZ-Eco	SZ-All	WP-Eco	WP-Top-1	WP-Top-3	WP-All	Zeit-All
TB	—	0.000,3	$1.080,0 \times 10^{-05}$	0.000,5	0.138,4	0.043,4	0.000,1	0.007,1
SZ-Eco	—	—	0.562,7	0.184,0	0.013,2	0.013,5	0.538,8	0.233,0
SZ-All	—	—	—	0.015,0	0.000,2	0.001,5	0.075,0	0.023,2
WP-Eco	—	—	—	—	0.049,1	0.096,5	0.468,1	0.099,5
WP-Top-1	—	—	—	—	—	0.966,6	0.008,1	0.069,5
WP-Top-3	—	—	—	—	—	—	0.005,6	0.047,5
WP-All	—	—	—	—	—	—	—	0.061,9
Zeit-All	—	—	—	—	—	—	—	—

TABLE 4 | *P*-values of the Kolmogorov-Smirnov goodness-of-fit test applied to the pairwise combinations of the distributions in **Figure 4** (compounding-specificity).

	TB	SZ-Eco	SZ-All	WP-Eco	WP-Top-1	WP-Top-3	WP-All	Zeit-All
TB	—	0.010,8	0.000,7	0.039,3	0.218,4	0.180,3	0.012,2	0.066,4
SZ-Eco	—	—	0.313,6	0.693,7	0.194,1	0.243,6	0.995,6	0.597,8
SZ-All	—	—	—	0.201,5	0.022,5	0.052,7	0.650,5	0.139,3
WP-Eco	—	—	—	—	0.435,4	0.593,0	0.980,7	0.484,3
WP-Top-1	—	—	—	—	—	0.996,2	0.188,1	0.525,9
WP-Top-3	—	—	—	—	—	—	0.307,9	0.598,5
WP-All	—	—	—	—	—	—	—	0.298,5
Zeit-All	—	—	—	—	—	—	—	—

FIGURE 3 | Rank-specificity distribution of threshold concepts based on lemma frequencies.

FIGURE 4 | Rank-specificity distribution of threshold concepts based on compounding frequencies.

When we look at **Tables 3**, **4**, we get the information that while the frequency distributions (sentence-specificity) tend to be distinguishable, the distinguishability of the compounding-specificities is much less: obviously, the frequencies of compounds to which our threshold concepts belong are more independent of the underlying corpus. Moreover, the distributions in **Figures 2**–**4** tend to be all Zipfian: although a lognormal distribution is also a good fit in 17 (of 24) cases, power law fitting is still a valid option (there is not a single significant p-value < 0.05 for any $R < 0$; note further that a lognormal distribution is a heavy-tailed distribution, too): the exponent α ranges from ≈ 1.3 to ≈ 2.8, where the minimum x value of the fit is given as "x-min" (see **Table 5**)[16].

From this perspective, we see the alternative of hypothesis H1a, which states that the use of threshold concepts in formal

---

[16]We apply the toolbox of Alstott et al. (2014) according to Clauset et al. (2009): Power laws (first) are compared to lognormal distributions (second): "R is the log likelihood ratio between the two candidate distributions. This number will be positive if the data is more likely in the first distribution, and negative if the data is more likely in the second distribution. The significance value for that direction is p." (Alstott et al., 2014, p. 5).

language corpora is neither more compounding-specific nor more sentence-specific than in informal language corpora, as being falsified.

Next we consider Hypothesis H1b. For this purpose, we compare the series of cohesion values induced by the series of alpha-cut graphs (see above) based on our eight different corpora. We start with exemplifying alpha-cut graphs based on three different corpora using the same set of threshold concepts and cutting for the same $\alpha = 0.7$: corpus SZ-All (**Figure 5**), corpus TB (**Figure 6**), and corpus WP-Eco (**Figure 7**). These graphs, which are all based on the same vertex set, illustrate a networking effect that is later confirmed by our analysis of the entire time series of alpha-cut graphs: Wikipedia-based corpora exhibit the densest networking, followed by textbook corpora and newspaper corpora. Threshold concepts associate more strongly and more often in the case of the former compared to the latter. Moreover, in the case of the newspaper corpus, the number of network components is highest (so that the number of isolated nodes is also highest), while in the case of the textbook corpus there is a unique dominant vertex (*costs/Kosten*) in terms of compounding- and sentence-specificity. But what exactly does the network density look like when we look at the entire time series of these alpha-cut graphs? **Figure 8** shows the corresponding distributions starting from the TCNs derived from word embedding similarities based on the skip-gram model of word2vec and thus for syntagmatic associations (starting from the respective seed word to the probable context in the sense of being defined by neighboring words). Very remarkably, all four Wikipedia corpora behave very alike: the cohesion values of the TCN series induced by these corpora decrease at the latest compared to all other corpora and their corresponding TCN series, i.e., they decrease for the comparatively highest α values. Conversely, the cohesion values of the corresponding TCN series induced by the newspaper corpora (SZ-All, ZEIT-All) decrease the fastest. In the middle of this spectrum we surprisingly observe two series of cohesion values: that for the textbook corpus and that for the economics-related SZ-Eco corpus, though rather in the neighborhood of the Wikipedia corpora than in the one of the newspaper corpora. At this point, we have to ask whether the distributions shown in **Figure 8** are actually different or not. For this purpose we again perform Kolmogorov-Smirnov tests of goodness-of-fit, but now separately for both axes from **Figure 8**. The reason is that neither axis is ordinal scaled, so we first perform a corresponding scaling before we can compare the corresponding feature distributions. As shown in **Table 6**, we get a mixed result: while the alpha-cuts of the individual distributions increase very differently (so that the distributions are mostly clearly distinguishable from each other), this does not apply to the decreases in cohesion values caused by the increasing alpha-cuts: here the distributions are all indistinguishable. For the distributions of the cohesion values this means that they are in fact all almost "identical" and therefore indistinguishable mirrored S-curves when being scaled appropriately.

From this spectrum of distributions, we get the following assessment: in Wikipedia-based corpora, the threshold concepts are most strongly associated with each other—metaphorically
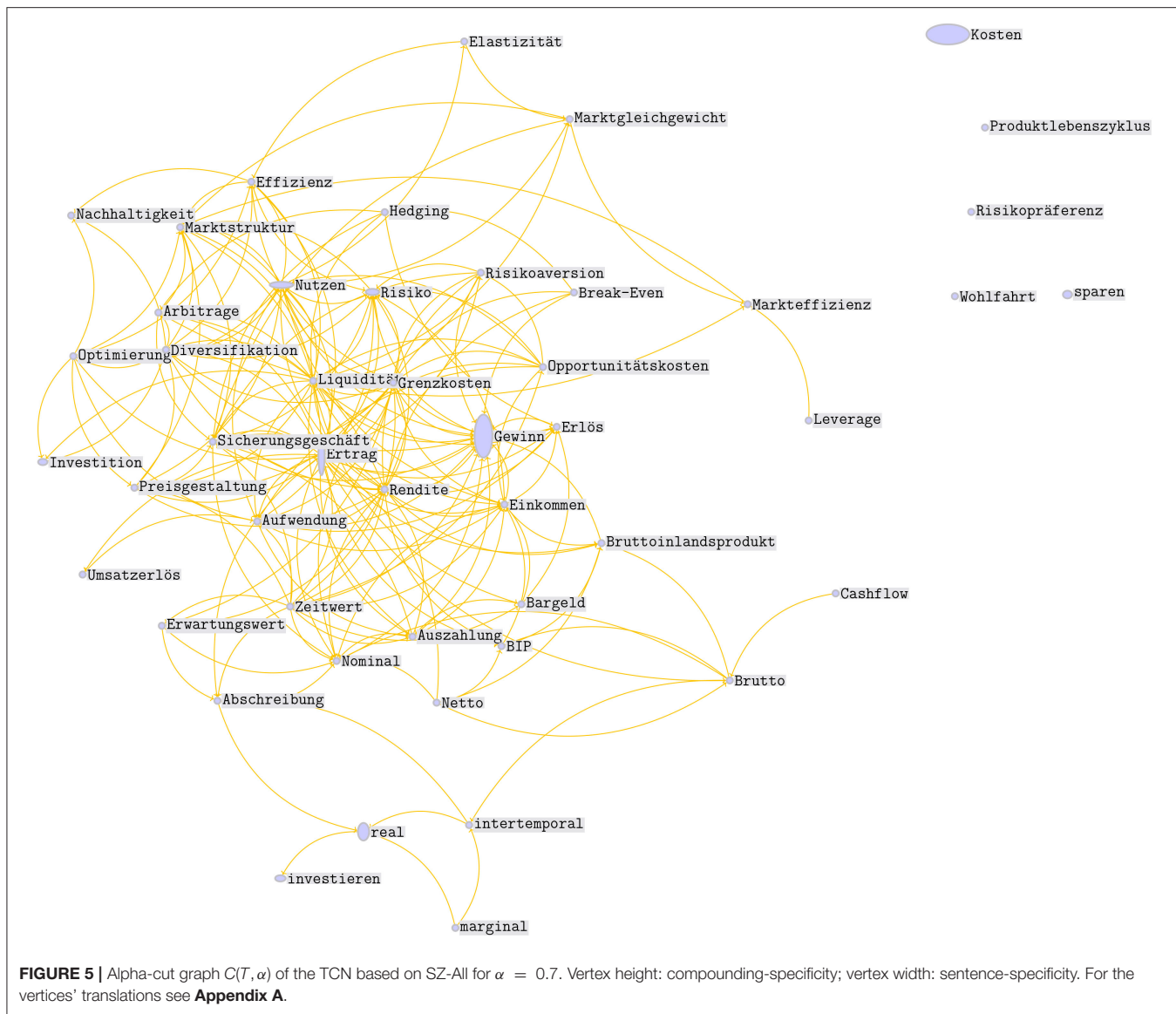
TABLE 5 | Power law goodness-of-fit tests for the distributions from **Figures 2–4**.

