Semantic Density and Gravity in Lay-oriented Medical Knowledge Communication II: The Case of Patient Information Leaflets

Aage Hill-Madsen (Aalborg University)

Abstract
This article investigates a lay-oriented medical genre, viz., the small type of brochure (the so-called *Patient Information Leaflet* (PIL)) that accompanies the packaging of medicinal products, providing basic information about the drug in question, particularly its use. The study draws a semantic profile of the texts in order to establish the character, and in particular the degree of challenge for lay readers, of the knowledge communication taking place in the genre. For analytical methodology, a framework derived from Legitimation Code Theory is used to analyze *semantic density* (complexity of meaning) and *semantic gravity* (strengths of context-dependence), on the assumption that semantically dense and context-independent texts may be challenging for certain groups of readers. Analytical results reveal that very different strengths of semantic density are represented in all texts, and the same is true for semantic gravity. Based on the analytical results, the PIL genre is then profiled in terms of its *semantic codes*. The conclusion is that PILs are a highly composite genre that integrates a multitude of codes, thus requiring a relatively high degree of versatility in coding orientation on the part of readers.

Keywords: medical discourse; expert-to-lay communication; knowledge mediation; semantic density; semantic gravity

1. Introduction

During the last couple of decades, healthcare systems (at least in the Western world) have become increasingly patient-oriented. Gone are the days when the role allocated to patients was that of passive receivers of treatment, dispensed by uncommunicative health practitioners. Today, much greater emphasis is placed on *patient engagement* and *patient-*

Hill-Madsen, Aage. 2025. 'Semantic Density and Gravity in Lay-oriented Medical Knowledge Communication II: The Case of Patient Information Leaflets.' *Nordic Journal of English Studies* 24(1): 41–74. https://doi.org/10.35360/njes.v24i1.41116.

centered care, making healthcare 'a partnership among practitioners, patients, and their families (when appropriate) to ensure that decisions respect patients' wants, needs, and preferences and that patients have the education and support they need to make decisions and participate in their own care' (Carman et al. 2013: 223–224). Research indicates that patient engagement enhances the quality of healthcare and leads to improved treatment outcomes (Coulter and Ellins 2007; Epstein and Street 2008). However, while, undoubtedly, greater patient involvement places higher demands on healthcare professionals, it can only add—*ceteris paribus*—to the demands made on patients also, since patient-centeredness presupposes *patient activation*, i.e., activation of 'an individual's knowledge, skill, and confidence for managing his/her own health and health care' (Hibbard and Mahoney 2010: 377). If, thus, treatment outcomes are taken to be, at least partially, contingent on the patient's own resources and active participation, then 'patienthood' becomes tinged with questions of failure or success. In other words, failure or success of treatment are outcomes for which the patient him-/herself is made partly responsible, and so the relevant question is, with a phrase borrowed from K. Maton (2014: 11): what are 'the rules of the game?', i.e., what does it take, or what sort of resources are required to achieve success, with 'success' to be understood as an optimal care trajectory?

Part of the answer to this question may be found in the concept of *health literacy*, defined by the US Government's Center for Disease Control and Prevention (2021) as 'the degree to which individuals have the ability to find, understand, and use information and services to inform health-related decisions and actions for themselves and others'. However, this definition (still) places the burden on patients, so to speak, and ignores the fact that different sources of information, even ones that are aimed at the same target audience, are likely to differ in the type of linguistic demands made on receivers. An obvious difference is the one between oral and written genres, in so far as certain groups of patients are likely to be far more comfortable with the semantic style of the former than the latter. However, even genres belonging to the same medium can be expected to vary. The present study, therefore, is part of a larger project, initiated with Hill-Madsen (2022), investigating the types of semantic resources presupposed in a number of written lay-oriented genres within the field of medicine. In the same way as Hill-Madsen (2022) investigated the *EPAR – summary for the public* genre from this perspective, the research object

of the present article is the *Patient Information Leaflet* (PIL) genre, i.e., the small brochure found inside the packaging of medicinal products. The two genres are to some extent parallel ones, in that for every medicinal product authorized for marketing in the EU, a text from each genre is published, containing different types of information about the drug.

In research literature, the PIL genre has often been investigated for potential comprehensibility problems (e.g., Askehave and Zethsen 2003; Hill-Madsen 2019a; Bjerrum and Foged 2003; Bradley et al. 1994; Clerehan and Buchbinder 2006; Hadjipavlou et al. 2013; Harwood and Harrison 2004; Pander Maat and Lentz 2010; Ross et al. 2006). Such research, however, has mostly taken a purely lexicogrammatical approach, e.g., by identifying surface forms such as lexical items deemed unfamiliar to target readers. In contrast, the present investigation is focused on the semantic stratum, aiming to identify the type(s) of semantic codes (see section 3) inherent in the texts.

Methodologically, the approach will be exactly the same as in Hill-Madsen (2022): A small selection of texts from the PIL genre (for selection criteria, see section 2) will be analyzed on the basis of the framework proposed by Maton and Doran (2017a; 2017b). The framework has been inspired by discussions between Legitimation Code Theory (LCT) and Systemic-Functional Linguistics (SFL) (see, e.g., Martin and Maton 2017; Maton et al. 2016), and is designed for the profiling of texts in terms of *semantic gravity* (SG) and *semantic density* (SD) (to be elaborated on in section 3). Analytical results will be presented in section 4 and compared with those relating to the *EPAR – summary for the public* genre. Based on these results, the PIL genre will be characterized in terms of semantic code(s) (also in section 4), and the implications for comprehensibility will be discussed in section 5.

2. Data: The PIL genre

Providing information and instructions relevant to the end users of medicinal products (EMA 2021), the PIL genre may be seen as one of the clearest manifestations of patient involvement in the written medium. In existence in the EU since 1993, the genre, along with the *EPAR – summary for the public*, is mandated by EU legislation (European Parliament and Council 2001), which prescribes a completely fixed move structure, each

with a clearly defined content (see EMA 2021). The individual sections are the following:[1]

0. *General introduction*
1. *What X is and what it is used for* (specifying the medical condition for which the drug is prescribed and how the drug works).
2. *What you need to know before you take X* (identifying circumstances (e.g., other medical conditions) under which the drug should not be taken or under which precautions should be taken if the drug is used, e.g., in connection with pregnancy).
3. *How to take X* (providing instructions in administration of the drug).
4. *Possible side effects*
5. *How to store X* (detailing storage conditions).
6. *Contents of the pack and other information* (specifying the chemical ingredients of the drug).

To enable comparison with the findings in Hill-Madsen (2022), PILs for the exact same medicinal products have been selected (six texts altogether) as those for which *EPAR summaries* were selected in Hill-Madsen (2022).[2]

3. Theoretical framework: A topology of semantics

As briefly mentioned in the Introduction, this article is concerned with profiling the PIL genre in terms of *semantic gravity* (SG) and *epistemic-semantic density* (SD). Since the two concepts have been relatively elaborately defined in Hill-Madsen (2022), the definitions provided here (3.1) will be slightly more condensed. Section 3.2 adds a further theoretical dimension, viz., the concept of *semantic codes*.

3.1. Semantic gravity and density

The concepts of SG and SD originate in *Legitimation Code Theory*, which is a sociological theory (of education), and not a linguistic theory. SG and

[1] To view the PILs analysed in this study, use the links under References, Primary sources, and select the final section of the pdf entitled B. PACKAGE LEAFLET.
[2] For bibliographical details of the sampled texts, see References, Primary sources. In the analytical section (section 4), the texts will be referred to by the name of the medicinal product.

SD are thus sociological concepts that can be used to analyse any kind of social practice, from body movements in ballet to verbal discourse. For every specific type of practice, a 'translation device' is needed, i.e., a framework that operationalizes the two concepts for the given type of analytical purpose (Maton 2014: 113). Therefore, for the investigation of English discourse specifically, two related frameworks for SD will be introduced in 3.1.1 and 3.1.2, and one for SG in 3.1.3.

