# English, German and the "Semi-Professional" International Translator: A Morphological Approach to Implied Connotative Features

**Christina Alexandris**
*National University of Athens*

## Abstract

The present approach targets the facilitation of translation and the correct transfer of opinions, style and overall spirit of written and spoken online journalistic texts. Here, we present the integration of an annotation strategy for written and spoken journalistic texts, detecting elements with explicit and implicit connotative features. The proposed annotation strategy is morphologically based, functioning as a checklist and targeting to address re-occurring problems encountered mainly by "semi-professional" translators, namely journalists, economists and other professionals working with multilingual written and transcribed journalistic texts available from the media and the Internet. Most of these professionals, usually having an above-average fluency of one or more foreign languages, often lack the necessary exposure to the culture(s) related to the concerned foreign language(s), especially due to distance or frequent changes of location. Thus, essential information is presented either in a subtle form or in an indirect way, constituting emotionally and socio-culturally "marked" elements, often going undetected. The present user-oriented design is aimed to be

integrated into an annotation tool targeting the indication of the largest possible percentage of the points in the texts signalizing "marked" information, alerting the user-translator to evaluate these expressions and, in the case of transcribed spoken journalistic texts, to allow a comparison of "marked" elements with prosodic and paralinguistic features in the respective multimedia files.

Keywords: "semi-professional" translators, explicit and implicit connotative features, emotionally and socio-culturally "marked" elements, morphological approach, annotation, journalistic texts

# 1. Introduction

For translators, terminology and idiomatic expressions, although infrequently occurring in most written and spoken texts, are easily retrievable in conventional and on-line dictionaries. The present approach intends to address the most typical and frequently occurring problems in the translation process in regard to connotative features related to speaker and culture-dependent elements constituting emotionally and socio-culturally "marked" elements. Although it is not always expected for the translator to render exactly the same style and "spirit" in the target language, these connotative features can be transferred or presented as additional information.

Characteristic examples of elements are assumed information (Alexandris 2003), expressions and the special usage of linguistic and paralinguistic elements. These features are essential for the detection and definition of speakers' or journalists' viewpoints, the perception of irony and the discovery of additional background information. The absence of such information hinders the full understanding of the content of the texts and in some cases may even create a distance between the international public and the news

related to the foreign language(s) concerned. The present approach concerns the integration of an annotation strategy in transcription tools for written and transcribed spoken journalistic texts. The present approach targets to facilitate the decision-making process regarding the translation as well as and the correct transfer of opinions and style in written and spoken online journalistic texts. Here, we present the integration of an annotation strategy for written and spoken journalistic texts detecting the highest possible percentage of elements with explicit and implicit connotative features within a context-free framework, namely, without the dependence on other texts or on other sources of information. The detection of elements with connotative features constituting emotionally and socio-culturally "marked" elements within a context-free framework is enabled with the use of the morphologically-based approach presented in the proposed strategy.

The proposed annotation strategy may function as a checklist or editor highlighting emotionally and socio-culturally "marked" elements in multilingual written and transcribed journalistic texts available from the media and the web.

# 2. "Semi-Professional" Translators and the International Public

## 2.1. User-translator Profile and Interests

The User-translator profile and interests are presented here as Requirements 1-5. These Requirements function as a basis for the design of the proposed annotation strategy presented in 6.2.

Most "semi-professional" translators such as journalists, economists and other professionals working with multilingual written and transcribed journalistic texts usually have a fairly good fluency of a foreign language (such as English or German) or more

than one foreign languages, but are either non-native speakers or lack the necessary exposure to the culture(s) related to the foreign language(s) concerned, especially if the place of work/ correspondence changes frequently (Requirement 1). Thus, essential information presented either in a subtle form or in an indirect way ("explicitation" De Silva 2006), such as text-in-context relationships and socio-textual practices is often undetected. This type of information, namely speaker and culture-dependent elements beyond the "Who/What-When-Where-(How)" framework or emotionally and socio-culturally "marked" elements, usually occurs in journalistic texts concerning foreign policy and financial/business news. These types of journalistic texts always were of interest to the international public. However, it is predominately in domestic politics where most such information occurs and the domestic politics of some countries is of almost equal interest to the international public as are foreign politics and financial/business news. The international public is generally interested to be able to have an insight in respect to a specific issue or event from an insider's or an outsider's point of view or from an international perspective. These points of view are not always explicitly expressed and are rather conveyed in an indirect manner (Requirement 2). Titles often do give a general idea of the degree of neutrality in a journalistic text, however this does not apply to all cases. To perceive implicit information and to understand the speaker's or author's spirit, the reader or viewer has to read between the lines (Requirement 3).

Furthermore, the international public often has to make choices with respect to what types of online spoken/written journalistic texts to use for professional purposes and what types of online spoken/written journalistic texts to discard (Requirement 4). In some cases, the user-translator seeks texts providing emotionally and socio-culturally "marked" elements beyond the "Who/What-When-Where-(How)" framework, and, in other cases, the user-translator seeks texts containing neutral, "unmarked" information. This

decision-making process may be time-consuming (Requirement 5), since the user-translator is required to read the text many times. Detailed reading is required to check whether the text contains speaker and culture-dependent emotionally and socio-culturally "marked" elements and implied points of view.

## 2.2. Problems Encountered

Typical indicators of the speaker's or author's intention and emotionally and socio-culturally "marked" elements are the use of fixed expressions and idioms in written and in spoken transcribed texts and the use of paralinguistic elements such as prosody and gestures in spoken texts and multimedia files. Text type can sometimes be a useful indicator of the presence or absence of "marked" elements beyond the "Who/What-When-Where-(How)" framework. However, all the above mentioned typical indicators are not reliable and consistent enough to apply in the majority of online spoken/written journalistic texts to be used by international professionals.

Journalistic text types can be differentiated according to the type of speech act performed in them. The basic speech acts in journalistic texts can be divided into four types, namely "Announcement", "Description", "Declaration" and "Expression" (Alexandris & Fotinea 2004). The "Announcement" speech act involves the announcement of news. This speech act is performed by journalists (anchormen and correspondents) only and constitutes the type of speech act performed in most online spoken/written journalistic texts constituting headline news. The "Description" speech act (performed by all speakers except politicians) involves the description of a situation. In most online spoken/written journalistic texts, this speech act is usually performed by correspondents. The "Declaration" speech act is related to the declaration of one's opinion, position or the announcement of facts.

The "Declaration" speech act is performed by politicians and by non-politicians, namely professionals such as government officials, police chiefs, doctors and experts in a scientific field. The "Declaration" speech act is also performed by journalists explicitly stating their opinions. The "Expression" speech act involves the expression of one's feelings and is usually limited to speakers that are neither journalists nor politicians. However, speakers or journalists may implicitly express their opinions or even their feelings in contexts appearing to be "Announcement" or "Description" speech acts. It is the information contained in the latter types of texts, where essential information, presented either in a subtle form or in an indirect way, is often undetected, especially by the international public.

