新しい数理・データサイエンス教育のために

一 人文科学研究のための LaTex2e による多言語処理入門

Toward New Mathematical and Data Science Education Program:
Multilingual Processing in LaTex2e for Humanities Research

石井正人

Masato ISHII

1. 始めに

2016年に文科省が策定した「大学の数理・データサイエンス教育強化方策」¹⁾に基づき、「数理・データサイエンス教育」のカリキュラムや教材開発を推進する東京大学など基幹6大学による「数理・データサイエンス教育強化拠点コンソーシアム」²⁾が創設され、さらに2019年に「大学における数理・データサイエンス教育の全国展開」のために、千葉大学を含む協力校20校が選定され³⁾、これらの大学において数理・データサイエンス教育のカリキュラムを先駆的に実施していくことになった。これに基づき、2020年度から先行的に千葉大学でも「普遍教育」(一般教養教育)が改組され、「数理・データサイエンス科目」が専門分野を問わず全学生に対して必修化される⁴⁾。千葉大学には元々、数学・統計学の専門教員によって豊富な数理・データサイエンス教育が行われているが、さらにこれを拡充し、これまで数理・データサイエンス教育に余り積極的に参加してこなかった専門分野の学生をも受け入れる体制を整えた。

数理・データサイエンス教育は統計学などの数理工学が中心になるが、数理・データサイエンス教育を強化し、裾野を広げるためにも、数理工学に加えて、コンピュータサイエンスの概略や、認知科学の諸分野、情報リテラシーとアカデミックスキルなど広い領域を統合するのが理想である。高校における「情報科」の内容もそのような高い理想によって構築されており5、大学における数理・データサイエンス教育もこれと密接に連動することが求められる。

数理科学を専門とするわけでもない私が、千葉大学における新しい数理・データサイエンス教育の一端を担おうとするのは、上のような広い文脈の中でなら、自分の経験を活かし、主に人文科学を専門とする自分の学生の役に立てるかもしれないと考えたからである。歴史言語学や中世ヨーロッパ文献学といういかにも古き良き時代の純然たる人文科学の研究分野を専門とする私も、LaTex2eを使ってサンスクリット語やドイツ語史の教材を開発し

¹⁾ https://www.mext.go.jp/b menu/shingi/chousa/koutou/080/gaiyou/ icsFiles/afieldfile/2016/12/21/1380788 01.pdf

²⁾ http://www.mi.u-tokyo.ac.jp/consortium/

³⁾ https://www.mext.go.jp/b menu/shingi/chousa/koutou/095/gaiyou/1412367.htm

⁴⁾ 概要はなお未公表。

⁵⁾ 水越敏行・村井純・生田孝至『新・情報の科学』平成31年1月、日本文教出版。

た経験からコンピュータサイエンスの成果(というには恥ずかしいほどの末端)の恩恵に浴 し、その有り難みを痛感した経験を持つ。

主に人文科学で直面する多言語処理の諸問題に、これほどスピーディーで美しいソルーションを提供してくれた人々に対し、いくら感謝しても足りない。とうに諦めていたデーバナーガリー文字における合字(ligature)や60、古代ギリシャ語におけるアクセント類(iota subscriptum も含めて70)の処理にこれほど簡便な処理方法が実現されるとは夢にも思わなかった80。この喜びは、インターネットと文書・書籍等の電子化が急速に進み、写本や古い印刷本の膨大な資料に、日本の自室に居ながらにして、いくらでも簡単にアクセスできるようになったことへの感激と相等しい。更にまたこれは、浩瀚な辞書類が、例えば数十巻に及ぶグリムの独独事典がネット上で利用可能となり90、分厚い Liddell/Scott のギリシャ語辞典、Lewis/Short、Georges、Gaffiot のラテン語辞典、Monier、Böthlingk のサンスクリット辞典等が、そしてその上 Du Cange の中世ラテン語辞典までもが電子辞書アプリとなり、小さなタブレットー台でどこでも簡便に利用できるようになった随喜の思いと同じである。広い意味でのコンピュータサイエンスやデータ通信技術の発達のおかげで、教育研究のインプットにおいてもアウトプットにおいても、人文科学は現在、古今未曾有の楽園状態にいると言っても過言ではない。

