Words and meaning: How the lexical encoding of technical concepts contributes to their mental representation

Permalink
https://escholarship.org/uc/item/6td9j0kr

Journal

ISSN
1069-7977

Authors
Jucks, Regina
Paus, Elisabeth

Publication Date
2009

Peer reviewed
Words and Meaning: How the Lexical Encoding of Technical Concepts Contributes to their Mental Representation

Elisabeth Paus (paus@paed.psych.uni-frankfurt.de)
Goethe University, Institute of Psychology, Senckenberganlage 15
60325 Frankfurt, Germany

Regina Jucks (jucks@paed.psych.uni-frankfurt.de)
Goethe University, Institute of Psychology, Senckenberganlage 15
60325 Frankfurt, Germany

Abstract
Some specialist concepts can be encoded either with words of Greek or Latin origin or in everyday language terms. These synonyms are expected to indicate to the same underlying meaning. This study investigates whether the origin of technical terms influences their mental representation as well as subjective feelings of knowing. The linguistic encoding of specialist concepts can be assumed to impact the connotation and complexity transmitted to the recipient. As expected, everyday language technical terms (ELTT) were perceived to be more easily defined, more familiar, and more easily accessible than technical terms of Greek and Latin origin (GLTT). Consequently, participants estimated ELTT to be better understood. However, there was no specific effect of encoding on the difference between subjective and objective learning parameters. Theoretical and practical implementations for communication in learning contexts are discussed.

Keywords: lexical encoding; technical terms; explanatory knowledge; mental representation; feeling of knowing

Introduction
The widespread use of new media (Goldberg, Russell, & Cook, 2003; MacArthur, 2006) provides laypeople with almost unrestricted access to specialist information from a multitude of fields. For instance, readers looking for health information on the internet are likely to come across words such as “serotonin” or “transmitters.” The presentation of knowledge is influenced by the selection and usage of appropriate specialist vocabulary. Technical terms can be considered the “building blocks of knowledge”; they form the core of its content.

In general, words may be more or less complex. Complexity can be defined in terms of the following features (Nückles & Bromme, 2002): Firstly, a more complex word is related to many other words. Secondly, a more complex word integrates many different aspects and, thirdly, it can be described on different levels. The last feature, in particular, implies that complex words can be used more or less specifically. A word can be described by its intension (defining features of the word) and/or its extension (concepts related to the term that share the same features) (Weingartner, 1973). Bromme, Rambow, and Wiedmann (1998) emphasize that specialist concepts are mostly both extensionally wide-ranging (a term applies to many different reference objects) and intentionally rich (embedded in complex theories). Accordingly, specialist concepts are at risk of being subjectively understood differently, dependent on the reader’s training and the term’s function, context, and usage (Bromme, Rambow, & Nückles, 2001). As findings in research on medical expertise show, the conceptual meaning underlying a technical term is differently specified by experts and laypersons: whereas experts know the complex specialist knowledge that is encapsulated by technical terms (Schmidt & Boshuizen, 1992), laypersons may have a rather vague or erroneous understanding of what technical terms mean (Gittelman, Mahabee-Gittens, & Gonzales-del-Ray, 2004).

Everyday language does not bear the same risks. Although the meaning underlying the term “man” may be represented differently between interlocutors in everyday communication, the shared meaning suffices for successful communication (“The man is standing on the street.”). The word “man” simply represents what it refers to: a man. In the very apt words of Gertrude Stein (1999): “Rose is a rose is a rose is a rose”.

In the context of technical terms, the underlying meaning and the interlocutors’ mental representations may differ. For example, a medical doctor talking about “migraine” may be referring to different aspects than a layperson (Jucks & Bromme, 2007). Concepts are the central components of thinking. Thus, the terms chosen to encode specialist concepts may be assumed to impact the connotation and complexity transmitted to the recipient. When references touch the core of concepts, differences in representation may become relevant (the vagueness of the linguistic surface).