Fixed expressions and idioms do not occur often enough in texts to signalize a point of view or attitude conveyed. Furthermore, fixed expressions and idioms are usually easily detected by a non-native speaker with a fair command of the language concerned. They are also easily accessed at specialized dictionaries for idiomatic expressions and other printed or online resources.

Paralinguistic elements such as prosody and gestures in written and transcribed texts and multimedia files may provide sufficient input for the detection and definition of speakers' or journalists' opinions and/or viewpoints and the perception of irony. However, paralinguistic elements may be absent in some multimedia files, for instance in cases where the speaker (for example, a correspondent) is not visible. Furthermore, in languages such as UK English or Standard German, it is observed that speakers do not always demonstrate obvious changes in prosody when making subtle ironic statements, changes not always easily detected by an international audience. Another problem encountered is that paralinguistic elements expressing speakers' or journalists' attitude may be misinterpreted or overlooked by an international audience, unless extreme emotions are expressed.

## 2.3. Error Data

The proposed specifications and related annotation module are language-specific, and its basic parameters are designed based on translational error-types retrieved from empirical data concerning the language pairs English, German and Greek. Most data involves mistranslations of spoken and written journalistic texts (data provided from graduate courses for professional translators), as well as failure to correctly interpret paralinguistic elements in online multimedia journalistic files by transcribers and professional journalists. The data here was collected from graduate courses for journalists (M.A in Quality Journalism and Digital Technologies, Danube University at Krems, Athena- Research and Innovation Center in Information, Communication and Knowledge Technologies, Athens, Institution of Promotion of Journalism Ath.Vas. Botsi, Athens) and from data transcribed from spoken journalistic texts in European Union projects, especially the CIMWOS Project (www.xanthi.ilsp.gr/cimwos/).

# 3. Connotative Features and Superfluous Information

## 3.1. Relation of Framework of Information Content and the Flouting of Grice's Cooperativity Principle

The overall framework of the information content of the journalistic texts may be described as the "Who/What-When-Where-(How)" framework, with the feature "How" constituting an optional component, depending on the type of content of each specific journalistic text. The presence of superfluous elements beyond the absolutely necessary "Who/What-When-Where-(How)" framework expressed by the "necessary" elements in a journalistic text are

indicative of the text's non-neutral character. Typical examples are texts with a large percentage of modifiers such as adjectives and adverbials, as well as characteristic types of expressions. This differentiation between "superfluous" and "necessary" elements is compatible with to the flouting of the maxims of quantity, quality and manner in Grice's Cooperativity Principle (Grice 1989; Hatim 1997). Grice's Cooperativity Principle is stated as follows:

Quantity: Make your contribution as informative as is required.

Quality: Do not say what you believe to be false.
　　　　Do not say that for which you lack adequate evidence.

Manner: Avoid obscurity of expression.
　　　　Avoid ambiguity.
　　　　Be brief (avoid unnecessary prolixity).
　　　　Be orderly.

(Grice 1975)

More specifically, expressions with "superfluous" information flout the Gricean Maxim of Quantity, namely "Make your contribution as informative as is required". Expressions with "superfluous" information mostly coincide with modifier categories such as adjectives and adverbs, as well as verbs whose semantic content involves modifying features related to concepts such as "mode", concepts of positive or negative nature or of emotional and ethical quality.

As a criterion for comparison, the framework of headline news is used, labelled as the "Who/What-When-Where-(How)" framework, supposedly aiming to the transfer of essential information in an objective manner. A "bare" transfer of essential information does not in all cases necessarily ensure objectivity of information transfer. For the evaluation of objectivity, a considerable set of factors should

be taken into account, namely, whether the entirety or only selected aspects of an event are presented, paralinguistic elements such as tone in spoken texts, as well as the types of other media and text forms the journalistic text is related to (for instance, appended pictures and videos or the case of important headline news delivered in the form of a very casual conversation).

In respect to the Gricean Maxim of Quality, "Do not say what you believe to be false" and "Do not say that for which you lack adequate evidence", with the exception of the case in which the maxim is not flouted for the purpose of propaganda, it is observed that information presented in a form that may be characterized as flouting the Gricean Maxim of Quality usually contains superfluous elements flouting the Gricean Maxim of Quantity. This relationship applies in a similar way to information flouting the Gricean Maxim of Manner. Specifically, in regard to the flouting of the Gricean Maxim of Manner, namely "Avoid obscurity of expression", "Avoid ambiguity" and "Be orderly", superfluous information (flouting the Gricean Maxim of Quantity) is often encountered in written and spoken texts containing socio-culturally and emotionally "marked" elements, unless propaganda or similar communicative targets are involved. The Gricean Maxim of Manner also includes the part "Be brief" (avoid unnecessary prolixity) which partially coincides with the Gricean Maxim of Quantity and is thus directly related to the avoidance of superfluous information.

In general, the overall context of the written and spoken online journalistic text may be described as containing a subset of word-types, coinciding with superfluous information in the text and indicating emotionally and socio-culturally "marked" elements, expressing style and overall spirit of the author, speaker or of the intended readership or audience.

In the present approach we address emotionally and socio-culturally "marked" elements namely evident connotative features in texts, related to typical grammatical categories such as adjectives,

adverbs, in addition to specific categories of nouns and expressions, as well as implied connotative features that are related to other types of grammatical categories and are either included in the lexical semantic features of these word-types or are related to their grammatical features. The implied connotative features are not fully sublanguage-independent. Furthermore, these implied connotative features typically occur in journalistic texts.

The evident and implied connotative features serve as a set of specifications that may be used as formal template and checklist, to facilitate the decision-making process regarding the translation, the detailed processing and the correct transfer of opinions, style and overall spirit of written and spoken online journalistic texts. These specifications serve as a strategy to capture the largest possible percentage of emotionally and socio-culturally "marked" elements within a context-free framework.

## 3.2. Evident Versus Implied Connotative Features

Word-Groups with Evident or Implied Connotative Features in their semantic content, that may be labeled as emotionally and socio-culturally "marked" elements, may be divided into two main categories. The first category involves specific word groups and expressions whose semantic content may contain connotative elements or information related to associations, such as text-in-context relationships and socio-textual practices. A typical characteristic of such word groups and expressions is that their semantics is defined by a relatively large number of semantic features and respective selectional restrictions within the framework of an ontology. A larger set of semantic features results to the location of such expressions in lower levels of the hierarchical structure of ontologies (Gayral et al. 2000; Wilks & Fass 1992).

Typical examples of Word-Groups with Evident Connotative Features (A-Type Categories) are the more typical grammatical

categories of (1) adjectives and (2) adverbials, containing semantic features related to (i) descriptive features (ii) mode (iii) malignant/ benign action or (iv) emotional / ethical gravity. Typical examples retrieved from English journalistic texts are the expressions "enormous", "gloomy", "sensibly", "absurd", "spiteful", "scornfully" and "disgraceful". From a morphological aspect, Word-Groups with Evident Connotative Features are detected by their word stems.