このような有益な技術的発展の成果を少しでも多くの人文科学専攻の学生が活用し、人 文科学の発展だけでなく、コンピュータサイエンスやデータサイエンスの更なる発展に寄 与できるようになるのが私の願いであり、千葉大学における新しい数理・データサイエンス 教育への期待である。

2. Word process ということ

何十年も前に私などが学生の頃にコンピュータサイエンスの初歩を学んだ時には、コンピュータの働きに calculation process と word process という 2 つがあると教わった記憶がある。符号化した電子信号の並びが、演算命令なのか、そのままの文字列なのか、それぞれで処理の仕方が異なる。今日プログラミングのレベルではなく日常的にこのような問題に遭遇するのは、Excel でセルに打ち込んだ文字が、そのまま文字列なのか関数なのか区別する必要が生じたときくらいであろうか。

1980年代からしばらく、日本ではワードプロセッサが発展し、普及した。コンピュータのワードプロセスにのみ特化したマシンであった。同じ頃私はドイツに留学したが、使われ

⁶⁾ त + व = で tva 等。

⁷⁾ ἦ, "A 等

⁸ 印刷・編集サイドの要求からこの原稿はやむなく Word で書いているが、上の注二つを見ると Word/IME でデーバナーガリー文字の合字も古代ギリシャ語の複雑なアクセントも処理できているではないか、何を苦しんでわざわざ LaTex2e を使わねばならないのかという疑問を持たれる向きもあるかもしれない。しかしこれらは筆者が、特殊文字から 1 文字ずつ呼び出したりするなど、かなりの手間をかけて処理している。テクストベースで処理できる tex の方が格 段に効率的である。

 $^{^{9)}\;\}underline{\text{http://www.woerterbuchnetz.de/cgi-bin/WBNetz/setupStartSeite.tcl}}$

ているのはコンピュータの一本槍で、ワードプロセッサという概念すら通じなかった。日本 語が諸言語と比較しても文字の種類が多く、日本語話者は文字への関心が高い。それゆえに 日本で起こりえた技術史上のエピソードであったろう¹⁰。

今日ではパソコンに Word などのワープロソフトを走らせてワードプロセスを利用するのが当たり前になった。けれども、本来表計算ソフトであるはずの Excel が日本では「公式書類作成ソフト」として役所や企業でゆがんだ広まり方をして生産性を下げていることが問題になったが¹¹⁾、これも私には、ワードプロセスに対する日本人の関心と要求の高さが現れた現象であると思える¹²⁾。

ワードプロセスというもの、そしてそれを体現するワープロソフトというものは、ただ文字を書くだけでなく、私見によれば少なくとも次の10の機能を果たしている。

- ① 文字コード Character Code
- ② 組版 Typesetting
- ③ 割付 Text Layout/Float
- ④ 章立て Thesis Composition/Table of Contents
- ⑤ 検索(索引・脚注・参考文献一覧)Retrieval/Index, Footnote, Reference
- ⑥ 校正 Proofreading
- ⑦ 校閱 Review and Emendation
- ⑧ ゲラ刷りとモニタリング Proof Print/Monitoring
- ⑨ 印刷 Printing
- ⑩ 製本 Bookbinding

何百年もかけて発達してきたこれらの技術が集大成される形で、DTP (Desk Top Publishing) が生み出された。DTP が当たり前になる前には、原稿を手書きで書くことが当たり前で、その後その手書き原稿を元に、印刷物ができあがるまでの編集・校正・印刷・製本のそれぞれのプロセスに、高い専門技術と、膨大な機械設備と人員をかけていた時代が長くあったわけである。これらが急速にコンピュータによって一元化され、劇的なコスト削減を実現していく技術発展の歴史を体験した筆者などにしてみれば、これだけの「本作り」のプロセスがまさか小ぶりなパソコン一台とプリンタにまで集約されるとは想像もしていなかった。技術といえども歴史的規定性があるから13、手動タイプライターなど見たことも

