PART II
FOUNDATIONS

2 Theoretical assumptions

2.1 Overview

This chapter is divided into three main parts, which together present a theoretical basis for our study of translational complexity in selected parallel texts of English and Norwegian. The analysis is focussed on the relation between original and translation, and the first part of this chapter, 2.2 with subsections, argues for the choice of a product-oriented approach to translation.

With reference to tokens of parallel texts instantiating specific text types, the principal aim of our analysis is to find out to what extent it is possible to predict, or compute, a certain translation on the basis of a given source expression and otherwise accessible linguistic information, and without the aid of a human translator. For this purpose, the second part, 2.3 with subsections, presents principles for drawing the limit of computability in the translational relation between a unit in the source text and its correspondent in the target text.[1]

In the third part, 2.4 with subsections, the basic notions of information, knowledge, and informational content are discussed before we present our typology of information sources for translation.

2.2 An objectivist approach to translation

On the background of the discussion of different approaches to the study of translation (cf. 1.4.1 with subsections), a relevant distinction is one made by Karl R. Popper between the *products* of behaviour and production *behaviour*. Its relevance follows from the fact that translation is a kind of human behaviour which results in a

---

[1] The notion of 'computability' is discussed in 3.2.1.

product. Popper's distinction is part of his objectivist approach to knowledge, which we will present in 2.2.1, and in 2.2.2 the phenomenon of translation is discussed in the light of his approach. In 2.2.3 we will relate certain concepts, categories, and methods of translation studies to Popper's framework, and in 2.2.4 comment on the approach taken in our own investigation.

2.2.1 Popper's objectivist view of knowledge

The distinction between the *products* of behaviour and production *behaviour* is presented in the essay "Epistemology Without a Knowing Subject" (Popper 1979: 106–152), upon which the following exposition is based. The topic of his essay is epistemology, understood by Popper as "the theory of *scientific knowledge*" (1979: 108). He starts by making certain fundamental distinctions: he divides reality into three domains of knowledge, and he draws the line between objective and subjective knowledge. Then, starting from a discussion of biological behaviour in general, he presents a model of the growth of knowledge, in which scientific knowledge is a special case of objective knowledge, and the distinction between products and production behaviour plays an important role in the model of knowledge growth.

Popper describes the three different domains of knowledge as "worlds or universes": in his words, the first world is "the world of physical objects or of physical states"; the second world is "the world of states of consciousness, or of mental states, or perhaps of behavioural dispositions to act"; the third world is "the world of objective contents of thought, especially of scientific and poetic thoughts and of works of art" (1979: 106). Popper does not claim this to be the only possible way of dividing reality into domains, but he finds this approach to be convenient. He argues for the independent existence of the third world through two thought experiments, in both of which he imagines a scenario where all machines and tools created by man are gone, together with human skills and knowledge of building and using the tools. In the first case books and libraries still exist, so that after some time human civilisation may be rebuilt through man's capacity to learn. In the second case all books and libraries are also destroyed, so that there are no pools of objective

knowledge to learn from, and hence our civilisation cannot be rebuilt until the knowledge itself has been rediscovered.

Popper claims there are two different senses of knowledge (1979: 108–109): Subjective knowledge is something located in the mind of an individual; it is a state of mind, a second world object. According to Popper, it is knowledge in the subjective sense that has been the concern of traditional epistemology. Objective knowledge, on the other hand, exists independently of any particular knowing sub- ject; it belongs to the third world, and consists of problems, theories, and arguments. Scientific knowledge falls within this domain, and hence it is third world objects that are of interest to the philosophy of science. Popper views the process of learning in humans as growth of subjective knowledge, and the second world as a medium between the physical first world and the abstract third world. He states that "all our actions in the first world are influenced by our second-world grasp of the third world" (1979: 148–149).

A prominent aspect of the third world is its autonomy, a point illustrated in several ways by Popper. For instance, he describes the content of a book as a third world object, and states that what makes it a book is something abstract, more specifically "its possibility or potentiality of being understood, its dispositional character of being understood or interpreted, or misunderstood or misinterpreted", and he claims that "this potentiality or disposition may exist without ever being actualized or realized" (1979: 116). In the same way, the abstract content of a book exists independently of its author, although there is (normally) not an arbitrary relationship between the book and its author.

Popper observes that although the third world has independent existence, it is a human creation (1979: 112–115). Objective knowledge is a product of human behaviour: it is a result of problem-solving and discovery carried out by humans in order to cope with the first (and possibly also the second) world. Moreover, the third world has an important "feed-back effect" upon our consciousness (1979: 112, 119, 147–148), and in that way the growth of objective knowledge is due to an interaction between humans and the third world.

Through a discussion of animal behaviour Popper arrives at his distinction between behaviour and the product, or structures, resulting from behaviour (1979: 112–114). The study of these structures gives rise to two types of problems: first, problems dealing with the methods used when producing the structures (e.g. the problems involved in a spider's act of weaving its web), and, second, problems dealing with the structures themselves (e.g. the problems related to the cobweb). Then, applying this distinction to human behaviour, especially to language and science, Popper takes an anti-behaviouristic and anti-psychologistic stance in stating that understanding the problems connected with the products is the basis for understanding the production problems. Moreover, he claims that "we can learn more about production behaviour by studying the products themselves than we can learn about the products by studying production behaviour" (1979: 114). If we relate this statement to Popper's conception of knowledge growth, we may see that the impact of objective knowledge on human behaviour can be greater than the impact of individual human behaviour on objective knowledge.

2.2.2 Translation in relation to Popper's theory

In our view it is highly interesting to discuss translation in the light of Popper's epistemological framework because of the two-sidedness of this phenomenon: translation consists of both a process and a product, and the two are mutually dependent. Having looked at Popper's theory, the return to translation brings forth the question of whether the study of the product of translation is basic to the study of the translation process, and the related question of whether it is fruitful to study the product of translation prior to a study of the translation process. Before trying to answer these questions in 2.2.4, we will here locate the objects involved in translation within Popper's different domains of reality, and then in 2.2.3 relate the different approaches to translation to Popper's framework.

The translator, as a physical object, naturally belongs to the first world. With respect to translation competence, we have in 1.2 described it as including the following components: knowledge of source and target language systems, and of how these systems are interrelated; background knowledge of various kinds; skill in

interpreting source language texts, which includes the recognition of pragmatic, stylistic and formal aspects of the texts, and skill in producing target language texts which satisfy relevant demands of equivalence.[2] The two skills mentioned are second world objects. With respect to the different knowledge components, their world status is not unique. The knowledge components have intersubjective existence insofar as they are independent of individual translators. Thus, as instances of objective knowledge they belong to the third world, on a par with the knowledge of a language system shared by the members of a language community. On the other hand, as components of the subjective knowledge of a specific translator the two skills belong to the second world of mental objects. The manner in which they are represented in the brain of an individual is a first world object.

The translation process, consisting of a series of information processing steps in the translator's mind, is a second world object, and so is each discovery, or creation, of a target expression in the translator's mind during the translation process. On the other hand, a particular translation strategy (such as the choice of resolving all reception problems before beginning to produce the target text), becomes a third world object if it is formulated and made intersubjectively available.[3] But as long as it remains an individual course of action, it is a second world object.

While the physical realisations of specific source and target texts belong to the first world, the product of an act of translation is, like the content of a book, a third world object, and so is the corresponding source text. After the product of the translation process is output, and thus in principle intersubjectively available, the relation between original and translation is an object of the third world. The set of translational interrelations between the source and target language systems is also a third world object, but holds between different types of entities than the translational relation between specific source and target texts do. While the former is a relation between linguistic types, the latter holds between linguistic tokens. This point is developed further in 2.3.1.

---

[2] The description given in 1.2 of a translator's ability to construct a target text has here been modified in accord with the discussion of translational equivalence in 1.4.1.1.

[3] Chesterman (1997: 91) makes a quite similar point regarding the world status of translation strategies.

2.2.3 Translation studies in relation to Popper's theory

Several different approaches to the study of translation were presented in 1.4.1.1–3. We will now return to some of the concepts, categories, and methods discussed in that connection in order to relate these to Popper's framework.

Starting with Koller's description of translational equivalence (cf. 1.4.1.1), we may observe that the equivalence relation, as a specification of the properties with respect to which original and translation should be equivalent, exists independently of individual text recipients, and is thus a third world object. However, not all of the properties involved belong to the third world. With respect to denotational equivalence, the extra-linguistic state of affairs described by the source text may be a physical object, a mental object, as well as a third world object, but the denotation relation between a linguistic expression and the described state of affairs belongs to the third world as a part of the language system. Both connotative effects, and pragmatic aspects, of source and target text are dependent on the subjective experience of, and understanding by, a text recipient. Connotative and pragmatic equivalence thus involve second world objects, although the links between certain linguistic expressions and specific connotative and pragmatic effects may belong to the domain of objective knowledge insofar as such links are shared by a community of language users. Text-normative and formal-aesthetic equivalence also pertain to third world objects, since the textual properties they involve exist independently of the individual language user.

Toury's notion of norms in translation was briefly commented on in 1.4.1.1, where we noted that according to Toury (1965: 65), the norms govern translation behaviour, but the norms themselves are not available for observation. If translation norms govern the production of translations, then they are included among the components of translation competence. Following the discussion in 2.2.2, it is our view that translation norms, as components of the subjective knowledge of a specific

translator, are mental objects of the second world, but insofar as the norms are shared by different translators, they are intersubjective entities of the third world.[4]

Turning to House's model of translation quality assessment (cf. 1.4.1.2), we may pass lightly over her concept of translational equivalence since it is, like Koller's, an objective relation between third world objects. With respect to her notion of a cultural filter involved in covert translation, this, too, belongs to the third world, as an over-individual entity. As regards the task of translation quality assessment, it applies to third world objects, i.e. source and target texts, while the evaluation itself takes place in the second world: the comparison of textual profiles is an instance of information processing in the mind of the evaluator. Once the evaluation is done, however, its result becomes an object of the third world, as a piece of objective thought content that may be discussed and criticised.

Finally, we may briefly consider the different models of the translation process (cf. 1.4.1.3). As we have seen, there is no isomorphy between Krings' three-phase model of the course of a translation task and earlier two- or three-phase models. Furthermore, in the earlier models, the three stages described as analysis, transfer, and synthesis are aspects of translation which have been abstracted away from the actual process, from that which happens in real time, and as abstractions made by translation researchers and integrated in theories of translation, they are third world objects. On the other hand, the three phases identified in Krings' model are psycho-logical processes, and hence objects of the second world. Consequently, the modern process-oriented studies differ from the earlier approaches with respect to the world-status of the object of investigation.

2.2.4 The present approach

The approach taken in our analysis of translational complexity conforms with Popper's epistemological framework, and we adhere to his view that the third world creates "its own *domain of autonomy*" (1979: 118). Our empirical point of departure is the translational relation as instantiated by intersubjectively available parallel texts.