3.1.1 Epistemic-semantic density of words

In relation to English discourse, SD is concerned with the complexity of meanings expressed by wordings. SD is in fact a superordinate concept comprising several different subcategories, of which the one relevant to present purposes is *epistemic-semantic density* (ESD), referring to meanings based on empirical descriptions and formal definitions of words and terms (when the concept of ESD is applied to linguistic symbols) (Maton and Doran 2017b). Important to present purposes is the fact that the ESD of words is a (highly) variable and register-dependent phenomenon: as already described in Hill-Madsen (2022), terms belonging to specialized registers tend to enter into elaborate taxonomic relations that confer highly complex meanings on individual terms. Words belonging to the ordinary vocabulary of everyday, non-specialized registers, on the other hand, tend to be much 'shallower' in meaning. As a 'translation device' designed to explore the ESD of words, Maton and Doran (2017b) propose a framework with eight different categories of strengths, set out in Table 1.

As already noted, specialized or technical terms ('technicals' in Table 1) are held to represent stronger ESD than everyday words ('everydays'). In the 'technicals' category, stronger ESD is represented by multi-part terms ('conglomerates') such as *pharmacokinetics*[3] (consisting of *pharmac-* and *(o)kinetics*) and *aminotransaminase* (*amino-* + *-trans-* + *-aminase*), whereas single-part technicals such as *lactation* (containing one lexical element only) represent weaker ESD. Also, words denoting actions (once again exemplified by *lactation*) or qualities (e.g., *renal*) are stronger in ESD than 'elements', i.e., 'thingy' referents such as *plasma* and *statins*

[3] Examples representing categories 4–8 have been taken from the *Nustendi* SmPC (EMA 2020a). Examples representing categories 1–3 are the present author's own inventions.

(for support of this claim, see Hill-Madsen 2022, Maton and Doran 2017b: 60–61).

Table 1: Complete overview of ESD strengths in words (adapted from Maton and Doran 2017b)

ESD	Type	Subtype	Sub-subtype	Definitions
+	***Technical***	***Conglomerate*** (multiple distinct parts)	**8. *properties***	actions/qualities
			7. *elements*	items, entities or things
		Compact (single part only)	**6. *properties***	actions/qualities
			5. *elements*	items, entities or things
	Everyday	***consolidated*** (nominaliza-tions)	**4. *specialist***	nominalizations located in texts dominated by technical
			3. *generalist*	nominalizations located in texts dominated by everydays
–		***common*** (congruent)	**2. *nuanced***	relatively more differentiated meanings (hyponyms/more specific)
			1. *plain*	relatively general and less differentiated meanings

In the 'everyday'-term section of the continuum, 'consolidated' words (grammatically realized in nominalizations) are associated with stronger ESD than 'common' terms. Stronger ESD is represented by everyday 'specialist' terms, i.e., nominalizations that are not in themselves technical terms, but appear in specialized texts. One example is *elevation*, which is not technical per se, but assumes more specific meaning when employed in a medical text to refer to a certain type of physiological process. Nominalizations appearing in non-specialized texts ('generalist' terms), on the other hand, indicate weaker ESD. Examples are words such as *inquiry*, *investigation* and *achievement.* Finally, 'common, everyday' terms divide into 'plain' and 'nuanced' ones, i.e., very simple/highly general everyday meanings such as *find*, *get* and *happen* vs. more specific terms from everyday vocabulary such as *hospital*, *doctor* and *needle*. To enable analytical distinction between the two 'common' categories, the criterion adopted here will be whether, for a given word, a hypernym exists. If a hypernym exists, this indicates that the word itself (as a hyponym) denotes a more specific ('nuanced') meaning, whereas if no hypernym can be found for a word, it will be classified as 'plain'.

3.1.2. Epistemological condensation: Clausing

Meaning, apart from being 'fossilized' in individual words, can also be actively *established* in discourse by relating meanings (of words) to each other. Such a process of active 'meaning creation' is termed *epistemological condensation* (EC) in Maton and Doran (2017a), referring to an active process of knowledge *formation* in discourse (Maton and Doran 2017a: 79). In English discourse, the realizational 'site' of this process is clauses. Here, therefore, the concept of EC will be used to explore the active knowledge building taking place in the clauses of PILs. In the framework proposed by Maton and Doran (2017a), eight different strengths of 'relation creation' are posited, listed in Table 2 below with short glosses (for more elaborate definitions, and for the rationale behind viewing the eight types as different EC *strengths*, see Hill-Madsen 2022).

Table 2: A taxonomy of EC strengths (adapted from Maton and Doran 2017a: 82)

EC	Type	Subtype	Sub-subtype	Gloss
+	***Connecting***	***Taxonomizing***	8. *Classifying*	a) identifying type-subtype relations (*X is a type of Y*) or b) naming/equating (*X is called a …*)
			7. *Composing*	identifying part-whole relations between items
		Coordinating	6. *Causing*	identifying causal relations between items
			5. *Correlating*	identifying a correlation or dependency between items, but not a causal relation.
–	***Augmenting***	***Characterizing***	4. *Displaying*	ascribing qualities and attributes to phenomena
			3. *Dramatizing*	describing or reporting physical, verbal or mental action
		Establishing	2. *Positioning*	identifying the location of a phenomenon in space and/or time

			1. *Positing*	reporting that an item/phenomenon exists

The glosses in Table 2 highlight the different *strengths* of semantic relation created in each type of clause. In the two top categories ('classifying' and 'composing'), a two-way relation is established, influencing both items: a clause like *herrings are fish* relate the two items (*herrings* and *fish*) in a mutually defining relation, viz., as species and genus. The 'causing' and 'correlating' categories also link items in a two-way relation, but the link is less pronounced. In a clause such as *smoking causes lung cancer,* it is established as a characteristic of *smoking* that it is associated with *lung cancer* and vice-versa, but the two are not mutually defining concepts. In 'displaying' clauses, the relation is one-way only: in a clause such as *smoking is bad for you,* the item *bad (for you)* is established as a characteristic of *smoking*, but *smoking* does not reflect back on *bad.* In the 'dramatizing' and 'positioning' categories, then, only relatively weak relations are established: a 'dramatizing' clause such as *the patient went to see the doctor* indicates *the patient's* ability to perform the action, but the verb and the other noun are not made to share any characteristics of *the patient* in either a one- or two-way relation (cf. Maton and Doran 2017a: 88). Finally, in 'positing' clauses, such as *there is so much snow!*, no relation at all is created with other items.

3.1.3 Semantic gravity

Where the concept of SD (as applied to English discourse) is concerned with the complexity of meaning in words and wordings, SG refers to strengths of context-dependence (Maton 2013, 2014: ch. 6). In Maton's own words, SG is defined as

> the degree to which meaning relates to its context, whether that is social or symbolic. Semantic gravity may be relatively stronger (+) or weaker (–) along a continuum of strengths. The stronger the semantic gravity (SG+), the more closely meaning is related to its context; the weaker the gravity (SG–), the less dependent meaning is on its context. (Maton 2011: 65)

For the analysis of SG in discourse, the framework adopted here as a 'translation device' will be Hasan et al. (2007), the reason being that the

categories of the framework may be interpreted as varying strengths of context-embeddedness/-independence in linguistic utterances or speech acts (termed *rhetorical units* in Hasan et al. (2007) and grammatically realized in clauses). Thus, the framework's 11 categories are ranged on a continuum according to their degree of situatedness in/distance from the 'here-and-now' of the communicative exchange. The continuum is represented in Figure 1 below, with categories towards the left-hand pole representing 'stronger' gravity/context-dependence (SG+), and those towards the left being 'weaker' (SG–).