# 4. Word Groups with Implied Connotative Features

## 4.1. Outline

Word-Groups with Implied Connotative Features (B-Type Categories) are the grammatical categories of (1) verbs (or nominializations of verbs) containing semantic features (including implied connotations in language use) related to (i) mode (ii) malignant/benign action or (iii) emotional/ethical gravity, as well as (2) nouns with suffixes producing diminutives, (3) derivational suffixes resulting to a (i) verbalization, (ii) an adjectivization or (iii) an additional nominalization process of proper nouns. From a morphological aspect, Word-Groups with Implied Connotative Features are detected by their word stems or by their suffixes, depending on their type.

## 4.2. Verb Stems

In respect to their semantic content, the verb-types of Category B containing semantic features related to (i) mode (ii) malignant/ benign action or (iii) emotional/ethical gravity may be categorized as semantic entities within an established ontological framework of verbs. For instance, in the example of the verb "to-delay" the more

restrictive concepts can be related to various additional features, corresponding to the interpretations "linger", "dally" and "tarry" (dated), according to the combination of features used. The features defining each verb, ranging from the more general to the more restrictive concept, are formalized from standard and formal definitions and examples from dictionaries, a methodology encountered in data mining applications (Kontos et al. 2000). The interpretation of these verbs is accounted for in the form of additional features located at the end-nodes of the ontology, functioning as additional semantic restrictions or "Selectional Restrictions" (Gayral et al. 2000; Wilks & Fass 1992), within an ontology of words. Therefore, the semantics are very restricted at the end-nodes of the ontology.

Selectional Restrictions, already existing in a less formal manner in the taxonomies of the sciences and in the sublanguages of non-literary and especially, scientific texts, are applied within an ontology-search tree which provides a hierarchical structure to account for the relation between the concepts with the more general semantic meaning and the concepts with the more restricted meaning. Subsequently, this mechanism can also account for the relation between expressions with the more general semantic meaning and the expressions with the more restricted meaning. Additionally, the hierarchical structure, characterizing an ontology, can provide a context-free framework for describing the (sublanguage-independent) word category.

An example of this is the English verb "plummet" (mode), which may be described as a form of falling and is thus related to the verb "fall". However it also has additional features signalizing a rather dramatic fall, which might be related to making noise such as a "thump". These features differentiating the verb "plummet" from the verb "fall" constitute Selectional Restrictions within an ontology of verbs, directing the search towards the verb "plummet", a more restrictive meaning of "fall". In turn, the verb "fall", constitutes a

more restrictive meaning of the verb "move", signalizing a downward movement. Concepts with a less restrictive meaning are located near the top of an ontology. Thus, for example, verbs with connotative features such as "plummeting" (mode) in English journalistic texts and "zurueckrudern" ("to row back") in German journalistic texts (defining mode) are related to the concepts "fall" and "go-back", that are located near the top of the ontological hierarchical structure as more general concepts that have the additional property of having relatively non-connotative and neutral characteristics.

In a similar way, verbs with connotative features such as "swear" in English (emotional/ethical gravity) and "fechten" in German ("fechten" defining malignant action) are related to the concepts "say-address" and "say-respond" respectively, which have relatively non-connotative and neutral characteristics and constitute general concepts. For instance, the English verb "swear", may be described as a form of making a declaration and is related to the verb "declare", with the additional features of making a declaration with an ethical or moral commitment, signalizing ethical or emotional gravity. These features constituting Selectional Restrictions, differentiate the concept "swear" form the concept "declare". In turn, the verb "declare" constitutes a more restrictive meaning of the verb "speak", signalizing a statement of formal character downward movement. A similar example related to emotional or ethical gravity is the verb "love" in English related to the relatively neutral and general concept "feel".

It is understood here that the detection and translation of verb types may be problematic in languages where most verbs are primitive verbs such as "go" and "do", and typically contain a comparatively smaller set of semantic features.

## 4.3. Suffixes

Derivational suffixes may be related to connotative features, at least for the language pairs English, German and Greek, however, with considerable variations, since this form of implied connotative features is not equally common in all languages concerned. The most typical categories of derivational suffixes where implied connotative features are likely to occur are diminutives (or in derivational suffixes describing maximalization) and suffixes used from a foreign language, as well as suffixes signalizing gender (Greek) and specific types of suffixes resulting to a nominalization (German).

From the overall category of derivational suffixes, diminutives typically occurring in German and Greek may be perplexing to "semi-professional" translators whose mother tongue makes a limited use of such elements. In particular, in German and in Greek, the boundaries between literal or of connotatively "marked" use of diminutives is not always clear. Furthermore, in Greek, diminutives are also used as a form of positive politeness (Sifianou 2001). Differentiations between literal and connotative features in derivational suffixes are depicted in (1).

(1) Greek:
   (i) -´aki/-´akis: gat-´aki = little cat, kitten (literal)
   (ii) -´aki/-´akis: giatroud´aki = young/umimportant doctor
      (connotative feature)

German:
  (iii) -chen: Kindchen = little child (literal)
  (iv) -chen: Professorchen = (little) professor (pejorative)
      (connotative feature)

A subset of connotatively "marked" derivational suffixes related to suffixes from foreign languages is typically observable in English.

In English, typical examples are the suffixes "-istas" from Spanish, as well as the derivational suffixes of French origin such as "-ette" and "-esque" which, in certain contexts may be considered to be connotatively "marked".

Stems of words of French origin related to German suffixes may be related to connotative features in German, as having a slight nuance of irony in some contexts. However, a more typical example of connotatively "marked" derivational suffixes in German is the derivational suffix "-erei" used for the nominalization of a verb or appended to a noun. In this case, the boundary between is neutral or connotatively "marked" use may not always be clear to a non-native speaker. For example, in the word "Malerei" ("a painting"), the "-erei" suffix nominalizes the verb "malen" ("to-paint"). However, in the case of the word "Schreierei", a nominalization of the verb "schreien" ("to-scream"), connotative features are added.

In the case of Greek, some general rules apply for certain types of connotatively "marked" derivational suffixes, such as derivational suffixes with maximalizing features ("-ar´as," "-´ara") ("alitar´as," = a big punk, "periptos´ara" = a big "basket case" (of a person)), or cases in which a suffix with a female grammatical gender is used for words for which the masculine grammatical gender is used as a default ("-a") ("dimosiogr´afa" (a female journalist, pejorative), as opposed to "dimosiogr´afos" (a male/female journalist) and "ethopoi´a" (a female actor, pejorative), as opposed to "ethopoi´os" (a male/female actor).

# 5. Word-Groups with Implied Connotative Features: Grammatical Aspects

## 5.1. Word-Group Types with Implied Connotative Features

The third category of word-groups with connotative features

involves words whose implied connotative features are interlinked with grammatical aspects and are not related to lexical semantics. These grammatical categories are typical of spoken texts and may be classified as "C-Type" categories. These implied connotative features are detected either on the morphological level on the word level. Implied connotative features on the morphological level concern (1) modal verbs and (2) verb suffixes with or without the language-specific presence of the verb's subject contained in the morphological features of the suffix. Implied connotative features on the word level concern the grammatical categories of particles and adverbials, in particular, subsets of particles and adverbials (3).