---

[4] The latter point is also made by Chesterman (1997: 78). See Chesterman (1997: 63–70) for a systematic over-view of norms in translation.

Thus, our object of study is the product of translation, a third world phenomenon. The translational relation will be discussed further in 2.3 with subsections.

Since the translational correspondences studied in our investigation are correspondences between third world objects, they are themselves of the third world and hence create an autonomous domain. In our view this domain is a pool of information about a part of the extension of the translational relation between Norwegian and English, and we regard this domain as something we may learn from. We will even claim that this pool of information shows that it is fruitful to study translational correspondences in relation to source and target language systems and independently of the cognitive capacity and choice of strategy of individual translators.

With respect to the question raised in 2.2.2 of whether the study of the product of translation is basic to the study of the translation process, it is our opinion that the opposite cannot be true: the study of the process cannot be basic to that of the product. Product-oriented works like Koller's typology of equivalence and House's model of translation quality assessment demonstrate that it is possible to discover facts about the translation product without studying the process. We will even regard certain findings of process-oriented translation studies as supportive of the popperian view. For instance, Krings' description of the translation process is full of references to the *result* of the translator's activity, and it is difficult to imagine how to categorise the different phases of the translation process without relating them to the product. In other words, it seems unlikely that the described process itself, a second world object, can be isolated as an object of study without considering third world objects, the products. Also, as pointed out in 1.4.1.3, TAP studies have revealed a great degree of heterogeneity among translators at work. This implies that the product of translation is at least to some extent independent of translation method. On the other hand, it does not imply that the translation process is independent of its intended product.

Rather, in the case of translation it is the product and its relation to the original text which gives the process its identity: unless a certain psycholinguistic process creates a translation, it cannot be identified as a translation process. We do not claim that the study of the translation process is unimportant, but we believe that even in process-oriented investigations of translation it is useful to consider the relations

between the product and the source text, and that is our answer to the question, also raised in 2.2.2, of whether it is fruitful to study the product of translation prior to a study of the translation process. Translation research has accumulated substantial knowledge about the product, and this knowledge seems a most advantageous point of departure for further explorations into the translation process.

2.3 The translational relation

We regard translational relations as correspondence relations holding between languages as well as between linguistic items of different languages.[5] In 1.1 we have described the translational relation between parallel texts of two languages as constituting parts of the extension of the translational relation between that pair of languages. This indicates that relations of translation exist on two different levels. On the one hand, they exist on the level of linguistic usage, i.e. between items of situated language, ranging from single word utterances to entire texts. On the other hand, they exist on the level where language is seen in abstraction from usage, i.e. between units of language systems as well as between entire language systems. This distinction is the topic of 2.3.1.

We will follow Dyvik (1998, 1999, 2005), who treats the translational relation as a theoretical primitive. Thus, the concept is "not to be defined in terms of other concepts, but assumed to be extractable from translational data by interpretive methods" (2005: 27), and the translational relation between two languages can be seen as given since it has an empirical basis "in the ubiquitous activity of practical translation" (2005: 27). The activity of translation takes place in a (cross-linguistic) language community, and bilingually competent informants may share judgments concerning the appropriateness of specific translations of given source texts. Such convergence among language users with respect to the acceptability of translations provides an empirical basis for identifying translational correspondence relations as part of the extension of the translational relation. For language pairs where modern

---

[5] Toury (1995: 77), on the other hand, claims that "translation relationships … normally obtain first and foremost between TEXTUAL SEGMENTS, very often even small-scale, rather low-level linguistic items."

parallel corpora (see 1.4.3 with subsections) are available, there are now excellent opportunities for investigating such correspondence relations.

Dyvik (1998, 1999, 2005) argues that translation is an important source of knowledge about the semantics of natural languages. Due to its empirical basis "the translational relation emerges as epistemologically prior to more abstract and theory-bound notions such as 'meaning', 'synonymy', 'paraphrase' and 'inference'" (2005: 27). In particular, translation is a normal type of language use, as opposed to meta-linguistic reflection, and its results are intersubjectively available (cf. Dyvik 1998: 51).

This is further developed in Dyvik (1999: 217–218), where he discusses the difference between meaning properties and translational properties. The observable relations between pairs of source and target texts allow us to discover translational properties of words and phrases in the texts. Those properties provide a key to meaning properties since the words and phrases of a language have translational properties in common "only if they share meaning properties" (1999: 218). As translational properties are observable in cross-linguistic data, they are "epistemo-logically more accessible" than meaning properties, which have traditionally been analysed through methods with elements of subjective judgment (1999: 218). Thus, the epistemological status of translational properties supports treating the trans-lational relation as a theoretical primitive. The translational relation between langu-ages is "assumed to be extractable from translational data by interpretive methods" (Dyvik 2005: 27), which involve distinguishing aspects of the language *system* from those of language *use* in the translational relation between texts.

2.3.1 A phenomenon of *langue* or *parole*?

Thus, a relevant point in connection with the translational relation is the saussurean distinction between the language system seen in abstraction from actual language use, *la langue*, and the language when used as a means of communication, *la parole*. Again, we adhere to Dyvik, who points out that as a relation between situated texts the translational relation holds between items on the level of *parole* (1998: 51–52). This follows from the fact that the translation of a specific source text is shaped not

only by the linguistic expressions used in the original, but also by "the context of utterance, the purpose of the utterance, and various other kinds of background knowledge" (1998: 52). Thus, translational correspondences between texts may be determined not only by information about the source and target language systems and their interrelations, but also by additional information sources.

However, the translational relation can also be seen as a relation between languages, and then holding between items on the level of *langue*. Dyvik argues that studying the translational relation as a *langue* phenomenon implies that we "disregard translational choices that can be motivated only by reference to the particular text and its circumstances", and this is the basis for isolating "translational correspondence relations between the sign inventories of the two languages — relations between words and phrases seen as types rather than textual tokens" (1998: 52).

The type-token distinction is important in our empirical investigation. When we analyse the product of translation instantiated as translationally corresponding strings of words, we regard the corresponding strings as linguistic types (cf. 4.3.6.2), but since the activity of translation applies to situated texts, we cannot account for the relation between a specific string and its correspondent without paying attention to the factors governing language use. In particular, these factors determine the possible interpretations that may be assigned to the corresponding strings, which again influence the analysis of translational complexity in the string pair.[6]

2.3.2 Predictability in the translational relation

As indicated in 1.1, our investigation aims at finding out to what extent it is possible to automatise translation in the case of selected English-Norwegian parallel texts representing two specific text types. This presupposes viewing the translation task as a kind of computation.[7] The problem may also be described as the following: given a certain source language expression, how far is it possible to predict its target language correspondent? We assume that if we could have access to information

---

[6] This point is discussed further in 4.3.6.2.

[7] Cf. section 3.2.2, which comments of the topic of viewing different kinds of human language processing as instances of computation.

about all factors that may influence the choice of target expression, then we would be able to predict the translation.

Prior to translation a source text is located in a domain of discourse. When a translator has created a target text that is regarded as an optimal translation, he or she has been as well informed as possible regarding the choice of target expression. That is, all necessary information has been available to the translator in the given domain of discourse. Likewise, in order to achieve automatic translation this information must be represented in an accessible format prior to the translation task. Hence we assume that the translational relation is predictable insofar as the source text together with *a pre-structured domain of information* can provide all the information needed to produce the target text.[8]

Is it then possible that this pre-structured domain can contain information about all factors which, in addition to the source text, have an influence on the choice of target expression? As discussed in 2.3.1, the translation of a specific source text is determined not only by the source and target language systems and their inter-relations, but also by "the context of utterance, the purpose of the utterance, and various other kinds of background knowledge" (Dyvik 1998: 52). We will assume it is possible to describe language systems and their interrelations and to include repre-sentations of such information in the pre-structured domain — i.e. to capture the domain of translationally relevant *linguistic* information.[9] By this assumption we follow Dyvik (1998, 1999) where the notion of 'linguistic predictability' is used to distinguish the translational relation between situated texts from the translational relation between the sign inventories of two languages. Dyvik's point is that to identify the translational relation on the level of *langue* is to isolate "the linguistically predictable translations" between two languages (1998: 52).

---

[8] In 2.4.2.1–3 we discuss the information sources which we assume to be included in this pre-structured domain, as well as sources falling outside of it.

[9] This assumption may appear to be in conflict with the point made in 1.4.2.3 that, so far, no natural language has yet been fully described in a computer-implementable format. However, that this has not yet been done, does not mean that it is theoretically impossible to provide a full-coverage computational grammar for a given language. Our assumption is that it is *in principle* possible to describe all parts of a language system, given that all parts of it are known and that there exists a grammar formalism in which those parts may be represented.

When interpreting a given source text, a translator will also exploit relevant non-linguistic information that he or she has access to. Thus, we regard such non-linguistic information as included in the domain of discourse of the source text, and our question is then to what extent this, too, can be represented in a pre-structured domain of information. We will assume it is possible to describe the information contained in *restricted* semantic domains of the world. This has been achieved in artificial intelligence systems and in various systems for natural language processing, of which automatic translation is an example. In such systems, knowledge modules represent restricted domains of technical information.[10] On the other hand, we assume it is not a manageable task to capture information about all possible domains of the world. Granted unlimited storage possibilities, the amount of world information that could be captured might be theoretically unlimited, but in practice it is necessary to draw a limit in order to secure tractability of the pre-structured domain of information.[11] Moreover, as parts of the world are unstructured, how would information about those parts be formalised?

Thus, an important property of the pre-structured domain of linguistic information is that it is finite. To be *finite* basically means to have an end or a limit. If information is represented in a finite way, it is contained in, or derivable from, a limited structure, and hence we may assume that it is in practice a feasible task to find and identify a particular informational element contained in, or derivable from, this structure.

In the present study of translation, our point of departure is not a restricted domain, but the domain of general language. Although we want to investigate whether translational complexity varies between pieces of general language texts and samples of domain-specific texts, we have chosen to limit the pre-structured domain to information about the source and target language systems and their interrelations. We regard this a necessary and helpful restriction as it provides a principled delimitation of the pre-structured domain, and also puts a theoretical limit on the extent to which the translational relation is predictable. Our analysis of translational

---

[10] Cf. the discussion of restricted semantic domains and sublanguages in 1.4.2.3.

[11] *Tractability* in a technical sense is explained in 3.2.1. Here the word is used in a general sense. According to the *Longman Dictionary of Contemporary English* (3rd ed.), the adjective *tractable* means 'easy to control or deal with'.

correspondences will demonstrate the consequences of this delimitation in relation to empirical data, and we shall see that the limit of predictability in the translational relation will be relative to certain presuppositions concerning the descriptions of the languages involved. In particular, this limit depends on where the division is drawn between linguistic and extra-linguistic information, to be described in 2.4.2.1.