^{10) 2003} 年にシャープがそのワープロ「書院」シリーズを生産中止にし、最終的にワードプロセッサの時代が終わった。https://jp.sharp/support/word/info050901-2.html

¹¹⁾ いわゆる「神エクセル問題」。三重大学名誉教授の奥村晴彦氏の論考を参照されたい、https://oku.edu.mie-u.ac.jp/~okumura/SSS2013.pdf。「神エクセル問題」が無駄な時間を取り、日本の生産性を下げているという問題提起を受け、自民党の河野太郎議員が行革推進本部から文科省に対し、科研費申請書類など公式書式に使っている現状を改めるように指示した、https://twitter.com/konotarogomame/status/793699820712054784

¹²⁾ ワープロソフトの世界標準は Word になったが、日本語日本文学研究や漢文研究におけるワードプロセスに対する要求に応える水準は、「一太郎」に遙かに及ばない、というのが筆者の個人的感想である。

¹³⁾ 良い例がキーボードの配列、いわゆる qwerty 配列である。安岡孝一の説明が簡にして要を得ている。

ない現代の若い学生たちであっても今後のイノベーションを担うために、この技術史の流れを理解しておく必要があるであろう。特に日本語を扱う市場では、ワードプロセスのこれらの諸機能に深い関心と強い要求がもたれているので、なおさらのことである。

それぞれに問題点を挙げておきたい。

3. Word process の諸機能における教育課題

① 文字コード Character Code¹⁴⁾

文字形の標準化・規格化ということの歴史的経過、教育的課題、メディア文化的意味についてこれからのグローバル化時代にはますます理解しておく必要がある。Unicode(UTF-8)によって、例えば漢字の字形は繁体・簡体・日本式とも収録され、どれがスタンダードとなって他を駆逐するという恐れはなくなった。しかし技術的に解決されたとしても、個々の使用者が場面に応じてこれを使いこなせる必要は依然として残る。「亞細亞・亚細亚・亜細亜」「勞動・劳动・労働」といった3種の漢字字体の区別は大きな問題となり得る。当然のことであるが、台湾だけでなく韓国においても繁体しか認められない150。しかも上の「労働」の例から分かるように、漢字の使用法そのものにまで問題は及ぶ。

いわゆる「幽霊文字」¹⁶⁾については、文字コード上は重要な問題であるにしても、 よほどの専門家・好事家でない限り詳しく知っている必要はないかもしれない。 しかしフォントの問題は一部ここに含められるだろう。

フォントが商業デザインの一部であって、第1に有料の商品であるということと(現在無料で使用可能なのは著作権が切れたものである)、第2に時代ごとの流行や意味付けがあるということを自覚する必要がある¹⁷⁾。このためには、例えば、日本語組版と日本語フォントの歴史を担ってきたといっても良い「株式会社モリサワ」の沿革¹⁸⁾

18) https://www.morisawa.co.jp/about/history/

http://www.asahi.com/special/kotoba/archive2015/moji/2011081800015.html.

4

https://kotobaken.jp/qa/yokuaru/qa-35/。人間工学的研究成果に基づいて決定された合理的なものでは決してないようだが、今更変えられないし、変えた場合のコストの方が高そうなので、積極的な改良提案も普及しない様子である。 14)深沢千尋『文字コード「超」研究 改訂第 2 版』(2011 年、ラトルズ)。矢野啓介『[改訂新版]プログラマのための文字コード技術入門(WEB+DB PRESS plus シリーズ)』(2018 年、技術評論社)。

¹⁵⁾ それどころか韓国においていわゆる「正字」や「異体字」が正式とされていて、繁体字より難しい場合がある。たとえば、繁体・簡体・日本式=青、韓国式=青。繁体・日本式=並、簡体=并、韓国式=並、等。