Language-specific grammatical categories which in some cases may be considered to be emotionally and socio-culturally "marked" may be regarded as the most problematic Word-Group with connotative features for the non-native speaker and "semi-professional" translator. The basic problem of the Word-Groups with implied connotative features concerning Grammatical Aspects (C-Type categories) is that their connotative use is "hidden" within the grammatical-morphological features and is highly context-dependent, especially in spoken texts. Since the semantics of this Word-Group is highly determined by its usage within the language community, information expressed by accompanying paralinguistic features such as prosody, tone and gestures may be especially helpful to the non-native speaker and "semi-professional" translator.

## 5.2. Modal Verbs

Modal verbs constitute a typical problem for "semi-professional" translators and non-native speakers, since the distinction between their literal semantic meaning and their polite or connotative use is not always visible to the recipient. The semantics of some modal verbs may be weaker or stronger. The multiple uses of modal verbs are typical of languages such as English and German. Modal verbs

may be used literally or may be used to express a suggestion or a prediction, an understatement or even irony.

In the context of specific text types, such as technical texts, the use of the modal verb is related to specific speech acts. For example, the German modal verbs "muss" ("must"), "darf" ("is-allowed"), "kann" ("can") and "soll" ("should"), are related to speech acts expressing "command/direction," "prohibition", "permission," or "possibility/recommendation" respectively (Lehrndorfer, 1996). A similar use in the respective speech acts in non-technical texts is presented in (2).

(2) (Translations are close to source text)

(i) The initiative must not be supported
(ii) Die Initiative muss nicht unterstuetzt werden (German)
(iii) The initiative is not allowed to be supported
(iv) Die Initiative darf nicht unterstuetzt werden (German)
(v) The initiative could be supported
(vi) Die Initiative koennte unterstuetzt werden (German)
(vii) The initiative should be supported
(viii) Die Initiative sollte unterstuetzt werden (German)

However, the boundaries of the possible differentiations in respect to the semantics of modal verbs become less evident in the context of journalistic texts, especially in spoken texts. For example, in English, the verbs "must", "should" and even "shall" are often used in a similar way, although the verb "must" is related to strongest and less complex semantics. Additionally, the semantic interpretation of the verbs "may" and "might" can be identical in some contexts. Thus in (iii), the modal verbs are used literally, to express a suggestion or a prediction or to express an understatement.

(3) (i) In this case, all negotiations must start within the next few hours

(ii) In this case, all negotiations should start within the
    next few hours
(iii) In this case, we may block all negotiations

The correct interpretation of the semantic content of modal verbs
is context-specific, however, pointers signalizing the above-
presented subset of modal verbs such as "should" and "darf" (in
German) containing multiple variations in respect to their semantics
may indicate fragments in spoken and in written texts where
connotative features and implied information are most likely to
occur.

## 5.3. Grammatical Categories with Verbal Features: Personal Versus Impersonal Verbal Features

In Greek, as a verb-framed and pro-drop language (like Spanish
or Italian), the suffix of the finite verb contains the features of the
verb's subject. In the case of transcribed spoken journalistic texts,
the verb's subject is the Speaker or the Receiver.  For example, the
Greek verb "΄kano" ("I-do/make") in the context of "make
questions" (in a literal transfer from Greek) is a primitive verb,
containing the features of the verb's subject and is not restricted to
the task-specific semantics of the English equivalent verb "ask".
Therefore, in the Greek verb "΄kano" ("I-do/make"), in the context
of "to make questions" (Example 4a), the relationship between the
Speaker or the Receiver is overtly expressed, in the sense that the
Speaker is "doing something" in respect to the Receiver, while the
respective English equivalent verb "ask" is identified as an "action-
type" expression. In a similar way, the German equivalent
expression "Fragen stellen" ("pose/put questions"), the verb
"stellen" contains more specific semantic features than the Greek
verb "΄kano" and can be classified as an "action-type" expression.
    In Example 4b, the Greek verb "olokli΄rosate" ("you-have-

completed") is equivalent to the verb "finished" in English, and with the expression "Sie sind" and the adjective "fertig" in German. We note that the semantics of the Greek verb "olokli΄rosate" ("completed") allows it to be classified as an expression where the relationship between the Speaker or the Receiver is overtly expressed, whereas the verb "finished" in English is classified as an "action-type" task-related expression. In a similar example (4c), in the utterance "Please proceed", the Greek verb "sine΄xiste" ("you-proceed") is equivalent to the respective verb "proceed" in English, and with expression "Bitte machen/sprechen Sie" and the adverb "weiter" in German.

Therefore, in contrast to verb-framed and pro-drop languages such as Greek, both in English and in German, the semantic feature of the subject is not included in the semantics of the verb, containing only the semantic features of the verb's content.

The expression of the verb's subject in the verb's suffix may in some contexts stress the importance of the relation of the verb's subject, namely the Speaker, towards the action being performed or the object of the action performed, depending on the verb's semantics. This feature may in some contexts be interpreted as an indicator of directness, politeness, personal involvement or solidarity towards the Recipient.

Differences in respect to the the languages Greek, English and German are presented in (4a), (4b) and (4c).

(4a) (i) you wish to *do* (English)
    (ii) moechten Sie […..] *machen* (German)
    (iii) (verb = "[you] do") = "΄kanete" ("do") (Greek)

(4b) (i) you have *finished* (English)
    (ii) Sie *sind* […..] *fertig*  (German)
    (iii) (verb = "[you] have-completed") = olokli΄rosate
       ("completed") (Greek)

(4c) (i) Please *proceed* (English)
    (ii) Bitte *machen* (=do)/sprechen(=speak) Sie *weiter* (German)
    (iii) (verb = "(you) proceed") = sine´xiste ("proceeed") (Greek)

To interpret the effect of verb-framed and pro-drop languages, the combination of the semantics of the verb stem with the subject feature of the suffix suffices to allude to the complete semantic interpretation of the verb. However, if the translator wishes to render a similar spirit in the target language, other linguistic means must be used. For example, in English and German, particular types of particles, adverbials and exclamations may be used to create a friendly, communicative effect.

## 5.4. Grammatical Categories: Adverbials and Particles:

C-type categories also include a specific set of particles, adverbials and exclamations. These types of words may constitute word-groups with implied connotative features and emotionally and socio-culturally "marked" elements and usually occur in spoken texts, including journalistic texts. Typical examples of such elements are the expressions "merely", "utterly", "surely", "doch" (German), "eben" (German) and temporal expression such as "now" and "gleich" (German). In spoken language, these adverbials and particles may either be used to emphasize the semantic content of the spoken phase or sentence (emphasis) or to allow a more casual or spontaneous effect of the overall spoken utterance (casual).

In the following utterances in (5), we note that in English and in German, the addition of adverbials, particles or exclamations (indicated in italics) are used to add emphasis or additional features such as a casual and communicative effect. In both languages, in some cases the same element may be used in both ways, namely,

both for emphasis or for a more familiar, or friendly effect.

