

# 語学学習支援のための言語処理

工学博士 奥村 学 監修  
博士(工学) 永田 亮 著

コロナ社

## 刊行のことば

人間の思考，コミュニケーションにおいて不可欠なものである言語を計算機上で扱う自然言語処理という研究分野は，すでに半世紀の歴史を経るに至り，技術的にはかなり成熟するとともに，分野が細かく細分化され，また，処理対象となるものも，新聞以外に論文，特許，WWW上のテキストなど多岐にわたり，さらに，応用システムもさまざまなものが生まれつつある。そして，自然言語処理は，現在では，WWWの普及とともに，ネットワーク社会の基盤を支える重要な情報技術の一つとなっているといえる。

これまでの自然言語処理に関する専門書は，自然言語処理全般を広く浅く扱う教科書（入門書）以外には，情報検索，テキスト要約などを扱う，わずかの書籍が出版されているだけという状況であった。この現状を鑑みるに，読者は，「実際にいま役に立つ本」，「いまの話題に即した本」を求めているのではないかと推測される。そこで，これまでの自然言語処理に関する専門書では扱われておらず，なおかつ，「いま重要と考えられ，今後もその重要さが変わらない」と考えられるテーマを扱った書籍の出版を企画することになった。

このような背景の下生まれた「自然言語処理」シリーズの構成を以下に示す。

1. 自然言語処理で利用される，統計的手法，機械学習手法などを広く扱う  
近年の自然言語処理は，コーパスに基づき，統計的手法あるいは機械学習手法を用いて，規則なり知識を自動獲得し，それを用いた処理を行うという手法を採用することが一般的になってきている。現状多くの研究者は，他の先端的な研究者の論文などを参考に，それらの統計的手法，機械学習手法に関する知識を得ており，体系的な知識を得る手がかりに欠けている。そこで，そのような統計的，機械学習手法に関する体系的知識を与える専門書が必要と感じている。
2. 情報検索，テキスト要約などと並ぶ，自然言語処理の応用を扱う  
自然言語処理分野も歴史を重ね，技術もある程度成熟し，実際に使えるシステム，技術として世の中に少しずつ流通するようになってきている

## ii 刊 行 の こ と ば

ものも出てきている。そのようなシステム、技術として、検索エンジン、要約システムなどがあり、それらに関する書籍も出版されるようになってきている。これらと同様に、近年実用化され、また、注目を集めている技術として、情報抽出、対話システムなどがあり、これらの技術に関する書籍の必要性を感じている。

### 3. 処理対象が新しい自然言語処理を扱う

自然言語処理の対象とするテキストは、近年多様化し始めており、その中でも、注目を集めているコンテンツに、特許（知的財産）、WWW上のテキストが挙げられる。これらを対象とした自然言語処理は、その処理結果により有用な情報が得られる可能性が高いことから、研究者が加速度的に増加し始めている。しかし、これらのテキストを対象とした自然言語処理は、これまでの自然言語処理と異なる点が多く、これまでの書籍で扱われていない内容が多い。

### 4. 自然言語処理の要素技術を扱う

形態素解析、構文解析、意味解析、談話解析など、自然言語処理の要素技術については、教科書の中で取り上げられることは多いが、技術が成熟しつつあるにもかかわらず、これまで技術の現状を詳細に説明する専門書が書かれることは少なかった。これらの技術を学びたいと思う研究者は、実際の論文を頼らざるを得なかったというのが現状ではないかと考える。

本シリーズの構成を述べてきたが、この構成は現在の仮のものであることを最後に付記しておきたい。今後これらの候補も含め、新たな書籍が本シリーズに加わり、本シリーズがさらに充実したものとなることを祈っている。

本シリーズは、その分野の第一人者の方々に各書籍の執筆をご快諾願えたことで、成功への最初の一步を踏み出せたのではないかと考えている。シリーズの書籍が、読者がその分野での研究を始める上で役に立ち、また、実際のシステム開発の上で参考になるとしたら、この企画を始めたものとして望外の幸せである。最後に、このような画期的な企画にご賛同下さり、実現に向けた労をとって下さったコロナ社の各氏に感謝したい。

2013年12月

監修者 奥村 学

# ま え が き

言語処理の最終目標は、人間の言語を理解し、人間の言葉を話すコンピュータを実現することにある。ざっくりと言い換えれば、人間の言語をコンピュータに学ばせるための方法を探求している。その結果、自動翻訳や文書要約などの技術が実現される。