As pointed out above, Dyvik (1998: 52) argues that to identify the translational relation between the sign inventories of two languages is to find the linguistically predictable correspondences of that language pair. Such sign correspondences are linguistically predictable because they hold between signs with shared meaning properties (cf. Dyvik 1999: 217). This should, however, not be understood as if our criterion for distinguishing between linguistically predictable and non-predictable correspondences is exclusively the presence or absence of shared meaning properties. Other properties than those related to meaning may also be shared in a linguistically predictable correspondence between an SL sign and a TL sign. E.g., syntactic properties may be shared between translationally corresponding phrases if source and target language are structurally related. The criterion of shared meaning properties specifies what must *at least* be present in a linguistically predictable correspondence.

The set of *linguistically predictable translations* of a source language sign, its LPT set, is the full set of target language signs sharing a maximum, given the TL, of meaning properties associated with the SL sign (cf. Dyvik 1998: 56–57). That is, since language systems are differently structured in terms of grammar and lexical inventory, we cannot, within the scope of general language, expect that all meaning properties associated with a given SL sign is present in each member of its LPT set.[12] Then, taking into account differences between the two language systems, the LPT set of a given SL sign is the set of TL signs exhibiting a maximum of the meaning properties of the former. In the case of specific translational correspondences, it is shared intuitions among bilingually competent language users which decide what properties are included in this maximum. Furthermore, to describe a given target

---

[12] This point is also made in 6.3.2 in connection with denotational non-equivalence in translation.

language expression as a linguistically predictable translation of a source expression means that the former is one of the members of the LPT set of the latter.

The LPT set of a given SL sign may have zero, one, or more than one, member. There may be cultural or other differences between source and target language causing the situation where there is no TL sign associated with the meaning properties expressed by the SL sign. In cases where the LPT set is empty, translators may solve the problem by paraphrasing the source expression, and *parole*-related factors such as the use of world knowledge or contextual information will contribute to the choice of target expression. Consider the following example, found among the recorded data:

(1a) Det var ikke skiføre lenger,   (BV)
      'It was not conditions-for-skiing longer.'
(1b) It was no longer possible to ski,

The Norwegian noun *skiføre* means 'conditions for skiing', and has no lexical correspondent in English. The source sentence (1a) describes the situation where it is impossible to ski because there are no longer suitable conditions for it. The English translation (1b) is a paraphrase of this, chosen on the basis of general world knowledge.

In cases where the LPT set has exactly one member, there is a one-to-one correspondence between source and target language sign. An example would be the relation between a technical term in the source language and its target language equivalent. In general language it is a more common situation that the LPT set includes more than one member, and in such cases translation involves making a choice between the alternative target expressions. However, it may depend on the circumstances which member will be the optimal translation among the predictable candidates (cf. Dyvik 1998: 56). Such *parole*-related factors may also motivate a translation which is not an LPT member.[13]

---

[13] This point is illustrated by several of the phenomena discussed in chapter 6. Cf. e.g. the analysis of example (21) in 6.3.1.3.

For the sake of illustration, consider the English noun *pencil*. In the general sense of 'writing instrument' its LPT set with respect to Norwegian is {*blyant*}. The following is an example where *pencil* is not translated into *blyant*:

(2a)   Got a pencil?[14]
(2b)   Har du noe å skrive med?
        'Have you something to write with?'

The source text (2a) is found in a dialogue context: two characters are talking on the phone; one of them has important information to share with the other, and the question (2a) is uttered when the former person wants to make sure that the listener is able to write down the details contained in the information. In this context it is possible to choose the translation *Har du en blyant?* ('Do you have a pencil?'), but instead the translator has picked the semantically less specific expression *Har du noe å skrive med?* ('Have you something to write with?'). Thus, *pencil* corresponds with *noe å skrive med*. The chosen translation may be said to be pragmatically equivalent with the source text, as there is focus on the fact that the addressee needs a writing instrument, and not necessarily a pencil. In this sense the textual context has motivated the choice of a translation of *pencil* which falls outside its LPT set.

On this background we may draw a distinction between predicting translations and generating specific target texts. To *predict* the translation(s) of a given source expression is to identify its LPT set; i.e. to find the set of target expressions sharing a maximum of meaning properties associated with the original. To *generate* a specific translation from an original may involve accessing other information sources than the information expressed linguistically in the source text, and it may involve making a choice between several alternative translations, among which some may be linguistically predictable, and some may be not.

The distinction between linguistically predictable and non-predictable translation can be related to the notion of 'computability', which will be discussed in chapter 3.

---

[14] The example is taken from Sue Grafton's novel *"D" is for Deadbeat*; see the list of primary sources. The novel is included in the ENPC (cf. 1.4.3.2).

As a first approximation, a 'computation' may be defined as a step-by-step procedure for solving a task in a specific way, and, thus, a *computable* task is a task that can be solved by a specifiable procedure. In the beginning of this section, we presented the assumption that the relation between a source language expression and its translation is *predictable* provided that the source expression together with a pre-structured domain of information can provide the translator with the information needed to produce the target text. Moreover, we have restricted this domain to include information about source and target language systems and their interrelations. A translation task, then, is computable if an automatic translation procedure is able to produce the target text correctly by exploiting the pre-structured domain of linguistic information. In this sense, we regard the computable part of the translational relation as identical to the linguistically predictable part.

To sum up, our investigation of how far it is possible to automatise translation in selected English-Norwegian parallel texts is based on assumptions regarding the limit of linguistic predictability in the translational relation. We assume that the linguistically predictable part of the translational relation is limited to the level of correspondences between *langue* units, and that it is computable from the information contained in the source expression, together with pre-structured information about the source and target language systems and their interrelations.[15]

2.3.3 The notion of 'literal translation'

Through the notion of linguistically predictable translations, Dyvik (1999: 217) explains a further notion of 'literal translation': "...the meaning properties of a sign are precisely the set of properties we want to capture, if we can, in literal translation." Thus, literal translation covers predictable correspondences between signs of two different languages; it deals with LPT sets, and it does not cover translations involving *parole*-related factors. In the present approach *literal translation* and *linguistically predictable translation* are synonymous expressions.

---

[15] This topic is revisited in 2.4.2.1, discussing the distinction between linguistic and extra-linguistic information, and in 3.2.5, describing computability in relation to translation.

When relating literal translation to meaning properties care must be taken to avoid circularity. If literal translation is defined in terms of meaning properties, then the translational relation is no longer a primitive, and our task is to clarify why it is *plausible* to assume that literal translation and meaning properties are related in the manner described above. In this respect, we have in 2.3 cited Dyvik (1998, 1999, 2005), who argues that since bilingually competent informants may share judgments on the appropriateness of given translations, there is an empirical basis for identifying the translational relation. Hence, the literal translational relation can be assumed to be elicitable from informants without resort to meaning descriptions. Then we can use the relations, given our plausibility arguments, to describe meaning properties.

Literal translation in the sense described here must not be seen as related to the notion of 'literal translation' defined by Vinay and Darbelnet (1995) as a translation method: "Literal, or word for word, translation is the direct transfer of a SL text into a grammatically and idiomatically appropriate TL text in which the translators' task is limited to observing the adherence to the linguistic servitudes of the TL."[16] The product of literal translation in the sense of Vinay and Darbelnet matches types 1 and 2 in our correspondence type hierarchy, whereas types 1, 2, and 3 are included in Dyvik's concept of a literal translational relation. Then, we find a closer match between Dyvik's notion and the product of Newmark's (1981: 39) concept of literal, or semantic, translation, which he has defined as the translation method that "attempts to render, as closely as the semantic and syntactic structures of the second language allow, the exact contextual meaning of the original."[17] Chesterman (1997: 12) sums up the various understandings of literal translation by observing that they have in common an emphasis on "closeness to the original form."

For the purposes of the present study, *literal translation* refers only to Dyvik's concept, which primarily serves to describe the relation between source and target text, and must not be associated with translation methods.

---

[16] The quotation is taken from Venuti (2000: 86). Pages 31–42 of Vinay and Darbelnet (1995) are reprinted in Venuti (2000: 84–93).
[17] Cf. Palumbo (2009: 49, 70, 167).

2.4 Information sources for translation

The topic of information sources for translation was introduced in chapter 1. Section 1.2 presented a tentative overview of our description of the types of information needed to produce a specific translation from a given source text, and information sources for translation were briefly mentioned in the context of automatic translation (cf. 1.4.2.3–4). The basic notions of information, informational content, and knowledge will be discussed in 2.4.1 with subsections, before we present our typology of information sources for translation in 2.4.2 with subsections.

2.4.1 Basic notions

In the preliminary version of the typology presented in 1.2 two important, basic notions are 'information' and 'knowledge'. These are concepts used by laypersons as well as by specialists. In non-technical discussions among laypersons these notions are normally taken for granted, as concepts that we all have an intuitive understanding of, whereas within a certain field of study, such as linguistics, information theory, or philosophy, the same concepts may be used in specific, technical senses. Our understanding of these, and related, notions are presented in 2.4.1.1–5.

2.4.1.1 Information

There is similarity, but also important differences, between Popper's concept of objective knowledge, which exists without a knowing subject, and the notion of information found in communication theory, i.e. information existing independently of any interpreting, cognitive agent. The work of Shannon and Weaver (1949) is commonly recognised as the origins of communication theory (also referred to as *information theory*). Our notion of 'information' is borrowed from this science, and the present discussion is based on Dretske (1981), whose project is "an attempt to develop an information-based theory of knowledge" (1981: 3), an attempt to apply the insights of communication theory in order to develop "a genuine theory *of information* as this is understood in cognitive and semantic studies" (1981: 4).

Within communication theory 'information' is understood as "an objective commodity, something whose generation, transmission, and reception do not require or in any way presuppose interpretive processes" (Dretske 1981: vii). Rather, what defines information are relations holding between distinct states, events, and structures (1981: x). In contrast to cognitive and semantic studies, communication theory treats information as a purely quantitative notion: the theory deals only with *amounts* of information, not with informational *contents* (1981: 3). Thus, information is either present or not; it is something that can be measured. Unlike notions like beliefs and propositions, information cannot be either true or false: its existence requires truth, and as Dretske points out, this property of information (in the technical sense) has the consequence that 'false information' or 'mis-information' are inconsistent concepts (1981: 45).

Further, "the amount of information associated with, or generated by, the occurrence of an event (or the realization of a state of affairs)" is measured in terms of "the reduction in uncertainty, the elimination of possibilities, represented by that event or state of affairs" (1981: 4). Thus, the emergence of a state or occurrence of an event for which there is an overwhelming probability represents very little information, whereas an unexpected state or event represents a relatively large amount of information (1981: 8–9).

