

Detecting Legal Definitions for Automated Style Checking in Draft Laws

Stefan HÖFLER¹, Alexandra BÜNZLI and Kyoko SUGISAKI
Institute of Computational Linguistics, University of Zurich, Switzerland

Abstract. The automatic detection of definitions in legal texts has mainly been considered from the perspective of legal information retrieval so far. In this paper, we demonstrate that the same techniques can also be employed for the purpose of a domain-specific automated style checking of draft laws. We report on the automatic identification of legal definitions in a large corpus of Swiss laws, and we show that the information provided by the pattern-oriented detection mechanism we apply suffices to allow for an automatic evaluation of a range of stylistic rules concerning the use of legal definitions in legislative texts.

Keywords. Definition extraction, Legislative drafting, Automatic Style Checking

Introduction

In recent years, there have been at least two major studies dealing with the automatic detection of definitions in legal texts [1,2]. The goal of these studies was to provide easier access to the definitions of concepts contained in large collections of legal texts and to facilitate the construction of legal ontologies. In this paper, we show how the detection strategies developed by these studies can be adopted for yet another purpose: to support domain-specific style checking in legislative drafts. We demonstrate how the said legal information retrieval techniques can be combined with the methods used in style checkers for technical writing to identify and assess legal definitions in drafts of Swiss laws.

The paper is organised as follows. We first introduce the method of error modelling used in commercial style checkers for technical writing and discuss what it means to adapt this method to the domain of legislative drafting (section 1). We then present legal definitions as one phenomenon that lends itself to being assessed by automated style checking (section 2). In the main part of the paper, we report on enhancing the method of error modelling with legal information techniques to detect legal definitions in a corpus of Swiss legislative texts (section 3). Afterwards, we briefly discuss the ramifications of our analysis for an automated style checking of legal definitions in draft laws (section 4) and conclude with a short summary and an outlook to future research.

¹Corresponding author: Stefan Höfler, Institute of Computational Linguistics, University of Zurich, Binzmühlestrasse 14, CH-8050 Zurich, Switzerland; E-mail: hoefler@cl.uzh.ch

1. Automated Style Checking for Legislative Drafting

1.1. Method

Automated style checking has primarily been employed in the domain of technical writing [3]. Companies often control the language used in their technical documentation in order to improve the understandability, readability and translatability of these texts. Controlled language checkers are tools that evaluate input texts for compliance with the style guidelines of a company. Examples of wide-spread commercial tools that offer such style checking for technical texts written in German are acrolinx IQ, developed by acrolinx², and CLAT, developed by IAI³.

State-of-the-art controlled language checkers work along the following lines. They first perform an automatic analysis of the input text (tokenisation, text segmentation, morphological analysis, part-of-speech tagging) and enrich it with the respective structural and linguistic information. They then apply a number of pre-defined rules that model potential “errors” (i.e. violations of individual style guidelines) and aim at detecting them in the analysed text. Most checkers give their users the option to choose which rules the input text is to be checked for. Once a potential violation of the company’s style guidelines has been detected, the respective passage is highlighted and an appropriate explanatory text is made available to the user.

1.2. Domain

Our research is concerned with adopting the method of automated style checking by error modelling to legislative drafting in Switzerland. In the federal administration of the Swiss Confederation, drafts of acts and ordinances go through several editorial cycles. They are, in most cases, originally written by civil servants in one of the federal offices concerned, and they are then reviewed and edited both by legal experts (at the Federal Office of Justice) and language experts (at the Federal Chancellery). While the former ensure that the drafts meet all relevant legal requirements, the latter are concerned with the linguistic quality of the texts. To support this task, the federal authorities have drawn up style guidelines specifically geared towards Swiss legislative texts [4].

These guidelines differ in two major aspects from the guidelines applied in technical writing. Firstly, while guidelines for technical writing usually consist of relatively simple linguistic rules (e.g. that measurement nouns should be abbreviated or that future tense should be avoided), a substantial number of legislative drafting rules are concerned with more complex domain-specific issues (e.g. that one sentence should not contain more than one norm, or that legal definitions must not contain normative elements). Secondly, while style guidelines for technical writing must be followed at all times, many style guidelines for legislative drafting are merely rules of thumb. Whether they must actually be applied in a specific case often depends on a variety of textual and extra-linguistic factors and must thus be left for the editors to decide. Automated style checking can merely point the editors to passages that *potentially* contain the violation of a style guideline.