In everyday German, many words of Latin or Greek origin – particularly those introduced into German over the last 400 years – have synonyms of German origin. Thus, in his categorization of technical terms in German, Bromme (1996) differentiates between everyday language terms in specific linguistic usage and loan words borrowed from classical or modern scientific languages. Laypersons recognize the need for a deep and thorough understanding of technical terms, particularly those with a Latin or Greek origin. However, experts and laypersons may fail to recognize that technical terms on the threshold to everyday language have different meanings in the specific field and in general (e.g., Schorling & Saunders, 2000).
Laypersons base their perceived comprehension of a text on surface characteristics such as the frequency of technical terms. They overestimate what they are able to understand. Research in the domain of text comprehension has shown that laypersons often have problems recognizing that they have failed to understand the content of a text correctly (illusion of knowing, Glenberg & Epstein, 1985; Glenberg, Wilkinson, & Epstein, 1982).

Whereas Glenberg and colleagues concentrated on how people deal with newly learned content, the study we present in this paper focuses on subjective evaluations of existing knowledge. It has been shown that laypersons tend to overestimate their understanding of texts including technical terms encoded in everyday language. Laypersons were asked to compare the comprehensibility of two versions of a text in which the key specialist concepts were encoded either in everyday language or in loan words. As expected, the version using loan words was judged to be less comprehensible (Clark, Weinberger, Jucks, Spitalnik, & Wallace, 2003). Furthermore, when asked to anticipate the knowledge of others, people tend to overestimate the generality of their own knowledge (overestimation hypothesis; see, e.g., Bromme et al., 2001; Hayes & Bajzek, 2008).

The extent to which people overestimate their knowledge depends on the type of knowledge concerned. Being able to explain the meaning of a certain term usually presupposes not only knowledge of facts, but also more complex patterns of knowledge – so-called explanatory knowledge (Rozenblit & Keil, 2002). Research has shown that people tend to overestimate their level of this kind of knowledge, in particular (illusion of explanatory depth, Mills & Keil, 2004).

Several factors influence subjective knowledge estimation (Rozenblit & Keil, 2002): In contrast to simple nomination or description of knowledge contents, explanations are complex hierarchical systems with no explicit starting or end points. Thus, there is no explicit criterion for evaluating the completeness and correctness of a given explanation. Moreover, the quality of an explanation is defined in terms of conclusiveness and traceability, two attributes that are difficult for a layperson to estimate. Furthermore, explanations of technical terms are reproduced less frequently then, for example, facts or stories; therefore, it is hard for laypersons to estimate their expertise. Finally, laypersons often acquire their knowledge in a specific context. Even if a term is understood correctly in that context, this does not automatically imply that it will be reliably understood and that it can be explained in other contexts. When a layperson is asked to assess his or her subjective comprehension of a term, this judgment is not usually based on strong evidence. All of these factors affect laypersons’ subjective estimations of their knowledge of technical terms.

How else can lexical encoding be expected to impact the ability to gauge one’s knowledge correctly? One approach that highlights the dissociation between subjective knowledge estimations and objective knowledge indications is feeling of knowing (FOK, e.g., Koriat & Levy-Sadot, 2001). According to this approach, persons have a feeling of whether or not knowledge contents are familiar to them, even if they cannot recall them at that particular moment. Koriat and Levy-Sadot (2001) introduced two models to explain FOK: the cue familiarity model and the cue accessibility model. The cue familiarity model postulates that the FOK estimation is determined by the familiarity of the term itself. The cue accessibility model postulates that it is determined by the accessibility of information relating to that term (whether or not that information is correct being largely irrelevant). The FOK literature focuses on the ability to recall knowledge of facts (e.g., recall of a name). However, the models can also be applied to estimations of knowledge of technical terms with more comprehensive underlying knowledge structures. Both the perceived familiarity of technical terms and the accessibility of related information can be expected to influence subjective estimations of knowledge (Rozenblit & Keil, 2002).