	Alpha	x-min	R	P
Lemma and compound, **Figure 2**				
SZ-All	1.711629	4.0	0.004996	0.737173
SZ-Eco	1.565592	2.0	−0.686457	0.305550
TB	2.825375	41.0	0.009203	0.853003
WP-All	1.593086	3.0	0.009178	0.719042
WP-Eco	1.613395	3.0	−0.568429	0.322225
WP-Top-1	1.522315	3.0	−0.002717	0.793323
WP-Top-3	1.485334	2.0	−0.242887	0.866997
Zeit-All	1.395048	1.0	−0.666070	0.403735
Only lemma, **Figure 3**				
SZ-All	1.735390	5.0	−0.126319	0.294422
SZ-Eco	1.548944	2.0	−0.689713	0.307030
TB	1.336778	1.0	−0.627145	0.792154
WP-All	1.593086	3.0	0.009178	0.719042
WP-Eco	1.483664	2.0	0.000933	0.964148
WP-Top-1	1.433428	2.0	−0.173451	0.273327
WP-Top-3	1.546151	4.0	0.015474	0.301308
Zeit-All	1.395048	1.0	−0.666070	0.403735
Only compound, **Figure 4**				
SZ-All	1.355914	1.0	−0.666272	0.268328
SZ-Eco	1.559961	3.0	0.028074	0.235693
TB	1.395212	2.0	−0.683652	0.307499
WP-All	1.395048	1.0	−0.666070	0.403735
WP-Eco	1.676775	4.0	−0.580707	0.369533
WP-Top-1	1.545465	3.0	−0.641716	0.320173
WP-Top-3	1.464465	2.0	−0.649692	0.317607
Zeit-All	1.418524	2.0	−0.080603	0.606211

speaking, they form a denser network of particles that are located much closer to each other. For much higher values than for any other corpus, the network cohesion (starting from a completely connected graph) takes a maximum value of 1; and for equally maximum values the cohesion is at least 50, 75%, etc.: the deletion of lower weighted edges in TCNs based on Wikipedia corpora is therefore more likely to lead to more cohesive networks compared to the other TCNs. In view of this finding, the textbook-based TCNs are surprisingly less cohesive. *Based on our cognitive model, this suggests that reading such textbooks makes stronger syntagmatic associations under threshold concepts less likely.* Wikipedia seems to write more densely about these concepts, in a way that makes their associations more probable and also more pronounced. This may be related to the text type of Wikipedia (*encyclopedic communication*) as opposed to textbooks, which may also contain longer motivational, exemplary or elaborating text passages. In any case, however, we see the hypothesis confirmed that formal language corpora make stronger associations between threshold concepts more likely than informal language corpora—this is

indirectly confirmed by the values of **Table 6** regarding the *x*-axis (formal language corpora are significantly "shifted" to the right compared to their newspaper-based counterparts, i.e., SZ-All and ZEIT-All). An extreme-value-forming special position of textbooks, however, cannot be confirmed. Moreover, the strengths of the associations of threshold concepts obtained by means of informal texts on topics related to economics (SZ-Eco) can hardly be distinguished from those obtained with the help of textbooks: *from this point of view, we do not see a special role for textbooks compared to quasi informal newspaper articles*. The only exception is Wikipedia—regardless of the topic of economics.

Figures 9 and **10** essentially confirm the results obtained so far. However, we now observe, for higher α values, that the cluster values of textbook-based networks become seemingly indistinguishable from those observable for Wikipedia corpora-based networks—the same observation concerns the SZ-Eco-based networks. Textbook-based TCNs are again hardly distinguishable from TCNs derived from informal language newspaper articles about topics related to economics (SZ-Eco).

FIGURE 5 | Alpha-cut graph $C(T, \alpha)$ of the TCN based on SZ-All for $\alpha = 0.7$. Vertex height: compounding-specificity; vertex width: sentence-specificity. For the vertices' translations see **Appendix A**.
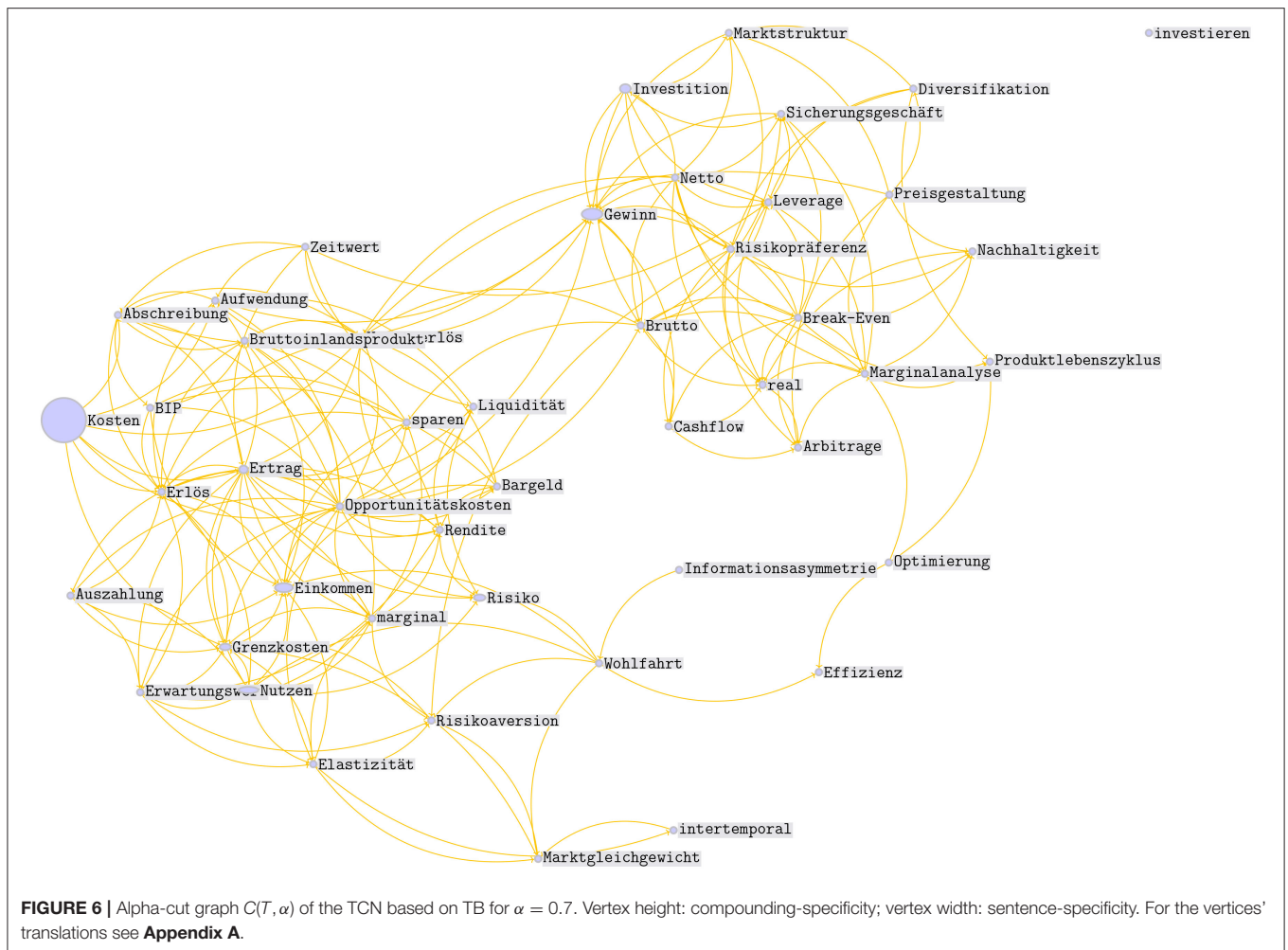
In any case, **Table 7** also shows that all value distributions along the x and y axis are now distinguishable with only three exceptions: the dynamics of clustering is obviously more corpus specific.

Any special role of textbooks almost completely disappears if we consider the cbow model of word2vec (i.e., associations starting from lexical contexts toward target words and thus paradigmatic associations) (see **Figure 8**). In other words, paradigmatic associations of the sort *Bruttoinlandsprodukt/gross domestic product* and *BIP/GDP* seem to be highest from the perspective of Wikipedia-based corpora and higher from the perspective of newspaper corpora than from the perspective of the textbook corpus, while syntagmatic associations of the sort *Gewinn/profit* and *marginal/marginal* are still highest in the case of Wikipedia-based corpora, but are more pronounced from the

perspective of textbooks than from newspapers. **Table 8** leads to an assessment similar to **Table 6**.

Note that in all these cases of cbow (**Figure 9**) and skip-gram-based (**Figure 10**) networks and their underlying embeddings we use standard parameter settings and especially a rate of five iterations: from this point of view, it could be that shorter corpora are more negatively affected by such iterations than longer ones. Scaling their size by increasing the number of iterations can lead to false dissociations of words (as a test of 100 iterations based on the textbook corpus actually suggests). Instead, the sizes of the larger corpora should be reduced to those of the smallest corpora, i.e., the corpus of textbooks—but the corresponding sampling routine and experimentation will be part of future work. In any case, it should be noted that our results are conditioned by the latter assessment. And this means that the alternative of

FIGURE 6 | Alpha-cut graph $C(T, \alpha)$ of the TCN based on TB for $\alpha = 0.7$. Vertex height: compounding-specificity; vertex width: sentence-specificity. For the vertices' translations see **Appendix A**.

Hypothesis H1b is only falsified if we compare Wikipedia-based corpora with newspaper corpora. However, in the case of WP-All, we must refrain from a focus on economics-related topics. The inclusion of the textbook corpus in the set of formal language corpora definitely does not allow such a falsification: so either H1b is wrong or our current measuring procedure does not allow yet for falsifying the alternative of H1b.
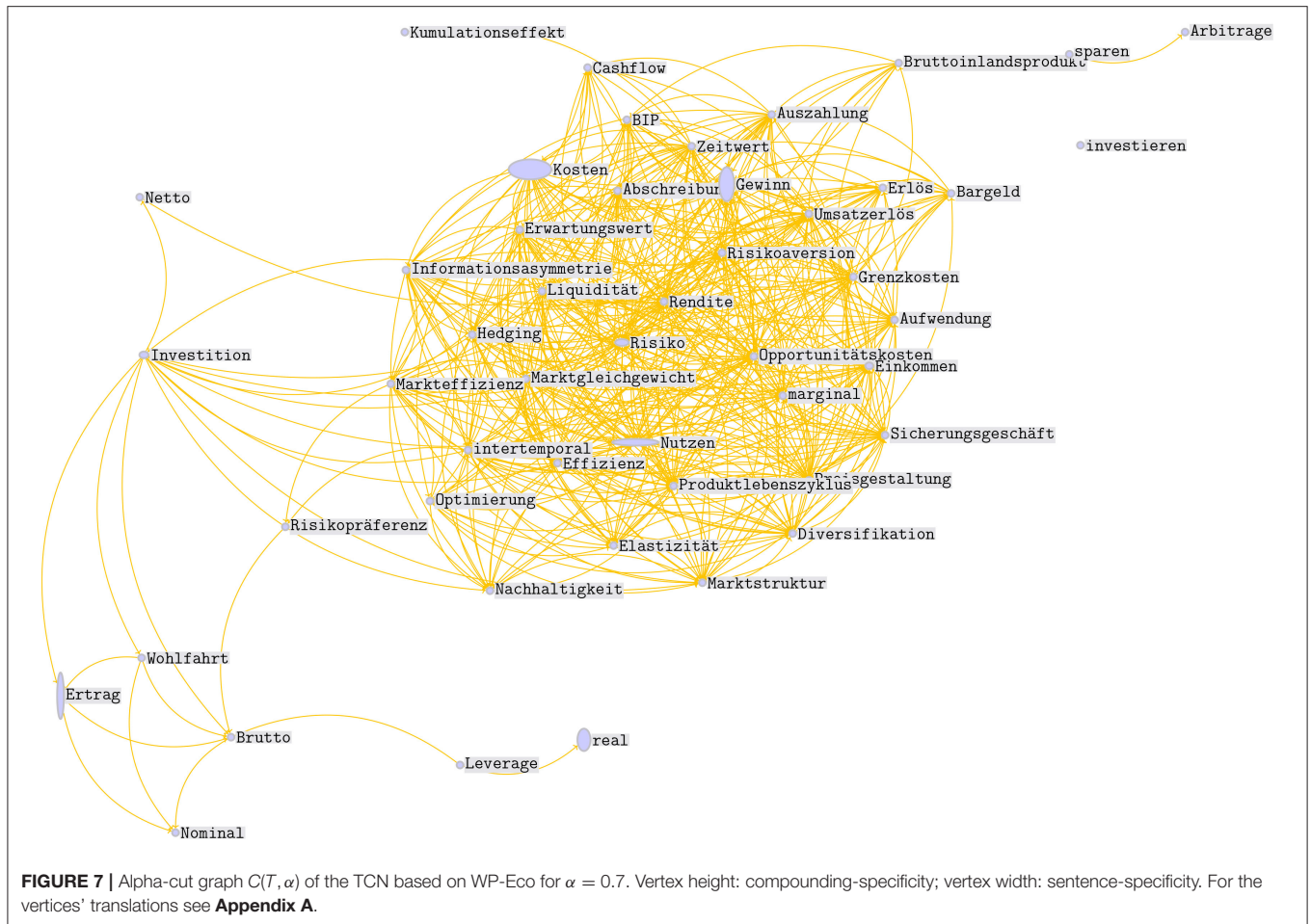
6. DISCUSSION