As a consequence, good precision is more important for an automated style checker than good recall: while false positives can be annoying for editors who use such a system

²<http://www.acrolinx.de>

³<http://www.iai-sb.de/iai>

and consequently damage its acceptability, true positives will be useful no matter how many other potential true positives the system has missed.

2. Legal Definitions

Legal definitions are a special kind of statements used in acts and ordinances [5,6,7]. Rather than conveying an actual norm, these statements explain the sense in which a specific term is used in the given text. Legal definitions are thus auxiliary constructs; they serve the purposes of (i) keeping the text short and easy to read, (ii) specifying linguistic elements contained in other, more immediately normative statements, and thus (iii) guiding the interpretation of those other statements. Legal definitions can be listed in a specific article at the beginning of a law or occur at any place throughout the text.

The following sentence represents a typical example of a legal definition. It defines the sense in which the term *clearing* is used in the respective text:

- (1) *Als Rodung gilt die dauernde oder vorübergehende Zweckentfremdung von Waldboden.*
'Clearing shall be deemed to be the permanent or temporary misuse of forest soil.'⁴

Bratschi [5] summarises the problem with using legal definitions in legislative texts as follows (original in German, translation ours):

Good legal definitions can, if they are employed economically, improve the comprehensibility of a legislative text, guide its interpretation and thus contribute to an optimisation of the application of the law. In practice, however, bad legal definitions are frequently found alongside good ones, that is, legal definitions that are completely superfluous or that are phrased in a manner that obscures the content of legal norms rather than clarifies it.

Drawing from official drafting guidelines and from professional experience, she develops a set of thirteen rules of thumb describing the requirements that good legal definitions should fulfil. At least five of them lend themselves to an automatic evaluation:

- (2)
 - a. If possible, legal definitions are to be avoided.
 - b. A term must usually only be defined if it occurs at least three times in the text.
 - c. A term must only be defined once within the same text.
 - d. Legal definitions must be placed where they can be easily found.
 - e. A term must normally not be defined by itself.

As a first step towards automatically assessing these style rules, the legal definitions contained in a draft law have to be detected. This includes the identification of the legal definition as a whole as well as the determination of its two main components: the term it defines (its *definiendum*) and the definition it gives for that term (its *definiens*). In the next section, we will describe how we approached the task of detecting legal definitions in Swiss legislative texts. Afterwards, we will discuss the strategies that can be devised to automatically assess whether the detected legal definitions comply with the style rules listed in (2).

⁴Article 4 of the Federal Act on Forest (SR 921.0)

3. Experiment: Detecting Legal Definitions by Shallow Pattern-Matching

3.1. Approach and Related Work

The approach we adopted in our experiment is similar to that used by Walter [1] and by de Maat et al. [2]. Walter's work deals with the automatic extraction of definitions from German court decisions, and de Maat et al. are concerned with the automatic classification of sentences in Dutch laws. In both cases, a set of patterns typical to sentences containing definitions is compiled and the texts are searched for occurrences of these patterns. While Walter includes syntactic information in their patterns, de Maat et al. work at the surface level of the texts only. Walter reports that with their method it was possible to extract definitions from German court decisions with a precision of above 70%; in the experiments of de Maat et al., 94% of the sentences in Dutch laws were classified correctly. We suspected that the difference was to be explained with the fact that the language of court decisions exhibit greater variation than the language of acts and ordinances and hypothesised that a similar approach would also be successful with Swiss legislative texts. Our experiments are thus aimed at exploring to what extent legal definitions in German-language Swiss laws can be detected merely by searching for typical surface patterns.

3.2. Pre-processing

To build a test environment for our study, we downloaded a corpus of 1,874 (out of a total of 1,915) Swiss legislative texts from the Federal Administration's website⁵ (federal and state constitutions, federal acts and ordinances) and converted these HTML documents into XML. Apart from line breaks, all formatting information was removed from the documents, and all special characters were rendered into a normal text-based representation. We then enriched the texts with explicit document structure information by performing an automatic pattern-based annotation of chapter, section, subsection, article, paragraph, sentence and enumeration item boundaries.