Dretske points out that "*any* situation may be regarded as a *source* of information" (1981: 9). The focus of communication theory is on information sources, on measuring average amounts of information available from such sources; the theory does not aim to describe particular pieces of information, which would be of interest in semantic studies (cf. Dretske 1981: 10–11, 47, 52–53). Thus, although we want to exploit the information concept, we do not share the focus of communication theory, as our analysis will deal with particular pieces of text.

Dretske observes that communication theory has by some been viewed as "a theory of *signal transmission*, a theory about those physical events (signals) that, in some sense, carry information" (1981: 40). This yields a mathematical theory of information which describes statistical and other properties of signals, but, as he further points out, "[a] genuine theory of information would be a theory about the

content of our messages, not a theory about the form in which this content is embodied" (1981: 40). The distinction between, on the one hand, the signal as a physical event governed by probabilities, and, on the other hand, the informational content carried by the signal emphasises Dretske's view that the study of information involves not only those properties of information that can be accounted for in terms of quantitative measures, but also properties pertaining to the content of a particular piece of information. Studying the latter falls, as we have seen, outside the scope of communication theory.

Although 'information' in the ordinary, non-technical sense may be viewed as a semantic notion, Dretske warns against merging it with the concept of 'meaning': "... there is no reason to think that every meaningful sign must carry information or, if it does, that the information it carries must be identical to its meaning" (1981: 42).[18] He thus keeps 'meaning' strictly apart from the communication-theoretic concept of 'information' (1981: 41–44), and in his view meaning is a product manufactured from information (1981: vii). It may seem that communication theory, with its quantitative focus, cannot contribute to the study of meaning. However, Dretske argues that it is misguided to assume that "*meaning* is the *only* semantically relevant concept" (1981: 46). Information, as "[a] commodity capable of yielding knowledge", is also semantically relevant, and for that reason Dretske finds it fruitful to apply the insights of communication theory also when studying the semantic aspects of information (1981: 46). The information concept is relevant to the present project because our focus is on the various pieces of information that contribute to the selection of a given translation, and not merely on describing the meaning of the corresponding source expression.

2.4.1.2 Informational content

An important part of Dretske's project is to exploit the insights of communication theory in order to give an account of 'informational content'. The basic difference

---

[18] For instance, if a small child says to his parent "I have a tummy-ache", then the meaning of that signal is that he has a tummy-ache. However, if it is the case that the child has no tummy-ache — only happened to utter this sentence to get attention — then the signal does not carry the information that he has a tummy-ache.

between the amount of information a signal carries and its informational content is that the latter cannot be quantified. While it makes sense to ask whether a certain signal carries more or less information than another signal, it does not make sense to ask whether the informational content of that signal is larger or smaller than the content of the other signal (cf. Dretske 1981: 47–48). This illustrates how a study of the semantic aspects of information necessitates a shift from the communication-theoretic focus on average amounts of information to a focus on particular pieces of information.

To phrase it in very general terms, informational content is the information that something is the case. Dretske uses *message* as a synonym of *informational content* (see e.g. 1981: 55), and in his notation informational content is the information "that s is F", where "s is F" is used as a shorthand for some state, event, or structure, the lowercase s indicating an information source (cf. 1981: 66). His explication of informational content involves describing what conditions must be satisfied when a signal r carries the information that s is F (1981: 63–65). Firstly, the signal cannot carry a smaller amount of information than the amount generated by the state of affairs described as "s is F". This is a purely quantitative condition, and it illustrates the point made by Dretske (1981: 60) that to communicate a specific informational content, i.e. to convey a particular message, requires that *all* the information behind that message, and nothing less, must be transmitted. The second condition on informational content states that the signal r cannot carry the information "that s is F" unless s really is F, and the third condition states that r must carry the same information as that generated by s's being F. The latter two restrictions are of a qualitative kind, or, in the words of Dretske, they "together constitute ... the semantic conditions on information" (1981: 64).

In addition to these three conditions, the informational content carried by a signal r is influenced by information already available to the recipient from other sources than r, in particular information about the conditions governing the probability of the informational content carried by r. When measuring the amount of information generated by some source, information is needed about the set of alternative possibilities existing at the source, the absolute probability of each of these possibilities,

and the probability of each of these possibilities relative to conditions governing the transmission of information from the source (cf. Dretske 1981: 43, 53–56). Hence, the amount of information gained by the recipient of a signal is influenced by information already available to the recipient with respect to the probabilities of the alternative possibilities, and in that way background information may determine the informational content that is transmitted by a specific signal to the recipient.

This may be illustrated by a simple example: if we already have the information that it is daytime, then receiving the signal of twelve bell strokes will tell us it is noon, because that is a far more probable state of affairs than the alternative of midnight. Thus, background information plays a part in Dretske's eventual definition of informational content (1981: 65): to say that a signal r carries the informational content "that s is F" means that there is a maximal probability for s being F, given r and available information concerning the possibilities existing at the information source, and that there would not have been such a maximal probability without the signal r. Thus, if s being F is the cause of the signal, then the signal has the informational content that s is F provided that there is no possible alternative cause of the signal, given available information about the possibilities. That is, something contains information about its cause only if other causes are impossible: frozen water tells us that the temperature in that water is below zero degrees Celsius, because temperatures above zero cannot cause water to freeze.

What is here referred to as "background information" is in Dretske's definition labelled "k" and described as "what the receiver already knows (if anything) about the possibilities that exist at the source" (1981: 65). In 2.4.1.3 we shall see that he conceives of knowledge as something existing within the mind of the cognitive agent. Thus, Dretske may seem to imply that informational content is dependent on the state of mind of the recipient — on how the signal is interpreted by the recipient. We prefer to regard this as an inaccuracy in his description of informational content, and we have chosen to read k as 'background information'. Elsewhere Dretske stresses that the conditional probabilities of the possibilities existing at the information source are objective features, that they are not determined by how likely the recipient believes each possibility to be, and that the amount of information carried by a signal

is independent of how much information the recipient is able to gain from it (1981: 55–57). Thus, background information influences the informational content of a signal regardless of whether the signal has been absorbed by the recipient or not, and informational content exists, like information, as an objective commodity, independent of interpretive processes.

In the present study we will relate the notion of informational content to the analysis of translational correspondences. More specifically, the concept will be applied when we describe semantic divergences between translational units in chapter 6. E.g., the discussion will show that differences with respect to amounts of information may have the effect that source and target text do not convey identical messages, and that a certain expression may carry different messages depending on whether specific background information is available or not.[19]

2.4.1.3 Knowledge

Dretske presents the traditional conception of 'knowledge' as "justified true belief" (1981: 85), and points out that as long as the notion of 'justification' is left unanalysed, this is not a satisfactory account. In his approach 'justification' is linked to information: the true belief that something is the case (s is F) counts as knowledge only if it is caused by the information that s is F (1981: 86). As described in 2.4.1.1, information, according to Dretske, requires truth, so that 'false information' becomes an inconsistency, and thus the causation of a belief by information amounts to a justification of that belief.

A consequence of this account is that instances of true belief do not necessarily count as 'knowledge'. It is possible to form a true belief without having received information supporting the belief. For instance, if Mary takes a look in the fridge and perceives some round fruits of red and yellow colour in a semi-transparent plastic bag, she may believe there are nectarines in the fridge. But the plastic bag contains apples, and she has mistaken the apples for nectarines. However, as there happen to be nectarines, too, in the fridge (hidden in a paper bag), her belief is true. But she has

---

[19] The former point is relevant to the discussions in 6.3.1 with subsections, and the latter point is illustrated by the analysis of example (28) in 6.3.2.3.

not received any information about the nectarines, and thus she does not have the knowledge that there are nectarines in the fridge.

Dretske underlines that his account of knowledge is intended as a description rather than as a definition of 'knowledge', as an explication of what 'knowledge' *is* rather than of what it *means* (1981: 91–92). To see knowledge as information-caused belief is to understand knowledge as a property of individual minds, as a state of mind of information-receiving cognitive agents.

There is a clear difference between Dretske's account of knowledge and Popper's concept of objective knowledge:[20] while the latter exists independently of particular knowing subjects, the former is understood as a state of mind of the individual. In Popper's terms, Dretske's 'knowledge' is a second world object, whereas 'objective knowledge' belongs to the third world. 'Knowledge' as described by Dretske corresponds, at least partly, with Popper's notion of 'subjective knowledge' (cf. Popper 1979: 108). Popper's 'objective knowledge' is of a more abstract kind than Dretske's 'knowledge': objective knowledge, being a result of human activity, presupposes past or present knowledge states in humans, but cannot be reduced to such knowledge states. Objective knowledge exists in the form of *shared content* of different knowledge states (perhaps caused in different ways) in human minds, and we have to ascribe a sort of intersubjective existence to this shared content in order to account for human interaction with it. Through this intersubjectivity the popperian 'objective knowledge' becomes a more abstract object than Dretske's 'knowledge', and it may seem as if Dretske, when viewing 'knowledge' as a cognitive object, does not draw the distinction made by Popper between the content of knowledge and how knowledge is represented in the mind of the individual.

2.4.1.4 Knowledge and information compared

It is clear from the preceding discussion that knowledge and information are different commodities, and a further comparison of these concepts is relevant for our later discussion of information sources for translation. Although we want to adhere to the

---

[20] Cf. the discussion of 'objective knowledge' in 2.2.1.

conception of 'information' as given by communication theory, Dretske's information-based description of 'knowledge' does not quite suit our purposes, since it is understood as a state of mind, and our object of study is a third world phenomenon (cf. 2.2.4).

We have previously stated that our investigation conforms with Popper's epistemological framework, and an important similarity between Popper's concept of objective knowledge and the notion of information found in communication theory has already been pointed out in 2.4.1.1: objective knowledge exists independently of the knowing subject, and information exists whether there is any interpreting agent or not.

There are also differences between the two notions, and a few of these could be mentioned. First, we have seen that Popper views objective knowledge as a product of human activity; the creation of objective knowledge requires knowledge acquisition in humans (cf. 2.2.1). Conversely, human activity is not a prerequisite for the creation or existence of information (although, of course, some information is information about humans and their activities).

Second, in Popper's concept there is focus on knowledge *content*, whereas information, as we have seen, is a quantitative notion. Objective knowledge is described as contents of thought, commodities that cannot easily be measured in the way that information is measured in terms of reduction in uncertainty.