(5) English:
(i) Let's take a look at it (neutral)
(ii) Let's take a look at it *now* (casual, communicative)
(iii) That is *so* "last year" (US)  (emphatic)
(iv) That is *quite* right (literal, or understatement)
(v) Ok, see you *then* (casual, familiar, communicative) (UK)

German:
(i) Reissen Sie *ja* kein Blatt heraus (emphatic)
(ii) Das ist *doch* selbstverstaendlich (emphatic or casual)
(iii) Das ist *eben* so (emphatic and/or cynical)
(iv) Das wollte ich *gerade* sagen (literal "right now" or emphatic)
(v) Das werde ich *gleich* wiederholen (literal "right now" or emphatic)

In spoken texts in languages such as Greek, differentiations are observed in respect to the use of the above-mentioned types of adverbials, since they are not used for conveying a casual, familiar, or communicative effect to the Receiver. In the following example, we note that in English and in German, the addition of the adverbial "now" and "gleich" (German: "(right) now") seems to soften any harsh effect (6a and 6b). In contrast, in the Greek example, friendliness does not appear to be effected by the presence or absence of the adverbial "now" (´tora) (6c).

(6a) English:
(i) I will ask you some more questions
(ii) will ask you some more questions *now* (Adverb)

(6b) German:
(i) Ich werde Ihnen einige weitere Fragen stellen

(ii) Ich werde Ihnen *gleich* (Adverb) einige weitere Fragen
stellen

(6c) Greek:
(i) Tha sas ´kano meri´kes ero´tisis a´koma
a. I will ask ("make") you some more questions
(ii) Tha sas ´kano *´tora* (=now) meri´kes ero´tisis a´koma
(iii) *´Tora* (=now) tha sas ´kano ´tora meri´kes ero´tisis
a´koma
a. I will ask ("make") you some more questions now

In spoken English, discourse markers comprising particles, adverbials and exclamations may be used in various ways, allowing multiple interpretations, according to the context in which they are produced. Typical examples are the expressions "Right" and "Well", used as exclamations and discourse particles. The expression "Right", originally an adjective, may be used literally to express agreement or any other form of confirmation and well as a reluctant or hesitant acceptance and, in some cases, even frustration, expressed from the side of the Speaker. The expression "Well", originally an adverb, is also used as discourse marker for the opening of sentence or the expression of hesitation or a feeling of awkwardness expressed by the Speaker. In some cases, the expression may also be used in a casual, communicative spirit, such as in the sentence "Well, hello to you too". The above-described differentiations are not always easily perceived by non-native speakers and "semi-professional" translators, often resulting to loss of implied information in a spoken journalistic text.

All word-groups with evident or implied connotative features are summarized in Table 1.

Table 1. Word-Groups with Evident or Implied Connotative Features in Journalistic Texts

| Semantic Content | Word-Groups with Evident Connotative Features |
|---|---|
| Word Level | A-Type Categories<br>Word-Groups with Evident Connotative Features |
| (1) Adjectives | adjectives containing semantic features related to<br>  (i) descriptive features<br>  (ii) mode<br>  (iii) malignant/benign action or<br>  (iv) emotional/ethical gravity |
| (2) Adverbials | (3) adverbials, containing semantic features related to<br>   descriptive features<br>  (i) mode<br>  (ii) malignant/benign action or<br>  (iii) emotional/ethical gravity |
| Morphological Level | B-Type Categories<br>Word-Groups with Implied Connotative Features |
| (3) Verb-stems | verbs containing semantic features (including implied connotations in language use) related to<br>  (i) mode<br>  (ii) malignant/benign action or<br>  (iii) emotional/ethical gravity |
| (4) Suffixes: Diminutives | nouns with suffixes producing diminutives |
| (5) Other Derivational Suffixes | derivational suffixes resulting to a<br>  (i) verbalization or<br>  (ii) adjectivization<br>  (iii) an additional nominalization of proper nouns |
| (6) Modal Verbs | (mainly encountered in languages such as English and German) |
| (7) Subject Feature in Verb-suffixes | (pro-drop languages such as Greek and Italian) |
| Word Level | |
| (8) Particles and Adverbials | (mainly encountered in languages such as English and German) |

# 6. Formalization of Evident and Implied Connotative Features

## 6.1. Framework

### 6.1.1. Appending Features to Word-Groups

The evident and implied connotative features of the above-described word categories related to emotionally and socio-culturally "marked" elements may be formalized as special features linked to the respective word category type. Specifically, the feature [expl-conn] may be appended to A-Type Categories, namely word-groups with evident (or explicit) connotative features such as the above-presented subsets of adjectives and adverbs. The feature [expl-conn] is directly matched on the word-level and requires no detection of the word's morphological features.

The feature [impl-conn] may be appended to B-Type Categories, namely word-groups with implied connotative features. The feature is [impl-conn] is matched on a morphological level, specifically, on the verb-stem of verbs containing semantic features related to mode, malignant/benign action or emotional/ethical gravity, and on the suffix of nouns with suffixes producing diminutives. Additionally, the [impl-conn] feature is appended to all derivational suffixes connected with verb stems, with the exception of the types of derivational suffixes producing participles and actor thematic roles, and to all derivational suffixes connected to word stems constituting proper nouns.

In C-Type Categories, namely word-groups with implied connotative features concerning grammatical aspects in spoken (journalistic) text, the feature [gram-conn:subj] is matched on a morphological level, specifically, the suffix of verbs of verb-framed and pro-drop languages. The feature [gram-conn:±emph] is matched on a word-level being directly matched to the adverbial or particle

used in spoken texts in languages such as English or German for an emphatic or casual/spontaneous effect. For modal verbs, the subset of modal verbs containing likely connotative features is signalized by the feature [gram-conn:modl] in transcribed spoken texts. In the case of the presentation of the [gram-conn:subj] and the [gram-conn:modl] annotation, the user-translator is alerted to decide whether in the context in question the annotated feature is related to connotative information.

## 6.1.2. Formalization

The formalization of evident and implied connotative features in written and transcribed spoken journalistic texts provide to the "semi-professional" translator a set of specifications for the in-depth understanding and successful management of information of the source language text and the successful transfer of the highest possible extent of information in the target language. These specifications may be described in Table 2. Their formalization allows their integration in an annotation tool for online written and transcribed spoken journalistic texts.

This formalization may function as a stand-alone framework, but it may also be applied as an annotation tool in combination with a statistical evaluation of texts. Thus, the translator receives information about the overall "marked" context of the journalistic text and is presented with the problematic points, with a high likelihood of connotative elements annotated with the respective tags in each text.

The relation of the formalized evident and implied connotative features in respect to the "Who/What-When-Where-How" basic content of journalistic texts is outlined in Table 2.