3.3. Patterns

By manually inspecting the texts, and by interviewing professional legal editors, we compiled, as our hypotheses, a set of patterns that are typical for legal definitions in Swiss acts and ordinances. We focused on legal definitions that appear in the textual part of a law and excluded legal definitions listed in tables and appendices. To make sure that we capture as many phenomena as possible, we started off with a relatively wide conception of what a legal definition is that includes any definitional statement with at least one of the three functions mentioned at the beginning of section 2 above. We found that, on the basis of their form, one can distinguish between three types of legal definitions: bracketed definitions, enumerated definitions, sentential definitions. We will briefly describe the characteristics of each type and report the main challenges we faced in the development of strategies to detect legal definitions of that type automatically.

⁵<http://www.admin.ch/ch/d/sr/>

3.3.1. Bracketed Definitions

Bracketed definitions are mainly used to introduce abbreviations. The abbreviation immediately follows the longer phrase it will henceforth be assumed to stand for. In the example below, the short term *Dienst* ('service') is introduced as an abbreviation for *Bundessicherheitsdienst* ('Federal Security Service'):

- (3) *Der Bundessicherheitsdienst (Dienst) übt die Aufgaben im Sinn von Artikel 1 aus.*
'The Federal Security Service (Service) performs the tasks according to article 1.'⁶

Bracketed definitions exhibit the following pattern (*X* = definiendum; *Y* = definiens):

- (4) *Y (X)*

The main difficulty with identifying bracketed definitions is that they must be distinguished from other contents that can also occur in parentheses, such as listings of specific examples for a term, or references to other legislative texts. We have developed a set of search patterns (implemented in the form of regular expressions) that explicitly exclude specific types of non-definitional content that also tend to appear in parentheses.

We have included proper abbreviations as shown in example (5) below in this search pattern as they are subject to the same style rules as legal definitions.

- (5) *Im Übrigen gelten die Vorschriften der Kernenergieverordnung vom 10. Dezember 2004 (KEV).*
'As for the rest, the provisions of the Nuclear Energy Ordinance of 10 December 2004 (NEO) apply.'⁷

A special sub-pattern occurs where the abbreviation of a specific law is introduced within a bracketed reference to that law:

- (6) *Für alle Gebäude, in denen Bundesbehörden untergebracht sind, wird das Hausrecht (Art. 14 des BG vom 26. März 1934 über die politischen und polizeilichen Garantien zugunsten der Eidgenossenschaft, GarG) von den Vorstehern der untergebrachten Departemente, Gruppen, Ämter oder andern Bundesbehörden ausgeübt.*
'For all buildings in which federal authorities are accommodated, the householder's rights (Art. 14 of the Federal Act of 26 March 1934 on Political and Police Guarantees towards the Confederation, GarA) are exercised by the heads of the accommodated departments, groups, offices or other federal authorities.'⁸

The following pattern applies to this sub-type of bracketed abbreviations:

- (7) (*<article_ref>R</article_ref> des/der Y, X*)

⁶Article 2 Paragraph 1 of the Ordinance on Security within Federal Responsibility (SR 120.71)

⁷Article 1 Paragraph 4 of the Ordinance on Containers and Pipes with Safety-Related Classification in Nuclear Facilities (SR 732.13)

⁸Article 23 Paragraph 2 of the Federal Act on Measures for the Preservation of Internal Security (SR 120)

With the described pattern-matching methods, we are able to precisely identify the abbreviation (the definiendum), but we are not always able to delineate with certainty the phrase the abbreviation refers to (the definiens). However, as we will show in section 4, the information the method can supply is in fact sufficient for the application of all but one of the style rules in question.

3.3.2. Enumerated Definitions

Definitions of related terms are frequently compiled in a single enumeration, where each item constitutes a legal definition of its own. The following paragraph is an example of such an enumeration of legal definitions:

(8) *In diesem Gesetz bedeuten:*

- a. *Museum des Bundes: Museum, das organisatorisch zur zentralen oder dezentralen Bundesverwaltung gehört;*
- b. *Sammlung des Bundes: Bestand an beweglichen Kulturgütern, der im Eigentum des Bundes oder einer Einheit der dezentralen Bundesverwaltung steht.*

‘In this act shall mean:

- a. museum of the Confederation: a museum affiliated to the central or decentralised federal administration;
- b. collection of the Confederation: a stock of mobile cultural goods in the possession of the Confederation or of a unit of the decentralised federal administration.’⁹