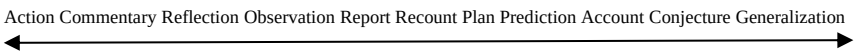

Figure 1: A cline of RU types (adapted from Hasan et al. 2007: 724)

Each category is determined by a combination of two semantic features of a given rhetorical unit: the type of *Central Entity* (grammatically the Subject of a ranking clause) and the type of *Event Orientation* (a feature of the grammatical Verb). Types of Central Entity vary according to their proximity to/remoteness from the 'here' of the communicative event, with four types distinguishable as:

1) Interactant (speaker/listener)
2) Co-present person/object
3) Absent person/object
4) Generalized person/object.

Types of Event Orientation are distinguished according to their closeness to/remoteness from, the 'now' of the speech event in terms of:

- Time: The event or action may be either a) concurrent with the 'now' or b) 'prior', i.e., situated in the past
- Habituality: The event or action may be a) non-habitual, i.e., located in time or b) habitual, i.e., timeless
- 'Realis'/'irrealis': The event or action may be a) real or b) imagined, e.g., forecast or hypothetical.

The 11 categories from the cline in Figure 1 each combine the following values (adapted from Hasan et al. 2007: 722).

Table 3: Types of rhetorical units defined and exemplified. Examples are the invention of the present author

	Type of Central Entity	**Type of Event Orientation**	**Examples**
1. Action	interactant	proposal (i.e., command)	*Remember to take your medicine!*
2. Commentary	interactant	concurrent, non-habitual	*You're looking pale and ill today.*
3. Reflection	interactant	concurrent, habitual	*As a patient, you have certain responsibilities for your own health.*
4. Observation	co-present person/object	concurrent, habitual	*[Pharmacist handing a patient/customer a package of medicine:] This medicine may have some unpleasant side effects.*
5. Report	absent person/object	concurrent, non-habitual	*[Pharmacist to patient/customer:] The side effects are listed in the information leaflet.*
6. Recount	[all types]	prior, i.e., situated in the past	*The clinical trials found a number of side effects associated with the drug.*

7. Plan	interactant	forecast, non-hypothetical, volitional	*[Doctor to patient:] I'll call the hospital later and book an appointment for you.*
8. Prediction	co-present, absent or generalized person/object	forecast, non-hypothetical, non-volitional	*[Pharmacist to patient/customer:] You will be rather certain to experience some not-so-pleasant side effects.*
9. Account	absent person/object	concurrent, habitual	*[Patient to doctor:] My wife always reminds me to take my medication, so that's how I remember.*
10. Conjecture	[all types]	forecast, hypothetical	*[Doctor to patient:] If you don't lose weight, your heart may suffer.*
11. General-ization	generalized person/object	concurrent, habitual	*The primary function of the immune system is to protect the body from disease.*

3.2. Semantic codes: A topology integrating SG and SD values

One aspect of LCT not taken into account in the investigation of *EPAR summaries for the public* in Hill-Madsen (2022), but to be incorporated in the present investigation, is the notion of *semantic codes* (Martin et al. 2020, on which the following account is based). Representing different *combinations* of SD and SG values, the notion of *codes* integrates the two types of meaning (SG and SD) and may be conceptualized as 'areas' or subsections of a semantic 'space' encompassing all the combinations of

SD and SG values possible. The spatial metaphor derives from the fact that since SD and SG are both scalar concepts, the two dimensions may be diagrammed as intersecting axes generating a plane with four quadrants, each representing a certain (type of) combination of SD and SG values:

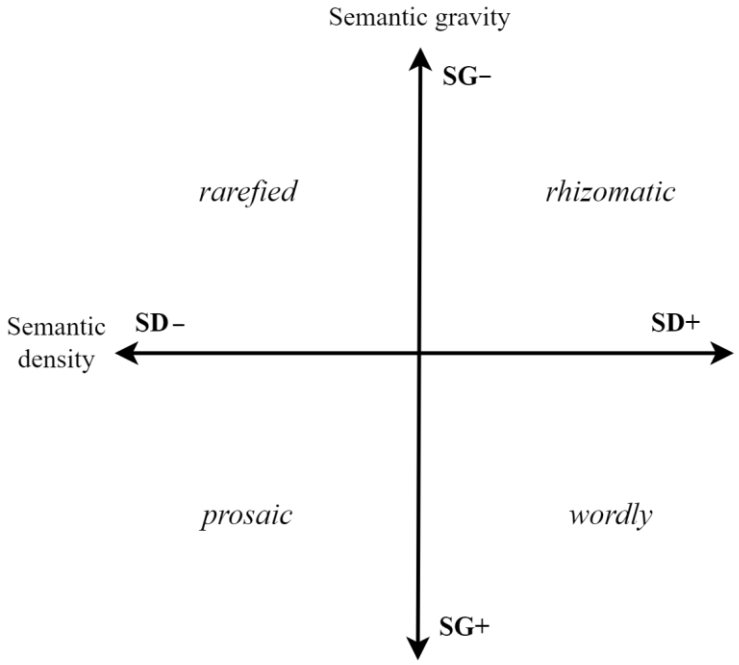

Figure 2. The semantic plane integrating SD and SG values (adapted from Maton 2016: 16)

In the semantic topology in Figure 2, the quadrants represent four principal types of *semantic code*, each characterized by a combination of (a certain range of) SD and SG values. As Figure 2 illustrates, *rhizomatic codes* (glossed as *theoretical codes* by Shay 2013: 574) are characterized by weaker SG and stronger SD, equalling the combination of more context-independent and complex meanings. Examples of texts with this profile are scientific literature setting out the natural laws of, e.g., physics or chemistry. As the diametrical opposite of *rhizomatic codes*, *prosaic codes* (or *practical codes* in Shay's 2013: 573 gloss) combine more context-dependent and simpler meanings (SG+, SD−). In other words, when

prosaic codes are realized in verbal discourse, they will, e.g., be manifested in the 'language-in-action' (Halliday and Hasan 1976: 34) accompanying practical, everyday activities (where the medium is typically oral). In the top left-hand quadrant, *rarefied codes* (or *generic*, in Shay 2013: 576), on the other hand, combine simpler with more context-independent meanings (SD−, SG−). Textbooks at the lower levels of schooling are likely to belong to this category (presuming that such literature does not yet introduce scientific and specialized terminology). Finally, *worldly codes* (glossed as *vocational/professional* by Shay 2013: 575) represent more context-dependent and complex meanings, linguistically manifested, e.g., in the oral discourse of two mechanics (as specialists in auto repair) collaborating on the repair of a car.

The value of Martin et al.'s (2020) theory is that, when applied to discourse, it enables the semantic profiling of any given text or text type, and the 'analysis of code shift and code drift, can be used for all kinds of practices, and enable[s] us to see code clashes and code matches' (Martin et al. 2020: 23). *Code clashes* and *code matches* are here taken to refer to the possibility of harmony or dissonance between the expectations and resources of readers and the actual semantic profile of a text and text type; and *code shift* and *code drift* are taken to refer to the possibility of *inconsistencies* of semantic code within one and the same text. As the analyses in section 4 will show, these possibilities are highly relevant to the present investigation.

4. Analytical approach and results

The analytical approach has been completely the same as in Hill-Madsen (2022), which means that the whole corpus has been manually analyzed by the present author, all based on the author's subjective judgments. More specifically, all lexical words and all clauses have been categorized in terms of SD, and all clauses, moreover, in terms of SG. Findings for ESD at word level will be presented in 4.1, for ESD at clause level (henceforth *EC* in accordance with Maton and Doran 2017a) in 4.2, and for SG in 4.3. Finally, based on the findings in 4.1–4.3, a profile of the PIL texts will be presented by identifying the semantic range of the genre on the SD-SG plane theorized in section 3. For comparison with the *EPAR summary* genre, quantitative results for both genres will be presented in tables, and the location and extent of the *EPAR summary* genre on the SD-SG plane will also be identified.

4.1. ESD of words in PILs

The distribution of lexical words in the PIL corpus across the eight ESD categories is as shown in Table 4. For the *EPAR summary* corpus, results are reproduced in Table 5.