^{16) 「}JIS 第 1・第 2 水準(JIS X0208=狭義の「JIS 漢字」)の中には、なぜこの字が入ってしまったのか、どういう字なのか分からない、というものがあります。」比留間直和「幽霊、ところにより出ます — 変換辞書のはなし4」 http://www.asahi.com/special/kotoba/archive2015/moji/2011072700001.html、

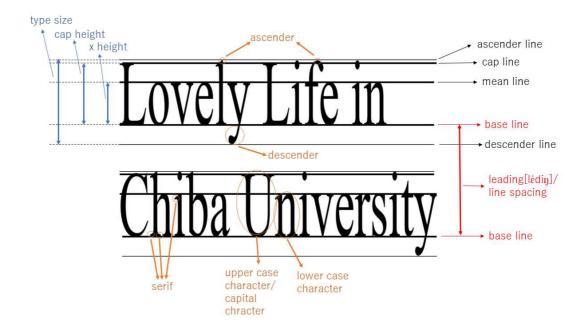
[「]幽霊文字の読み、どこから - 変換辞書のはなし5」

マニア向きであって、本質的ではない問題だと思われるかもしれないが、貴重な番号がいくつか「無駄に」割り当てられているだけでなく、JIS 漢字の選定過程自体を見直して、修正する、という日本の国語政策の根幹に係わる問題でもあるのである。

ID 字体の文化史的意味についての研究はまだ少ない。ドイツではフラクトゥーアとナチスプロパガンダとの関係についての研究があるようだ。以下の論文はこのテーマについての日本における数少ない成果の1つである。深井里奈子「【研究ノート】ドイツにおける筆記体の変遷と手書き文字に求められる役割—初等学校の書字教育を参考に一」、千葉大学大学院人文社会科学研究科編『千葉大学大学院人文社会科学研究科研究プロジェクト報告書』299 巻 86-91 頁(2016 年 2 月)、https://opac.ll.chiba-u.jp/da/curator/100347/BA31027730 299 p086 FUK.pdf。

を概観してみせること(写植とは何かということの説明も含めて)自体に意味があると思われる。

また漢字書体が元来「筆遣い」を元に作られているもので、「止め払い」の問題の理解もそこを基礎にする。欧文書体も同様に手書きのカリグラフィーを基本にして作られている。これはさらにタイポグラフィーでも重要な指針として引き継がれる。ベースラインの重要性から、セリフやループがなぜ存在するのかとか、イタリック体と斜体とは本来異なるというようなことの理解も同じく手書きの感覚を基礎にする。



今日ではほとんどの学生がアルファベットの筆記体を習得せず、使用する習慣もないし¹⁹⁾、そもそもこのITの時代に筆記体を身につける必要があるかという議論もあるようだ。それももっともだと筆者も考えるが²⁰⁾、欧文書体をより深く理解するため

¹⁹⁾ 現行の「中学校学習指導要領」の第 2 章「各教科」、第 9 節「外国語」、第 2 「各言語の目標及び内容等」、「英語」、3 「指導計画の作成と内容の取扱い」の、

^{「(1)} 指導計画の作成に当たっては、次の事項に配慮するものとする。」の項に、

[「]エ 文字指導に当たっては、生徒の学習負担に配慮し筆記体を指導することもできること。」とあって、筆記体は必修ではないとされている。https://www.mext.go.jp/a_menu/shotou/new-cs/youryou/chu/gai.htm。

これは平成 10 年(1998 年)告示、平成 14 年(2002 年)施行の学習指導要領から始まったことのようである。

https://www.mext.go.jp/a menu/shotou/cs/1320086.htm.