Enumerated definitions exhibit the following pattern:

(9) <enumeration_item>**X**: **Y**</enumeration_item>

The main difficulty here is to distinguish such definitions from other enumeration items that contain a phrase followed by a colon. To gear our search patterns towards actual definitions, we have implemented additional restrictions based on the following observations. Firstly, phrases introducing enumerations of legal definitions usually contain the phrases *im Sinne* (‘in the sense’), *in dieser Verordnung* (‘in this ordinance’) or *in diesem Gesetz* (‘in this act’) and the verbs *gelten als* (‘be deemed’) or *bedeuten* (‘mean’); cf. example (8) above. Secondly, if legal definitions are collected in a separate article, the title of that article usually contains the words *Begriffe* (‘terms’) or *Definitionen* (‘definitions’). Only enumeration items that exhibit pattern (9) and fulfil at least one of these two additional criteria are marked as legal definitions by our system.

3.3.3. Sentential definitions

Legal definitions can also be phrased as full sentences. An example of a sentential legal definition is the passage cited in (1) above. We have derived a set of five main patterns that such sentences usually follow:¹⁰

⁹Article 3 of the Federal Act on the Museums and Collections of the Confederation (SR 432.30)

¹⁰Approximate translations: (10) **X** is/are deemed to be(:) **Y** – (11) ‘**X** comprises/comprise(:) **Y**’ – (12) ‘**X** is/are present if **Y**’ – (13) ‘**X** is/are to be understood as **Y**’ – (14) ‘**X** is/are(:) **Y**’.

- (10) <sentence>Als *X* gilt/gelten(:) *Y*</sentence>
 (11) <sentence>*X* umfasst/umfassen(:) *Y*</sentence>
 (12) <sentence>*X* liegt/liegen vor, wenn *Y*</sentence>
 (13) <sentence>Unter *X* ist/sind *Y*₁ zu verstehen(, *Y*₂)</sentence>
 (14) <sentence>*X* ist/sind(:) *Y*</sentence>

Optionally, the phrase *im Sinne dieser Verordnung/dieses Gesetzes/von ...* ('in the sense of this ordinance/of this act/of ...') can be added after the element *X* in each of the above patterns, as illustrated by the following example:

- (15) *Als Handel im Sinne dieses Gesetzes gilt jedes gewerbsmässige Anbieten, Erwerben oder Weitergeben von Kriegsmaterial.*
 'Trade in the sense of this act shall be deemed to be any commercial advertising, acquiring or passing on of war material.'¹¹

Syntactic variants with the phrase *im Sinne (von)* in sentence-initial position are also possible:

- (16) *Im Sinne der folgenden Bestimmungen gelten als rheumatische Krankheiten:*
 a. *Chronische Polyarthritis*
 b. *Spondylarthritis ankylopoetica [...]*
 'In the sense of the following provisions, rheumatic diseases shall be deemed to be:
 a. chronic polyarthritis
 b. ankylosing spondylitis [...]'¹²

Pilot studies we performed showed early on that pattern (14) cannot be sufficiently restricted by means of regular expressions operating on the surface of the text alone: even if the phrase *im Sinne (von)* is added as an additional restriction, the number of false positives still outnumbers by far the number of correctly recognised instances. We thus decided to exclude pattern (14) from the present investigation and to deal with it in a follow-up study, in which we intend to parse the texts and enrich them with explicit syntactic information.

3.4. Results

To determine the recall that our search patterns exhibit we had 17 legislative texts manually annotated for legal definitions. The texts were selected from across the domains of law defined in the federal administration's systematic collection¹³: 2 texts were selected from constitutional law, 2 from private law, 2 from criminal law, 2 from education, science and culture law, 2 from national defence law, 2 from finance law, 3 from energy and transport law, and 2 from economy law. The annotators were told to mark whatever statement they deemed a legal definition. For the reasons explained in the previous sec-

¹¹Article 6 Paragraph 2 of the Federal Act on War Material (SR 514.51)

¹²Article 3 Paragraph 1 of the Federal Act on Federal Subsidies for Fighting Rheumatic Diseases (SR 818.21)

¹³<http://www.admin.ch/ch/d/sr/>

Table 1. Precisions of the individual search patterns. The numbers for the listed sentence patterns also include the syntactic variations discussed in section 3.3.3. For each pattern, 150 randomly chosen positives were evaluated (or fewer if a smaller total number of positives were returned by the system).