Table 4: Distribution of ESD categories for words in the PIL corpus

	Beovu (%)	Inbrija (%)	Kaftrio (%)	Nustendi (%)	Recarbrio (%)	Rinvoq (%)	Average (%)	St. dev. (%)
8. Tech.: congl: properties	2.1	0.6	1	1.4	1.6	3	1.6	0.8
7. Tech.: congl.: elements	1.9	3.6	7.6	9.9	3.8	3.3	5	3.1
6. Tech.: comp.: properties	6.5	3.5	3.2	3.4	4.3	6.7	4.6	1.6
5. Tech.: comp.: elements	8.2	3.7	6.9	8.7	6.2	9	7.1	2
4. Everyday: cons.: spec.	1.1	5.6	1.3	2.7	4.1	2	2.8	1.6
3. Everyday: cons.: gen.	15	13	6.5	5.3	11.5	5.8	9.5	4.2
2. Everyday: comm.: nuanc.	32.7	34.7	39.3	32.1	35.5	34.8	34.8	2.5
1. Everyday: comm.: plain	32.7	35.5	34.3	36.5	33.1	35.4	34.6	1.5

Table 5: Distribution of ESD categories for words in the EPAR summary corpus (Hill-Madsen 2022)

	Beovu (%)	Inbrija (%)	Kaftrio (%)	Nustendi (%)	Recarbrio (%)	Rinvoq (%)	Average (%)	St. dev. (%)
8. Tech.: congl.: properties	2.8	0.9	0	1.2	1.7	3.9	1.8	1.4
7. Tech.: congl.: elements	0.9	2.3	3.8	1.9	0.8	1.3	1.8	1.1
6. Tech.: comp.: properties	5.4	5.3	10.2	5.3	8.1	6	6.7	2
5. Tech.: comp.: elements	17	17.9	17.3	23.9	25.6	14.8	19.4	4.3
4. Everyday: cons.: spec.	6.6	12.6	10.7	9.5	12.9	13.2	10.9	2.6
3. Everyday: cons.: gen.	0	1.1	0.6	0.5	0	1.3	0.6	0.5
2. Everyday: comm.: nuanc.	25.6	26.4	25.7	25.3	23	31.6	26.3	2.9
1. Everyday: comm.: plain	35.1	33.5	31.7	32.5	27.8	29.7	31.7	2.6

Table 4 reveals a highly consistent picture, with standard deviation values (abbreviated *St. dev.* in the table) generally very low. The frequencies show a sizeable majority of items (an average of almost 80%) belong to the three lowest ‘everyday’ categories, indicating that the PILs constitute a predominantly non-specialized register. Nevertheless, a considerable minority (around 20% on average) are shown to belong to the top 5 strengths, revealing a marked presence of the type of complex meanings associated with specialized registers also. Thus, like the *EPAR summary* genre, the PILs integrate two disparate registers, though the proportion of ‘technicals’ is somewhat higher in the EPARs and the proportion of ‘everydays’ somewhat lower. This reveals the PILs to be

slightly closer to an overall ESD profile to be expected from non-specialized, quotidian discourse.

What Table 4 does not reveal is that the registerial hybridity characterizing the texts as a whole (the macro level) is manifested at the micro level also, within individual sentences. This means that there is a general tendency for individual sentences to exhibit the same kind of marked ESD oscillation from word to word (exemplified below) as observed in the EPAR summaries (see Hill-Madsen 2022). This is illustrated in example (1).

(1) Nustendi is given
- if you have been using a statin (such as simvastatin, a commonly used medicine that treats high cholesterol) together with ezetimibe and this does not lower your LDL-cholesterol sufficiently. (*Nustendi* PIL, section 1)

Example (1) is a very typical illustration of the sentence-internal co-occurrence of words with both stronger and weaker ESD value. The items *simvastatin*, *cholesterol*, *ezetimibe* and *LDL* all belong to category 7, 'technical: conglomerate: elements', being specialized biochemical terms for substances in the blood (*LDL* and *cholesterol*) and pharmaceutical terms for active substances in drugs (*simvastatin* and *ezetimibe*). All are 'conglomerate' terms consisting of multiple parts. Thus, as an abbreviation, *LDL* stands for the three words *low – density – lipoprotein* (with the latter even covering the two parts *lipo-* and *-protein*), and *simvastatin* consists of *simva-* and *-statin*, the combination of which refers to a particular subtype of statins (cholesterol-lowering drugs). Similarly, *ezetimibe* consists of *ezet-* and *-imibe*, each referring to certain very specific chemical characteristics. Even *cholesterol* turns out to be a multi-part word, composed of *chole-* and *-sterol*, signifying that *cholesterol* is a particular subcategory of *sterols* (a superordinate biochemical category of fatty substances). The two terms *Nustendi* and *statin*, on the other hand, while also technical (*Nustendi* because it is the brand name of a specific *type* of drug with its particular pharmaceutical characteristics), both belong in category 5, 'technical: compact: elements', by having a single part only. Intermixed with the technical items are everyday terms from both of the two bottom-most categories: The category 'everyday: common: plain' comprises the items *given*, *using* + *used*, *commonly*, *high*,

together, *lower* and *sufficiently*, which must all be categorized as belonging to the very core English vocabulary of words with absolutely basic and non-specific denotations.[4] The words *medicine* and *treats*, on the other hand, carry more elaborate ('nuanced') meanings, in that the semantic compositionality of *medicine* can be broken down into elements such as 'substance', 'chemical', and 'having curative properties' (see also Hill-Madsen 2022), and *treats* is analysable as a combination of elements such as 'act on', 'medically' and 'with a curative purpose'.

While a definition of the pharmaceutical term *ezetimibe* is provided earlier in the PIL text (*Ezetimibe works in your bowel by reducing the amount of cholesterol absorbed from food*, *Nustendi* PIL, section 1), the term *simvastatin* is defined in the reduced relative clause *a commonly used medicine that treats high cholesterol* in example (1), immediately after the term itself. Apart from the specialized term *cholesterol* (whose specific, biochemical meaning is unlikely to be known to lay readers, but likely to be known in the simpler sense of 'a substance in the bloodstream with long-term health-threatening effects'), the definition consists of everyday terms only, thus 'translating' the specialized term into lay meanings.

Another example of intrasentential registerial hybridity is (2):

(2) The other ingredients are:
Core tablet: microcrystalline cellulose, mannitol, tartaric acid, hypromellose, silica colloidal anhydrous, magnesium stearate (*Nustendi* PIL, section 6)

While the words *ingredients*, *are*, *core* and *tablet* in example (2) are everyday terms from the two bottom-most categories, all the remaining words are technicals, spanning all the four top categories: As adjectival, multi-part terms, *microcrystalline* and *anhydrous* both belong to category 8 ('technicals: conglomerate: properties'), while *cellulose*, *hypromellose* and *stearate* are composites denoting 'elements' (category 7). *Tartaric* and *colloidal* are adjectival, but single-part terms (category 6: 'technicals:

[4] It may be noted that while *sufficiently*, because of its Romance origin, may carry slightly more formal connotations than the other (Germanic) words in the list, in terms of epistemic semantics the word must be considered completely synonymous with *enough*, and thus still be categorized as belonging to the bottommost ESD category for words.

compact: properties'), and *mannitol*, *acid* and *silica* are all single-part terms denoting 'elements' (category 5).

What is important to note about example (2) is that—unlike in example (1)—nothing is done to ensure a lay understanding of the technical terms, e.g., by means of definitions. This is the main point of *difference* regarding the intrasentential fusion of registers characteristic of both the PIL and EPAR summary genre: In the latter, the fusion in most cases manifested itself as an active process of linkage through definitions (see Hill-Madsen 2022), which was to a large extent what gave the genre its 'textbook' quality. In the PILs, the same phenomenon occurs (as illustrated in example (1)), but very often the registerial fusion takes the form of a mere juxtaposition of elements with very disparate ESD values as seen in example (2), with no active textual measures being taken to mediate between the two registers.

4.2. EC: clausing

Table 6 sets out clausing type frequencies in the PIL corpus. For the EPAR summaries, the EC values are those shown in Table 7.