A third difference between information and knowledge pertains not only to the popperian 'objective knowledge', but also to Dretske's 'knowledge': Dretske makes the point that knowledge and information belong to different "orders of intentionality" (1981:171–175). Physical structures and signals represent intentional states of the lowest order. When a signal carries information about its source, it occupies an intentional state relative to the source (cf. Dretske 1981: 172). While signals exhibit low-order intentionality, beliefs, knowledge, and meaning represent higher-order intentional states. According to Dretske (1981: 172) it is the ability to occupy higher-order intentional states that distinguishes information-processing systems with cognitive attributes from those that are unable to perform cognition. He explains this in a way which highlights the *selective* character of knowledge (higher-order) as

opposed to information (lower-order). It is not possible for a system exhibiting low-order intentionality to carry the information that p without also necessarily carrying all information that follows from p, either analytically or by natural law.[21] For instance, the information that a given amount of water freezes necessarily includes the information that the water is expanding. This property of information is described by Dretske as "nesting" (1981: 71, 179). However, it is possible for a system with cognitive attributes (e.g., a person) to have the *knowledge* that p without necessarily having the knowledge of everything that follows from p. Dretske's example is that it is possible to know that the solution to an equation is 23 without knowing that the solution is also the cube root of 12167 (1981: 173).

Then, how is the selective character of knowledge related to Popper's concept of 'objective knowledge'? It seems clear that the property of knowing p without knowing everything that follows from p pertains to the cognitive agent rather than to the objectivised knowledge content. Moreover, in arguing for the separate existence of objective knowledge, Popper makes the point that a theory may have consequences which nobody has discovered yet (1979: 116). That is, the content of a theory comprises everything that follows from it, whether anybody has realised it yet or not. This indicates that objective knowledge does not have the same selective quality as subjective knowledge; and it indicates a further similarity with 'information' in the technical sense. It seems that if we may assume the existence of the objective knowledge that p, Popper would also assume the existence of at least all analytic consequences of p. Still, Popper's 'objective knowledge' is distinct from the concept of information: because objective knowledge originates in subjective knowledge states in human minds, objective knowledge inherits a higher order of intentionality than that of information.

2.4.1.5 The knowledge of translators

As previously accounted for, our object of study is the product of translation, which, in our view, may serve as a reflection of translation competence.[22] In 1.2 we presen-

---

[21] For the sake of convenience "p" is used, like "s is F", as a shorthand for some state, event, or structure.
[22] Cf. 1.2 and 2.2.4.

ted our conception of translation competence as a combination of several factors: knowledge of source and target language systems, and of how these systems are interrelated, various kinds of background knowledge, and skills in interpreting and producing text in context. The mentioned skills involve knowledge of the pragmatic factors governing the interplay between linguistic forms and textual contexts.

It was pointed out in 2.2.2 that when these types of knowledge belong to a particular translator, they fall under the notion of subjective knowledge. Similarly, the skills mentioned are also second world objects and cannot be common objects of knowledge. However, when we, in this study, analyse translational relations between texts, we observe the product independently of its production, and we assume that a certain set of translational correspondences may be produced by different translators and by various translation strategies. We think it is safe to assume this because, as already pointed out in 2.3, different bilingually competent informants may share judgments concerning the appropriateness of specific translations of given source texts. Further, this assumption presupposes the existence of objective knowledge about translational relations between texts, knowledge which can be shared by different translators and which can be exploited by different translation methods. We aim to study this objective knowledge content insofar as it is detectable by analysing translational correspondence relations in our empirical data, and we will mainly disregard the possibly varying strategies or mental procedures of individual trans-lators, although the recorded data can to some extent indicate differences concerning translators' preferences.[23] These strategies are of course legitimate and worthwhile objects of study in other contexts. Our focus is on the objective knowledge of *translators*, not on *the* translator's knowledge, and this is what we refer to when stating in 1.2 that our study is neither a cognitive nor a psycho-linguistic investigation of translation.

Then it is our task to try to find out more about the content of the objective knowledge presupposed by translational relations between texts of two languages.[24] We are interested in what is *implied* in the knowledge of translators: we do not

---

[23] The latter point will be illustrated by discussions in chapters 5 and 6.
[24] Cf. the description in 2.4.2.2 of given, general information sources for translation.

assume that actual translators use all available knowledge in every translation task, but we assume that given the existence of this knowledge there is the potential for performing the amount of analysis and inference required by each translation task. This resembles the property of information described by Dretske as "nesting" (cf. 2.4.1.4): embedded in the knowledge of source and target language and their interrelations is the knowledge required to analyse a particular piece of source text and produce a linguistically predictable translation of it. It may also be compared to the work of a grammarian: the grammarian explores and systematises what is involved in the knowledge of a given language, without assuming that the individual language user, whose knowledge the grammarian describes, is able to produce the same kind of systematisations. Explicating what is involved in a given body of knowledge is not the same as making claims about the inferences actually made by people having the knowledge. Our study of the objective knowledge of translators is a similar kind of explication, and, in line with the view taken in 2.2.4, we think that such explication can and should have its empirical basis in the observed products of the knowing subjects, which are, in our case, actual translations.

2.4.2 Typology of information sources

Sections 2.4.2.1–3 present a typology of information sources for translation, defined for the purpose of measuring translational complexity in terms of how much and what kinds of information are needed in translation. It is intended as one possible way of describing information sources for translation, and the typology is motivated by the nature of our object of study. The classification is not done according to criteria related to the cognitive equipment of individual translators, as our approach is to draw distinctions reflecting the types of information sources we assume are relevant in order to account for the observable relations between originals and their translations.

As presented in 1.1 and 1.3–1.3.2, translational complexity in our approach is associated with the need for information in translation tasks. In chapter 3 the structure of translation tasks will be described in terms of how much, and what kinds, of information are needed in order to translate. For those purposes the information

typology will be applied, as well as in the discussion of the empirical results in chapters 5 and 6. Given the analytical framework to be described in chapter 3, it will not be possible to quantify the need for information in mathematical terms; it can be analysed only insofar as each of the four correspondence types represents an upper and a lower bound on the required amount of information within its class.[25] More-over, in chapter 3 the need for information is related to two questions raised in 1.2: to what extent can the various information sources for translation be represented in a finite way, and what is the amount of effort required in order to access and process them? With respect to the issue of finiteness, it is appropriate, in this chapter, to consider whether the various information types are included in the pre-structured domain of linguistic information introduced in 2.3.2 as defining the limit on predictability (and, hence, also on computability) in the translational relation.

In the information typology, distinctions are drawn along three different dimensions. Along the first dimension we assume a division between linguistic and extra-linguistic information sources. Previously, in 2.3.2, the limit of predictability in the translational relation is associated with a pre-structured domain containing infor-mation about the source and target language systems and their interrelations. Thus, extra-linguistic information is not included in this domain. It is, however, debatable to what extent it is possible to distinguish between purely linguistic information and world information, and it is especially difficult to draw a line between the linguistic and the extra-linguistic when we enter the fields of semantics and pragmatics, which will be discussed in 2.4.2.1.[26]

Second, we assume a division between general and task-specific information sources. General information is given prior to the translation activity; it includes information about source and target languages and their interrelations, and various types of information about the world. Task-specific information comprises the

---

[25] This point is explained towards the end of section 3.2.4.

[26] In 1.2 we have indicated a preliminary tripartite division into (a) purely linguistic, (b) pragmatic, and (c) extra-linguistic information sources. In 2.4.2.1 we will argue that pragmatic information may occur in the linguistic as well as in the extra-linguistic domain of information, thus advocating a binary main division between linguistic and extra-linguistic information sources.

different kinds of information associated with a particular piece of source text and the concrete task of translating it into a given target language.

Third, we distinguish between mono- and bilingual information sources: monolingual information includes information about source and target language respectively, and the information coded linguistically in the source text. Bilingual information deals with how the two languages correspond translationally.

It should be noted that we do not assume that every one of these three dimensions is necessarily crossed by each of the other two. That is, we do not assume that the domain of information sources for translation has a geometric structure like that of a cube with three axes crossing each other. In particular, the distinction between mono- and bilingual sources is only relevant within the domain of linguistic information.

As we have made clear in 1.2, as well as above, the study of translation competence is not part of our investigation, although the information that is accessible through the competence of translators is naturally included in the typology of information sources for translation. There are certain points of relatedness between our typology and models of translation competence that have been developed within process-oriented translation studies. Hurtado Albir and Alves (2009: 63–68) provide an overview of such models. In general, translation competence models have in common that they are divided into components, and that they distinguish between knowledge modules and skills, or abilities. Further, certain distinctions seem to be shared by several of them, in particular the opposition between linguistic and extra-linguistic knowledge, and the division between general and specialised skills. According to Hurtado Albir and Alves (2009: 64), most of these models still lack empirical validation.

The three dimensions of our information typology are the concern of sections 2.4.2.1–3. It is not our ambition to provide exhaustive descriptions of these dimensions, but rather to clarify the distinctions we want to draw along them, since these distinctions are exploited in the empirical analysis of translational correspondences. Moreover, we do not assume that each and all of the information types to be discussed are available in any case of translation, although some of them, such as information about SL and TL and their interrelations, are necessarily required.

2.4.2.1 Linguistic versus extra-linguistic information sources

In our framework, this is an opposition between information derived from the source and target language systems and information about the extra-linguistic world. With respect to translational complexity, the division between linguistic and extra-linguistic information sources is closely related to the limit of predictability in observed translational correspondences in parallel texts. As previously discussed in 2.3.2, we assume that given a specific source expression, it is possible to predict a translation insofar as information about the factors that determine the translation is available in a pre-structured domain of linguistic information. Further, we argued that language systems and their interrelations can be described in a finite way,[27] and that these are the information sources included in the pre-structured domain. On the other hand, we have pointed out that to include extra-linguistic information about the world in the pre-structured domain will yield intractability, and that there must be a principled limit on the amount of information it may contain.[28]

Thus, granted that the domain of extra-linguistic information is infinite, we assume that linguistic and extra-linguistic information will show different properties in relation to translational complexity. More specifically, we assume that the degree of complexity is higher in translational correspondences involving extra-linguistic information than in cases involving purely linguistic information (cf. 1.3.1–2). But we do not *a priori* assume that processing information about the extra-linguistic world will be more complex than processing linguistic information, simply because it is non-linguistic. Intuitively, it seems reasonable to assume that there can be pieces of linguistic information which lead to greater complexity, and are harder to make representations of, than certain pieces of information about the extra-linguistic world. It also seems reasonable to assume that there can be many instances of extra-linguistic information which may readily be represented in a finite way.

---

[27] In this context we disregard the phenomenon of type 0 grammars, a class of formal grammars which are assumed to be finite, but for which there exists no known procedure for distinguishing the set of structures generated by a grammar of this kind from structures that cannot be generated by it (cf. Partee et al. 1990: 519–520.) Natural languages are generally seen as falling outside of this class, as a language user is normally able to decide whether a given expression belongs to the language or not.

[28] *Intractability* in a technical sense is explained in 3.2.1. Here the word is used in a more general sense. According to the entry for *intractable* in the *Longman Dictionary of Contemporary English* (3rd ed.), "an intractable problem is very difficult to deal with or find an answer to."