²⁰ 関連するので、学生の欧文手書き文字に関しては次のような問題をこの場で指摘しておきたい。試験の答案は今日でも手書きであることが多いので、ベースラインを理解していない答案は、g や y がベースラインの上に乗るなど、見た目が異様なものになる。また、キャピタルラインとミーンラインを理解せず、K と k といった大文字と小文字の区別を意識して書き分けた様子がない答案も多い。後者は学校文法に違反し、減点の対象となる。日本人学生の手書きアルファベットに大きな問題が現れていることに外国人教師たちが気付いていて、日本人の外国語習得に影響しているのではないかという観測もあるほどである。以下の書を参照のこと:林 メーナー エルケ『アルファベットの正しい書き 方――ドイツ語を例にとって』上智大学出版、2008 年。――アルファベットの正しい書き方などと難しいことを言わず、そんな些末な問題にこだわらず、要するに通じれば良いのだ、という立場が多いとは思うが、それで全て済むなら楽な話で、文字コードに関する技術的諸問題などは最初から存在しないのである。悪筆な筆者がこう主張しても説

には最低限の知識は必要であろうと思う。

- ② 組版 Typesetting²¹⁾
- ③ 割付 Text Layout/Float

宮沢賢治『銀河鉄道の夜』おいて、ジョバンニが「活版所」で働く描写を理解することが今の若い人にはかなり難しくなっている。額に汗して印刷物を作り出す労働の意味がこの作品の中でしめる位置は小さくないと思うが、それは適切な解説があれば良いことだ。(ますむらひろしのマンガ版22)ならより理解が早かろう。)

肝心なことは、コンピュータによる文書作成において、思い通りの文字を一つ一つ画面 (紙面) に置いていくというプロセスと、テクスト全体のレイアウトのプロセスは、不可分のものであるけれども、本来独立したプロセスであるということを理解することである。(別のプログラムになる)

欧文レイアウトで問題になるのは、アルファベット活字の横幅がそれぞれ微妙に違 っているということだ。これをそのままにする(いわゆる「ベタ組み」)と、単語の見 た目に影響し、ひどい場合には単語の切れ目さえ不分明になりかねず、さらにジャス ティフィケーションができなくなる。この微妙な字間を調整するために、カーニング (kerning)²³⁾という高度な技術が必要であった。タイプライター機においては、均質な 歯車の動きで全ての印字を送っていかねばならないから、今日タイプライター書体と 呼ばれる、あの特別な書体、全てのアルファベットに同じ幅を与える書体を考案した。 読みづらいものであるが、20 世紀半ばくらいまでの公文書がみなタイプライターに よって作られ、この独特な書体で書かれているため、時代的・文化的な記号性を帯び るに至った。今日でもタイプライター書体がわざわざ使われることのあるゆえんであ る。古いワープロソフトではこの点が上手く処理できず、全ていわゆる「半角文字」 に作り替えた奇妙な書体で切り抜けていた。タイプライター書体の劣化版とでも言え るが、今日でも緊急避難で使用可能な方法ではある。他方で、和文の中に欧文書体を 入れると、ジャスティフィケーションが効かなくなったこともあった。これらは今自 動的に調整してくれるが、多言語処理の場合、いつ壊れるかも分からない。読みやす い組版という問題はどうやって作り出されたか、活字のレベルから理解しておくのは 無駄ではないであろう。

得力がないのは承知しているが。

²¹⁾ ジェイムズ・クレイグ (組版工学研究会監訳)『欧文組版入門』(朗文堂、1989年)

 $^{^{22}}$ マンガ家のますむらひろしが、宮沢賢治『銀河鉄道の夜』の 35 年ぶり 3 回目のマンガ化「銀河鉄道の夜 最終形前編」を 2020 年新年号から「しんぶん赤旗日曜版」に連載している。その第 $6\sim10$ 回では、山形県米沢市の伊藤活版所がクレジットされ、これまで詳しく描かれてこなかった活版所の内部と作業が描写されている。

また youtube 等にも、古典的な本作りを紹介する教育用動画がアップされている。

²³⁾ Adobe 社のタイポグラフィーに関するチュートリアルページにある以下の記載がとても参考になる。

https://www.adobe.com/jp/print/tips/indesign/category9/page3 1.html.