As evidenced in section 5.4, threshold concepts occur significantly more frequently in formal textbook corpora than in Wikipedia and newspaper corpora, both with respect to the naming variants investigated here and with respect to their frequencies as components of compounds: In line with Hypothesis H1a, the textbook corpora examined had a higher density of compounds and unique sentences than all other corpora investigated here. However, we have also shown that their surrounding networks are not exceptional (in terms of stronger syntagmatic or paradigmatic connections). Regarding the network structure we observed Wikipedia to be exceptional, and this observation is independent of the topic of economics

as it holds for the non-economic corpora, as well. This finding points to a special role of encyclopedic communication as a representative of formal language communication, a role that may have been underestimated in educational sciences until now. However, based on our experiments we must also note that we could not confirm H1b (or falsify its alternative hypothesis).

6.1. Limitations and Suggestions for Future Research

There are several points of departure for improving the procedure we have developed for measuring the usage frequencies of threshold concepts in corpora of formal and informal language:

- We observe that Wikipedia stands out in terms of networking of threshold concepts. Since this observation extends beyond the domain of economics, Wikipedia seems to be characterized by a rather high level of density of specialized language terms in general. As indicated in subsection 5.4, this property is likely due to the encyclopedic genre of Wikipedia, which raises issues

FIGURE 7 | Alpha-cut graph $C(T, \alpha)$ of the TCN based on WP-Eco for $\alpha = 0.7$. Vertex height: compounding-specificity; vertex width: sentence-specificity. For the vertices' translations see **Appendix A**.

FIGURE 8 | Network cohesion as a function of the minimum weight per α-cut of word embedding networks of threshold concepts according to the skip-gram model.

of comparability. In order to test for genre-specificity, comparison with a further encyclopedic resource is a point of departure for future work. We elaborate on this in subsection 6.3 below.

- So far, we analyzed usage regularities of threshold concepts in such a way that we assumed a one-to-one mapping between selected words and the corresponding concepts: For example, the lemma /cost/ then stands directly and uniquely for the corresponding concept of cost. This is where we can start and develop a more general two-step procedure that assumes that concepts can be lexically named by groups of words that form a sort of paradigm of lexical paraphrases of the same concept. This view locates lexical naming alternatives for concepts above the level of lexeme groups, for instance, synonymy clusters, but below the level of word fields. Using the apparatus of word embeddings, such lexeme clusters can be computed as cliques of words with very high cosine similarities of their embeddings in the surrounding co-text, that is, clusters of paradigmatically strongly associated words. However, one should not underestimate the amount of post-correction required to clean up such clusters, for example to sort out highly associated words that do not designate the concept underlying the cluster. In any case, such a procedure makes it possible to identify further co-texts that make reference to the same threshold concept. This would mean to considerably enlarge the database of threshold concept research. Ideally, this approach would also include non-lexical paraphrases.

TABLE 6 | *P*-values of the Kolmogorov-Smirnov goodness-of-fit test applied to the pairwise combinations of the *x* and *y* values of the distributions in **Figure 8**.

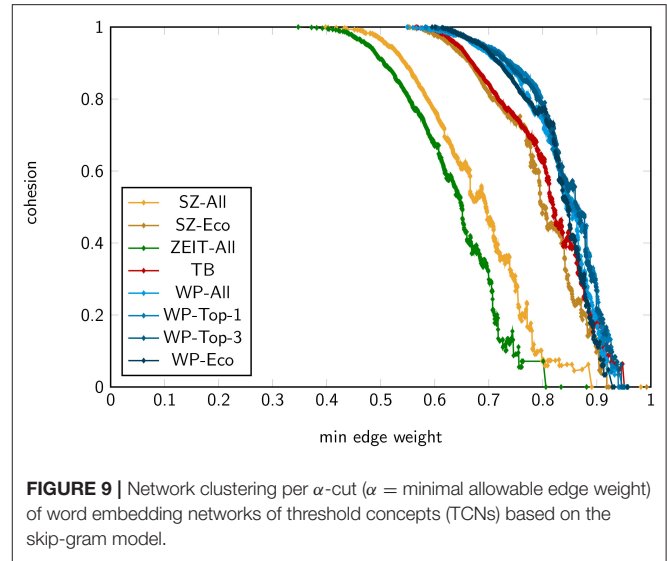|	SZ-All	SZ-Eco	TB	WP-All	WP-Eco	WP-Top-1	WP-Top-3	Zeit-All
x-values								
SZ-All	—	$1.332,3 \times 10^{-15}$	$6.661,3 \times 10^{-16}$	$3.774,8 \times 10^{-15}$	$3.774,8 \times 10^{-15}$	$3.254,7 \times 10^{-300}$	$2.164,6 \times 10^{-299}$	$1.443,3 \times 10^{-15}$
SZ-Eco	—	—	$2.133,0 \times 10^{-05}$	$1.554,3 \times 10^{-15}$	$1.554,3 \times 10^{-15}$	$1.332,3 \times 10^{-15}$	$1.332,3 \times 10^{-15}$	$1.110,2 \times 10^{-16}$
TB	—	—	—	$1.554,3 \times 10^{-15}$	$1.554,3 \times 10^{-15}$	$6.661,3 \times 10^{-16}$	$6.661,3 \times 10^{-16}$	$1.554,3 \times 10^{-15}$
WP-All	—	—	—	—	$0.304,7$	$9.832,6 \times 10^{-07}$	$0.000,2$	$2.109,4 \times 10^{-15}$
WP-Eco	—	—	—	—	—	$1.718,7 \times 10^{-07}$	$6.944,0 \times 10^{-05}$	$2.109,4 \times 10^{-15}$
WP-Top-1	—	—	—	—	—	—	$0.477,5$	$1.443,3 \times 10^{-15}$
WP-Top-3	—	—	—	—	—	—	—	$1.443,3 \times 10^{-15}$
Zeit-All	—	—	—	—	—	—	—	—
y-values								
SZ-All	—	$1.000,0$	$1.000,0$	$1.000,0$	$1.000,0$	$1.000,0$	$1.000,0$	$1.000,0$
SZ-Eco	—	—	$1.000,0$	$1.000,0$	$1.000,0$	$1.000,0$	$1.000,0$	$1.000,0$
TB	—	—	—	$1.000,0$	$1.000,0$	$1.000,0$	$1.000,0$	$1.000,0$
WP-All	—	—	—	—	$1.000,0$	$1.000,0$	$1.000,0$	$1.000,0$
WP-Eco	—	—	—	—	—	$1.000,0$	$1.000,0$	$1.000,0$
WP-Top-1	—	—	—	—	—	—	$1.000,0$	$1.000,0$
WP-Top-3	—	—	—	—	—	—	—	$1.000,0$
Zeit-All	—	—	—	—	—	—	—	—

FIGURE 9 | Network clustering per α-cut (α = minimal allowable edge weight) of word embedding networks of threshold concepts (TCNs) based on the skip-gram model.

- A second extension concerns the detailed consideration of *basic*-, *discipline*-, and *procedural*-level concepts. More specifically, formal language corpora could be divided into subsets of texts depending on their learning level, which are either at the basic, disciplinary or procedural level. In this way, we gain access to contexts of use of threshold concepts that allow us to assign them to one of these levels or to determine linguistic evidence of what was described above as conceptual change, i.e., the transition in the use of a concept between these levels that might indicate a higher dynamics relevant to formal learning contexts.

- A third extension concerns the broadening of the basis of comparison of threshold concepts. That is, instead of just networking them with each other, we could additionally examine how they network with non-threshold concepts or with concepts that belong to one of the three basic, disciplinary or procedural learner levels. In any event, this should again be done in such a way that each of these reference sets is small and selected in advance in order to allow transparent comparisons.

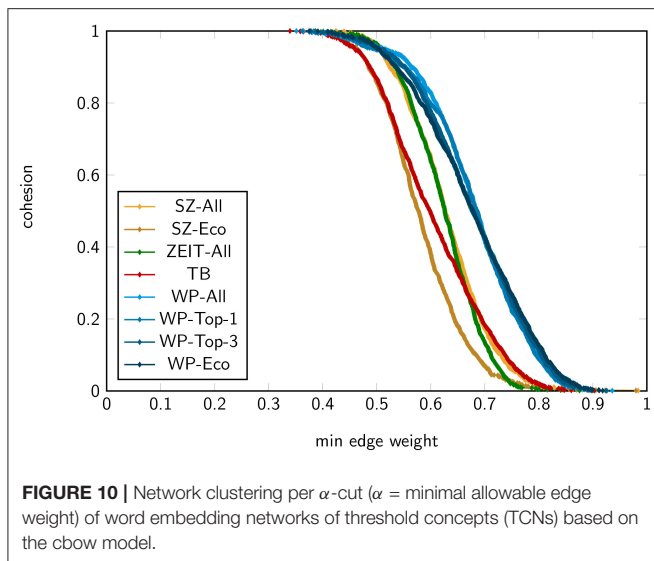## 6.2. Implications of Learning Media for Learning Assessment

Different resources can be interpreted to make different claims about the relations between threshold concepts. For the 63 threshold concept expressions $t_1, \ldots, t_{63}$ under consideration, this claim can be represented in the form: "sense (t_1) is related to sense (t_2), sense (t_2) is more related to sense(t_8)," and so on, where the degree of relatedness differs between the corpora (cf. subsection 4.1.3). That is, different resources express a different "take on threshold concepts." This in turn leads to the question whether the different resources also imply or lead learners to assume a significantly different understanding of threshold concepts, or consequently whether different learning media might be appropriate in different contexts (e.g., depending

TABLE 7 | P-values of the Kolmogorov-Smirnov goodness-of-fit test applied to the pairwise combinations of the x and y values of the distributions in **Figure 9**.