本書が対象としているのは、語学学習支援のための言語処理である。すなわち、言語処理を用いて人間が言語を学習するのを支援しようという分野である。コンピュータが言語を学習するための言語処理は、少し考えただけでも、語学学習支援と相性がよさそうである。実際、言語処理と語学学習支援は相性がよく、さまざまな研究・開発がなされている。

一方で、一見相性がよさそうな両者であるが、気を付けないといけない落とし穴がいくつもある。場合によっては、相性が悪い部分もある。また、言語処理の他の分野とはいろいろな点で異なる部分が多いため、すでに言語処理の経験があると、逆に妨げになる場合もある。この部分が語学学習支援のための言語処理の難しさであり、面白さでもある。

本書で伝えたいのは正にこの点である。本書では、語学学習支援のための言語処理における難しさを紹介し、その解決法をできるだけわかりやすく説明することを試みた。また、語学学習支援のための言語処理において重要であるが、あまり知られていない情報やティップスを紹介することにも努めた。

今後、言語処理が語学学習支援に大きく貢献すること、また、その成果が実用化され社会に役に立つことを期待している。本書がその一助を担うことができれば幸いである。

本書を著すにあたり、多くの方々にお世話になった。東京工業大学の奥村学先生には、本書を監修していただくと共に、本書への助言をいただいた。首都

大学東京の小町守先生には、本書の構成、内容、章末問題などについて多くの助言をいただいた。また、東京工業大学の高村大也先生と東京大学の川崎義史先生に原稿のチェックをお願いし、両先生から多くの改善案をいただいた。コロナ社様には、本書の刊行のためにたいへんご尽力をいただいた。これらの方々に改めてお礼を申し上げたい。最後に、裕福でない家庭環境にもかかわらず博士課程まで進学させてくれた両親に感謝したい。

2017年9月

永田 亮

コロナ社

# 目 次

## 1. イントロダクション

1.1 語学学習支援の目的 .....	1
1.2 本書の目的と読者対象 .....	3
1.3 本書の構成と読み方 .....	5
章 末 問 題 .....	8

## 2. 処理の対象となるデータ

2.1 処理の対象となるデータの概要 .....	9
2.2 学習者コーパス .....	12
2.2.1 学習者コーパスの重要性と変遷 .....	12
2.2.2 学習者コーパスの構築 .....	18
2.2.3 学習者コーパスの特徴 .....	24
2.3 その他の関連するデータ .....	32
2.4 この章のまとめ .....	33
章 末 問 題 .....	33

## 3. 語学学習支援のための言語処理を支える要素技術

3.1 概 要 .....	36
3.2 文 分 割 .....	36

3.2.1	タスク概要	36
3.2.2	性能と実例	37
3.2.3	学習者の文章を対象にした処理	38
3.3	トークン同定	41
3.3.1	タスク概要	41
3.3.2	性能と実例	41
3.3.3	学習者の文章を対象にした処理	42
3.4	品詞解析と形態素解析	43
3.4.1	タスク概要	43
3.4.2	性能と実例	45
3.4.3	学習者の文章を対象にした処理	47
3.5	句 解 析	49
3.5.1	タスク概要	49
3.5.2	性能と実例	49
3.5.3	学習者の文章を対象にした処理	51
3.6	構 文 解 析	51
3.6.1	タスク概要	51
3.6.2	性能と実例	53
3.6.3	学習者の文章を対象にした処理	56
3.7	この章のまとめ	57
	章 末 問 題	58

## 4. ライティング学習支援

4.1	文法誤り検出と訂正	60
4.1.1	和文英訳における文法誤り検出/訂正	60
4.1.2	自由記述英文における文法誤り検出/訂正	67

4.2 キーワード推薦	90
4.3 この章のまとめ	100
章 末 問 題	101

## 5. リーディング学習支援

5.1 難解語の同定	103
5.2 難解語の言い換え	108
5.3 この章のまとめ	115
章 末 問 題	115

## 6. 教材作成支援

6.1 スラッシュ・リーディング教材生成	117
6.2 英語リズム学習用教材の生成	131
6.3 語彙問題生成	135
6.4 この章のまとめ	144
章 末 問 題	144