Table 6: Distribution of EC ('clausing') types in the PIL corpus

	Beovu (%)	Inbrija (%)	Kaftrio (%)	Nustendi (%)	Recarbrio (%)	Rinvoq (%)	Average (%)	St. dev. (%)
8. classifying	5.9	9.1	6.7	8.4	11.7	5.1	7.8	2.4
7. composing	4.7	2.8	4.9	3.4	4.7	4.8	4.2	0.9
6. causing	10	4	4.9	6.9	5.3	8.9	6.7	2.4
5. correlating	0.6	0.4	0	0	0.6	0	0.3	0.3
4. displaying	14.7	11.1	16.4	15.8	19.3	12.8	15	2.9
3. dramatizing	58.2	69	64.9	63.5	56.7	59.7	62	4.6
2. positioning	3.5	1.6	1.8	0.5	0.6	0.9	1.5	1.1
1. positing	2.4	2	0.4	1.5	1.2	1.1	1.4	0.7

Table 7: Distribution of EC ('clausing') types in the EPAR summary corpus

	Beovu (%)	Inbrija (%)	Kaftrio (%)	Nustendi (%)	Recarbrio (%)	Rinvoq (%)	Average (%)	St. dev. (%)
8: classifying	15.3	15.6	9.7	9.7	6.0	10.2	11.1	3.7
7: composing	1.8	5.2	2.9	2.2	1.2	0.9	2.4	1.6
6: causing	6.3	8.3	15.5	16.1	10.7	9.3	11.0	4.0
5: correlating	0	0	0	0	0	0	0	0
4: displaying	22.5	15.6	27.2	19.4	20.2	16.7	20.3	4.2
3: dramatizing	53.2	54.2	40.8	52.7	57.1	68.5	54.4	8.9
2: positioning	0.9	0	1.0	0	2.4	3.7	1.4	1.5
1: positing	0	1.0	2.9	0	1.2	0	0.9	1.1

Once again, the picture is very consistent in the PIL genre, with standard deviation values that are in most cases very low. Table 6 shows the 'classifying', 'causing' and 'displaying' categories all to be noticeably represented, but the most frequent type by far is the 'dramatizing' category, accounting for more than 6 out of ten clauses on average. The 'correlating', 'positioning' and 'positing' categories, on the other hand, are either virtually or completely absent. Therefore, only the above-mentioned four categories with a clear representation in the PIL corpus will be exemplified and their function in the texts commented on in the following.

Category 8, 'classifying' clauses:
'Classifying' clauses, which are concerned with linking superordinate and subordinate classes of phenomena, are represented first and foremost in the form of definitions (see also 4.1) of key medical/pharmaceutical terms in the texts, as in example (3).

(3) Beovu contains the active substance brolucizumab, [1] which belongs to a group of medicines [2] called antineovascularisation agents. (*Beovu* PIL, section 1)

In the example, the two clauses marked [1] and [2] represent the 'subtyping' and the 'equating' subcategory of the 'classifying' clause type, respectively. Clause [1] relates the referent of *which*, i.e., *the active substance brolucizumab* to the higher-order class of *medicines*. Clause [2] (a reduced relative clause), then, names a subclass of active substances (*antineovascularisation agents*) at an intermediate level of generality

between *brolucizumab* and *medicines*. Together, the two clauses establish a three-tiered taxonomy, as illustrated in Figure 3.

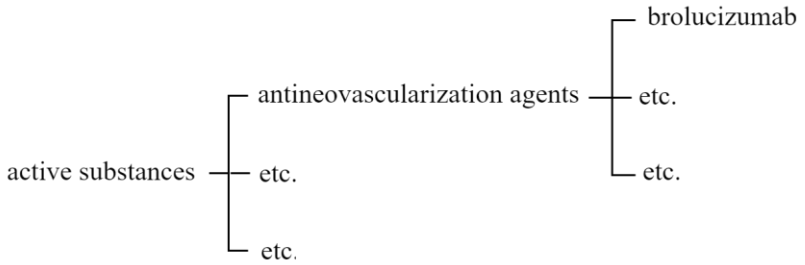

Figure 3: The three-level taxonomy established in example (3)

Example (3) may be characterized as an overt process of classification, insofar as the two clauses in question are both fully-fledged Subject–Verb–Complement clauses[5] which thus explicitly relate two items to each other in an 'is-a' relationship. Frequently, however, the linkage is only implicit, which is the case when a non-technical synonym is merely provided in brackets after a specialized term, or vice-versa, as in examples (4) and (5).

(4) [...] new or increased abnormal body movements (dyskinesia) (*Inbrija* PIL, section 4)

(5) [...] melanoma (a type of skin cancer) (*Inbrija* PIL, section 2)

A completely similar phenomenon was registered in the EPAR summaries, i.e., a number of bracketed elements which, as in (4) and (5), are here interpreted as non-restrictive relative clauses with both Subject and Verb implicit. In other words, it is assumed here that the bracketed element in Example (4) should be interpreted as *which are called dyskinesia* and the one in (5) as *which is a type of skin cancer*. Example (4) is thus a case of the 'equating' subcategory of 'classifying' clauses, and (5) an instance of 'subtyping'.

[5] In clause [2], which is a reduced relative clause, the Subject is implicit, of course, but clearly linked with *medicines* as the antecedent of the implicit relative *which/that*.

Category 6: 'causing' clauses:
'Causing' clauses are found across the PIL sections, but especially in Section 1, as in example (6).

(6) Ivacaftor causes the protein to work better, while tezacaftor and elexacaftor increase the amount of protein at the cell surface. Kaftrio taken with ivacaftor helps your breathing by improving your lung function. (*Kaftrio* PIL, section 1)

When occurring in section 1 of the PILs, 'causing' clauses serve to detail the drug's mechanism of action, as in the three underlined clauses in (6), which all refer to the positive causative effects of active substances. In other sections, the focus is typically on negative effects (side effects).

(7) RINVOQ can reduce your body's ability to fight infections and so may worsen an infection that you already have [...] (*Rinvoq* PIL, section 2)

In the preceding example (6), the lexical item *cause* was an obvious indicator of a 'causing' clause, but both examples (6 and 7) reflect that other verbs such as *increase*, *reduce*, *lower*, *improve* and *worsen* lend themselves to the same interpretation.

Category 4: 'displaying' clauses:
'Displaying' clauses are concerned with attributing properties or traits to entities (Maton and Doran 2017: 88). Not surprisingly, in the PIL corpus the two entities to which characteristics are ascribed are the patient and the drug. In the most obvious cases, clauses in this category feature a linking verb, such as *be*, in combination with an adjectival Subject Complement, as in the first two underlined clauses in (8).

(8) If you are pregnant or breast-feeding, think you may be pregnant or are planning to have a baby, ask your doctor or pharmacist for advice before taking this medicine. (*Inbrija* PIL, section 2)

Category 3: 'dramatizing' clauses:
Since the 'dramatizing' category is the most frequent one in the corpus, it is not surprising that it is also the most diversely manifested. In the corpus,

the clauses 'dramatize' the activities (practical/physical, verbal and mental) of the human agents involved (primarily the patient but in a minority of cases also those of the physician). In this respect, the PILs are different from the EPAR summaries, which are much less concerned with human agents, with 'dramatizing' clauses being mainly used to describe the drug's mechanism of action (see Hill-Madsen 2022). From the PILs, a typical example is (9).

(9) [α] Tell your doctor right away [β1] if you have an allergic reaction, convulsions (seizures or fits), diarrhoea, [β2α] or develop kidney problems [β2β] while receiving Recarbrio [...] (*Recarbrio* PIL, section 2)

In (9), the patient is 'dramatized' as being (potentially) involved in verbal and practical activity (*tell*[*ing*] and *receiving*), as well 'physiological activity' in the references to the potential experience of side effects (*allergic reactions* etc.).

4.3. Semantic gravity

The statistics regarding frequencies of the different types of semantic gravity in the PIL corpus are shown in Table 8.