Type (Pattern)	Total Returned	Total Evaluated	True Positives	False Positives	Precision
Bracketed Definitions	7,691	150	141	9	0.94
Enumerated Definitions	1,072	150	149	1	0.99
Sentential Definitions:					
– <i>Als X gilt/gelten Y</i>	1,498	150	144	6	0.96
– <i>X umfasst/umfassen Y</i>	713	150	121	29	0.81
– <i>X liegt/liegen vor, wenn Y</i>	116	116	116	0	1.00
– <i>Unter X ist/sind Y zu verstehen</i>	23	23	23	0	1.00
– <i>X ist/sind Y</i>	NA	NA	NA	NA	NA

tion, 12% of the annotated sentences were excluded from the evaluation because they exhibited pattern (14). Of the remaining 79 paragraphs that the annotators had marked as containing legal definitions, our system recognised 74, which amounts to a recall of 94%.

Precision was evaluated for each pattern individually. The developed search strategies were applied to all 1,874 texts contained in our corpus. For each pattern, we evaluated a set of 150 randomly chosen instances returned by the system – or the total number of instances returned if it was less than 150. The results are detailed in Table 1. Precision was at 94% or above for all but one of the evaluated patterns: sentential definitions with *umfassen* (‘comprise’) ranged slightly below at only 81% precision. Most of the patterns we devised thus proved to be fairly reliable indicators for the presence of a legal definition.

Our results show that a substantial number of legal definitions can be detected automatically with relatively shallow pattern-matching methods. In that regard, they confirm what de Maat et al. [2] found for Dutch laws. To identify legal definitions of pattern (14), we will have to parse the texts and enrich them with explicit syntactic information .

4. Discussion: Assessing the Detected Definitions

We will now discuss to what extent the presented detection methods will allow us to assess legal definitions for their compliance with the style rules introduced in section 2:

- (17) a. If possible, legal definitions are to be avoided.
- b. A term must usually only be defined if it occurs at least three times in the text.
- c. A term must only be defined once within the same text.
- d. Legal definitions must be placed where they can be easily found.
- e. A term must normally not be defined by itself.

Rule (17a) can be evaluated merely by detecting legal definitions in the text: if, in general, any legal definitions should be avoided, any legal definition that can be found in a text

potentially violates the rule and should thus be highlighted to the user. As no further processing is needed, the described detection mechanisms are sufficient to assess a text for its compliance with this very crude rule of thumb.¹⁴

Rule (17b) can be assessed by searching the text for occurrences of the terms defined by the detected legal definitions. A legal definition violates the rule if less than three occurrences of its definiendum can be found in the remainder of the text.

We found, however, that this rule does not apply to all types of legal definitions that our detection strategies identify. It is often the case that a legal definition is used not to provide an abbreviation but only to specify more precisely the meaning of a term that was used in an immediately preceding sentence or paragraph:

- (18) ¹ *Bei angeordneter Nacht- und Sonntagsarbeit wird die Arbeitszeit mit dem Faktor 1,25 multipliziert; sie wird durch Freizeit ausgeglichen.*
² *Als Nachtarbeit gilt die zwischen 22 und 6 Uhr geleistete Arbeit.*
⁴ For ordered night and Sunday shifts, the working hours are multiplied by a factor of 1.25; they are compensated with leisure time.
² A night shift shall be deemed to be any work carried out between 22 and 6 o'clock.¹⁵

We thus evaluate rule (17b) only for a specific subset of the legal definitions, namely for (i) bracketed legal definitions and abbreviations, (ii) enumerated legal definitions, and (iii) legal definitions that are comprised in a special definitions article, i.e. an article whose title contains the keywords *Begriffe* ('terms') or *Definitionen* ('definitions').