Table 2. Who/What-When-Where-How of Journalistic Texts

| | | Text | | |
|---|---|---|---|---|
| Basic Content: | Who/What | When | Where | How |
| Additional Features and Information: | evident/ implied connotative features | evident /implied connotative features | evident/ implied connotative features | evident /implied connotative features |
| Features:<br><br>Word level: (evident or "explicit") | [expl-conn] | [expl-conn] | [expl-conn] | [expl-conn] |
| Morphological Level: (implied) | [impl-conn] [gram-conn:subj] | [impl-conn] [gram-conn:subj] | [impl-conn] [gram-conn:subj] | [impl-conn] [gram-conn:subj] |
| Word level: (implied) | [gram-conn:±emph] [gram-conn:modl] | [gram-conn:±emph] [gram-conn:modl] | [gram-conn:±emph] [gram-conn:modl] | [gram-conn:±emph] [gram-conn:modl] |

## 6.2. Design and Implementation

### 6.2.1. General Description of Specifications

The above-described formalization of evident and implied connotative features in written and spoken journalistic texts allows its implementation in an annotation tool for written and transcribed spoken journalistic texts available online. Since userfriendliness has become a necessary feature in software systems (Constantie & Lockwood 1999; Nielsen 2000), a basic requirement for software development is the knowledge of what the user wants to do with the

system (Wiegers 2005). This knowledge of user requirements is of crucial importance if applications are targeted towards a broader and international public, such as the "semi-professional" translators as a user-group.

According to all the above presented User Requirements (Requirements 1-5) in 2.1 the user-translator has to perform a decision-making process in respect to the translation of the online spoken/written journalistic text. Specifically, the user-translator is required to read the chosen text, detect and evaluate the parts of the text where emotionally and socio-culturally "marked" elements are present.

The aim of the present user-user-oriented design is its integration in a annotation tool targeting  (1) to provide the "semi-professional" translator with indications of the emotionally and socio-culturally "marked" elements, namely, the evident and implied connotative features in the texts and, specifically, the features [expl-conn], [impl-conn], [gram-conn:subj], [gram-conn:±emph] and [gram-conn:modl], (2) to facilitate the user-translator's decision of actually using the text, depending on its level of emotionally and socio-culturally "marked" elements, (3) to allow the user-translator to compare "marked" elements with prosodic and paralinguistic features in the respective multimedia files and (4) to indicate the largest possible percentage of the points in the texts signalizing "marked" information. The interface of the annotation module is designed to annotate and highlight all the points possibly containing "marked" information, alerting the user-translator to evaluate the parts of the text containing these expressions. Incoming text may be transcribed data from written and transcribed journalistic texts, downloaded written texts from the internet (blogs, pdfs etc.) or scanned files from newspaper articles and journals. In respect to transcribed spoken journalistic texts, the module is designed to be able to be compatible and to operate with most commercial transcription tools, some of which are available online.

Specifically, the user-translator activates the proposed module when viewing journalistic texts online. The online journalistic texts are scanned by the proposed module and the user-translator is presented with an output constituting the actual online journalistic text with all the instances of "marked" information highlighted with a different color, format or any other type of indication. The user-translator is presented with the output and related information on the screen

In the written and transcribed spoken journalistic texts, the emotionally and socio-culturally "marked" elements are highlighted in the text. Additionally, the "semi-professional" translator is provided with indications of the evident and implied connotative features in the texts, namely the features [expl-conn], [impl-conn], [gram-conn:subj] and [gram-conn:±emph]. The highlighted and annotated features alert the user-translator to evaluate them and to determine their connotative function within the text, if applicable.

For the achievement of speed and efficiency, the annotation module operates on keyword detection at morpheme-level or word-level and on the interaction with a database constructed on ontological principles, which may be made accessible to the translator-user.

Keyword detection at morpheme-level or word-level and interaction with a database is a strategy typically employed in Spoken Dialog Systems, where processing speed both by the Speech Recognizer and any modules interacting with the speech recognition process is of crucial importance for task success and the overall efficiency of the system.

## 6.2.2. Framework of Implementation

The detection-strategies are in the form of parallel and complementary program functions (Functions 1 -5) and are defined within a pragmatic framework, namely Grice's cooperative principle

(Grice 1975). "Marked" elements concerning implicit expression of opinions and/or viewpoints, the perception of irony and general context, spirit and background information are be related to the flouting of the Maxims of Quantity, Quality and Manner in Grice's Cooperativity Principle. Specifically, the detection-strategies are based on the parameters of the flouting of Grice's Cooperative Principle (Grice 1989; Hatim 1997) presented in 3.1.

### 6.2.3. Morphological Processing on a Word basis

A more obvious connection to the Maxim of Quantity in the proposed approach is the strategy of counting specific groups of words as predefined sets of strings. In other words, the strategy involves morphological processing on a statistical basis. This concerns the surveyance of any repetitions of the same word in short intervals (i.e. in each sentence or in every other sentence) (Function 1) and of any high frequency of adjectives and adverbials, containing semantic features related to (i) descriptive features (ii) mode (iii) malignant/benign action or (iv) emotional/ethical gravity (Function 2). Function 2 may result to a slight overgeneration of "marked" adjectives and adverbs (see examples (9) and (10) below), however it helps "trap" and sustain passages in texts with understatements and subtle puns that are difficult to process by typical systematic strategies.

A typically high frequency of adjectives and adverbials was encountered in the online journalistic videos in (7a) and (7b).

(7a) …"dirty little (secret)", "(how) ludicrous", "big (lie)", "frankly", "just", "really", " real"…

*Speaker: Kathryn Serkes, Assn. of American Physicians and Surgeons*
Lou Dobbs Tonight, CNN, "Is universal health care the

solution for the United States health care system?"[1]

(7b) …"laessig" (casual, cool), "bequem" (comfortable, convenient), "pur" (pure, a word of foreign origin in German),"frischen" (fresh), "pompoesen" (pompous)...

*Speaker: Female Correspondent (name not available)*, Spiegel Online
"Merkels Wahlkampf auf Russisch: Medwedew, Machismo und Menschenrechte"[2]

Additionally, the detection of elements with a typically pragmatic function (for example, emphasis), namely of specific sets of particles, adverbials and exclamations are included in the operations of the morphological processing task (Function 3). Typical examples of such elements are the expressions "merely", "utterly", "surely", "gee", "oops" and "oh". In German, typical examples are the exclamations "paperlapap" (= nonsence) (in transcribed file from Spiegel Online, CDU/CSU-Wahlprogramm: Wer glaubt an Steuergeschenke? 29.06.2009) and "doch bitte" („doch" = emphasis element, „bitte" (=please, if you please: here, slightly ironic) (in context: „sich doch bitte (um die Menschenrechte in seinem Land zu kuemmern", in a transcribed file from Spiegel Online, Merkels Wahlkampf auf Russisch: Medwedew, Machismo und Menschenrechte 14.08.2009).