As previously mentioned, many specialist concepts can be encoded in either everyday language or in loan words. For laypersons, the perceived familiarity of ELTT can be assumed to be higher, given that these are also used in general and colloquial language. The term Zuckerkrankheit (literally, “sugar disease”), for example, is listed in medical encyclopedias (Häcker & Stapf, 2004), but also used in everyday discourse, mostly without explicit knowledge of its medical specifics. As such, laypersons are more familiar with this term than with its technical synonym Diabetes. Koriat and Levy-Sadot (2001) showed that the models of cue familiarity and cue accessibility, originally thought to be mutually exclusive, are in fact interconnected. If a term is perceived to be highly familiar, information relating to it is expected to be more easily accessible. Based on the cue accessibility model, it can therefore be predicted that information relating to an ELTT is also more accessible for laypersons than is information relating to a GLTT.

In this study, we examine the extent to which the lexical encoding of a specialist concept influences its mental representation. We tested the following hypotheses:

(I) Technical terms of Greek or Latin origin (GLTT) are more clearly categorized as specialist concepts than are everyday language technical terms (ELTT).

(II) ELTT are perceived to be more easily defined than GLTT.

(III) ELTT are judged to be more familiar than GLTT.

(IV) Information relating to ELTT is more easily accessible than information relating to GLTT.

According to Koriat and Levy-Sadot (2001), information relating to a term is more easily accessible if the term is perceived to be familiar. Consequently, FOK should be stronger for ELTT than for GLTT. Koriat and Levy-Sadot (2001) postulate that FOK is influenced solely by the number of associations activated and not by the accuracy of
those associations. Therefore, we expected participants to produce longer – but not more accurate – explanations for ELTT than for GLTT. Moreover, we are interested in the accuracy of FOK judgments. First, we expected participants to overestimate their knowledge concerning the underlying meaning of specialist terms in general. Additionally, based on the idea that FOK is higher for ELTT, but that objective explanations of these terms are not of better quality, we predicted participants’ confidence judgments to be even less accurate when concepts were encoded in ELTT.

(V) FOK is higher for ELTT; therefore, the explanations produced for ELTT will be longer (but not of higher quality) than those produced for GLTT.

(VI) Knowledge concerning the meaning of specialist terms is overestimated; therefore, FOK judgments are less accurate for ELTT.

Method

Participants and Design

Participants in this first data collection within an ongoing study were 23 psychology students (17 female, 6 male) at the University of Frankfurt with a mean age of 25.57 years (SD = 3.09). Of the 23 participants, 18 were German native speakers, 2 had spoken German as their second native language since childhood, and 3 had spoken German for more than 5 years. The independent variable “encoding of technical terms” was conceptualized on two levels: “everyday language” versus “foreign language.”

Material

Selection of Technical Terms 17 technical terms of Greek or Latin origin were selected and matched with everyday language synonyms (see Table 1). Based on the entries in technical encyclopedias (Häcker & Stapf, 2004; Hildebrandt, 2004; Wilpert, 2001; Woll, Vogl, & Weigert, 2000), all term pairs were defined as synonyms encoding the same underlying concepts. Terms were selected according to the following criteria: First, both terms in each pair of synonyms had separate entries in the technical encyclopedias specified. To avoid ambiguity, we ensured that the terms selected could be explicitly assigned to one technical area. Terms included in a dictionary of foreign words were categorized as GLTT (Knauf & Löcher, 2004; Währing, 2007; Wermke, Klosa, Kunkel-Razum, & Scholze-Stubenrech, 2001). We also checked that all terms featured in the major German-language spelling dictionary, the Duden (Wermke, Kunkel-Razum, & Scholze-Stubenrech, 2004); this applied to all but one of the terms. The 17 pairs of synonyms comprised 5 from the field of medicine, 4 from psychology, 4 from business administration, and 5 from German language and literature studies.

Questionnaire Construction Two versions of the questionnaire were constructed to ensure that GLTTs and their ELTT synonyms were not presented simultaneously, potentially influencing the evaluation. In version A, the order of words (9 GLTTs and 8 ELTTs) was randomized. In version B, the respective synonyms were presented in parallel order.