ワープロソフトにおける「レイアウト」や「ページ設定」関連の指定には、伝統的な組版や割付の用語が今日でも使用されている。紙の寸法のA列・B列等の指定から、「段組」や「ヘッダ、フッタ」、「禁則処理」、「縦中横」、「回り込み」など。一度でも元の意味に立ち返り、組版を理解することは無駄では無いと思う。

特に欧文の組版では、「ハイフネーション」の理解が必要である。筆者が経験した範囲であるが、高校で「ハイフネーション(分綴法)」を教わった学生に出会ったことはない²⁴。今日ではワープロソフトの性能が上がり、場合によっては英語以外の分綴をも自動的に行ってくれるので、分綴法について詳しく知る必要はないという意見もあり、筆者ももっともだと思う。手書きで間違いを犯すことに注意すれば十分であるかもしれない²⁵)。ただ、和文欧文混在の組版では、自動的なハイフネーションが効かないことが多いので注意を要する。失敗すると、特に縦組だと悲惨な結果になる。校正の段階でこれを手作業で発見・修正していくことになると、ハイフネーションの知識は不可欠となる。

④ 章立て Thesis Composition/Table of Contents

ワープロソフトにおいて、階層化して箇条書きや段落分けや章立てを行う機能が発達し、かつそれを「目次」として一望できるようにする機能も進んだ。こんなことが特別な「機能」であると、ここで論じること自体が奇異に思われるかもしれない。それほど「章立て」とその管理はワープロソフトにおいて当たり前の機能になっている。しかし章立てについては、次の2つの面において、固有の課題を有する独立した問題として理解する必要がある。

第1に、章立てとナンバーリング等は、法律などの公文書や取扱説明書・仕様書等 の体裁を決める「国際標準」に関わる問題である、ということである。

章立てのナンバーリングの方法は、俗に「公用文方式」と「理科系方式」²⁶⁾と呼ばれる 2 つが行われている。

「公用文方式」とは、章立ての階層が変わる毎に、番号表記を変え、文章構成を明確にしようとするものである。現代日本の文章におけるこの方式の根拠は、内閣官房長官による依命通知『公用文作成の要領〔公用文改善の趣旨徹底について〕』(昭和27年4月4日 内閣閣甲第16号依命通知)にあり27、その中で次のように指示されて

²⁴ 見落としかもしれないが、「学習指導要領」の中にも、「分綴法」「ハイフネーション」への言及は見当たらないようである。

²⁵⁾ 注 20) で指摘したのと同様、手書きの試験答案で分綴法に違反すると、正書法に違反したのに準じ、広義の文法の誤りであるから、減点の対象となる。

²⁶⁾ 確証はないが、この名称は次の書の影響の元にできたもののようである。木下是雄『理科系の作文技術』中公新書624、1981年。この書は今日でもアカデミック・ライティングの指導にとって、「理科系」「文化系」を問わず必須である。

²⁷⁾ 文化庁 HPの「国語施策・国語教育、公用文に関する諸通知」に全文が公開されている。常用漢字の改訂等で一部変更はあったが、全体として現在も効力のある文章である。

いる:



「理科系方式」とは、章立ての階層関係を 1.2.1, 1.2.2.1 等という方法で、全ての箇条番号において明示するものである。この方式の根拠は、ISO 2145:1978 (Second Edition) "Documentation – Numbering of divisions and subdivisions in written documents"28) である。

すなわち、日本の公用文においては、「理科系方式」の章立てナンバーリングは未だに推奨されていないのである。しかるに「公用文方式」は直ぐに分かる限界がある。「アイウ」や「ABC」などの文字を使っていると、数に限りがあるので行き詰まる可能性がある、ということである。筆者の知る例では、あるドイツの古い詳しい文法書が、ギリシャ文字や aaa/bbb/ccc など並べてみるナンバーリングでも足りなくなり、ヘブライ文字まで動員していた。その点「理科系方式」にこのような限界はないが、たとえば第6階層まで数字が並んでいると、単なる数字の並びから構造の直感的な理解はかなり難しい。どちらの方式にも一長一短あるということである。