	SZ-All	SZ-Eco	TB	WP-All	WP-Eco	WP-Top-1	WP-Top-3	Zeit-All
***x*-values**								
SZ-All	—	$1.332,3 \times 10^{-15}$	$6.661,3 \times 10^{-16}$	$3.774,8 \times 10^{-15}$	$3.774,8 \times 10^{-15}$	$3.254,7 \times 10^{-300}$	$2.164,6 \times 10^{-299}$	$1.443,3 \times 10^{-15}$
SZ-Eco	—	—	$2.133,0 \times 10^{-05}$	$1.554,3 \times 10^{-15}$	$1.554,3 \times 10^{-15}$	$1.332,3 \times 10^{-15}$	$1.332,3 \times 10^{-15}$	$1.110,2 \times 10^{-16}$
TB	—	—	—	$1.554,3 \times 10^{-15}$	$1.554,3 \times 10^{-15}$	$6.661,3 \times 10^{-16}$	$6.661,3 \times 10^{-16}$	$1.554,3 \times 10^{-15}$
WP-All	—	—	—	—	$0.304,7$	$9.832,6 \times 10^{-07}$	$0.000,2$	$2.109,4 \times 10^{-15}$
WP-Eco	—	—	—	—	—	$1.718,7 \times 10^{-07}$	$6.944,0 \times 10^{-05}$	$2.109,4 \times 10^{-15}$
WP-Top-1	—	—	—	—	—	—	$0.477,5$	$1.443,3 \times 10^{-15}$
WP-Top-3	—	—	—	—	—	—	—	$1.443,3 \times 10^{-15}$
Zeit-All	—	—	—	—	—	—	—	—
***y*-values**								
SZ-All	—	$1.000,0$	$6.661,3 \times 10^{-16}$	$3.774,8 \times 10^{-15}$	$3.774,8 \times 10^{-15}$	$1.086,4 \times 10^{-54}$	$2.445,3 \times 10^{-71}$	$1.443,3 \times 10^{-15}$
SZ-Eco	—	—	$1.221,2 \times 10^{-15}$	$1.554,3 \times 10^{-15}$	$1.554,3 \times 10^{-15}$	$1.332,3 \times 10^{-15}$	$1.332,3 \times 10^{-15}$	$1.110,2 \times 10^{-16}$
TB	—	—	—	$1.176,7 \times 10^{-09}$	$5.162,5 \times 10^{-14}$	$1.477,7 \times 10^{-11}$	$4.872,1 \times 10^{-10}$	$1.767,3 \times 10^{-10}$
WP-All	—	—	—	—	$6.728,1 \times 10^{-05}$	$3.734,6 \times 10^{-05}$	$0.012,6$	$7.908,8 \times 10^{-12}$
WP-Eco	—	—	—	—	—	$4.078,8 \times 10^{-11}$	$3.815,2 \times 10^{-06}$	$2.109,4 \times 10^{-15}$
WP-Top-1	—	—	—	—	—	—	$0.000,9$	$1.042,1 \times 10^{-12}$
WP-Top-3	—	—	—	—	—	—	—	$3.153,0 \times 10^{-14}$
Zeit-All	—	—	—	—	—	—	—	—

TABLE 8 | P-values of the Kolmogorov-Smirnov goodness-of-fit test applied to the pairwise combinations of the x and y values of the distributions in **Figure 10**.

	SZ-All	SZ-Eco	TB	WP-All	WP-Eco	WP-Top-1	WP-Top-3	Zeit-All
***x*-values**								
SZ-All	—	$1.332,3 \times 10^{-15}$	$1.776,4 \times 10^{-15}$	$3.774,8 \times 10^{-15}$	$3.774,8 \times 10^{-15}$	$3.835,6 \times 10^{-41}$	$2.193,2 \times 10^{-33}$	$0.199,8$
SZ-Eco	—	—	$5.596,8 \times 10^{-11}$	$1.554,3 \times 10^{-15}$	$1.554,3 \times 10^{-15}$	$1.332,3 \times 10^{-15}$	$1.332,3 \times 10^{-15}$	$1.110,2 \times 10^{-16}$
TB	—	—	—	$1.554,3 \times 10^{-15}$	$1.554,3 \times 10^{-15}$	$6.661,3 \times 10^{-16}$	$6.661,3 \times 10^{-16}$	$3.108,6 \times 10^{-15}$
WP-All	—	—	—	—	$0.000,6$	$0.595,3$	$0.041,9$	$2.109,4 \times 10^{-15}$
WP-Eco	—	—	—	—	—	$0.001,7$	$0.471,6$	$2.109,4 \times 10^{-15}$
WP-Top-1	—	—	—	—	—	—	$0.047,2$	$1.443,3 \times 10^{-15}$
WP-Top-3	—	—	—	—	—	—	—	$1.443,3 \times 10^{-15}$
Zeit-All	—	—	—	—	—	—	—	—
***y*-values**								
SZ-All	—	$1.000,0$	$1.000,0$	$1.000,0$	$1.000,0$	$1.000,0$	$1.000,0$	$1.000,0$
SZ-Eco	—	—	$1.000,0$	$1.000,0$	$1.000,0$	$1.000,0$	$1.000,0$	$1.000,0$
TB	—	—	—	$1.000,0$	$1.000,0$	$1.000,0$	$1.000,0$	$1.000,0$
WP-All	—	—	—	—	$1.000,0$	$1.000,0$	$1.000,0$	$1.000,0$
WP-Eco	—	—	—	—	—	$1.000,0$	$1.000,0$	$1.000,0$
WP-Top-1	—	—	—	—	—	—	$1.000,0$	$1.000,0$
WP-Top-3	—	—	—	—	—	—	—	$1.000,0$
Zeit-All	—	—	—	—	—	—	—	—

FIGURE 10 | Network clustering per α-cut (α = minimal allowable edge weight) of word embedding networks of threshold concepts (TCNs) based on the cbow model.

on current level of a learner). Such questions obviously pertain to the encompassing methodological structure outlined in **Figure 1**, not just to the learning media. Studies combining educational and computational linguistic methods, as presented here, make it possible to derive assumptions of the effects of different types of texts used for learning on the learning outcome, include individual learner-internal influence factors, as most explicitly formulated in the *offer-use model* (Helmke, 2009). Accordingly, a most straightforward continuation of our approach is to implement an educational assessment of student learning, related to the examined threshold concepts. The computational linguistic assessment very likely has implications for text comprehension (Kintsch, 1988) and domain learning (Alexander, 2018). For instance using educational assessment, we can link findings on threshold concept profiles in texts to findings on learners' understandings of these threshold concepts, as evidenced in their assessment performance (Brückner and Zlatkin-Troitschanskaia, 2018). Such an assessment is necessary since there is no straightforward mapping between dictionary concepts and indexed concepts (a student's private understanding of lexical meanings), as mentioned in sections 3 and 4. Although such studies are future work (but see Mehler and Ramesh, 2019 for a formal learning assessment framework), a few points of departure can already be considered, as we do in the following.

As observed by Rincke (2010), there is a striking similarity of the acquisition of a special language with language acquisition in general (both include terminological and conceptual change). Hence, we might expect to find some empirical evidence relevant for the acquisition of threshold concepts in language acquisition studies. In this regard, Oakhill et al. (2003) show in a study on language development that word reading and text comprehension are dissociated. This implies that text comprehension and word decoding follow different developmental trajectories and can be taught at least to some degree independently. The acquisition of threshold concepts

proceeds at least on these two routes, meaning that developing respective understandings draws on text comprehension as well as on lexical definitions. This line of thought emphasizes the need for a semantic analysis of threshold concepts in business education which, as far as we know, is missing (see section 4.1.3). Furthermore, we may hypothesize that denser networks of threshold concepts pose higher requirements on word decoding, while looser networks pose higher requirements on text comprehension (very likely there is an interaction with text type which is discussed in subsection 6.3 below). Now given that from learning assessments (cf. **Figure 1**) we know that a student cohort has a better developed definitional than applicational competence in dealing with threshold concepts, a deliberate choice of learning media can foster or balance this asymmetry in competence.

Text comprehension is not only based on memory processes but also on constructionist processes (van den Broek et al., 2005). The latter can, for instance, arise due to associations bound up with readers' indexed concepts. This includes personal preferences as well as all sorts of top-down processes. Constructionist aspects of comprehension are bound up with learners' everyday language and prior knowledge and experience. If we liken the acquisition of a specialized language to second language acquisition, this implies that also the first or prior language(s) should be taken into account (cf. Shanahan et al., 2006). Here we meet advice from educational research, namely *to regard everyday language and specialized language as respectively developable in their own right and to address them both in class* ("Alltags- und Fachsprache als je für sich entwicklungsfähig anzusehen und im Unterricht zu thematisieren" [translated by AL]) (Rincke, 2010, p. 235). The everyday language competence can only be tapped in the classroom, if one can classify the text resources according to their language level. On a large scale this can only be done with the help of automatic methods.

The prior knowledge of learners also plays a role in reading hypertexts, such as Wikipedia articles. Interestingly, using hypertexts as a learning resource can be advantageous in particular for informed learners, since the hypertext structure allows them to exert a strategic reading processes (Salmerón et al., 2006). That is, online (hypertext) resources, such as Wikipedia can enrich the learning landscape for formal education (as they already do as a matter of fact by virtue of student selection, cf. subsection 2.2).

This poses the question of reliability of Wikipedia[17]. Wikipedia articles seem to be reliable in general (e.g., Wilkinson and Huberman, 2007). However, with regard to specific topics, such as *respiration in medicine* Wikipedia turned out to be an insufficiently reliable resource for learning (Meier, 2008; Azer, 2015)[18]. Accordingly, a qualitative assessment is needed in order to find evidence on which of these

---

[17]We are thankful to an anonymous reviewer for emphasizing this issue.

[18]To be fair, since Wikipedia (and related specialized Wikis) is a highly dynamic resource the situation may have changed already since the time of publication of the study. We know of no recent replication, however.

opposing sides Wikipedia's economic articles belong—there is no *a priori* reason to exclude newspapers from such an assessment[19].

6.3. Text Types in and for Learning

Wikipedia, newspapers and textbooks are all examples of different text types. These text types differ with respect to narrative structure, content, target group, and many more properties. In particular the didactic structure of different learning media genres interact with conceptual change, as discussed in the following for each text type used in our study.

1. *Interpretation variant I: Textbooks are optimal.* Assuming that textbooks are optimized for the transfer of specialized information in higher education, the question arises as to the significance of our findings. Section 5 indicates that the network of threshold concepts based on the usage regularities confirmed in Wikipedia is much *denser* than in the case of newspaper corpora or even textbooks. Metaphorically speaking, the encyclopedia-based association network manifests a more densely distributed "matter" of much more closely associated conceptual units. Higher density and stronger associativity also mean a higher degree of confirmation and thus stability, because the underlying associative relations are more strongly confirmed by co-occurrences that can actually be observed in many sentence windows. Stability is here a simple consequence of the fact that a change of such strongly confirmed associations would require a higher amount of textual information contradicting the already confirmed associations by aiming into other directions of associations instead. This amount probably would be equal to the amount of the original textual information, which underlies the stabilized associations. According to this interpretation we may say that *encyclopedic textual information seems to over-confirm the associations of threshold concepts.* Conversely, the network of associations based on newspaper corpora seems to be *under-confirmed* and therefore too unstable: by positioning the same concepts in ever new contexts, their association relations virtually fan out, so that each individual association is far less confirmed. New textual information then does not necessarily confirm what already exists, but rather refers to ever new possibilities of association. In the middle of these two extreme cases we find the association network resulting from textbook corpora. Under the interpretation that this network is optimally organized, we find that textbooks balance the under-confirmation induced by newspaper corpora with the over-confirmation by encyclopedias in terms of a fluent equilibrium: an optimum, so to speak, as a balance of firstness and confirmation according to the notion of pragmatic information (von Weizsäcker, 1974). Textbooks are organized around threshold concepts in such a way that their readers can learn

the targeted concepts with sufficient conceptual density (outcome perspective), but not in such a way that they would not be able to recontextualize them or transfer them between different contexts (process perspective) whereby these recontextualizations do not excessively disturb and consequently do not dissolve the previously confirmed associations of threshold concepts. This is supported by the directional way in which textbooks guide learners through the learning process by providing them with an epistemic structure of the discipline (Dalimunte and Pramoolsook, 2020).