Table 8: SG clause type frequencies in the PIL corpus

	Beovu (%)	Inbrija (%)	Kaftrio (%)	Nustendi (%)	Recarbrio (%)	Rinvoq (%)	Average (%)	St. dev. (%)
11. Generalization	39.1	19	30.8	31.7	29.6	34.6	30.8	6.7
10. Conjecture	12.8	22.3	15.7	13.4	13.6	21.8	16.6	4.3
9. Account	1.9	1.6	1	1.8	0	0.4	1	0.8
8. Prediction	10.9	6.5	6	6.1	8.8	6.2	7.4	2
7. Plan	0	0	0	0	0	0	0	0
6. Recount	2.6	3.3	6	4.3	6.4	3.5	4.4	1.5
5. Report	0	2.2	1.5	0	0	0.4	0.7	0.9
4. Observation	2.6	3.8	2.5	3	3.2	1.2	2.7	0.9
3. Reflection	4.5	2.7	4	3.7	4.8	4.3	4	0.7
2. Comment	7.7	8.2	6.6	11.6	11.2	6.2	8.6	2.3
1. Action	17.9	30.4	26.8	24.4	19.2	21.4	23.4	4.8

Whereas the *EPAR summary* genre was largely centred around two types of SG in clauses,[6] viz., Generalization and Recount (Hill-Madsen 2022), Table 8 above reveals a much more varied picture for the PILs, with all categories except 7 (Plan) being represented, albeit with markedly diverse frequencies. Five categories dominate the genre, viz., Generalization, Conjecture, Prediction, Commentary, and Action, whereas Account and the four mid-level categories (Recount, Report, Observation and Reflection) have markedly lower representations. The statistics thus reveal the PIL genre to be a surprising constellation of very different strengths of context-dependence. In the following, only the five dominant categories will be exemplified.

Generalization:
Clauses in this category occur in all sections of the texts, but with the highest frequencies in section 1, which sets out the most general characteristics of the drug.

(10) [9] Nustendi is a medicine that lowers levels of 'bad' cholesterol [...] [11α] Nustendi contains two active substances, [11β] which reduce your cholesterol in two ways: [12.1] bempedoic acid decreases the production of cholesterol in the liver [12.2] and increases the removal of LDL-cholesterol from the blood; [...] [15α] Nustendi is given to adults with primary hypercholesterolaemia or mixed dyslipidaemia, [15β] which are conditions that cause a high cholesterol level in the blood. (*Nustendi* PIL, section 1)

All seven clauses in (10) feature the combination of Central Entities (realized by the grammatical subjects) with generic reference (primarily the brand name of the drug and the names of the active substances—as types—in question) in combination with timeless Event Orientation, realized in the verbs *is/are*, *contains*, *reduce*, *decreases* and *increases.* Example (10) is a highly representative example, insofar as the clauses serve to detail the type of medical condition that the brand of drug is used to treat and its mechanism of action. Especially clauses [12.1], [12.2] and [15β] bear witness to a didactic element being also present in the PILs,

[6] Frequencies were not given in tables in Hill-Madsen (2022).

though less pronounced than in the EPAR summaries: In the three clauses, the reader is provided with abstract knowledge about certain medico-pharmaceutical concepts, which is a momentary intrusion of a 'textbook' register.

Conjecture:
By definition, clauses of Conjecture are concerned with future ('forecast'), hypothetical situations relating to any type of Central Entity, specific or general. Even when specific CEs are involved, the strong context-independence of Conjecture resides in the 'irrealis' of these utterances. In the corpus, a central manifestation of Conjecture is in connection with warnings to the patient about possible side effects, as in clauses [17.1] and [17.2].

(11) ([15] Do not use Inbrija: [...])
[17.1] if you get blurred vision, red eyes, severe eye and head pain, halos around lights, pupils of your eyes that are larger than normal size [17.2] and feel sick. (*Recarbrio* PIL, section 2)

All the negative effects referred to in (11) are only potential, future occurrences.

Prediction:
Being by definition concerned with non-hypothetical, future situations, Predictions in the PILs typically prepare the patient for what might be termed 'concomitant circumstances or events' in connection with the treatment that is about to be initiated.

(12) [52.1α] Your doctor will test you for tuberculosis [52.1β] before starting RINVOQ [52.2] and may retest during treatment.
(*Rinvoq* PIL, section 2)

The first two clauses in (12) are both non-hypothetical, whereas the third clause [52.2] is Conjecture, referring to something that is a future possibility only.

Commentary:
Commentary is defined as a type of speech act very close to the 'here-and-now' of the communicative situation, having an interactant or a co-present object or person as the Central Entity and having a concurrent, non-habitual Event Orientation. In the corpus, they manifest themselves as statements about the patient's current doings or circumstances, but typically with a conditional element involved, as in the *if*-clauses in (13).

(13) (Do not take Nustendi:)
- if you are pregnant
- if you are breast-feeding
[...]
- with a statin if you currently have liver problems (*Nustendi* PIL, section 2)

Action:
As the statistics in Table 5 show, Action is the second-most frequent category, making the PILs a highly instructional genre. In terms of grammatical realization, these speech acts almost invariably manifest themselves in imperatives, as in (11) and (13) (*Do not use Inbrija / Do not take Nustendi* etc.). The instantiations of Action in (11) and (13) are also very typical, reflecting the way the imperatives are in many cases linked with some kind of conditionality (*Do not take X if …* or *Tell your doctor if …*). Thus, even in these most context-dependent of speech acts, an element of 'irrealis' intrudes itself.

The element of abstractness noted in connection with the 'Action' speech acts are in fact indicative of a more general tendency in the PILs, viz., a frequent ambiguity regarding the strength of context-(in)dependence in individual utterances, as in (14):

(14) Beovu is given as an injection into your eye (intravitreal use) by an eye doctor. (*Beovu* PIL, section 3)

Example (14) illustrates a characteristic mixture of generic and very specific references. In terms of SG, the statement in (14) belongs to the category of Generalization, with the Central Entity realized by the brand name (denoting a *type* of drug) and the Event Orientation being timeless. Similarly, the NP *an eye doctor* must be interpreted as carrying generic

reference, since no specific physician is intended here. *Your*, on the other hand, is very specific, referring exophorically to the reader. The intrusion of highly specific reference into a sentence that is otherwise completely generic creates a certain degree of registerial dissonance that is characteristic of the texts as a whole. In a number of cases, the dissonance is *inter*sentential, when sentences of very different SG strengths are juxtaposed, as in (15).

(15) [84α] Before starting Inbrija, [84β] you must be taking regular treatment for Parkinson's disease combining a socalled dopa-decarboxylase inhibitor with levodopa. [85] The recommended dose of Inbrija is 2 capsules to treat each off period. [86] Do not use more than 2 capsules for each off period. (*Inbrija* PIL, section 3)

In (15), two instances of Action (clauses [84β] and [86]) are interrupted by Generalization (in clause [85]).

The conflation of the generic and the particular as a key feature of the PIL genre may also be taken to explain the relatively high frequency of Conjecture, which mostly takes the shape of warnings that certain adverse reactions to the drug (or other types of negative circumstances) *may* occur. The warnings cover a wide range of circumstances that are known to occur in the patient cohort *as a whole*, but the hypothetical nature of the warnings reflects that they are only potentially applicable to the *individual* patient. The implications of this conflation will be further discussed in section 5.