## 7. 学習者の能力/特徴の分析

7.1 言語能力の自動評価	146
7.2 学習者コーパスからの特徴表現抽出	159
7.3 母語干渉の分析	171
7.4 この章のまとめ	179
章 末 問 題	180



付 録 .....	181
A.1 コーパス処理のための便利なコマンド類 .....	181
A.2 語学学習に関する文献リスト .....	183
引用・参考文献 .....	184
章末問題解答 .....	200
索 引 .....	209

目次社

# 1

## イントロダクション

語学学習支援のための言語処理と聞いて、読者の方々はどのようなことを思い浮かべるであろう。本書を読み進めていけば自然と感じとれると思うが、語学学習支援のための言語処理は、他の言語処理の分野と比べると一風変わったところがある。おかしいたとえかもしれないが、右側通行の外国の道路を運転しているときに、ふと微妙な感覚を覚えることがある。ブレーキもハンドルもある。信号が赤になれば車は止まる。それでも、左折した途端、普段とは異なる風景が広がり、ちょっとした注意を迫られることになる。語学学習支援のための言語処理には、そんなところがある。

本章では、そんな一風変わったところがある語学学習支援のための言語処理を入口からのぞいてみよう。これから進む道筋の大まかな方向性を示し、道中の約束事や気を付けるべき点を述べようと思う。

### 1.1 語学学習支援の目的

本書を始めるにあたり、語学学習支援の目的を明確にしておこう。後述するように、語学学習支援では、他分野で扱うユーザ支援とは性質や目的が大きく異なる。そのため、語学学習支援の目的を初めに確認しておくことは有益である。

では、語学学習支援の目的とはなんであろう。語学学習の第一義的な目的は言語能力を高めることであろう。そうであれば、語学学習支援の目的は学習者の言語能力を高めるために支援を行うことにある（多くの場合、なんらかの情報を与えることになる）。これを支援対象の観点から大きく分類すると、学習者自身を支援する直接的支援と教育者などを支援する間接的支援の2種類がある。

このとき、他の支援と大きく異なるのは前者の直接的支援である。他分野の

## 2 1. イントロダクション

支援では支援対象にかかる負荷を極力減らすことを目的とする。一方、語学学習における直接的支援では、そのかぎりでない。むしろ、その逆である。なぜなら、適切な負荷なしには学習が行われないからである。このことは一般に、学習支援に当てはまる。

このことをより明確にするために、つぎのような具体的な例を考えてみよう。いま、語学学習支援として、英文誤り訂正を考える。また、他分野の支援として（英文誤り訂正と一見よく似た）英文校閲支援を考える。両者とも、与えられた英文中の誤り箇所を見つけ出し、訂正候補をユーザに与えるという点では似ている。しかしながら、後者の場合、仮に訂正候補の正しさが保証されていれば（すなわち完全な英文校閲支援システムが実現できれば）、ユーザは訂正箇所や訂正内容を吟味する必要はなく、英文の訂正は自動的に行ってよい。したがって、ユーザにかかる負荷はゼロである。これが理想的な英文校閲支援である。一方、これでは語学学習の支援とはならない。ユーザ（学習者）が訂正箇所や訂正内容を吟味しなければ、学習が行われないのはある意味自明である。言い換えれば、完全な英文校閲支援システムが実現したとしても、語学学習支援システムとしては不十分である。有効な語学学習支援となるためには、学習者になんらかの認知的負荷をかける必要がある。例えば、訂正箇所だけを提示し、どのように訂正すべきかを学習者に考えさせるなどの工夫が必要となる。このとき、「学習者に考えさせる」という部分が負荷となる<sup>†</sup>。

語学学習支援のための言語処理を始めたばかりの人が陥りやすい間違いは正にこの点である。よく見かけるのは、技術的に可能なことをすべて実現してしまい、（明示的ではないが）学習者へかかる負荷をすべて取り除くことを目指してしまう研究である。そうではなく、適切な負荷とはなにかをつねに念頭に置くことが大切である。場合によっては、技術的には支援可能でも、あえて支援しないという選択も必要になるかもしれない。逆に、本質的でない負荷は極

---

<sup>†</sup> ただし、これが「適切な」負荷かどうかは、また別の問題であり、慎重に検討する必要がある。ここで伝えたいのは、少なくとも負荷がかからなければ学習は行われない、ということである。