2. *Interpretation variant II: Encyclopedic texts are optimal.* As conjectured in subsection 5.4, the dense network of threshold concepts observed in Wikipedia is probably due the fact that Wikipedia is an encyclopedic resource and as such introduces special threshold concept terms by means of definitions. In this sense, Wikipedia represents the result state of threshold concept knowledge. In contrast, a textbook often *develops* a concept and takes a more process-oriented approach (see, e.g., Dalimunte and Pramoolsook, 2020). The semantic flavors of both approaches have already been observed in the sample sketch in section 3. Such differences, we argue, are finally reflected in different network densities. Teleologically understood, as in domain learning (Alexander et al., 1995), a result is the goal of a process. A process can be conceived as a succession of (intermediate) states (cf., e.g., Fernando, 2011). In section 3 we suggested to connect conceptual change in particular to the update operations *greater level of abstraction* and *shifted vantage point* (Chi and Ohlsson, 2005). That is, the intermediate states are related in terms of semantic update operations. A consequence then is that a successor state is more developed than its predecessor state. Furthermore, the hypothesis is that update operations apply at a larger range to looser linked concept networks than to denser linked ones. To put it another way: denser networks are closer to a result state and make further conceptual changes more unlikely, and rightly so since result state are closer to an optimum. In order to make this line of argumentation and modeling more precise, however, the need for a semantic characterization of mental updates and the differences between different kinds of updates are required.

3. *Interpretation variant III: Newspaper articles are optimal.* Newspapers also offer a wide range of potential for learning despite the low frequency of compound-specific and sentence-specific threshold concepts and a lower semantic density compared to textbooks and Wikipedia. Depending on the curricular goal, e.g., whether the focus is on economic education in the sense of general maturity for social participation or on the professional expertise of an economist, alternative uses may be suitable.   A lower density of the threshold web in learning media, as was evident in the present findings, leaves room for a more in-depth examination of individual threshold conceptions by learners and can promote their motivation and understanding. Newspapers are by no means only complementary materials in economics courses. They offer the possibility of an active application

---

[19]Since textbooks are submitted to a quality control procedure, there is an *a priori* reason to exclude textbooks from a further quality assessment.

of what has been learned in the course due to the potential of the articles' alternative interpretations, current topics, events, and ever new contexts (see the articles in Hoyt and McGoldrick, 2012). The looser density of threshold concepts promotes newspapers to be used as an introductory learning opportunity (Helmke and Schrader, 2008). As Dalimunte and Pramoolsook (2020) note, despite the central structure textbooks provide for teaching, their texts are often more difficult for novices to read, so that newspapers can provide a first access to subject-specific learning in economics. Depending on the objective, e.g., a critical examination of the definition of concepts in newspaper articles, a certain amount of prior knowledge of the learning group is required. Newspapers corpora of SZ and *Zeit* in particular often require reading skills and prior subject-specific knowledge, so that they can also be used effectively by lecturers during their classes (McEachern, 2012). Newspapers can also be useful for cooperative forms of learning (McGoldrick et al., 2010), e.g., for jointly comparing and evaluating threshold concepts in different newspaper articles from different corpora. In addition, however, they can not only serve as exemplary texts and information materials, but can also be a central object for the design of lessons. For example, in his conclusion on the analysis of learning media, that are not originally developed for educational purposes, Croushore (2012, p. 636) writes: "[…] instructors of money and banking must be on constant alert for changes in the material. While this may seem difficult, these constant changes actually make the course easy to teach because nearly every day's newspaper provides new course material." Therefore, assuming that newspapers have a lower threshold concepts density, it seems reasonable to expect that more diverse associations are possible for the learner. In other words: Since the network is less stable, teachers have more freedom to design their courses.

4. *Interpretation variant IV: Synthetic view, or mixture model.* The previous bullet points provided reasons that each learning text type can be considered "optimal." But optimal with reference to what? Adopting the view that learning is a process (for a recent affirmation of this (somewhat obvious) view see Dalimunte and Pramoolsook, 2020) that conceptually develops in the triangle between lexical, dictionary and indexed concepts, among others, as outlined in section 4, one also adopts a *dynamic* rather than a *static* perspective. A dynamic perspective allows for a synthetic view on learning media since it conceives learning in its ecological niche. In relation to students' prior knowledge, current interest, curricular goals and teachers' content-related and pedagogical focus each text type can be used for its respective strengths. In the end, thus, a synthetic view amounts to an adjusted and combined approach. However, in order to be of value, it needs to be complemented with an assessment of learning situations in order to gain evidence about the most suitable learning resources for a given learning situation. Since this issue leads to the topic of this special issue, we want to elaborate on it in the subsequent section.

6.4. Comparative Media Analysis of Threshold Concept Webs and (Online) Information Processing and Learning

We argued that a dedicated linguistic analysis of textual learning media used in economics education is necessary due to the increasing digitalization of teaching and learning in economics. Digitalization is constantly increasing the range of learning media that can potentially be used by teachers and lecturers (Johinke and Di Lauro, 2020). The aforementioned, more frequent use of Wikipedia by students in economic learning contexts (Freire and Li, 2016) and the increasing digitization of textbooks and distribution as Open Educational Resources (Fischer et al., 2017) are facilitating computer-based and internet-based learning. These multimedia environments afford learning based on multiple representations (Mayer, 2014). In order to support teachers in their decisions for selecting media for teaching-and-learning purposes, it is necessary to apply a content quality criterion to compare media used for learning. For this purpose, the linguistic properties of threshold concepts were compared between several corpora from Wikipedia, the business-related newspaper sections of *Süddeutsche Zeitung* and *Zeit*, and 14 business and economics textbooks. Given the large amounts of learning resources in digital media and the associated comprehensive (corpus) data sets, the computational linguistic approach is advantageous in comparison to already established qualitative content-analytical procedures, in order to provide teachers with general and innovative information on the usefulness of media for learning in a condensed form. Linguistic procedures, which explore the morpho-syntactic structure of the underlying threshold concepts are particularly suitable, since the primary access of novices and beginners to economics is text-based. Text-based introduction to threshold is more commonly used than via diagrams and other visualizations (Tinkler and Woods, 2013). With the exception of studies on readability (Tinkler and Woods, 2013), word frequency counts (Leet and Lopus, 2003) or genre-specific analyses of a few textbooks (Dalimunte and Pramoolsook, 2020), there are no comparisons of digital media of different types and genres in economics education. The present analysis thus provides an important comparison of different media types and implications for their use in digital learning contexts. The use of threshold concept webs for the comparative media analysis of learning sources (Helmke and Schrader, 2008) is often a prerequisite for studies on learning success which rely on mental association patterns, such as those found in the studies by Davies and Mangan (2007), Vidal et al. (2015), and Ivan Montiel and Antolin-Lopez (2020). In this study, we found that the lexical and semantic density of threshold concepts is higher in Wikipedia than in textbooks and newspapers. This analysis of subject-specific concepts goes beyond the density analyses of pronouns found in textbook analyses of foreign language research (e.g., Kong, 2009) and offers a number of implications for the initiation and design of learning processes. On the one hand, threshold concept density can be an advantage for students who want to learn about a content area in a short time (Meier, 2008; Freire and Li, 2016), on the other hand, students need not only basic skills for

researching and evaluating web resources, but, especially with the difficulty of learning new threshold concepts (see section 1), they also need prior subject knowledge (Sender, 2017). Nevertheless, especially in introductory economics courses, Wikipedia can also stimulate creative learning processes, because on the one hand the platform includes references to external literature or alternative perspectives and thus can generate interest in different topics (Meier, 2008). On the other hand, due to the density of concepts and the editability of content, it also offers opportunities for students to critically reflect on content, to review existing articles (Johinke and Di Lauro, 2020) or to check their own misconceptions (Freire and Li, 2016). In turn, a lower density of threshold concepts, as is the case with newspapers in particular, does not imply a lower quality of newspapers for didactic purposes. Threshold concepts are special subject-specific concepts that require a gradual development of expertise over several phases (Davies and Mangan, 2007). This development requires examples, practical and professional applications in which the concepts are didactically embedded. The more variable the context is regularly updating newspapers reporting on changing topics, the more application possibilities are offered to the learner for an in-depth examination of acquired threshold concepts. In addition, the disciplinary and semantic density is not too high, so that even learners with little previous knowledge can approach the threshold concepts and develop initial ideas, which may need to be corrected or refined over time. Furthermore, lower specialization and stronger contextualization as well as a change of media from textbook to newspaper afford didactic advantages and enable learning through multiple-representations, which can be used in a targeted manner, especially in phases of learner activation and topic introducing, to cognitively activate and motivate learners. However, it should be kept in mind that newspapers are subject to daily change. The presented study is fundamental for future research on information and learning processes. It offers a number of links for further research that can be taken into account in conventional educational assessments. For example, it could be investigated how the density of threshold concepts between the different media types affects learning success, whether students with varying levels of prior knowledge benefit more or less from certain media, or which specificity (e.g., compound or sentence specificity) affects the learning process and how. These central linguistic characteristics can in turn help to determine how textbooks could be structured, which language use would support teaching or how closely threshold concepts should be linked to be as conducive to learning from learning media as possible.