The application of the rule is further complicated by the fact that the defined term does not necessarily occur in the inflectional form that it has in the definition. We use the TreeTagger [8] to lemmatise our texts and then search for the lemma of the definiendum rather than the word form found in the definition. However, a case that this approach does not solve yet is the problem that if legal definitions are meant to determine the meaning of a verb, the defined verb appears in its nominalised form in the legal definition but as a verb in the remaining text. Nominalised verbs thus need to be de-nominalised before the text can be searched. As an example, the legal definition in (19) specifies the sense in which the verb *konsolidieren* ('consolidate') is used in the text by referring to that verb in its nominalised form *Konsolidierung* ('consolidation'):

- (19) *Die Informatikplanungen werden auf Stufe VBS konsolidiert. Die Konsolidierung im Sinne dieser Verordnung umfasst: [. . .].*
'The IT strategies are consolidated at the level of the DDPS. Consolidation in the sense of this ordinance comprises: [. . .].'¹⁶

Rule (17c) can be checked by automatically assessing if any two definitions detected in the text define the same term. Since the inflectional form of the definiendum depends on

¹⁴Note again that even though the technical term "error modelling" is used to denote the method we apply, a style checker for draft laws cannot identify "errors" in the proper sense of the word but merely draws the editor's attention to passages that *may* be violating a style guideline or general drafting principle.

¹⁵Article 19 of the Ordinance on the Personal Statute of the Federal Institute of Intellectual Property (SR 172.010.321)

¹⁶Article 5 Paragraph 3 of the Ordinance on IT in the Federal Department of Defence, Civil Protection and Sports (SR 510.211.2)

the syntactic pattern realised in the definition, here too, lemmas rather than actual word forms have to be compared. The rule is *not* violated if a legal definition is followed by an exception, i.e. by another legal definition that has the same definiendum but additionally contains a negation.

To assess rule (17d), the position of the legal definition and the positions of the occurrences of the defined term in the text must be determined, and the distance between them must be measured. If the distance exceeds a certain pre-defined threshold, or if the legal definition appears after (but not immediately after) the first occurrence of the term it defines, a warning message must be issued.

The evaluation of rule (17e) involves checking if the definiendum of a definition also appears in its definiens. In contrast to the other rules at hand, both the definiendum *and* the definiens are required for this check. The presented detection methods thus do not allow us to evaluate it in bracketed definitions. As we have pointed out in section 3.3.1, the definiens of bracketed definitions cannot be delineated with certainty by the shallow pattern-matching mechanisms used in our system. Here again, it will be necessary to parse the texts and enhance it with explicit information on the boundaries of the syntactic constituents.

5. Conclusion

In this paper, we have discussed how definition extraction methods developed for legal information retrieval can be employed in the automated style checking of draft laws. We have demonstrated that relatively simple pattern-matching techniques suffice to detect a fair number of legal definitions in Swiss legislative texts, and we have shown how the information thus obtained can be used to check the definitions for their compliance with style rules specifically geared towards them. In future work, the coverage of the presented detection mechanisms and the outlined style checking strategies can be improved by parsing the texts and enhancing them with explicit information on their syntactic structure.

References

- [1] S. Walter, Definition extraction from court decisions using computational linguistics technology, in: G. Grewendorf and M. Rathert (Eds.), *Formal Linguistics and Law*, de Gruyter, Berlin and New York, 2009, 183–224.
- [2] E. de Maat, R. Winkels and T. van Engers, Making sense of legal texts, in: G. Grewendorf and M. Rathert (Eds.), *Formal Linguistics and Law*, de Gruyter, Berlin and New York, 2009, 225–256.
- [3] S. Geldbach, Neue Werkzeuge zur Autorenunterstützung, *MDÜ* 4, 2009, 10–19.
- [4] Swiss Federal Office of Justice, (Ed.), *Gesetzgebungsleitfaden: Leitfaden für die Ausarbeitung von Erlassen des Bundes*, 3rd edition, Bern, 2007.
- [5] R. Bratschi, “Frau im Sinne dieser Badeordnung ist auch der Bademeister.” Legaldefinitionen aus redaktioneller Sicht, *LeGes. Gesetzgebung und Evaluation* 20(2), 2009, 191–213.
- [6] S. Weber-Lejeune, *Legaldefinitionen unter besonderer Berücksichtigung des Umweltrechts*, Nomos, Berlin/Baden Baden, 1997.
- [7] V. Knapp, Einige Fragen der Legaldefinitionen, *Archiv für Rechts- und Sozialphilosophie* LXVI(4), 1980, 511–533.
- [8] H. Schmid, Probabilistic part-of-speech tagging using decision trees, in: *Proceedings of the International Conference on New Methods in Language Processing*, 1994, 44–49.