4.4. Locating the range of PILs on the semantic plane

To sum up, the analyses above evidenced an ESD profile for words in the PILs that indicated the texts to be a mixture of specialized and non-specialized discourse, though relatively firmly grounded in the latter. The hybridity was shown to manifest itself at both the macro and the micro (i.e., sentence) level. As for EC values, the PILs also showed quite marked diversity, encompassing clause types of both weaker and stronger condensation. However, especially the frequency of 'dramatizing' clauses was shown to be high, revealing the texts to be very much concerned with the activities of human agents. In terms of SG, a somewhat similar picture of oscillation emerged, with clauses of very different strengths of context-(in)dependence intermingled throughout the texts. Thus, from a panoptic

perspective, the analytical findings indicate an overall semantic range that extends all along both axes and thus into all four quadrants of the semantic plane theorized in 3.2, revealing the PIL genre to be a highly composite one that draws on all four types of semantic code (*rhizomatic*, *rarefied*, *wordly* and *prosaic*), albeit far from equally. Given the predominance of ESD categories at the lower end of the scale, the *rarefied* and *prosaic* codes have the strongest representation by far, but the *rhizomatic* and the *worldly* codes are represented also, though to a markedly lesser extent. Thus, example (3) (*Beovu contains the active substance brolucizumab, which belongs to a group of medicines called antineovascularisation agents*) is a sentence integrating elements from both the *rhizomatic* and the *rarefied* codes, being a mixture of lexical items with both stronger and weaker ESD and containing clauses that belong to the top EC category, as well as by representing Generalization, i.e., the highest category of context-independence on the SG scale (as operationalized for present purposes). An example like (16) spans the *prosaic* and the *wordly* codes, representing stronger context-dependence and weaker EC (the 'dramatizing' category) and fusing mostly weaker-ESD items with one stronger (*ganciclovir*).

(16) Tell your doctor about all the medicines you take, especially if you take medicines that contain ganciclovir … (*Recarbrio* PIL, sentence 38)

An example such as (8), on the other hand (*If you are pregnant or breast-feeding, think you may be pregnant or are planning to have a baby, ask your doctor or pharmacist for advice…*), belongs to a minority of sentences that are wholly confined to the *prosaic code* (with consistently weaker SD values and stronger context-dependence).

In comparison with the PILs, the EPAR summary genre is narrower in semantic range, being more or less confined to the two top quadrants of the plane, i.e., the *rarefied* and the *rhizomatic* codes. This was because, on the one hand, the full range of SD values were represented, but only categories from the top half of the SG scale (Generalization and Recount) on the other (see Hill-Madsen 2022).

5. Discussion and conclusion

While the PILs have been confirmed to contain elements that are atypical in lay-oriented genres (especially the *rhizomatic* codes), the investigation does not allow judgment of whether this feature in itself makes the texts too demanding for lay readers. What *can* be established, on the other hand, is that the semantic profile of the texts presupposes some degree of orientation to all four semantic codes. Ideal PIL readers, therefore, are those who are able to find their bearings in a highly diverse semantic landscape, where elements from the four codes constantly blend in varying constellations.

One important reason for the high degree of semantic diversity in the PILs must be found in the texts' central purpose of transmitting medico-pharmaceutical knowledge and information across an expert-lay knowledge asymmetry. This in itself does not explain the diversity in meanings, however, since, in theory at least, the expert knowledge and information could have been 'translated' into lay terms, i.e., have been rendered in lay terms *only*. What the actual textual practices in the PILs evidence, however, is that the texts are not only *products* of such 'translation', with source texts to be found in a different genre altogether, viz., the *Summary of Product Characteristics* (see, e.g., Hill-Madsen 2015a, 2015b, 2019b); the 'translation' is as much a text-*internal* process, with 'source' and 'target' elements embedded side by side in the same text. In Maton's (2014: 130) terms, this process of simplification or 'dilution' of stronger-ESD meanings equals *rarefaction*, which is a feature that the PILs share with the EPAR summaries to some extent. In both genres, the element of rarefaction is what contributes a didactic element to the texts, pragmatically taking the shape of repeated definitions and explanations of specialized concepts, and semantically manifesting itself in the close intertwining of complex and simple meanings, even at sentence level. This process of mediation, then, operates along the SD dimension.[7]

[7] It should be noted that, while the mediational 'direction' is in most cases from high to lower SD, the opposite direction was also seen, when occasionally lay terms are 'translated' into specialized ones, as noted in 4.2.

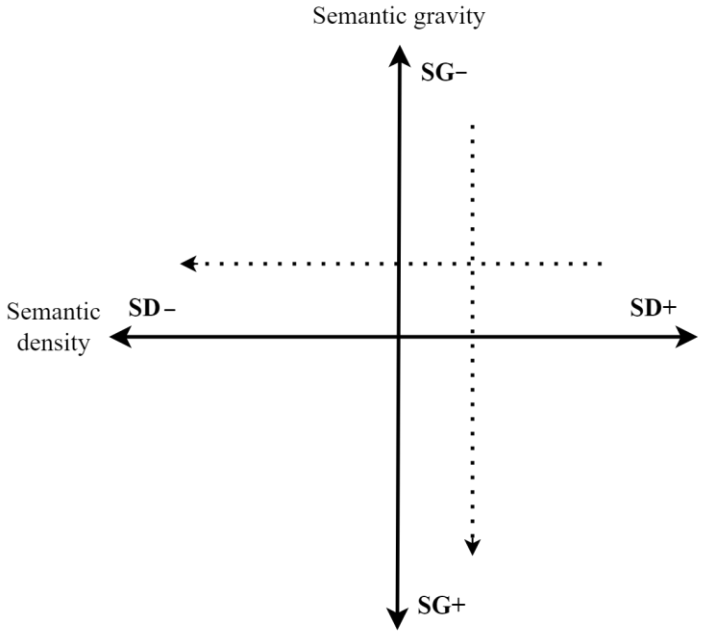

Figure 4: Axes of mediation in the PILs

However, the element of *rarefaction* only explains one dimension of the semantic diversity recorded. In fact, whereas the EPAR summaries feature only this single type of mediation, the PILs feature two. In somewhat the same way as *rarefaction* operates along the SD axis, this other type of mediation, which may be termed *concretization*, operates along the SG axis (see illustration of both in Figure 4), consisting in textual strategies aimed at making the general facts and observations regarding the drug relevant to the individual reader, or rather by *pretending* to personalize the information. An alternative (and oxymoronic) term for the phenomenon would thus be *personalized generalization*. In the analyses, the phenomenon was observed in the intrasentential mixture of generalized and very specific (exophoric) references, and in the juxtaposition of clauses and sentences with very different SG values, especially clauses of Generalization or Conjecture followed by Action clauses. The juxtaposition of weaker- and stronger-SG elements may be interpreted as the textual correlate of a different kind of incongruity between sender and receiver perspectives, viz., one pertaining to the scope

of relevance of the reported facts, especially those regarding side effects. For the specialist senders, the only possibility is to report facts as these apply to the whole of the patient population taking the drug, and so all the side effects and their frequencies of occurrence (observed in the clinical trials preceding authorization of the drug) are reported. To the individual patient, on the other hand, only a small subset of those facts will turn out to be relevant. Ideally, the individual patient would be told which particular side effects to expect in his/her specific case, which is obviously impossible. Instead, the texts repeatedly seek to establish a link of relevance between the general findings and the individual reader's situation through repeated exhortations to take action (typically by informing a healthcare professional) if side effects are experienced. Thus, given the apparent necessity of the dual type of mediation (involving both semantic axes) of medico-pharmaceutical facts and knowledge, the overall conclusion must be that the codal diversity of the PILs is unavoidable.

References

Primary sources

Inbrija = EMA. 2019. Inbrija: EPAR – Product information. https://www.ema.europa.eu/en/documents/product-information/inbrija-epar-product-information_en.pdf. Accessed 7 August 2021.

Rinvoq = EMA. 2019. Rinvoq: EPAR – Product information. https://www.ema.europa.eu/en/documents/product-information/rinvoq-epar-product-information_en.pdf. Accessed 7 August 2021.

Beovu = EMA. 2020. Beovu: EPAR – Product information. https://www.ema.europa.eu/en/documents/product-information/beovu-epar-product-information_en.pdf. Accessed 7 August 2021.

Kaftrio = EMA. 2020. Kaftrio: EPAR – Product information. https://www.ema.europa.eu/en/documents/product-information/kaftrio-epar-product-information_en.pdf. Accessed 7 August 2021.