7. CONCLUSION

The computational linguistic perspective adopted in the present contribution pursues an orientation which, in terms of educational research on threshold concepts, has two special features. On the one hand, it complements content analyses, which are classically used to analyze textbooks, protocols, or other textually and graphically represented materials in order to work out education-related meanings from the materials

(e.g., Krippendorff, 2013). The often tedious and lengthy manual evaluation with only a limited number of documents and the corresponding susceptibility to errors is as a matter of fact limited to a small amount of data. Computational linguistic analyses, to the contrary, can process huge corpora. Secondly, so-called *utilization-of-learning-opportunities* models are used to model the mechanisms of action of teaching-learning arrangements in educational research (e.g., Braun et al., 2014). These models show the interactions between learning-relevant aspects in terms of input-process-output paths. Very often learning outcomes are analyzed in connection with different input factors (e.g., socio-economic status, gender, intelligence, self-assessed use of learning media). Significantly less frequently, however, the learning potentials of the respective learning environments or learning materials are considered independently of a learner's assessment. With the computational linguistic approach presented here, especially the learning media that are used as input into the learning processes are processed on a large scale and thus a description of the learning environment is presented that can be considered in informal as well as formal learning processes. Ultimately learning, the meaning of threshold concept expressions and their use in text resources are embraced within the contour of an emerging research program—encompassing specialized vocabularies, learning and education, and computational linguistics—in terms of mental, referential and differential meanings. The latter two (referential and differential meanings) are used in order to derive hypotheses concerning formal and informal learning contexts with respect to a special class of expressions, *viz.* threshold concepts. A second focus was the development of a computational linguistic model for operationalizing threshold concepts for the analysis of learning resources. In this context, we developed the notion of a Threshold Concept Network (TCN) and quantified it by means of alpha-cuts, taking into account the "web of threshold concepts" (Davies and Mangan, 2007). In this way, we were able to prove an exceptional status of threshold concepts in textbooks, at least at the node level. The main result was that formal and informal resources can indeed be distinguished in terms of their threshold concepts' profiles. Furthermore, Wikipedia turns out to be a first class formal learning resource. Continuing this line of research will include at least the following steps: the methodological considerations discussed in subsection 6 are to be addressed. A lexical semantic analysis of threshold concepts is due. And, most importantly, our findings have to be tied back to education assessments of learners. Furthermore, experimental studies have to be designed that investigate systematically the impact of different resources on learning. Very often experimental studies are developed on assumptions that have not been tested themselves. On the basis of the computational linguistic assessment, however, it is possible to develop more specific questions. Most notably, the threshold concept acquisition of learners can be compared depending on the media to learn (e.g., Wikipedia vs. textbook vs. daily newspaper, and their interaction and complementary uses)—whereby, of course, the corresponding media competencies and information literacy or other (intellectual) characteristics must also be controlled (Vernooij, 2000). The assessments from the study presented here

provide a starting point for such experiments which in turn would round out the emerging research program we sketched.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available due to copyright restrictions, the newspaper and the textbook corpora are not publicly available. Wikipedia can be obtained via Wikipedia dumps. Further queries regarding the material and analysis presented here should be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

AL mainly has written subsection 2.2, and sections 3 and 4, and performed the Kolmogorov-Smirnov goodness-of-fit tests and the power law fitting (**Tables 2**–**8**). AM designed the computational linguistic measurement procedure for threshold concepts, implemented the corresponding network analyses, has written almost all parts of section 5 and generated **Figures 5**–**10**. Sections 6 and 7 have been jointly written by SB, AM, and AL, with the exception of subsection 6.4, which is mainly due to SB. GA carried out the preprocessing of the corpora and the word embeddings. SB selected the textbook corpus and mainly has written section 1, the preamble of section 2 and subsection 2.1. TU calculated the compound distributions and the threshold

concept networks and produced **Table 1**. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2020.578475/full#supplementary-material

REFERENCES

Alexander, P. A. (2018). "Into the future. A prospective look at the model of domain learning," in *The Model of Domain Learning: Understanding the Development of Expertise, Chapter 10*, eds H. Fives and D. L. Dinsmore (New York, NY: Routledge), 195–214. doi: 10.4324/9781315458014-12

Alexander, P. A., Jetton, T. L., and Kulikowich, J. A. (1995). Interrelationship of knowledge, interest, and recall: assessing a model of domain learning. *J. Educ. Psychol.* 87, 559–575. doi: 10.1037/0022-0663.87.4.559

Alstott, J., Bullmore, E., and Plenz, D. (2014). powerlaw: A Python package for analysis of heavy-tailed distributions. *PLoS ONE* 9:e95816. doi: 10.1371/journal.pone.0085777

Asher, N. (1993). *Reference to Abstract Objects in Discourse. Number 50 in Studies in Linguistics and Philosophy*. Dordrecht: Kluwer Academic Publishers.

Asher, N., and Lascarides, A. (2003). *Logics of Conversation*. Cambridge: Cambridge University Press.

Azer, S. A. (2015). Is Wikipedia a reliable learning resource for medical students? Evaluating respiratory topics. *Adv. Physiol. Educ.* 39, 5–14. doi: 10.1152/advan.00110.2014

Barsalou, L. W. (1993). "Flexibility, structure, and linguistic vagary in concepts: manifestations of a compositional system of perceptual symbols," in *Theories of Memory, Chapter 3*, eds A. F. Collins, M. A. Conway, and P. E. Morris (Hillsdale, NJ: Lawrence Erlbaun Associates), 29–101. doi: 10.4324/9781315782119-3

Barsalou, L. W. (1999). Perceptual symbol systems. *Behav. Brain Sci.* 22, 577–660. doi: 10.1017/S0140525X99002149

Barsalou, L. W., Yeh, W., Luka, B. J., Olseth, K. L., Mix, K. S., and Wu, L.-L. (1993). "Concepts and meaning," in *Chicago Linguistics Society 29: Papers from the Parasessions on Conceptual Representations*, eds K. Beals, G. Cooke,

D. Kathman, K. E. McCullough, S. Kita, and D. Testen (Chicago, IL: Chicago Linguistics Society), 23–61.

Barwise, J., and Perry, J. (1983). *Situations and Attitudes. The David Hume Series on Philosophy and Cognitive Science Reissues*. Stanford, CA: CSLI Publications.

Blum, U. (2017). *Grundlagen der Volkswirtschaftslehre*. Berlin; Boston, MA: De Gruyter Oldenbourg.

Braun, E., Weiß, T., and Seidel, T. (2014). "Lernumwelten in der Hochschule," in *P¨adagogische Psychologie: Mit Online-Materialien zum Download, 6th Edn.*, eds T. Seidel and A. Krapp (Weinheim: Beltz), 433–454.

Brooks, D. C. (2016). *ECAR Study of Undergraduate Students and Information Technology*. Available online at: https://er.educause.edu/~/media/files/library/2016/10/ers1605.pdf?la=en

Brückner, S., and Lücking, A. (2019). "Computerlinguistische Analyse des Schwellenkonzeptansatzes in der Wirtschaftsdidaktik," in *Talk at the Jahrestagung Sektion Berufs- und Wirtschaftsp¨adagogik* (Graz: Karl-Franzens-University).

Brückner, S., and Zlatkin-Troitschanskaia, O. (2018). "Threshold concepts for modeling and assessing higher education students' understanding and learning in economics," in *Assessment of Learning Outcomes in Higher Education: Cross-National Comparisons and Perspectives*, eds O. Zlatkin-Troitschanskaia, M. Toepper, H. A. Pant, C. Lautenbach, and C. Kuhn (Wiesbaden: Springer International Publishing), 103–121. doi: 10.1007/978-3-319-74338-7_6

Bücking, S. (2010). "German nominal compounds as underspecified names for kinds," in *New Impulses in Word-Formation*, ed S. Olsen (Hamburg: Buske), 253–281.

Chi, M. T. H., and Ohlsson, S. (2005). "Complex declarative learning," in *The Cambridge Handbook of Thinking and Reasoning, Chapter 16*, eds K. J. Holyoak and R. G. Morrison (Cambridge, NY: Cambridge University Press), 371–399.

Clauset, A., Shalizi, C. R., and Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Rev.* 51, 661–703. doi: 10.1137/070710111

Cohen, A., and Erteschik-Shir, N. (2002). Topic, focus, and the interpretation of bare plurals. *Nat. Lang. Semant.* 10, 125–165. doi: 10.1023/A:1016576614139

Cook, S. W., and Goldin-Meadow, S. (2006). The role of gesture in learning: do children use their hands to change their minds? *J. Cogn. Dev.* 7, 211–232. doi: 10.1207/s15327647jcd0702_4

Cousin, G. (2008). "Threshold concepts: old wine in new bottles or a new form of transactional curriculum inquiry," in *Threshold Concepts Within the Disciplines*, eds R. Land, J. H. Meyer, and J. Smith (Rotterdam: Sense Publishers), 261–272. doi: 10.1163/9789460911477_020

Croushore, D. (2012). "Using real-world applications to policy and everyday life to teach money and banking," in *International Handbook on Teaching and Learning in Economics. Elgar Original Reference*, eds G. M. Hoyt and K. McGoldrick (Cheltenham: Edward Elgar), 628–637. Retrieved from: http://www.elgaronline.com/view/9781848449688.xml

Crystal, D. (1997). *The Cambridge Encyclopedia of Language, 2nd Edn.* Cambridge: Cambridge University Press.

Dalimunte, A. A., and Pramoolsook, I. (2020). Genres classification and generic structures in the English language textbooks of economics and Islamic economics in an Indonesian university. *Lang. Educ. Acquisit. Res. Netw.* 13, 1–19.

Davies, P., and Mangan, J. (2007). Threshold concepts and the integration of understanding in economics. *Stud. High. Educ.* 32, 711–726. doi: 10.1080/03075070701685148

de Ruiter, J. P. (2007). Postcards from the mind: the relationship between speech, imagistic gesture, and thought. *Gesture* 7, 21–38. doi: 10.1075/gest.7.1.03rui

de Saussure, F. (1916). *Course de linguistique générale*. Lausanne; Paris: Payot.

Demberg, V., Keller, F., and Koller, A. (2013). Incremental, predictive parsing with psycholinguistically motivated tree-adjoining grammar. *Comput. Linguist.* 39, 1025–1066. doi: 10.1162/COLI_a_00160

Devetak, I., and Vogrinc, J. (2013). "The criteria for evaluating the quality of the science textbooks," in *Critical Analysis of Science Textbooks*, Vol. 10, ed M. S. Khine (Dordrecht: Springer Netherlands), 3–15. doi: 10.1007/978-94-007-4168-3_1

Fernando, T. (2011). Constructing situations and time. *J. Philos. Logic* 40, 371–396. doi: 10.1007/s10992-010-9155-1

Fillmore, C. J., Lee-Goodman, R. R., and Rhomieux, R. (2012). "The FrameNet construction," in *Sign-Based Construction Grammar, Number 193 in CSLI Lecture Notes, Chapter 7*, eds H. C. Boas and I. A. Sag (Stanford, CA: CSLI Publications), 309–372.

Fischer, L., Ernst, D., and Mason, S. (2017). Rating the quality of open textbooks: how reviewer and text characteristics predict ratings. *Int. Rev. Res. Open Distrib. Learn.* 18, 142–154. doi: 10.19173/irrodl.v18i4.2985

Frege, G. (1892). Über Sinn und Bedeutung. *Z. Philos. Philos. Krit.* 100, 25–50.

Freire, T., and Li, J. (2016). Using Wikipedia to enhance student learning: a case study in economics. *Educ. Inform. Technol.* 21, 1169–1181. doi: 10.1007/s10639-014-9374-0

Ginzburg, J. (2012). *The Interactive Stance: Meaning for Conversation*. Oxford: Oxford University Press.