As we distinguish the linguistic from the extra-linguistic sources of information present in the discourse domain of a given source expression, the source text is considered on the level of *parole*. Within the *linguistic information sources* for translation there is, firstly, the information supporting the translator's knowledge of source and target language systems and their interrelations. Secondly, these sources include the information that is linguistically encoded in the source expression. This covers information about the situation type described by the source text, information about the linguistic structure of the source expression, as well as information about relations of reference holding between expressions in the source text and extra-linguistic entities. The latter is derivable from the source language expression as it is interpreted in a specific context. Thirdly, the linguistic sources also include information available in the linguistic context of the source string.

The *extra-linguistic information sources* for translation comprise general background information about the world, information about particular technical domains, information about textual norms, and information derivable from previous translation training and practice. They also cover information about the utterance situation of the source text, and about the translation situation. These types may include elements such as information about the sender, about the purpose(s) of original and translation, about temporal and geographical location, etc. Another extra-linguistic information source may be information derived by applying different kinds of background information in common-sense reasoning about facts described by the SL text. It may appear surprising that information about textual norms is regarded as extra-linguistic; we will argue below that this is a consequence of the way in which we distinguish between linguistic and extra-linguistic information.

The fact that we have listed different types of information sources classified respectively as linguistic and extra-linguistic does not imply that it is always clear where to draw the line between them. However, there are certain kinds of information that we regard as purely linguistic. Traditionally, a language system is seen as a structure divided into four levels: phonology, morphology, syntax, and semantics. At each level the language system specifies an inventory of units, or building blocks, together with a set of rules for how these units may be combined. In addition, the

language system includes a lexicon, which is an open set of lexical units, and each such unit contains information from all the four different levels of the language system. Descriptions of the phonological, morphological, and syntactic structures of a language appear as plausible examples of purely linguistic information.

With respect to semantic phenomena, on the other hand, extra-linguistic pieces of information are not always easily distinguished from the linguistic, and it seems difficult to find a principled way of doing so. Considering a lexical unit, such as *apple*, it seems reasonable that information about its meaning falls within the domain of linguistic information. But how are the meaning properties of *apple* identified? Knowing the meaning of *apple* implies knowing that apples are a kind of fruit, normally round, which is good to eat, and it may also include knowledge of what different colours apples may have, how they taste, etc. All these pieces of knowledge are supported by information available from the extra-linguistic world, but it is not necessary to have all this information about apples in order to understand the meaning of the word *apple*. In our opinion neither the language system itself, nor the extra-linguistic world, can offer a definitive principle for sorting the meaning properties of a lexeme from extra-linguistic properties associated with its denotata; there is no *a priori* basis for a sorting of that kind.

But the fact that a conceptual distinction cannot be drawn in a unique way *a priori* does not imply that it is meaningless.[29] We have argued that the linguistic domain is limited, and that this determines the limit of predictability in the translational relation. In our study of translational correspondences the division between the linguistic and the extra-linguistic is often a question of distinguishing between semantic information derived from the language system and extra-linguistic information sources that also contribute to the interpretation of a given source text. This depends on how

---

[29] Pustejovsky (1995: 232–233) arrives at a similar position in a discussion of how to draw the border between "linguistic or lexical knowledge" and "commonsense knowledge". In his view this is a continuum rather than a dichotomy, but he still finds it fruitful to maintain the distinction because there are "clear cases of paradigmatic linguistic behaviour that are better treated as language specific knowledge, rather than in terms of general inferencing mechanisms."

the given language system is delimited, and thus we relate the distinction to the way in which language systems are conceptually individuated.[30]

Since there is no objective answer to where the limit is drawn, there is an element of choice here. The choice will be influenced by the purpose for which the language description is meant to be applied, and by empirical facts about language use. Also, there are restrictions on what may be conceived of as a language system. As it is unmotivated to include large amounts of world information in the semantic component (cf. the discussion of *apple*), there is an upper bound on this, and a lower bound follows from the fact that there must be a reasonable amount of language users sharing a certain inventory of signs as the means of communication within their community. Given these constraints, a certain textual token may be seen as an instance either of general language or of a certain sublanguage, possibly depending on the purpose of the analysis.

Hence, the distinction between linguistic and extra-linguistic information must be recognised as relative to certain chosen presuppositions concerning the descriptions of the language systems involved. A translation example from a text dealing with a restricted domain may illustrate this relativism.[31] In the *Agreement on the European Economic Area* (*AEEA*) the English expression *competent authority* corresponds translationally with the Norwegian expression *vedkommende myndighet*. An example of the correspondence is shown in (3):

(3a) *The competent authority* shall take the necessary decisions within the
      framework of its internal rules.   (*AEEA*)
(3b) *Vedkommende myndighet* skal treffe de nødvendige beslutninger innen
      rammen av sine interne regler.

When analysing the correspondence with respect to translational complexity, we treat the expressions as system units, or signs (cf. 4.3.6.2). The target sentence (3b) is not glossed, since we regard it as semantically equivalent with the source sentence (3a),

---

[30] This is in accord with Dyvik (2003: 9), who points out that the distinction "between instances of literal and instances of non-literal translation … must be drawn relative to the delimitation of the languages (general languages, sublanguages etc.) in which we assume that the texts are composed."

[31] The relativism is also discussed in chapter 6; cf. the analyses of (13) in 6.3.1.2, and of (20) in 6.3.1.3.

except for the pair of NPs in italics. *The competent authority* is translated as *vedkommende myndighet* ('the authority concerned'), and these two expressions deviate with respect to denotation: in the English text the property of having competence is attributed to *authority*, whereas in the Norwegian text the property of being concerned is attributed to *myndighet* ('authority').[32] Seen as system units, then, we do not regard *vedkommende myndighet* as a linguistically predictable translation of *competent authority* since certain meaning properties are not shared. However, this NP correspondence is recurrent among the data compiled from the *AEEA* and its Norwegian translation, which raises the question whether it is after all a linguistically predictable correspondence within the domain dealt with in the agreement text. Expressions of general language frequently acquire specialised meanings in texts pertaining to restricted, technical domains. If it is the case that within the domain of the *AEEA*, 'authority concerned' is one of the identifiable meanings of the expression *competent authority* when considered in isolation and independently of context, then the Norwegian expression *vedkommende myndighet* is a literal, linguistically predictable translation. This is an analysis which relies on the assumption that the *AEEA* is written in a domain-specific sublanguage with its own specialised vocabulary, so that the use of certain expressions in that sublanguage will be regulated by other conventions than those governing the use of general English. As those conventions will be shared by a community of sublanguage users, they are part of a language system, and we may assume that the expression *competent authority* is here a term-like lexical unit in English, and hence the italicised NP correspondence in (3) is predictable from linguistic information available prior to the translation task.

However, at least one instance of *competent authority* in the *AEEA* is not translated as *vedkommende myndighet*. In Article 58 *the competent authorities* is translated as *de kompetente organer* ('the competent institutions/bodies'):

(4a)   With a view to […] , *the competent authorities* shall cooperate in accordance with the provisions set out in Protocols 23 and 24.   (*AEEA*)

---

[32] This pair of NPs illustrates denotational non-equivalence between translationally corresponding, and co-referential, noun phrases; cf. 6.3.2.3.

(4b)   *De kompetente organer* skal samarbeide i samsvar med bestemmelsene i protokoll 23 og 24 med sikte på […] .

We will not regard (4) as a counterexample indicating that *vedkommende myndighet* is after all a non-predictable translation of the phrase *competent authority* in the domain of the *AEEA* text. Rather, we will analyse *kompetent organ* as a member of the set of linguistically predictable translations of *competent authority*. We regard the italicised NP correspondence in (4) as a case where not only the translational relation between the phrases as units is linguistically predictable, but where also each lexical component within the target expression *de kompetente organer* is a predictable translation of its correspondent in the source expression. The Norwegian lexeme *organ* may not at first glance seem a plausible member of the LPT set of the English lexeme *authority*, but within the given textual domain this is a recurrent lexical correspondence.[33] Thus, with respect to the restricted domain of the Agreement text, both phrases *vedkommende myndighet* and *kompetent organ* are predictable Norwegian translations of the English phrase *competent authority*.

On the other hand, if we analyse (3) and (4) in relation to the domain of general language use, we will conclude that both translations of *competent authority* are cases falling outside the linguistically predictable. This presupposes an analysis where the expression *competent authority* is not treated as a unit of the language system, and where we assume that when it is translated into Norwegian, the choice of target expression is determined by information about the world. In this particular case such information may be derived through the following inference based on world knowledge: an authority concerned with making certain decisions is required to have the necessary competence for that task, and will hence be the competent authority.

Bhatia (1997) presents a genre-based approach to legal translation which may support the choice of ascribing information about these lexical correspondences to the extra-linguistic domain. A technical field, or specialist discipline, of which the law is an example, is associated with what Bhatia (1997) refers to as a "disciplinary culture". Members of a specialist discipline communicate by using specialist genres,

---

[33] In the *AEEA* and its Norwegian translation this lexical correspondence is elsewhere found in the recurrent pair of compound nouns *surveillance authority – overvåkningsorgan*.

and these genres are shaped by conventions determined within the disciplinary culture, described as *generic conventions* by Bhatia (1997). He explains the necessity of learning these conventions for anyone who wants to produce, or translate, texts in these genres (1997: 206–208). With respect to the legal discipline, these conventions are described as "expectations about the way in which language operates in legal contexts, but such expectations are never explicitly stated anywhere but in legal culture" (Bhatia 1997: 208). Information about such conventions of the legal culture is derived from technical knowledge, and although it pertains to the linguistic form of law texts, it belongs to the domain of extra-linguistic information.[34] In our view, the information that *competent authority* corresponds translationally with *vedkommende myndighet*, as well as with *kompetent organ*, is an example of a convention specific to the genre in which the different language versions of the Agreement text are written.[35]

Bhatia's notion of generic conventions may clarify what we understand by information about textual norms, identified above as a subtype within the extra-linguistic information sources for translation. Textual norms, or conventions, control or influence *parole*-related factors such as lexical choices, style, and textual structure. We regard information about such norms as extra-linguistic since information about the characteristic features of specific genres, or text types, is not part of a language system: textual norms are distinct from the conventions that constitute a language system and are shared by the members of the language community.[36] But as this information type deals with linguistic usage, we want to keep it apart from world information, whether general or technical. The distinction is motivated since information about textual norms may account for other aspects of a linguistic expression than those determined by information about facts of the world. In general terms, this is a distinction between information about extra-linguistic states of affairs, and information about norms controlling the use of language describing those states of affairs. With respect to law text, the division is clear: the former kind of information

---

[34] Cf. the discussion of norms in law texts in 5.4.2.1.

[35] This point is also illustrated by example (20) in 6.3.1.3.