---

[1] CNN, Added On August 19, 2009 <http://www.cnn.com> (Average time of speaker: 2: 00 mins)
[2] 14.08.2009 http://www.spiegel.de/video/video-1016699.html (Length of video: 02:35 mins)

### 6.2.4. Morphological Processing on a Word-stem Basis and on a Word-suffix Basis

Morphological processing operating on a word-stem and suffix basis is a strategy with a more obvious connection to the flouting of the Maxim of Quality in Grice's Cooperativity Principle. Morphological processing operates on a word-stem basis involving the detection of verb-stems from verbs containing semantic features (including implied connotations in language use) related to (i) mode (ii) malignant/benign action or (iii) emotional/ethical gravity. These verb stems are located low in an ontology, approachable with Selectional Restrictions of the modifying type (and/or Wordnets). Other verb-stems, such as verb-stems of primitive verbs are discarded, with the exception of the defined subset of modal verbs (Function 4). Some verb stems with non-neutral semantic features may be processed by default by Functions 2 and Function 3 detecting adjectives, adverbs and exclamations (see text in 7a and 7b).

Morphological processing also operates on a suffix basis involving the detection of suffixes producing diminutives and a predefined set of derivational suffixes resulting to a nominalization of verbs (derivational suffixes producing participles and actor thematic roles are excluded), verbalization or adjectivization of proper nouns (Function 5). Thus, a specified and finite set of "marked" suffixes such as ("-istas", in American English, "-erei" in German) is defined, as well as a defined group of "marked" verb-stems, such as "swear","struggle", "schuften" (partially equivalent to "to slave away") (German).

Verb stems constituting "marked" elements in entire written and transcribed online texts are be presented to the translator, such as the fragment from the transcribed spoken text in (8a, b) (names of countries and politicians are omitted). In the transcribed Spoken Text with a medium percentage of instances with "marked"

information of the following example (8a), the verbs "vowed" (note the repetition of the same word), "loves", "dumbfolded" and "scratching" are perceived by the proposed module as verb stems constituting "marked" elements (indicated in bold). In the fragment of the text in (8a) the expression "absurdity" is marked by default, since it contains a word-stems related to an adjective "absurd" processed by the module We note that the fragment of the text in (8a) also contains expressions not incorporated in the morphological processing module, such as the expressions "hullabaloo", "in a bubble" and "scratching our heads", as well as the adverbial "just" (indicated in italics)[3].

(8a) (Fragment)

> GEORGE WILL (ABC 10/28/01): The administration knows he's **vowed**, […] has **vowed** revenge, he has anthrax, he **loves** biological weapons, he has terrorist training camps,including 747's to practice on...
> BILL MOYERS: It was proving difficult to distinguish the opinion of the pundits form the policies of the administration…but as the *hullabaloo* over […] grew in Washington, Bob Simon of CBS News 60 Minutes was **dumbfolded**..[.......]
> BILL MOYERS: **Absurdity**. The Washington press corps cannot question an absurdity?
> BILL MOYERS: *In a bubble.*
> BOB SIMON: Yeah *in a bubble*. Where as we who've spent weeks *just* walking the streets of […] and in other situations in […] just were **scratching** *our heads*. In

---

[3] Complete online transcribed text "Buying the War", April 25, 2007, PBS available at http://www.pbs.org/moyers/journal/btw/transcript1.html

ways that perhaps that the Washington press corps could not.

Typical examples in German are the verbs "zurueckegerudert" (= to row back) and "belehren" (= to take lessons from, to be patronized by) (in context:„sich von der Deutschen Kanzlerin belehren zu lassen)" [4] This text is also dotted with various expressions as well as "marked" elements processable by the proposed module.[5] Typical elements from this text are depicted in (8b).

The finite set of the "marked" elements processable in Functions 1-5 by the proposed module can, therefore, function as pointers to texts with non-neutral content, even if these "marked" elements do not constitute fixed expressions nor can fully capture all pragmatic information signalizing the intentions and/or overall spirit of the speaker or the author. Furthermore, for the international audience, these "marked" elements can be "safer" pointers to the intentions and/or spirit of the speaker than most culture-specific paralinguistic elements related to prosody and gesture.

(8b) (Fragment)
….."(Rekordverschuldung) egal" ((record of debt), who cares about that/ that's (totally) indifferent), "paperlapap" (= nonsence), "(ein) schoenes (Wahlversprechen)" ("that's some (ironic: great) pre-election promise), "(die Chefin will es so und damit) basta" (that's what the boss wants and that's final), "zurueckegerudert (= to row back)…..

---

[4] In a transcribed file from Spiegel Online, Merkels Wahlkampf auf Russisch: Medwedew, Machismo und Menschenrechte 14.08.2009
[5] Complete online spoken text available on Spiegel Online, CDU/CSU-Wahlprogramm: Wer glaubt an Steuergeschenke? 29.06.2009

> *Speaker: Male Correspondent,Authors/ Correspondents:*
> *Martin Heller, Martin Suemering*
> CDU/CSU-Wahlprogramm: "Wer glaubt an Steuerges-
> chenke?" 29.06.2009 [6]

The difference in the overall content and degree of neutrality between texts containing "marked" elements processable by the proposed module and texts not containing such elements is also shown in online written journalistic texts. The examples in (9) and (10) depict fragments from online written journalistic texts with a medium percentage of instances with "marked" information. In both texts the content is related to "Announcement" and "Description" Speech Acts respectively. The text in (9), whose content is related to the "Announcement" Speech Act, contains elements related to the author's style of writing and does not provide an entirely neutral description of facts. The content of the text in Example 10 is related to the "Declaration" Speech Act and explicitly stated point of view. We note that the fragment of the text in (9) also contains expressions not incorporated in the morphological processing module, but with word-stems related to adjectives processed by the module and marked by default, such as the expressions "gloom(y)" and "woe(ful)" (indicated in italics).

(9) Online written journalistic text with a medium percentage of instances with "marked" information and fragment from text

Title: "Danes ponder joining the euro, Icelanders wish they could"[7]

---

[6] Length of video: 02:05 mins.
[7] "Nordic nerves Denmark and the euro" Jan 29th 2009 | COPENHAGEN, From *The Economist* print edition.

"DENMARK'S story has an Old Testament resonance **just** now. Seven years of **plenty** since Anders Fogh Rasmussen became prime minister in 2001 are now being followed by **leaner** times. Danes **hardly** face seven years of famine, but jobs are being **axed**, property prices are **plummeting** and belts are being tightened everywhere. The government has already produced two bank rescue packages and it is now pondering the notion of a **much larger** fiscal stimulus.

…Yet the *gloom* has not, so far, **enveloped** the prime minister…..

…"It is **well** known that the international financial crisis came washing over us, as it did a lot of other countries," Mr Fogh Rasmussen said this week. He added that, thanks to **prudent** husbandry, Denmark is in a **good** position ….

…The prime minister's belief that his government "behaved **sensibly**" and is accordingly without **blame** for the country's economic *woes* is not universally accepted.

….Mr Fogh Rasmussen will **surely** then have won enough kudos to secure his place in posterity—and perhaps a **plum** foreign job such as NATO secretary-general or, if the Lisbon treaty is ratified, the EU's first permanent president. At least he would not then have to worry about seven **lean** years."