Haab, T. C., Schiff, A., and Whitehead, J. C. (2012). "Economics blogs and economic education," in *International Handbook on Teaching and Learning in Economics. Elgar Original Reference*, eds G. M. Hoyt and K. McGoldrick (Cheltenham: Edward Elgar), 167–173. Retrieved from: http://www.elgaronline.com/view/9781848449688.xml

Hatt, L. (2018). Threshold concepts in entrepreneurship-the entrepreneurs' perspective. *Educ. Train.* 60, 155–167. doi: 10.1108/ET-08-2017-0119

Heim, I. (2002). "File change semantics," in *Formal Semantics: The Essential Readings, Number 1 in Linguistics: The Essential Readings, Chapter 9*, eds P. Portner and B. H. Partee (Oxford; Malden, MA: Wiley-Blackwell), 223–248.

Helmke, A. (2009). *Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze: Klett-Kallmeyer.

Helmke, A., and Schrader, F.-W. (2008). Merkmale der Unterrichtsqualität: Potential, Reichweite und Grenzen. *Seminar* 14, 17–47.

Hemati, W., Uslu, T., and Mehler, A. (2016). "TextImager: a distributed UIMA-based system for NLP," in *Proceedings of the COLING 2016 System Demonstrations. Federated Conference on Computer Science and Information Systems* (Osaka).

Hjelmslev, L. (1961). *Prolegomena to a Theory of Language, 7th Edn.* Madison, WI: University of Wisconsin Press.

Hoadley, S., Tickle, L., Wood, L. N., and Kyng, T. (2015). Threshold concepts in finance: conceptualizing the curriculum. *Int. J. Math. Educ. Sci. Technol.* 46, 824–840. doi: 10.1080/0020739X.2015.1011244

Hoffmann, L., Kalverkämper, H., Wiegand, H. E., Galinskim, C., and Hullen, W. (Eds.). (1998). *Fachsprachen/Languages for Special Purposes: Ein Internationales Handbuch Zur Fachsprachenforschung und Terminologiewissenschaft*. Berlin; Boston, MA: De Gruyter, Inc.

Hofhues, S. (2016). "Informelles Lernen mit digitalen Medien in der Hochschule," in *Handbuch Informelles Lernen*, ed M. Rohs (Wiesbaden: Springer Fachmedien), 529–546. doi: 10.1007/978-3-658-05953-8_28

Hoyt, G. M., and McGoldrick, K. (Eds.). (2012). *International Handbook on Teaching and Learning in Economics. Elgar Original Reference* (Cheltenham: Edward Elgar). Retrieved from: http://www.elgaronline.com/view/9781848449688.xml

Hu, C., and Gao, H. (2019). Nouns and nominalizations in economics textbooks. *Lang. Context Text Soc. Semiot. Forum* 1, 288–312. doi: 10.1075/langct.00012.hu

Ivan Montiel, P. J. G., and Antolin-Lopez, R. (2020). What on earth should managers learn about corporate sustainability? A threshold concept approach. *J. Bus. Ethics* 162, 857–880. doi: 10.1007/s10551-019-04361-y

Jackendoff, R. (1983). *Semantics and Cognition*. Cambridge, MA: MIT Press.

Jackendoff, R. (1991). *Semantic Structures. Number 18 in Current Studies in Linguistics*. Cambridge, MA: MIT Press.

Jackendoff, R. (2002). *Foundations of Language. Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.

Jadin, T., and Zöserl, E. (2009). Informelles Lernen mit Web-2.0-Medien. *Bildungsforschung* 6, 41–61.

Johinke, R., and Di Lauro, F. (2020). Wikipedia in higher education: practice what you teach. *Stud. High. Educ.* 45, 947–949. doi: 10.1080/03075079.2020.1763014

Kamp, H., and Reyle, U. (1993). *From Discourse to Logic*. Dordrecht: Kluwer Academic Publishers.

Karttunen, L. (1969). "Discourse referents," in *Proceedings of the 1969 Conference on Computational Linguistics, COLING '69* (Stockholm), 1–38. doi: 10.3115/990403.990487

Keller, R. (1995). *Zeichentheorie. Zu einer Theorie semiotischen Wissens*. Tübingen: Francke.

Kilgour, P., Reynaud, D., Northcote, M., McLoughlin, C., and Gosselin, K. P. (2019). Threshold concepts about online pedagogy for novice online teachers in higher education. *High. Educ. Res. Dev.* 38, 1417–1431. doi: 10.1080/07294360.2018.1450360

Kintsch, W. (1988). The role of knowledge in discourse comprehension: a construction-integration model. *Psychol. Rev.* 95, 163–182. doi: 10.1037/0033-295X.95.2.163

Klein, U., and Kracht, M. (2014). "Notes on disagreement," in *Approaches to Meaning. Composition, Values, and Interpretation, Number 32 in Current Research in the Semantics/Pragmatics Interface*, eds D. Gutzman, J. Köpping, and C. Meier (Leiden: Brill), 276–305. doi: 10.1163/9789004279377_013

Knight, C., and Pryke, S. (2012). Wikipedia and the university, a case study. *Teach. High. Educ.* 17, 649–659. doi: 10.1080/13562517.2012.666734

Köhler, R. (1986). *Zur linguistischen Synergetik: Struktur und Dynamik der Lexik*. Bochum: Brockmeyer.

Kong, K. (2009). A comparison of the linguistic and interactional features of language learning websites and textbooks. *Comput. Assist. Lang. Learn.* 22, 31–55. doi: 10.1080/09588220802613799

Kricks, K., Mittelstädt, E., and Liening, A. (2013). Schwellenkonzepte und Phänomenografie. Explorative Studie zur Messung von Unterschieden im ökonomischen Verstehen. *Z. ökonom. Bild.* 2, 17–41. doi: 10.7808/zfoeb.1.2.74

Krifka, M. (2003). "Bare NPs: Kind-referring, indefinites, both, neither?" in *Proceedings of the 13th Conference on Semantics and Linguistic Theory, SALT XIII*, eds R. Young and Y. Zhou (Ithaca, NY: Cornell University), 180–203. doi: 10.3765/salt.v13i0.2880

Kripke, S. A. (2011). *Philosophical Troubles: Collected Papers*, Vol. 1. New York, NY: Oxford University Press.

Krippendorff, K. (2013). *Content Analysis: An Introduction to Its Methodology, 3rd Edn.* Thousand Oaks, CA: SAGE Publications, Inc.

Lamb, P., Hsu, S.-W., and Lemanski, M. (2019). A threshold concept and capability approach to the cross-cultural contextualization of Western management education. *J. Manag. Educ.* 44, 101–120. doi: 10.1177/1052562919851826

Langacker, R. W. (2013). *Essentials of Cognitive Grammar*. Oxford, NY: Oxford University Press.

Leet, D. R., and Lopus, J. S. (2003). *A Review of High School Economics Textbooks*. Available online at: https://ssrn.com/abstract=381760

Lenci, A. (2008). Distributional semantics in linguistic and cognitive research. *Riv. Linguist.* 20, 1–31.

Levy, O., Goldberg, Y., and Dagan, I. (2015). Improving distributional similarity with lessons learned from word embeddings. *Trans. Assoc. Comput. Linguist.* 3, 211–225. doi: 10.1162/tacl_a_00134

Lewis, D. (1969). *Convention: A Philosophical Study*. Cambridge, MA: Harvard University Press.

Lim, S. (2009). How and why do college students use Wikipedia? *J. Am. Soc. Inform. Sci. Technol.* 60, 2189–2202. doi: 10.1002/asi.21142

Link, G. (1983). "The logical analysis of plurals and mass terms: a lattice-theoretic approach," in *Meaning, Use, and Interpretation of Language, Grundlagen der Kommunikation und Kognition/Foundations of Communication and Cognition*, eds R. Bäuerle, C. Schwarze, and A. von Stechow (Berlin: de Gruyter), 302–323. doi: 10.1515/9783110852820.302

Linsky, L. (1971). "Reference and referents," in *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics and Psychology*, eds D. D. Steinberg and L. A. Jakobovitz (Cambridge: Cambridge University Press), 76–85.

Lodico, M. G., Spaulding, D. T., and Voegtle, K. H. (2006). *Methods in Educational Research: From Theory to Practice*. San Francisco, CA: Jossey-Bass.

Lucas, U., and Mladenovic, R. (2009). The identification of variation in students' understandings of disciplinary concepts: the application of the SOLO taxonomy within introductory accounting. *High. Educ.* 58, 257–283. doi: 10.1007/s10734-009-9218-9

Lücking, A. (2019). "Dialogue semantics: from cognitive structures to positive and negative learning," in *Frontiers and Advances in Positive Learning in the Age of InformaTiOn (PLATO)*, ed O. Zlatkin-Troitschanskaia (Cham: Springer), 197–205. doi: 10.1007/978-3-030-26578-6_15

Maurer, M., Schemer, C., Zlatkin-Troitschanskaia, O., and Jitomirski, J. (2019). "Positive and negative media effects on university students' learning: preliminary findings and a research program," in *Frontiers and Advances in Positive Learning in the Age of InformaTiOn (PLATO)*, ed O. Zlatkin-Troitschankskaia (Cham: Springer), 109–119. doi: 10.1007/978-3-030-26578-6_8

Mayer, R. E. (Eds.). (2005). "Introduction to multimedia learning," in *The Cambridge Handbook of Multimedia Learning* (Cambridge, NY: Cambridge University Press), 1–16. doi: 10.1017/CBO9780511816819.002

Mayer, R. E. (Eds.). (2014). "Cognitive theory of multimedia learning," in *The Cambridge Handbook of Multimedia Learning* (Cambridge, NY: Cambridge University Press), 43–71.

McEachern, W. A. (2012). "Macroeconomic principles are still relevant and still important," in *International Handbook on Teaching and Learning in Economics. Elgar Original Reference*, eds G. M. Hoyt and K. McGoldrick (Cheltenham: Edward Elgar), 413–422. Retrieved from: http://www.elgaronline.com/view/9781848449688.xml

McGoldrick, K., Rebelein, R., Rhoads, J., and Stockly, S. (2010). "Making cooperative learning effective for economics," in *Teaching Innovations in Economics*, eds M. Salemi and W. Walstad (Cheltenham; Northampton, MA: Edward Elgar), 65–94. doi: 10.4337/9780857930620.00011

Mehler, A., Gleim, R., Gaitsch, R., Uslu, T., and Hemati, W. (2020a). From topic networks to distributed cognitive maps: Zipfian topic universes in the area of volunteered geographic information. *Complexity* 4, 1–47. doi: 10.1155/2020/4607025

Mehler, A., Hemati, W., Welke, P., Konca, M., and Uslu, T. (2020b). Multiple texts as a limiting factor in online learning: quantifying (dis-)similarities of knowledge networks. *Front. Educ.* 5:206. doi: 10.3389/feduc.2020.562670

Mehler, A., Jussen, B., Geelhaar, T., Henlein, A., Abrami, G., Baumartz, D., et al. (2020c). The Frankfurt Latin Lexicon. From morphological expansion and word embeddings to SemioGraphs. *Stud. Saggi Linguist.* 58, 121–155. doi: 10.4454/ssl.v58i1.276

Mehler, A., and Ramesh, V. (2019). *TextInContext: On the Way to a Framework for Measuring the Context-Sensitive Complexity of Educationally Relevant Texts–A Combined Cognitive and Computational Linguistic Approach*. Cham: Springer International Publishing.