力減らすべきである。例えば、教材が整理されておらず、学習者が自分の能力に適したものを探すのに手間取るというのは、本質的な負荷ではないので取り除くべきであろう。

一方、間接的支援の場合、他分野の支援と同様にユーザ（例えば語学教師）の負荷を減らすことが目的となる。そのため、他分野の支援の考え方が基本的に適用可能である。ただし、2章で述べるように、語学学習支援で取り扱うデータは特殊である場合があるので、注意が必要である。

## 1.2 本書の目的と読者対象

本書の目的は大きく二つある。一つ目は、語学学習支援のための言語処理をこれから始める人向けに、必要となる知識や技術をわかりやすく解説することにある。そのような読者にとって本書がよい出発点となるよう心がけた。二つ目の目的は、関連する分野の人達に、語学学習支援のための言語処理の大まかな全体像を示すことにある。ここで関連する分野とは、言語処理だけではなく、語学教育、言語学、教育工学なども含む。

一つ目の目的を念頭において、本書は、言語処理の基礎的な事項の解説から始める。まず、語学学習支援で処理の対象となるデータについて詳細な議論を行う。つぎに、言語処理の要素技術のうち特に関連が深いものを導入する。これらの部分を読むことで、言語処理に馴染みのない読者でも本書が理解できるような構成とした。ただし、基礎的な事項とはいっても、膨大な知見があり本書ですべてをカバーすることは不可能であるので、ごく入門的な内容のみを紹介する。本書でカバーできなかった内容については関連する良書を紹介するにとどめる。その中でも、「自然言語処理の基礎」<sup>133)</sup>と「言語処理のための機械学習入門」<sup>168)</sup>は、本書を読み進める上で大きな助けになると思うので、適宜参照されたい。

第一の目的に関する読者対象は、情報科学に関連した学部学科の大学3年生以上の人達である。ただし、上述のとおり、言語処理の知識を前提とはしない。

必要となるのは、情報科学と数学（線形代数，微分積分，確率）のごく基礎的な知識である（数学については，高校までの知識で大部分を理解できる）<sup>†1</sup>。

言語処理の基本事項を説明した後は，語学学習支援のための言語処理における代表的なタスク<sup>†2</sup>を取り上げ，標準的な解法を解説する。その際には，語学学習支援のための言語処理とその他の言語処理における差異に注意し，語学学習支援に特有な問題をいかにうまく解決するかの説明に重点を置いた。両者の差異は，主に，語学学習支援の取り扱うデータの特殊性に起因する。例えば，非母語話者が算出する言語は，母語話者の言語とは異なる（典型例として，文法誤りがある）。そのため，通常言語処理では想定しない言語現象に対処しなくてはならない。また，母語話者が作成する文書でも，例えば教科書などでは，学習者向けに語彙や文法項目が制限されていることがあり，通常言語処理で想定する言語データとは性質が異なる場合もある。

二つ目の目的については，現状の言語処理技術で語学学習支援としてなにがどこまでできるのかを解説する。単に手法やアルゴリズムの性能を示すだけでなく，各タスクでできることがイメージしやすいよう，具体例を多く取り入れた。また，実用システムについてもできるだけ多く取り上げるようにした。これらの情報は，実際に語学学習支援システムの開発や学習者コーパスの分析を行う人に有益であろう。この部分については，情報科学と数学の知識を前提としない。

第二の目的に関する読者対象は，第一の目的における読者対象に加えて，語学学習に関係する分野の人達も含まれる。例えば，第二言語習得の研究者や語学教師が，実践，データ収集，データ分析を行う際に有益と思われる情報を取り入れている。

<sup>†1</sup> 情報科学と数学の知識がなくとも読めるよう極力簡潔にした。

<sup>†2</sup> ここで，タスクとは解くべき問題，課題というような意味である。例えば，学習者の書いた英文中の誤りを計算機で見つけ出すという問題（誤り検出）は，語学学習支援のための言語処理のタスクの一つである。

### 1.3 本書の構成と読み方

本書の構成は、図 1.1 に示すように、基礎、学習者支援、教師支援の三部に大きく分かれる。基礎では、まず、処理の対象となるデータを概観する。その後、文分割、形態素解析、構文解析など言語処理の要素技術を導入する。また、語学学習支援を対象とした場合、各要素技術においてどのような問題を考慮する必要があるかも説明する。つづいて、学習者支援と教師支援では、関連したタスクを章ごとにまとめた。基本的に、一つの節または項で一つのタスクを取り上げ、詳細に解説する。