(10) Fragment from a text with a medium percentage of instances with "marked" information and explicitly stated point of view

"Britain must tackle its energy gap," By Andrew Duff[8]

---

[8] Published February 4, 2009 19:39 | Last updated: February 4, 2009 19:39.

"At a time when confidence is **so low**, when investment is **so desperately** needed and when finance is **so challenging**, commentators seem to **demonise** even **modest** investor returns in the **most** strategically **important** part of the economy. What is needed to stimulate the **extraordinary** levels of investment required is **simply** encouragement. We have a **real** desire to work in partnership with government and communities to deliver billions of pounds of infrastructure, along with the jobs and technologies that will **surely** follow(*The Financial Times Limited 2009*)"

As an additional component to the annotation module, the [gram-conn:subj] feature may be added to verbs in transcribed spoken texts of pro-drop languages such as Greek. As an optional component to the annotation module, an interaction of the module with a lexicon involves the mapping of noun-stems or noun-stem + verb-stem pairs in respect to a lexicon containing fixed expressions or a database constructed from available specialized dictionaries for fixed expressions (Function 6). These are usually stems of words, not directly related to political, financial and journalistic sublanguages, such as "run-gauntlet" and "Kakao-ziehen" in German (partially equivalent to "to run somebody down"). However, the optional component may entail more processing time and/or computational cost.

### 6.2.5. Evaluation

The initial implementation and test-versions (core module and "toy" program) of the proposed module, along with user-translator-tests, will give the first concrete indicators of the usability and fine-tuning of the module's design. However, there are three factors to be

taken into account for the module's evaluation, namely (1) speed, (2) correctness and (3) avoidance of overgeneration. In other words, effectiveness and usability of the module is achieved when the speed of the output provided allows the user-translator to check multiple online journalistic texts within a short time span and when in the average online journalistic text the indicated "marked" elements and the parts of the text containing them are correctly selected but do not exceed 50% of the elements in the text, due to overgeneration.

However, since emotionally and socio-culturally "marked" elements most often do not occur in every sentence and are usually dispersed throughout the text, in most cases, a relatively small number of them may be enough to signalize speaker's or author's intentions. Due to its context-free specifications, the morphologically-based approach may not provide a 100% coverage of all implied information, however, the above-described categories of "marked" elements show to be able to function as indicators of texts with a non-neutral, emotionally and socially "marked" content.

## 6. Summary and Further Research

Although an above-average or even a fairly good fluency of a foreign language (such as English or German) is a typical characteristic of most "semi-professional" translators working with multilingual written and transcribed journalistic texts, the lack of the necessary exposure to the culture(s) related to the foreign language(s) concerned, especially due to distances and frequent change of location, often hinders the detection of essential information presented either in a subtle form or in an indirect way, such as text-in-context relationships and socio-textual practices.

Typical indicators of the presence of such elements are not reliable and consistent enough to apply in the majority of online spoken/written journalistic texts to be used by international

professionals. The target of the present approach is to address the most typical and frequently occurring problems in the translation process in regard to connotative features related to speaker and culture-dependent elements constituting emotionally and socio-culturally "marked" elements. These connotative features can be transferred or presented as additional information, even in cases where it is not always expected for the translator to convey exactly the same style and "spirit" in the target language.

A set of morphologically-based and context-free specifications is provided to the "semi-professional" translator for the detection of the largest possible percentage of evident and implied connotative features in written and transcribed spoken journalistic texts. The implementation of these specifications targets to the in-depth understanding and successful management of information of the source language text, and also targets to the successful transfer of the highest possible extent of information in the target language. A proposed set of morphologically-based and context-free specifications may function as a stand-alone framework, but they may also be applied as an annotation tool for online written and transcribed spoken journalistic texts in combination with a statistical evaluation of the texts. Thus, the translator receives information about the overall "marked" context of the journalistic text and is presented with the problematic points annotated with the tags in each text.

In respect to the designed module, further research will investigate two basic issues. The first issue to be investigated is (I) the extent to which approaches involving "marked" elements related to the flouting of the Maxims in Grice's cooperativity principle can apply to other languages and/or language families with similar or different morphological parameters. The second issue under investigation is a more rigorous statistical analysis related to the extent to which the processable "marked" elements functioning as pointers to non-neutral content can (II) compensate for and/or balance all unprocessable pragmatic information signalizing the

intentions and/or overall spirit of the speaker or the author.

# References

Alexandris, C. & S-E. Fotinea. 2004. Reusing Language Resources for Speech Applications involving Emotion, *Proceedings of the 4th LREC Conference*, Lisbon, Portugal 9, 1383-1386.

Alexandris, C. 2003. Translational Issues in the Sublanguage of Written and Spoken Journalistic Texts in Modern Greek in *Proceedings of the International Conference on Choice and Difference in Translation,* Athens, 2003, 287-307.

Constantie L. & L. Lockwood. 1999. *Software for Use: A Practical Guide to the Models and Methods of Usage-Centered Design*, Reading, MA: Addison-Wesley.

De Silva, R. 2006. Explicitation in Media Translation: English Translation of Japanese Journalistic Texts, in *Proceedings of the 1st Portuguese Translation Conference,* Caparica, Portugal, 2006.

Gayral, F,. et al. 2000. On Verb Selectional Restrictions: Advantages and Limitations. in Christodoulakis, D.N. (ed.): *Lecture Notes in Artificial Intelligence*, Berlin: Springer, 57-68.

Grice, H.P. 1989. *Studies in the Way of Words*, Cambridge, MA: Harvard University Press.

___. 1975 Logic and conversation, in Cole, P. & Morgan, J. (eds.) *Syntax and Semantics*, 3. New York: Academic Press.

Hatim, B. 1997. *Communication Across Cultures: Translation Theory and Contrastive Text Linguistics*, Exeter: University of Exeter Press.

Kontos, J,. et al. 2000. Greek Verb Semantic Processing for Stock Market Text Mining, in Christodoulakis, D.N. (ed.): *Lecture Notes in Artifical Intelligence*, Berlin: Springer, 395-405.

Lehrndorfer A. 1996. *Kontrolliertes Deutsch: Linguistische und*

*Sprachpsychologische Leitlinien für eine (maschniell) kontrollierte Sprache in der technischen Dokumentation*, Narr, Tuebingen.

Nielsen, J. 2000. *Designing Web Usability*, Indianapolis, IN: New Riders Publishing.

Sifianou, M. 2001. *Discourse Analysis. An Introduction*, Athens: Leader Books.

Wiegers, K.E. 2005. *Software Requirements*, Redmond, WA: Mircosoft Press.

Wilks, Y., & D. Fass. 1992. The Preference Semantics Family in *Computers Math. Applications* 23.2-5, Amsterdam: Pergamon Press, 205-221.

CIMWOS project 2000-2003. available at URL <www.xanthi.ilsp.gr/cimwos/>

Christina Alexandris
National University of Athens
Department of German Studies
School of Philosophy
Panepistimioupol,
GR-15784 Zografou, Athens
Greece
Phone +30-210-7277307; Email: calexandris@gs.uoa.gr