First, the participants were asked to rate 6 items for each of the 17 terms on a 5-point scale (agree strongly – disagree strongly) to evaluate the terms’ categorization, familiarity, perceived comprehension, and definitional context. Second, the participants’ objective knowledge of the terms was assessed through two open questions (for a detailed description of all variables, see the Dependent Measures section below).

Finally, demographic data were obtained: age, gender, degree program, number of semesters studied, occupational and educational qualifications, knowledge of foreign languages, and whether German was the native language.

Table 1: Examples of technical term pairs

<table>
<thead>
<tr>
<th>ELTT</th>
<th>GLTT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zuckerkrankheit (“sugar disease”)</td>
<td>Diabetes</td>
</tr>
<tr>
<td>Kleinhirn (“little brain”)</td>
<td>Cerebellum</td>
</tr>
<tr>
<td>Unternehmensführung (“company leadership”)</td>
<td>Management</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Procedure

Data were collected in a lecture room at the University of Frankfurt. Version A was randomly distributed to 11 participants; version B to 12 participants. The participants completed the questionnaires without any time constraints. An investigator was present throughout data collection to answer questions.

Dependent Measures

A survey was developed to access mental representation and feeling of knowing. Items measuring definitional context and subjective estimation of comprehension were formulated in such a way as to ensure that participants’ responses referred to explanatory knowledge (Rozenblit & Keil, 2002).

Mental Representation. The first dependent variable, mental representation, reflected the categorization, definitional context, familiarity, and accessibility of the terms.

Categorization. Participants were asked to judge whether or not each term was technical.

Definitional Context. Participants’ perception of the definitional context in general was assessed by the item “needs to be explained to be understood.” Two items assessed the range of the definitional context, i.e., whether a term can “be described on several levels” (Nückles & Bromme, 2002). A low defining context was operationalized
by the statement “can be explained in a few words”; a higher defining context by the statement “needs to be understood well to be explained.”

**Familiarity.** Participants were asked to rate their perceived familiarity with each term. Based on the cue familiarity model (Koriat & Levy-Sadot, 2001), this item allows us to examine whether participants indeed judged ELTT to be more familiar.

**Accessibility.** Participants were asked to note down all terms they could think of that related to each term given. According to the cue accessibility model (Koriat & Levy-Sadot, 2001), the accessibility of terms can be determined by the number of related terms identified.

**Feeling of Knowing.** The second dependent variable, feeling of knowing (Koriat & Levy-Sadot, 2001), was assessed in two ways:

**Subjective Estimation of Comprehension.** Participants were asked to rate their perceived comprehension of the terms.

**Explanations of specialist terms.** The participants were prompted to explain each specialist term as follows: “What does the term XY mean? Please give a brief explanation.” The number of words used to explain each term was ascertained. The completeness and the correctness (i.e., quality) of participants’ answers were assessed by two raters (following the definitions in the technical encyclopedias used). The raters also took into account whether and how participants illustrated the concepts by providing examples. Interrater reliability for the quality of explanations of ELTT (K = 0.95, p < .001) and of GLTT (K = 0.95, p < .001) was satisfactory.

**Results**

Unless otherwise indicated, analyses were performed using SPSS and the statistical assumptions were met. For all further statistical analyses, we used the aggregated means of participants’ responses for each term. We were thus able to compare the term pairs directly.

**Mental Representation**

A multivariate analysis yielded a large main effect of type of encoding on mental representation, $F(1,17) = 5.23, p < 0.05, \eta^2_p = .74$. (following Cohen, 1988, we interpreted the effect sizes as follows: $\eta^2_p < .06$ represents a small effect, $.06 < \eta^2_p < .13$ a medium effect, and $\eta^2_p > .13$ a large effect).

Univariate analysis showed a main effect of encoding on categorization, $F(1,17) = 26.71, p < 0.001, \eta^2_p = .63$, with GLTT being more clearly categorized as technical terms ($M = 4.27, SD = 0.81$) than ELTT ($M = 2.82, SD = 1.02$).