Meier, S. (2008). Is Wikipedia a credible source for undergraduate economics students? *Major Themes Econ.* 10, 79–105.

Meyer, J. H., and Land, R. (2005). Threshold concepts and troublesome knowledge (2). Epistemological considerations and a conceptual framework for teaching and learning. *High. Educ.* 49, 373–388. doi: 10.1007/s10734-004-6779-5

Meyer, J. H., and Land, R. (2013). "Threshold concepts and troublesome knowledge (1): linkages to ways of thinking and practising within the disciplines," in *Improving Student Learning. Theory and Practice-Ten Years On*, ed C. Rust (Oxford: Oxford Centre for Staff and Learning Development), 412–424.

Meyer, J. H., and Land, R. (Eds.). (2006). "Threshold concepts and troublesome knowledge: an introduction," in *Overcoming Barriers to Student Understanding: Threshold Concepts and Troublesome Knowledge* (London: Routledge), 3–18. doi: 10.4324/9780203966273

Meyer, J. H., and Shanahan, M. (2003). "The troublesome nature of a threshold concept in economics," in *Paper presented to the 10th Conference of the European Association for Research on Learning and Instruction (EARLI)* (Padova).

Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013). Efficient estimation of word representations in vector space. *CoRR* abs/1301.3781.

Miller, G. A., and Charles, W. G. (1991). Contextual correlates of semantic similarity. *Lang. Cogn. Process.* 6, 1–28. doi: 10.1080/01690969108406936

Millikan, R. G. (1984). *Language, Thought and Other Biological Categories*. Cambridge, MA: MIT Press.

Motos, R. M. (2011). "The role of interdisciplinary in lexicography and lexicology," in *New Approaches to Specialized English Lexicology and Lexicography*, ed I. B. Fernández (Newcastle upon Tyne: Cambridge Scholars Publishing), 3–13.

Mumm, M. (2015). *Kosten- und Leistungsrechnung. Internes Rechnungswesen für Industrie- und Handelsbetriebe, 2nd Edn*. Berlin; Heidelberg: Springer Gabler.

Murez, M., and Recanati, F. (2016). Mental files: an introduction. *Rev. Philos. Psychol.* 7, 265–281. doi: 10.3164/016-0314-3

Murphy, M. L. (2010). *Lexical Meaning. Cambridge Textbooks in Linguistics*. Cambridge, NY: Cambridge University Press.

Nagy, I. K. (2014). English for special purposes: specialized languages and problems of terminology. *Acta Univ. Sapient. Philol.* 6, 261–273. doi: 10.1515/ausp-2015-0018

Neurath, O. (1932). Protokollsätze. *Erkenntnis* 3, 204–214. doi: 10.1007/BF01886420

Neuweg, G. H. (2000). Mehr lernen, als man sagen kann: Konzepte und didaktische Perspektiven impliziten Lernens. *Unterrichtswissenschaft* 28, 197–217.

Newman, M. E. J. (2010). *Networks: An Introduction*. Oxford: Oxford University Press.

Nicola-Richmond, K., Pépin, G., Larkin, H., and Taylor, C. (2018). Threshold concepts in higher education: a synthesis of the literature relating to measurement of threshold crossing. *High. Educ. Res. Dev.* 37, 101–114. doi: 10.1080/07294360.2017.1339181

Oakhill, J., Cain, K., and Bryant, P. (2003). The dissociation of word reading and text comprehension: evidence from component skills. *Lang. Cogn. Process.* 18, 443–468. doi: 10.1080/01690960344000008

Pustejovsky, J. (1995). *The Generative Lexicon*. Cambridge, MA: MIT Press.

Reimann, N., and Jackson, I. (2006). "Threshold concepts in economics: a case study," in *Overcoming Barriers to Student Understanding. Threshold Concepts and Troublesome Knowledge, Chapter 8*, eds J. H. Meyer and R. Land (Abingdon; New York, NY: Routledge), 115–133.

Richardson, P. W. (2004). Reading and writing from textbooks in higher education: a case study from economics. *Stud. High. Educ.* 29, 505–521. doi: 10.1080/0307507042000236399

Rincke, K. (2010). Alltagssprache, Fachsprache und ihre besonderen Bedeutungen für das Lernen. *Z. Didaktik Naturwissensch.* 16, 235–260.

Roberts, C. (2019). "Contextual influences on reference," in *The Oxford Handbook of Reference, Chapter 13*, eds B. Abbott and J. Gundel (Oxford: Oxford University Press), 260–282. doi: 10.1093/oxfordhb/9780199687305.013.13

Roelcke, T. (2010). *Fachsprachen. Number 37 in Grundlagen der Germanistik, 3rd Edn*. Berlin: Schmidt.

Russell, B. (1919). *Introduction to Mathematical Philosophy, 2nd Edn*. London: George Allen & Unwin.

Sahlgren, M. (2008). The distributional hypothesis. *Ital. J. Linguist.* 20, 33–54.

Salmerón, L., Kintsch, W., and Cañas, J. J. (2006). Reading strategies and prior knowledge in learning from hypertext. *Mem. Cogn.* 34, 1157–1171. doi: 10.3758/BF03193262

Schuhen, M., and Kunde, F. (2016). "Informelles Lernen und ökonomische Bildung," in *Handbuch Informelles Lernen*, ed M. Rohs (Wiesbaden: Springer Fachmedien), 455–466. doi: 10.1007/978-3-658-05953-8_35

Schumann, S., Eberle, F., Oepke, M., Pflüger, M., Gruber, C., Stamm, P., et al. (2010). *Inhaltsauswahl für den Test zur Erfassung ökonomischen Wissens und Könnens im Projekt "ökonomische Kompetenzen von Maturandinnen und Maturanden (oekoma)"*. Technical report, Universität Zürich, Institut für Gymnasial- und Berufspädagogik.

Searle, J. R. (1969). *Speech Acts*. Cambridge: Cambridge University Press.

Searle, J. R. (2006). Social ontology: some basic principles. *Anthropol. Theory* 6, 12–29. doi: 10.1177/1463499606061731

Sender, T. (2017). *Wirtschaftsdidaktische Lerndiagnostik und Komplexität*. Lokalisierung liminaler Unsicherheitsphasen im Hinblick auf Schwellenübergänge. Komplexität, Entrepreneurship und Ökonomische Bildung (Zugl. dissertation). Springer Gabler, TU Dortmund, Wiesbaden, Germany.

Shanahan, M. P., Foster, G., and Meyer, J. H. (2006). Operationalising a threshold concept in economics: a pilot study using multiple choice questions on opportunity cost. *Int. Rev. Econ. Educ.* 5, 29–57. doi: 10.1016/S1477-3880(15)30119-5

Simkins, S. P. (1999). Promoting active-student learning using the world wide web in economics courses. *J. Econ. Educ.* 30, 278–287. doi: 10.1080/00220489909595990

Spärck Jones, K. (1972). A statistical interpretation of term specificity and its application in retrieval. *J. Document.* 28, 11–21. doi: 10.1108/eb026526

Steffens, Y., Schmitt, I. L., and Amann, S. (2017). *Mediennutzung Studierender: über den Umgang mit Medien in hochschulischen Kontexten-Systematisches Review nationaler und internationaler Studien zur Mediennutzung Studierender*. Köln: Universität zu Köln.

Strevens, P. (1977). Special-purpose language learning: a perspective. *Lang. Teach. Linguist. Abstr.* 10, 145–163. doi: 10.1017/S0261444800003402

Tinkler, S., and Woods, J. (2013). The readability of principles of macroeconomics textbooks. *J. Econ. Educ.* 44, 178–191. doi: 10.1080/00220485.2013.770345

Tuldava, J. (1998). *Probleme und Methoden der quantitativ-systemischen Lexikologie*. Trier: Wissenschaftlicher Verlag.

Uslu, T., Mehler, A., and Baumartz, D. (2019). "Computing classifier-based embeddings with the help of text2ddc," in *Proceedings of the 20th International Conference on Computational Linguistics and Intelligent Text Processing*, (*CICLing 2019*), *CICLing 2019* (La Rochelle).

van den Broek, P., Rapp, D. N., and Kendeou, P. (2005). Integrating memory-based and constructionist processes in accounts of reading comprehension. *Discourse Process.* 39, 299–316. doi: 10.1080/0163853X.2005.9651685

van Mourik, G., and Wilkin, C. L. (2019). Educational implications and the changing role of accountants: a conceptual approach to accounting education. *J. Vocat. Educ. Train.* 71, 312–335. doi: 10.1080/13636820.2018.1535517

Vaňková, L. (2018). "Fachsprachen und der Alltag: eine Untersuchung anhand der deutschen Tagespresse," in *Zentrum und Peripherie: aus sprachwissenschaftlicher Sicht*, eds V. Kotůlková and G. Rykalová (Opava: Slezská Univerzita v Opavě), 51–64.

Vernooij, F. (2000). "Tracking down the knowledge structure of students," in *Business Education for the Changing Workplace, Number 5 in Educational Innovation in Economics and Business*, eds L. Borghans, W. H. Gijselaers, R. G. Milter, and J. E. Stinson (Dordrecht: Kluwer), 437–450.

Vidal, N., Smith, R., and Spetic, W. (2015). Designing and teaching business & society courses from a threshold concept approach. *J. Manag. Educ.* 39, 497–530. doi: 10.1177/1052562915574595

von Weizsäcker, E. U. (1974). "Erstmaligkeit und Bestätigung als Komponenten der pragmatischen Information," in *Offene Systeme I. Beiträge zur Zeitstruktur von Information, Entropie und Evolution*, ed E. U. von Weizsäcker (Stuttgart: Klett), 82–113.

Weber, W., Kabst, R., and Baum, M. (2014). *Einführung in die Betriebswirtschaftslehre, 9th Edn*. Wiesbaden: Springer Gabler.

Widdowson, H. (1974). Literary and scientific uses of English. *English Lang. Teach. J.* 28, 282–292. doi: 10.1093/elt/XXVIII.4.282

Wierzbicka, A. (1996). *Semantics: Primes and Universals*. Oxford: Oxford University Press.

Wilkinson, D. M., and Huberman, B. A. (2007). "Cooperation and quality in Wikipedia," in *Proceedings of the 2007 International Symposium on Wikis, WikiSym '07* (Montreal, QC), 157–164. doi: 10.1145/1296951.1296968

Wittgenstein, L. (1984). *Tractatus Logico-Philosophicus; Tagebücher 1914–1916; Philosophische Untersuchungen. Number 1 in Werkausgabe*. Frankfurt am Main: Suhrkamp.

Zeevat, H. (2018). "Interpreting dependent NPs," in *Proceedings of Cognitive Structures: Linguistic, Philosophical and Psychological Perspectives, CoSt'18* (Düsseldorf).

Zipf, G. K. (1949). *Human Behavior and the Principle of Least Effort. An Introduction to Human Ecology*. Cambridge, MA: Addison-Wesley.

Zlatkin-Troitschanskaia, O., Toepper, M., Pant, H. A., Lautenbach, C., and Kuhn, C. (Eds.). (2018). *Assessment of Learning Outcomes in Higher Education: Cross-National Comparisons and Perspectives*. Wiesbaden: Springer International Publishing.