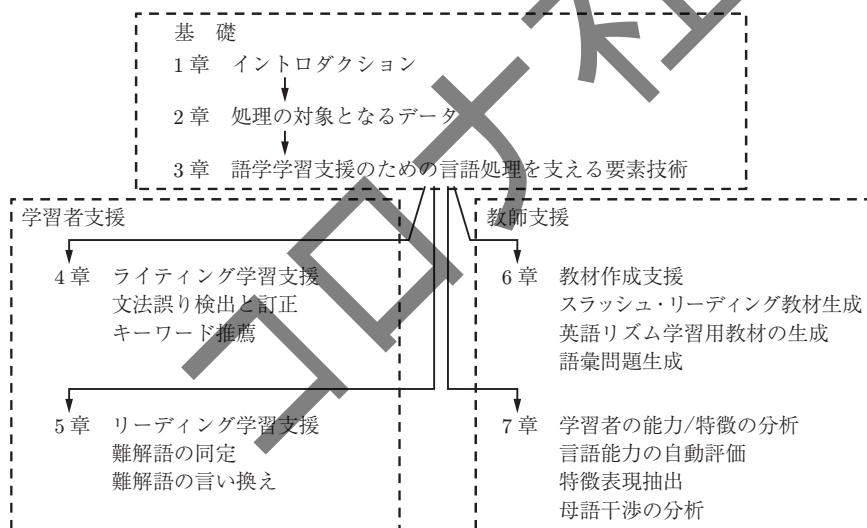


図 1.1 本書の構成

本書は、基礎、学習者支援、教師支援の順に通読することを念頭においているが、図 1.1 に示すとおり、必要な箇所だけを選択して読み進められるようになっている。例えば、基礎を読んだ後に、学習者支援を飛ばして教師支援を読むこともできる。同様に、基礎の後に、興味のある章もしくは節を読み進めていくことも可能である。また、言語処理の知識がある読者は、学習者支援また

は教師支援から始めてもよい。上述のどの読み方をしても、問題なく読み進められるよう工夫したつもりである。ただし、基礎では、言語処理の要素技術だけでなく、語学学習支援のための言語処理に特有な問題も取り上げたため、一度は目を通しておくことをすすめる。

基本的に、各節/各項目は

- (1) タスク概要 (目的, なにを対象とするのか)
- (2) 性能と実例 (どこまでできるのか)
- (3) 理論と技術 (どのように実現するか)
- (4) 実際の情報 (あれば, ティップスや注意点などを紹介)
- (5) 発展的内容と残された問題

という流れで進む。これは、読者ごとに必要な情報が異なる可能性を考慮したためである。例えば、研究者やシステム開発者は(1)~(5)のすべてが必要となるであろう。一方、第二言語習得/外国語習得の研究者や語学教師は、主に、(1), (2), (5)が興味の対象となるであろう。

本書では、(1)~(5)を通じて、わかりやすさを重視することとした。そのため、説明が幾分冗長なところもある。また、数学的な厳密さよりもわかりやすさを重視した部分もある。特に、(1), (2), (5)では、数式の使用は避けるようにした。

最後に、本書で使用する用語と表記に関する約束事について述べておきたい。語学学習支援のための言語処理において混乱を招く恐れのある用語がいくつかある。また、数式の表記にも注意を要するものがある。以下では、それらの使用についての約束事を述べて本章を締めくくることにしたい。

本書では、**学習**とは人間の学習を意味することとする。特に断らないかぎり、単に「学習」と表記した場合、人間の学習のことを指し、**機械学習** (machine learning) とは区別する (機械学習を意味する場合には「機械学習」と表記する)。また、機械学習という文脈では「学習」の代わりに「訓練」という用語を使用する。例えば、「分類器の学習」の代わりに「分類器の訓練」と表現する。同様に、「学習データ」の代わりに「訓練データ」を用いる。

機械学習では特徴 (attribute) および特徴量 (attribute value) という用語が用いられることが多い。しかしながら、本書では、言語処理の慣例に従い、素性 (feature) および素性値 (feature value) という用語を用いる (素性と素性値がどのような概念を表すかは文献168)などを参照のこと)。