A further univariate analysis revealed a main effect of the item “the term needs to be explained to be understood,” $F(1,17) = 12.67, p < 0.05, \eta^2_p = .44$, with the need for explanation being rated higher for GLTT ($M = 4.06, SD = 0.77$) than for ELTT ($M = 3.07, SD = 0.99$).

Results for the two items “can be explained in a few words” and “needs to be understood well to be explained” were mixed. There was no difference for the first item, representing a low defining context, $F(1,17) = 0.282, ns$; the mean value for all participants was $M = 3.66, SD = 0.41$. However, for the second item, representing a high defining context, there was a main effect, $F(1,17) = 5.31, p < 0.05, \eta^2_p = .25$, with higher endorsements of the item for GLTT ($M = 3.57, SD = 0.51$) than for ELTT ($M = 3.11, SD = 0.80$).

Univariate analysis revealed a main effect of encoding on familiarity, $F(1,17) = 3.59, p < 0.05, \eta^2_p = .18$, as well as on accessibility, $F(1,17) = 4.40, p < 0.05, \eta^2_p = .22$, of technical terms. As predicted, ELTT were rated to be more familiar ($M = 4.71, SD = 0.70$) than GLTT ($M = 4.35, SD = 0.83$) as well as more easily accessible ($M = 3.02, SD = 1.12$) than GLTT ($M = 2.31, SD = 1.31$).

**Feeling of Knowing**

In line with the findings of higher familiarity and accessibility of ELTT, there was a main effect for the subjective estimations of comprehension, $F(1,17) = 5.52, p < 0.05, \eta^2_p = .26$, with ELTT being rated as better understood ($M = 4.60, SD = 0.72$) than GLTT ($M = 4.11, SD = 1.02$). Furthermore, a main effect emerged for length of explanation, $F(1,17) = 3.87, p < 0.05, \eta^2_p = .20$, with more words being produced to explain ELTT ($M = 10.76, SD = 3.63$) than GLTT ($M = 9.01, SD = 3.00$). Explanations of ELTT were also of better quality ($M = 2.82, SD = 0.69$) than explanations of GLTT ($M = 2.36, SD = 0.74$), $F(1,17) = 6.11, p < 0.05, \eta^2_p = .28$. Furthermore, we examined accuracy of FOK judgments for specialist terms in general as well as the effects of condition on accuracy of FOK judgments in particular. A MANOVA revealed a main effect for encoding of specialist terms, $F(2,16) = 7.31, p < 0.05, \eta^2_p = .31$. Furthermore, there was a main effect for understanding, with participants rating their comprehension of specialist terms subjectively higher than it objectively was, $F(2,16) = 266.02, p < 0.001, \eta^2_p = .94$. However, no interaction effect emerged between encoding of specialist terms and understanding, $F(2,16) = 0.02, ns$.

**Discussion**

Overall, the empirical findings supported our hypotheses concerning the impact of the linguistic encoding of specialist concepts on different aspects of their mental representation and perceived comprehension.

**Mental Representation**

Our findings show that GLTT are more clearly categorized as technical terms than are ELTT. Furthermore, the GLTT were perceived to be less easily defined. Although ratings of a low definitional context were equal for both types of encoding, a main effect was found for high defining context. In other words, people are better able to recognize that the connotation and complexity transported by terms can vary for GLTT terms than for ELTT terms.

Based on the cue familiarity and cue accessibility models introduced by Koriat & Levy-Sadot (2001), we hypothesized that ELTT would be rated as more familiar...
and – provided that this was indeed the case – would be more easily accessible. Our findings confirmed these hypotheses. However, it remains unclear which additional factors – beyond word origin – contribute to these results. One explanation may be frequency of word use: Because people use ELTT more often (e.g., in private conversations), or are at least repeatedly exposed to these terms by diverse media in everyday life, they are more present, and people are thus more likely to think that they know their deeper meaning. Future research should also investigate the role of phonological familiarity, i.e., whether the sounds used in ELTT terms are more familiar and therefore more easily encoded.