[36] The notions of 'genre' and 'text type' are discussed in 4.2.1.2, where we explain why we prefer to speak of *text type*.

is derived from the legal domain, whereas the latter type is derived from the domain of law writing. The distinction may apply also in non-technical settings, as there are numerous contexts, written as well as spoken, where ordinary language users follow shared conventions governing their linguistic behaviour (e.g., dinner conversation, the writing of personal letters, etc.).

Thus, the notion of information about norms controlling language use is a very wide category, which may be refined by identifying types of norms included in this kind of information source. One possible subdivision is between norms applying to texts of general language and those that control specialised, technical language.[37] Norms of the first kind will be shared knowledge among general language users, whereas the second kind will be known by specialists within technical fields. Another distinction may be drawn between norms that influence the characteristics of various text types, and norms that govern the translation of the same types. The latter kind of norms is acquired by translators through translation instruction and practice. We assume that they largely correspond with the concept of norms in translation (cf. 2.2.3), although that notion may include more than textual norms.[38] As regards text-type specific norms, these may be different in, respectively, SL and TL, since text type characteristics are not always identical across languages.[39] Hence, the source text author is subject to the norms applying to the given text type in the source language, and the translator likewise to the corresponding textual norms of the target language.

It may seem arbitrary to relate the distinction between the linguistic and the extra-linguistic to the delimitation of language systems when the latter issue is, as we have seen, to some extent a matter of choice. In particular, as the distinction plays an important part in our analysis of translational complexity, it may seem as if the outcome of that investigation is determined by the way in which we choose to delimit

---

[37] Cf. the definition of *language for special purposes* (LSP) in 5.4.2.3.

[38] This indicates a certain degree of overlap between information about norms governing translation, and information derivable from previous translation training and practice. The latter type is identified above as a separate subtype within the extra-linguistic information sources. We return to this point in 2.4.2.2. Toury's (1995) translation norms have previously been mentioned in 1.4.1.1 and 2.2.3.

[39] This is e.g. shown by Nordrum's (2007) study of how English nominalisations are translated into Norwegian and Swedish in texts of popular science. Her results indicate that the norms of this text type are language-specific, since one fifth of the analysed English nominalisations were found to corresponds with finite con-structions in the Norwegian and Swedish texts. The language-specificity of textual norms is also illustrated in the discussions of examples (17) and (20) in 6.3.1.3.

the languages represented in our empirical data. But arbitrariness may be avoided. Firstly, it is a prerequisite in our analysis to be consistent with respect to the chosen presuppositions concerning the description of the languages. Secondly, arbitrariness can be avoided if the conceptual individuation of language systems is based on empirical facts about language use. Such facts are available through text corpora, dictionaries, and linguistically competent informants, and enable us to conceive of what information it is *reasonable* to include in a language system, given the purpose of its description. In particular, when working with empirical data we find that it is frequently quite possible to determine whether extra-linguistic information has contributed to an interpretation, and subsequent choice of translation, or not. To illustrate this, we may again consider an example discussed in 1.3.1, repeated in (5):

(5a)   Her kunne de snakke sammen uten å bli ropt inn for å gå i melke-butikken eller til bakeren.   (BV)
'Here could they talk together without to be called in for to go in milk-shop.DEF or to baker.DEF'

(5b)   They could talk here without being called in to go and buy milk or bread.

The example has previously been used to illustrate semantic divergence in a translational correspondence: the expressions *for å gå i melkebutikken eller til baker-en* ('to go to the shop selling milk or to the bakery') and *to go and buy milk or bread* do not denote the same activities, but both activities may have the same result, the purchase of milk or bread.[40] Otherwise, we consider string pair (5) to be a linguistically predictable correspondence. In the case of the Norwegian sequence *for å gå*, the English sequence *to go* is a linguistically predictable translation, and the pair of substrings *for å gå – to go* is a correspondence between system units, derivable from information about the lexicons and grammars of SL and TL and about their interrelations. Then, the Norwegian NPs *melkebutikken* and *bakeren* have no direct translational matches in the English sentence. Suggested LPT sets (cf. 2.3.2) in English of the Norwegian nouns *melkebutikk* and *baker* are given in (6):

---

[40] Cf. the discussion of example (5) in 1.3.1.

(6) *melkebutikk*:{*dairy, dairy shop, milk shop*}
   baker: {*baker, baker's, baker's shop, bakery, bakery shop, bakehouse, bakeshop*}

Thus, one literal translation of the Norwegian expression *for å gå i melkebutikken eller til bakeren* could be *to go to the milk shop or to the baker's*, but the translator has chosen the non-literal translation *to go and buy milk or bread*. We assume that through general world knowledge the translator will have been aware that the story from which (5a) is extracted takes place in a time when milk and bread were normally sold through specialised shops in Norway, while, at least in a certain part of the English-speaking world, milk would typically be delivered at people's homes. Thus, background information provides the motivation for disregarding the linguistically predictable *go to the milk shop* as an optimal translation of *gå i melkebutikken*. Then, applying common-sense reasoning to the described facts of the world makes it seem obvious that the purpose of going to the places described in (5a) would be to buy milk and baker's products, and this is the information that gets the focus in the chosen English translation: *to go and buy milk or bread*.

Example (5) thus illustrates the distinction between meaning and context-induced interpretation. The pre-structured domain of linguistic information available prior to translation contains information about the meaning properties of the words in the source text, and is thus the basis for identifying predictable translations. But the pre-structured domain is only a subset of the discourse domain of a source text, and, as (5) shows, extra-linguistic information present in the source text context may induce an interpretation which disfavours the use of a linguistically predictable translation.

In the discussion of examples (3)–(5) we have several times referred to reasoning, or inferencing, about extra-linguistic pieces of information. Such matters fall within the field of pragmatics, which concerns the relationship between linguistic expressions and the situations in which they occur, and studies how discourse-related factors influence the interpretation of linguistic expressions.[41] Pragmatic phenomena are of interest to our investigation of translational correspondences as translation

---

[41] Huang (2007: 2) defines pragmatics as "the systematic study of meaning by virtue of, or dependent on, the use of language." Leech (2008: 88) defines it as "the study of meaning in speech situations."

applies to situated texts, and is typically done to serve a communicative purpose. It is not an aim to make pragmatic factors in translation the centrepiece of our study, but to consider certain relevant phenomena. In particular, we are interested in how the information available to discourse participants influence the production and interpretations of situated expressions, since a text or an utterance is the product of information processing performed by the sender, and its interpretation is the result of information processing on the part of the recipient.[42] To interpret a source expression prior to translation involves finding its propositional content, and identifying its illocutionary force, or type of speech act performed. The notion of 'proposition' is normally associated with sentences; it designates "what a sentence says about the world" (Allwood et al. 1977: 20).[43] A speech act is "the type of action the speaker intends to accomplish in the course of producing an utterance" (Huang, 2007: 102). Type of speech act, or illocutionary force, is commonly attributed also to written statements.[44] In the task of interpreting a situated expression, pragmatic factors contribute to finding the propositional content as well as to identifying the speech act, and an important part of our analysis of translational correspondences involves comparing the respective interpretations of source and target text (cf. 4.3.6.2).

How are pragmatic factors then related to the division between linguistic and extra-linguistic information sources for translation, or to what extent is pragmatic information part of the language system? This pertains to how far the interaction between discourse participants is expressed through linguistic conventions shared by the members of a language community. For instance, in English it is a convention that both the imperative and the interrogative may be used to express the speech act of requesting something, as illustrated by (7) and (8), respectively:

---

[42] Cf. the discussion in 2.4.1.2 of how available background information may determine the informational content of a specific signal transmitted to a recipient.

[43] Löbner (2002: 23–24) defines the proposition of a sentence as its "descriptive meaning", i.e. the set of situations it may refer to, but this does not capture the difference between sentence and utterance. The notion of 'proposition' is also commented on in 6.3.2.

[44] According to Huang (2007: 106), the most infuential approach to the classification of speech acts is the "neo-Austinian typology of speech acts", based on Searle (1975). In this taxonomy, there are five main categories of speech acts: assertives, directives, commissives, expressives, and declaratives, and each main category is further divided into subtypes. E.g., typical examples of directives are advice, orders, questions, and requests; cf. Huang (2007: 106–108).

(7) Please close the door!

(8) Would you close the door?

Given an appropriate context, such as the situation where some people are having a conversation in a room where a door has been left open to a noisy corridor, a similar request could be made by uttering the indicative sentence in (9):

(9) Excuse me, I find that noise on the corridor a bit disturbing.

If someone utters (9) in that context, an addressee would most likely infer that the speaker wants some action to reduce the disturbance, such as closing the door, and the speaker's intention would probably be exactly to achieve that. The relevant difference between, on the one hand, examples (7) and (8), and, on the other hand, (9), is that the piece of information through which a speech act is performed, is available in the linguistic expressions in (7) and (8), whereas in (9) it is not linguistically encoded, but derivable from the extra-linguistic context of the utterance. These examples illustrate that pragmatic information may be linguistically encoded and it may be not, partly depending on the speaker's choice of expression, and partly on the extent to which a language system exhibits conventionalised ways of encoding pragmatic constraints on the use of language in context.

To sum up, linguistic information sources for translation firstly include information about the source and target language systems and about their interrelations, seen in abstraction from the utterance situation of the source text. These sources constitute the pre-structured domain of information which defines the limit of predictability in the translational relation, and which is a subset of the wider domain of discourse in which the source text is located. Further, the linguistic information sources include the information coded in the source text expression, i.e. information about the situation type described by the source text, about the linguistic structure of the source expression, and about reference relations derived by interpreting the source text in context. They also cover information available in the linguistic context of the source expression. The extra-linguistic information sources for translation include general

and technical information about the world, information about textual conventions, information about the utterance situation of the source text, and information derived by reasoning about facts described by the source text.

2.4.2.2 General versus task-specific information sources

In 2.4.2 we have described the division between general and task-specific information sources for translation as a division between information available prior to the translation activity and information associated with a particular piece of source text and the concrete task of translating it into a given target language.