本書では、第二言語習得 (second language acquisition) と語学学習 (language learning) という二つの用語を明確に区別することとする<sup>†1</sup>。第二言語習得とは、対象言語が日常生活で使用されている環境での学習を指すこととする。例えば、アメリカで英語の非母語話者が英語の学習を行う場合がこれに当たる。一方、語学学習とは、対象言語が日常的に使用されていない環境での学習を指すこととする。例えば、日本人が日本の学校で英語を学習する場合である。英語圏で英語を学ぶ場合と日本の教室で英語を学ぶ場合では、必要となる支援が異なることは想像に難くないであろう<sup>†2</sup>。

本書は、基本的に、語学学習の支援を対象としている (一部、母語の学習を取り扱う部分もある)。また、対象とする言語は主に英語と日本語である。特に、日本語を母語とする人が英語を学習する場合を想定する。これは、本書の読者に最も馴染みが深いと思われる二つの言語を選択した結果である。また、両言語とも語学学習のための言語処理で最も研究が進んでいる言語の一つである。

本書での数式の表記方法の基本はつぎのとおりである。文字の使い方としては、 $x$ ,  $y$ ,  $z$ などを優先的に変数として用いる。また、ベクトルは通常、縦ベクトルを想定し、小文字のアルファベットを太字で表記する (例： $\mathbf{x}$ )。行列は、大文字のアルファベットを太文字で表記する (例： $\mathbf{A}$ )。また、確率変数と集合は (太文字でない) 大文字のアルファベットで表す (例： $X$ )。確率変数と集合のどちらを表すかが文脈から読み取れないときは、導入の際に明確に言及する。それ以外の表記については、順次、各章で説明していく。

<sup>†1</sup> 語学学習の代わりに外国語学習という言葉が使われることもあるが、本書では前者を用いることにする。

<sup>†2</sup> このことを明示的に区別した言語処理の文献は少ないように思われる。第二言語習得として一括りで扱われることが多い。



# 索引

<b>【あ】</b>	<b>【け】</b>	
曖昧性 37	形式 (エッセイの) 153	素性値 7
アノテーション 11	形態素解析 43	<b>【た】</b>
アライメント 112	系列ラベリング 120	第二言語習得 7
<b>【い】</b>	検出精度 78	談話 154
言い換え精度 109	検出率 78	<b>【ち】</b>
言い換え率 109	<b>【こ】</b>	直接的評価 158
依存構造解析 51	構造 (エッセイの) 153	<b>【と】</b>
一致率 24	構文解析 51	統語解析 51
意味の保持 109	コーパス 10	到達度評価 158
<b>【え】</b>	語学学習 7	トークン 41
エッセイ 10	語義曖昧性解消 139	トークン同定 41
——の自動採点 147	コロケーション 140	<b>【な】</b>
<b>【か】</b>	混同行列 78	内容 (エッセイの) 153
回帰問題 150	<b>【し】</b>	難解語 103
学習 6	自由記述作文 9	難解語同定 104
学習者コーパス 11	熟達度評価 158	<b>【は】</b>
格フレーム 87	主名詞 74	パーズング 51
隠れマルコフモデル 45	条件付き確率場 45	<b>【ひ】</b>
過剰使用 160	<b>【す】</b>	<b>【ひ】</b>
間接的評価 158	スラッシュ・リーディング 117	品詞解析 43
<b>【き】</b>	<b>【せ】</b>	品詞タグ 43
機械学習 6	正解率 79	<b>【ふ】</b>
技巧 (エッセイの) 153	潜在的意味インデキシング 156	複合語 107
強勢タイミング言語 131	<b>【そ】</b>	フラグメント誤り 53
<b>【く】</b>	総合的評価 147	プロファイル 16
句解析 49	組織化 (エッセイの) 153	分析的评价 147
句構造解析 51	素性 7	文分割 36
		文法誤り検出 67
		文法誤り訂正 67