**Feeling of Knowing**

Given that ELTT are judged to be more familiar and are more easily accessible, we expected their subjective comprehension to be higher. Additionally, because comprehension of a technical term can be represented by how well people think they can explain its underlying meaning, we compared the number of words generated to explain each term. This approach is based on the idea that a person might think they have to produce more words to explain a term that seems to be more familiar. Our findings support our hypothesis that ELTT are perceived to be better understood and revealed that participants – in accordance with their subjective FOK – were able to produce more words in explanation of ELTT. However, in contrast to our predictions, the quality of explanations of ELTT was in fact higher than that of GLTT. People may be more confident about their knowledge of ELTT and therefore dare to write down everything they know, but more critical of their knowledge of GLTT, such that their writing is inhibited. Finally, our results show that people tend to overestimate their knowledge concerning the meaning of specialist terms in general. However, whereas findings concerning text comprehensibility confirm an influence of lexical encoding on perceived comprehension (Clark et al., 2003), the results of this study show no effects. This finding can be interpreted as follows: In measures of text comprehensibility, perceived comprehension of specialist terms mainly becomes relevant in reference to the understanding of the whole sentence – influenced by several other features. In contrast, in our study, we focused exclusively on the comprehension of specialist terms without a textual embedding. The context in which a technical term is used seems to crucially influence the effect of encoding on the accuracy of comprehension judgments. In a textual context, guessing the (superficial) meaning of terms may be facilitated. As a result, comprehensibility of the whole text is often overestimated. In contrast, when people focus on their subjective and objective understanding of specialist terms per se, they might reflect their specific knowledge more in detail (see above). In conclusion, further research can benefit from embedding technical terms in more realistic settings, such as learning scenarios at university, to declare their specific role in regard to comprehension judgments.

**Summary**

In summary, the findings of this study serve two purposes:

(I) They place the FOK model within a broader frame of reference. Whereas most recent studies on FOK have investigated the retrieval of knowledge of facts, our results indicate that the processes underlying FOK can be transferred to terms that transport more complex knowledge contents (e.g., explanatory knowledge; Rozenblit & Keil, 2002).

(II) Our findings provide insights into the usage of words in texts and lectures, for example. The results of this study indicate that the lexical encoding of specialist concepts, as well as the context in which they are embedded, influences people’s estimations of their knowledge. When ELTT are used in specialist contexts, laypeople may not be aware of the deeper meaning underlying a term. However, using GLTT may complicate or even hinder understanding. Hence, research investigating the specific contexts in which each kind of encoding makes sense is warranted.

Our findings need to be replicated in larger samples of participants from different social backgrounds. Further, more technical terms from different technical areas need to be investigated.

Another field influenced by the encoding of technical concepts is (computer-supported) collaborative learning. Virtual discourse plays an important role in knowledge co-construction in these learning environments (Häkkinen & Järvelä, 2006; Mäkitalo, Weinberger, Häkkinen, Järvelä, & Fischer, 2005). When communicating virtually or face-to-face, learning partners tend to use recently introduced words; in other words, they adapt linguistically to each other (Jucks, Becker, & Bromme, 2008). This alignment of terminological vocabulary between the interlocutors during communication is called lexical alignment (Pickering & Garrod, 2004). As shown in this study, the meaning transported by a technical term may depend on its encoding. Using the same technical terms in collaborative learning situations may therefore hold the risk that meaning is exchanged only on a superficial level, and that the underlying deeper meaning is not discussed: ELTT may fail to transport the complexity of the underlying meaning, whereas GLTT may imply a level of expert knowledge that does not exist.

To conclude, the linguistic encoding of specialist concepts plays a pivotal role in individual knowledge representation and can thus be expected to influence learning engagement and processes.

**References**


Acknowledgments

The authors are thankful to Lena Opitz, Stefanie Diener, and Franziska Thon for their support with data collection and analysis and to Susannah Goss for English language editing.