Thus, *general information sources* exist independently of specific translation tasks, and through the distinction between linguistic and extra-linguistic information they can be divided into information about source and target language and their interrelations, and information about the world derivable from the translator's background knowledge. The former corresponds with the pre-structured domain of linguistic information discussed in 2.3.2 and 2.4.2.1. General, extra-linguistic information sources cover information available through the general world knowledge of ordinary language users, as well as information about restricted, technical domains, which is required in the translation of technical texts. They also include information about textual norms, and information derivable from previous translation training and practice.[45]

As mentioned in 1.2, the information needed to produce a specific translation from a given source expression includes the types of information that are accessible through translation competence. Thus, the given, general information sources correspond with a translator's competence. In 2.4.1.5 we have argued for the existence of objective knowledge about translational relations between texts. The fact that translational relations hold between texts of two languages presupposes knowledge of how source and target languages are interrelated. Thus, we abstract away from individual translators and assume that prior to any translation activity,

---

[45] This is only one suggested way of dividing world information into subcategories. For one thing, restricted domains of information need not be technical. E.g. within a group of persons who have a certain "history" together, knowledge about shared experiences will constitute a restricted domain that may serve as a frame of reference influencing the interpretation of utterances made within that group.

there is a certain body of knowledge functioning as a pool of given information. Although we have pointed out similarities between information and objective knowledge, we have argued that these are not the same notions (cf. 2.4.1.4) and would thus avoid viewing the objective knowledge of translators as information. But since objective knowledge has the potential for being made intersubjectively available, it is our opinion that the objective knowledge of translators functions as an information source for translation. By regarding it as something that supplies given information, we assume that it is accessible when required for specific translation tasks, and as translations cannot be produced without a necessary amount of previously acquired knowledge, the objective knowledge of translators must exist prior to a translation activity. This is not to say, of course, that an individual translator possesses a constant body of knowledge which must exist before that translator is able to produce any translations — the knowledge of a translator normally grows through practice.[46]

In 2.4.2.1 we pointed out that there is some degree of overlap between two of the notions identified among the general, extra-linguistic information sources, i.e. information about textual norms, and information derivable from previous translation training and practice. The notions are clearly interconnected as a translator may acquire knowledge about the former through translation practice. Still, we keep the distinction, since textual norms apply to texts of individual languages independently of translation. Moreover, given our product-oriented approach, it is not relevant in the present study, whether information about textual conventions that have contributed to the choice of specific translations is derivable from a translator's general knowledge of text types, or from experience with translation.

Task-specific information sources for translation are available, or derivable, only in connection with specific translation tasks. These, too, may be sorted according to the distinction between linguistic and extra-linguistic information (cf. 2.4.2.1). Task-specific, *linguistic* information sources cover the information coded in the source language expression, as well as information available in its linguistic context. The

---

[46] Cf. Popper's view of knowledge growth, described in 2.2.1.

former includes information about the situation type described by the source text, about the linguistic structure of the source expression, and about relations of reference holding between expressions in the source text and extra-linguistic entities. The latter are derivable when the source text is interpreted relative to a specific utterance situation. The division between the information within the source expression and the information contained in its context reflects the fact that the information encoded in a linguistic expression is normally insufficient to determine the intended interpretation of a given utterance of that expression. Kay et al. (1994: 20) describes this interpretation task as "the resolution problem": in order to determine the intended interpretation it is necessary to merge the linguistically encoded information with information derived from the context, or utterance situation, in which the expression is located.[47] With respect to accessibility, we assume that the information coded in the SL expression is easier to access than contextual information: the former is directly available through general knowledge of the source language, whereas the derivation of the latter requires a greater amount of processing effort.

Task-specific, *extra*-linguistic information is derived from world knowledge possessed by, or given to, the individual who interprets, and translates, the source text. Pieces of task-specific, extra-linguistic information have been mentioned in connection with examples (3), (4), and (5) in 2.4.2.1, in order to illustrate how the task of interpreting a source expression may involve reasoning about the facts described in the source text, or in its context. Such reasoning may thus supply information which is not linguistically encoded in the source text. Furthermore, task-specific, extra-linguistic information includes information related to the utterance situation of the source text, such as information about the sender, about the purpose of the source text, and about its spatial and temporal location. It may also cover information about various aspects of the translation situation itself, such as information about the purpose of the translation, which is not necessarily the same as the purpose of writing the original.

---

[47] This is described by Huang (2007:5) as "linguistic underdeterminacy": "… the linguistically encoded meaning of a sentence radically underdetermines the proposition the speaker expresses when he or she utters that sentence."

A certain understanding of the notion of 'translation task' lies behind the present description of task-specific information sources. A translation task may involve translating anything from a single lexical item, or a sentence, to an entire document, such as a handbook or a novel.[48] As stated above, the very characteristic of task-specific information sources is that they are available only in connection with specific translation tasks, and this sets them apart from the general information sources, which are given prior to the translation activity. However, information about the sender, location, and purpose of the source text pertains to the text on a macrolevel, and it will thus be given prior to a concrete translation activity in the case where the task is to translate a subpart of a larger document for which the mentioned information types are known to the translator. Still, we do not find it appropriate to regard these types as general information, as they are associated with specific texts, and are not derivable from translation competence as such.

Perhaps the most important difference between general and task-specific information sources pertains to accessibility: we assume that information available prior to translation is easier to access than information that must be derived during the translation task. In chapter 3 this topic is developed further in discussions of the efforts required to solve translation tasks.

2.4.2.3 Mono- versus bilingual information sources

The third dimension identified in our typology cuts across only a subset of the other information types. Firstly, with respect to the opposition between linguistic and extra-linguistic information, it does not make sense to classify information about the extra-linguistic world as either mono- or bilingual.[49] Secondly, the distinction between general and task-specific information is relevant in the case of monolingual information sources, but not in the case of the bilingual, which we will comment on below.

Monolingual information sources for translation may be divided into those that are given prior to the translation task, and those associated with the translation of a

---

[48] The notion 'translation task' is further discussed in 3.2.4 and 3.3.1.1.

[49] On the other hand, translation competence models may include components described as "intercultural", or "bicultural" (cf. Hurtado Albir and Alves 2009: 65, 66), but the present typology applies to information, not to knowledge modules.

specific piece of text. Thus, general, monolingual information sources include information about source and target language systems, respectively; they are located, as discussed in 2.4.2.1, in the pre-structured domain of linguistic information. Task-specific, monolingual information sources, on the other hand, fall outside the pre-structured domain; as explained in 2.4.2.2, they cover the information coded in the source language expression, as well as information available in its linguistic context.

We assign only one type of information to the category of *bilingual* information, i.e. information about how source and target language are interrelated with respect to grammars and lexicons. It is our view that bilingual information for translation is located on the level of *langue* — it covers relations between linguistic signs — and this is a consequence of our delimitation of the finite, pre-structured domain of linguistic information (cf. 2.3.2). Since this is determined by the delimitation of language systems, and the distinction between mono- and bilingual information applies only to the linguistic domain, then bilingual information is limited to the correspondence relations between source and target language systems. Thus, we assume that there are no task-specific, bilingual information sources for translation, only general, bilingual information, which, together with general, monolingual information, constitute the pre-structured domain of linguistic information.

At one point we need to make an exception from our principle that the distinction between mono- and bilingual information does not apply to the extra-linguistic domain. As regards textual norms, we explained in 2.4.2.1 that they are not part of language systems, and hence information about textual norms are, in our approach, classified as extra-linguistic. However, since this is information about language use, and since the realisations of textual norms are language-specific, it makes sense to treat information about the textual norms of, respectively, source and target language as monolingual information, and information about how corresponding norms of the two languages differ, can be seen as bilingual.

2.5 Summary

As the present project investigates relations between translationally corresponding texts, a product-oriented approach is necessary. In this chapter, 2.2 with subsections

is a discussion of Karl R. Popper's distinction between the products of behaviour and production behaviour, and its relevance to the study of translation. Following Popper (1979), we have argued that with respect to translation, the study of its products is primary to the study of the translation process, in particular because it is the product and its relation to the original text that gives the process its identity.

The main objective of the present project is to investigate to what extent it is possible to automatise translation in selected English-Norwegian parallel texts instantiating two specific text types. In 2.3 with subsections we have, in accord with Dyvik (1998, 1999, 2005), described a principled limit on predictability in the translational relation. The notion of 'translational relation' covers correspondence relations between language systems as well as between texts and utterances of different languages. We assume that the linguistically predictable part of the translational relation exists on the level of correspondences between *langue* units, and that it is computable from pre-structured information about the source and target language systems and their interrelations. Then, with reference to specific original texts and their translations, the computability issue is a question of to what extent the translational correspondences contained in that body of parallel texts fall within the set of linguistically predictable correspondences between the given source and target language. In line with Dyvik (1999), we have defined 'literal translation' to be the same as 'linguistically predictable translation'.

For the purpose of developing a typology of information sources for translation, we have discussed certain basic concepts in 2.4.1 with subsections. 'Information', in the sense of communication theory, is a purely quantitative notion, something that is either present or not, and it exists independently of interpretive processes (Dretske 1981). 'Informational content', or the message carried by a specific signal, is of a different kind: it is determined by the existence of the information to be transmitted, and by the amount of information carried by the signal; it demands that the information transmitted is identical to the information generated at the source, and it is influenced by background information available to the recipient of the signal (Dretske 1981). Further, 'knowledge' is described by Dretske (1981) as information-supported belief, an account which makes knowledge a property of individual minds.

In our investigation of translation, we have rather put emphasis on Popper's notion of 'objective knowledge' (cf. 2.2.1). Objective knowledge exists in the form of shared content of different knowledge states in different human minds, and hence it may in principle exist independently of individual knowing subjects. Thus, in 2.4.1.5 we have argued that since different bilingual informants may share judgments concerning the appropriateness of specific translations of given source texts, we assume the existence of objective knowledge which can be shared by different translators and which can be exploited by various translation strategies. Moreover, we regard the objective knowledge of translators as a pool of information that is available prior to translation.

Our typology of information sources for translation is presented in 2.4.2 with subsections. The information sources are sorted along three different dimensions, each containing a binary division. Firstly, we distinguish between linguistic and extra-linguistic information; secondly, between general and task-specific information, and, thirdly, within the linguistic domain, between mono- and bilingual information. Figure 2.1 presents an overview of the information typology.

The most important distinction in the typology is that between linguistic and extra-linguistic information as it is associated with the limit of computability in the translational relation. The pre-structured domain of information about the source and target language systems and their interrelations, which defines the linguistically pre-dictable part of the translational relation, is a subpart of the linguistic information sources for translation. In 2.4.2.1 we have tied the limit of the linguistically predict-able to the delimitation, or individuation, of language systems, and we have further argued that the conceptual individuation of a language system relies on empirical facts about language use, and the delimitation of the relevant language community, together with certain chosen presuppositions regarding the purpose for which the description of the language system is meant to be used.

	linguistic		extra-linguistic
task-specific	• information coded in the source text, i.e. information about the described situation type, about the linguistic structure of the source string, and about reference relations holding between expressions in the source text and extra-linguistic entities • information available in the linguistic context of the source text		• information derived by reasoning about facts described by the source text, or in its context • information about the utterance situation of the source text • information about the translation situation
general	**monolingual** • information about the SL • information about the TL	**bilingual** • information about interrelations between source and target language systems	• general background information about the world • domain-specific technical information • information about textual norms • information derivable from translation training and practice

Figure 2.1. A summary of the typology of information sources for translation. The shadowed boxes indicate what is included in the pre-structured domain of linguistic information.