文法誤り検出/訂正	60, 67
文法性	109
文末記号	36
<b>【へ】</b>	
平易さ	109
編集距離	29, 81

<b>【ほ】</b>	
補完類似度	163
母語推定	178
<b>【み】</b>	
未知語	28

<b>【も】</b>	
モーラタイミング言語	131
<b>【わ】</b>	
分かち書き	43

<b>【A】</b>	
accuracy	79
achievement assessment	158
alignment	112
ambiguity	37
analytic score	147
annotation	11
automated essay scoring	147
<b>【B】</b>	
basic universal description	63
bi-gram	40
BUD	63
<b>【C】</b>	
case frame	87
chunking	49
collocation	140
comma splices	39
complementary similarity measure	163
complex word identification	104
conditional random field	45
confusion matrix	78
corpus	10
CRF	45
CSM	163
CWI	104

<b>【D】</b>	
dependency parsing	51
discourse	154
<b>【E】</b>	
edit distance	29
essay	10
<b>【F】</b>	
<i>F</i> 値	79
false negative	78
false positive	78
feature	7
feature value	7
<i>F</i> -measure	79
<b>【G】</b>	
grammaticality	109
grammatical error correction	67
grammatical error detection	67
grammatical error detection and correction	60, 67
<b>【H】</b>	
head noun	74
hidden Markov model	45
HMM	45
holistic score	147
<b>【I】</b>	
IDF	97

inter-annotator agreement rate	24
inversed document frequency	97
<b>【K】</b>	
$\kappa$ 統計量	24
<b>【L】</b>	
language learning	7
latent semantic indexing	156
learner corpus	11
Levenshtein distance	29
LSI	156
<b>【M】</b>	
machine learning	6
meaning preservation	109
mora-timed language	131
multi-word expression	107
<b>【N】</b>	
native language identification	178
NLI	178
<i>n</i> -gram	40
<b>【P】</b>	
parsing	51
part-of-speech tagging	43
phrase structure parsing	51
POS タグ	43
POS tagging	43

precision	78	second language acquisition	7	true positive	78
proficiency assessment	158	sequence labeling	120	<b>[U]</b>	
profile	16	simplicity	109	uni-gram	40
<b>[R]</b>		stress-timed language	131	unknown word	28
recall	78	<b>[T]</b>		<b>[W]</b>	
<b>[S]</b>		tokenization	41	word sense disambiguation	139
S 式	52	tri-gram	40	tion	
		true negative	78		

日本語学

—— 監修者・著者略歴 ——

奥村 学 (おくむら まなぶ)	永田 亮 (ながた りょう)
1984年 東京工業大学工学部情報工学科卒業	2000年 明治大学理工学部電気工学科卒業
1989年 東京工業大学大学院博士課程修了 (情報工学専攻) 工学博士	2002年 三重大学大学院博士前期課程修了 (情報工学専攻)
1989年 東京工業大学助手	2005年 三重大学大学院博士後期課程修了 (システム工学専攻)
1992年 北陸先端科学技術大学院大学助教授	博士 (工学) (三重大学)
2000年 東京工業大学助教授	2005年 兵庫教育大学助手
2007年 東京工業大学准教授	2007年 兵庫教育大学助教
2009年 東京工業大学教授	2008年 甲南大学講師
現在に至る	2012年 甲南大学准教授
	2017年 理化学研究所客員研究員兼任 現在に至る

## 語学学習支援のための言語処理

Natural Language Processing for Language Learning Assistance

© Ryo Nagata 2017

2017年11月10日 初版第1刷発行

検印省略

監修者 奥村 学  
著者 永田 亮  
発行者 株式会社 コロナ社  
代表者 牛来 真也  
印刷所 三美印刷株式会社  
製本所 有限会社 愛千製本所

112-0011 東京都文京区千石 4-46-10

発行所 株式会社 コロナ社  
CORONA PUBLISHING CO., LTD.

Tokyo Japan

振替 00140-8-14844・電話 (03) 3941-3131(代)

ホームページ <http://www.coronasha.co.jp>

ISBN 978-4-339-02761-7 C3355 Printed in Japan

(金)



JCOPY < 出版者著作権管理機構 委託出版物 >

本書の無断複製は著作権法上での例外を除き禁じられています。複製される場合は、そのつど事前に、出版者著作権管理機構 (電話 03-3513-6969, FAX 03-3513-6979, e-mail: info@jcopy.or.jp) の許諾を得てください。

本書のコピー、スキャン、デジタル化等の無断複製・転載は著作権法上での例外を除き禁じられています。購入者以外の第三者による本書の電子データ化及び電子書籍化は、いかなる場合も認めていません。落丁・乱丁はお取替えいたします。