Nustendi = EMA. 2020. Nustendi: EPAR – Product information. https://www.ema.europa.eu/en/documents/product-information/nustendi-epar-product-information_en.pdf. Accessed 7 August 2021.

Recarbrio = EMA. 2020. Recarbrio: EPAR – Product information. https://www.ema.europa.eu/en/documents/product-information/recarbrio-epar-product-information_en.pdf. Accessed 7 August 2021.

Secondary sources

Askehave, Inger, and Karen K. Zethsen. 2003. Communication barriers in public discourse: The patient package insert. *Document Design* 4(1): 22–41. https://doi.org/10.1075/dd.4.1.03ask.

Bjerrum, Lars, and Annette Foged. 2003. Patient information leaflets – helpful guidance or a source of confusion? *Pharmacoepidemiology and Drug Safety* 12(1): 55–59. https://doi.org/10.1002/pds.795.

Bradley, B., M. Singleton, and A. Li Wan Po. 1994. Readability of patient information leaflets on over-thecounter (OTC) medicines. *Journal of Clinical Pharmacy and Therapeutics* 19(1): 7–15. https://doi.org/10.1111/j.1365-2710.1994.tb00802.x.

Carmen, Kristin L., Pam Dardess, Maureen Maurer, Shoshanna Sofaer, Karen Adams, Christine Bechtel, and Jennifer Sweeney. 2013. Patient and family engagement: A framework for understanding the elements and developing interventions and policies. *Health Affairs* 32(2): 223–231. https://doi.org/10.1377/hlthaff.2012.1133.

Centers for Disease Control and Prevention. 2021. What is health literacy? https://www.cdc.gov/healthliteracy/learn/index.html. Accessed 27 January 2022.

Clerehan, Rosemary, and Rachelle Buchbinder. 2006. Towards a more valid account of functional text quality: The case of the patient information leaflet. *Text & Talk* 26(1): 39–68. https://doi.org/10.1515/TEXT.2006.003.

Coulter, Angela, and Jo Ellins. 2007. Effectiveness of strategies for informing, educating, and involving patients. BMJ 335(7609): 24–27. https://doi.org/10.1136/bmj.39246.581169.80.

EMA. 2021. *Quality review of documents human product information annotated template v10.2rev2 – highlighted*. https://www.ema.europa.eu/en/documents/template-form/qrd-product-information-annotated-template-english-version-102-rev1-highlighted_en.pdf. Accessed 1 February 2022.

European Parliament and Council. 2001. Directive 2001/83/EC. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02001L0083-20121116&qid=1472567249742&from=EN. Accessed 1 February 2022.

Epstein, Ronald M., and Richard L. Street Jr. 2008. *Patient-centered care for the 21st century: Physicians' roles, health systems and patients' preferences*. Philadelphia: ABIM Foundation.

Hadjipavlou, Marios, Sharid Khan, and Abhay Rane. 2013. Readability of patient information leaflets for urological conditions and treatments. *Journal of Clinical Urology* 6(5): 302–305. https://doi.org/10.1177/2051415813489554.

Halliday, M. A. K., and Ruqaiya Hasan. 1976. *Cohesion in English.* London: Longman.

Harwood, A., and J. E. Harrison. 2004. How readable are orthodontic patient information leaflets? *Journal of Orthodontics* 31(3): 210–219. https://doi.org/10.1179/146531204225022425.

Hasan, Ruqaiya, Carmel Cloran, Geoffrey Williams, and Annabelle Lukin. 2007. Semantic networks: The description of linguistic meaning in SFL. In *Continuing discourse on language: A functional perspective*, edited by Ruqaiya Hasan, Christian Matthiessen, and Jonathan J. Webster, 697–738. London and Oakville, CT: Equinox.

Hibbard, Judith H., and Eldon Mahoney. 2010. Toward a theory of patient and consumer activation. *Patient Education and Counseling* 78(3): 377–381. https://doi.org/10.1016/j.pec.2009.12.015.

Hill-Madsen, Aage. 2015a. Lexical strategies in intralingual translation between registers. *Hermes − Journal of Language and Communication in Business* 54: 85–105. https://doi.org/10.7146/hjlcb.v27i54.22949.

Hill-Madsen, Aage. 2015b. The 'unpacking' of grammatical metaphor as an intralingual translation strategy: From de-metaphorization to clausal paraphrase. In *Translation and comprehensibility*, edited by Karin Maksymski, Silke Gutermuth, and Silvia Hansen-Schirra, 195–226. Berlin: Frank & Timme.

Hill-Madsen, Aage. 2019a. Coherence in hybrid texts: The case of the Patient Information Leaflet genre. *Hermes − Journal of Language and Communication in Business* 59: 153–187. https://doi.org/10.7146/hjlcb.v59i1.117038.

Hill-Madsen, Aage. 2019b. The heterogeneity of intralingual translation. *Meta* 64(2): 537–560. https://doi.org/10.7202/1068206ar.

Hill-Madsen, Aage. 2022. Semantic density and gravity in lay-oriented medical knowledge communication: The case of the EPAR summary for the public. *Nordic Journal of English Studies* 21(2): 195–225.

Martin, James R., and Karl Maton. 2017. Systemic Functional Linguistics and Legitimation Code Theory on education: Rethinking field and knowledge structure. *Onomázein – Journal of Linguistics, Philology*

and Translation Special Issue II: 12–45. https://doi.org/10.7764/onomazein.ne2.02.

Maton, Karl. 2011. Theories and things: The semantics of disciplinarity. In *Disciplinarity: Systemic functional and sociological perspectives*, edited by Frances Christie, and Karl Maton, 62–84. London: Continuum.

Maton, Karl. 2013. Making semantic waves: A key to cumulative knowledge-building. *Linguistics and Education* 24(1): 8–22. http://dx.doi.org/10.1016/j.linged.2012.11.005.

Maton, Karl. 2014. *Knowledge and knowers: Towards a realist sociology of knowledge*. London and New York: Routledge.

Maton, Karl. 2016. Legitimation Code Theory: Building knowledge about knowledge building. In *Knowledge-building: Educational studies in Legitimation Code Theory*, edited by Karl Maton, Susan Hood, and Suellen Shay, 1–24. London: Routledge.

Maton, Karl, and Yaegan J. Doran. 2017a. Condensation: A translation device for revealing complexity of knowledge practices in discourse, Part 2—Clausing and sequencing. *Onomázein – Journal of Linguistics, Philology and Translation* Special Issue II: 77–110. https://doi.org/10.7764/onomazein.sfl.04.

Maton, Karl, and Yaegan J. Doran. 2017b. Semantic Density: A translation device for revealing complexity of knowledge practices in discourse, Part 1—Wording. *Onomázein – Journal of Linguistics, Philology and Translation* Special Issue II: 46–76. https://doi.org/10.7764/onomazein.sfl.03.

Martin, James R., Karl Maton, and Yaegan J. Doran. 2020. Academic discourse: An inter-disciplinary dialogue. In *Accessing academic discourse: Systemic-Functional Linguistics and Legitimation Code Theory*, edited by James R. Martin, Karl Maton, and Yaegan J. Doran, 1–32. London: Routledge

Karl Maton, James R. Martin, and Erika Matruglio. 2016. LCT and systemic functional linguistics: Enacting complementary theories for greater explanatory power. In *Knowledge-building: Educational studies in Legitimation Code Theory*, edited by Karl Maton, Susan Hood, and Suellen Shay, 93–113. London: Routledge.

Pander Maat, Henk, and Leo Lentz. 2010. Improving the usability of patient information leaflets. *Patient Education and Counseling* 80(1): 113–119. https://doi.org/10.1016/j.pec.2009.09.030.

Ross, Beth Stearman, Linda S. Potter, and Kay A. Armstrong. 2004. Improving patient educational literature: An understandable patient package insert for 'the pill'. *Journal of Obstetric, Gynecologic, and Neonatal Nursing* 33(2): 198–208. https://doi.org/10.1177/0884217504262969.