第2に、章立ては文章の内容的にも文章構成を考える上で重要な問題であるから、 アカデミック・ライティングの授業において重要なテーマである。

今ではワープロソフトで章立てを行うと、自然に「目次」が作られ、文章作成の作業に対応して、「目次」自体も改編されていく。しかもその「目次」において、大項目 (大きな見出し)の中に小項目が「たたみ込める」ようになっていて、大項目だけで構成を考えたり、たたんでいた小項目をクリック一つで「開き」、その都度立ち入って

https://www.bunka.go.jp/kokugo nihongo/sisaku/joho/joho/kijun/sanko/koyobun/index.html.

²⁸⁾ この文章は ISO の HP からダウンロード可能ではあるが、全文 2 ページながら高額の有料の文書であり、利用規約も厳しいので、引用紹介ができない。本稿執筆のために筆者は JSA を通じて私費で購入の上閲覧した。2.1.1.は"two, one, one"と読めと、"spoken form" まで指定してあるのが印象的であった。この文書の趣旨は、企画書の作成に関する ISO/IEC の規則集 "Principles and rules for the structure and drafting of ISO and IEC documents" (ISO/IEC Directives Part 2) https://www.iec.ch/members experts/refdocs/iec/isoiecdir-2%7Bed7.0%7Den.pdf と、それを受けた JIS 8301 に生かされているということである。これらとは別にアメリカ合衆国国防総省の物資調達に係わる標準「MIL 規格」のなかの文書作成標準や、アメリカ合衆国政府の Printing Office Style Manual 等があるそうだ。次の論文を参照した。貝島良太「取説やビジネス文書の項目番号の付け方--数字の他、アルファベット、片仮名などは?-」(TCシンポジウム 2017 発表) http://www.bow-wow.jp/sht3/pdf/028 symposium2017(R2).pdf。

一章だけの全体を参照したりできるのである。

これは決して当たり前の機能ではない。章立てや目次の必要性を「ワープロソフト 以前」から理解し、ワープロソフトにも章立てや目次を支援する機能を求めた人々が いたからこそ生まれ出た機能なのである。文章を作成するに場合でも、読み取って活 用する場合でも、章立てや目次が果たす役割の重要性を理解していて始めて、このよ うな機能と技術が発達した。(ろくな章立ても目次もなく、索引もない大著というも のがしばしば存在するのである。)

この機能は、当然のこととしてアカデミックスキルの中のライティングの中で重要な位置を占める、文章構成の技能と密接に関連する²⁹⁾。アウトライン・プロセッシングという作業効率化の重要なテーマである³⁰⁾。その機能がなぜ・どのように生まれてきたのか、その機能をどう使うか、という点を合わせて学ばないと、技術の習得は進まない。こうしてワードプロセッシングの技術は、結局のところ言葉の技術、ライティングやリテラシーを含むアカデミックスキルへとつながっていく。

⑤ 検索(索引・脚注・参考文献一覧)Retrieval/Index, Footnote, Reference

学術テクストに対して、これを読む側に検索をかける要求があり、この検索の要求 に応えるために、他方で書いた側に索引・脚注・参考文献一覧(そもそも目次がここ にも含まれるが)を付ける義務が生まれる。書いた側も直ちに読む側に回るのが学術 テクストを生産・流通・消費活用する循環の習いであるから、検索要求と索引(注・ 文献一覧)作成義務とは表裏一体となる。

テクスト検索も、索引・注・参考文献一覧作成といった検索関連機能も、それを使 う必要と経験があって、初めて使いこなすことができる。引用形式や注の付け方、文 研リストの作り方は論文の書き方と共に学ぶべきことで、専門分野ごとに微妙に違う 作法を後からいくらでも学ぶことができるが、そもそも検索機能が何のためにあるか、 注はなんのために付けるのか³¹⁾は、アカデミックスキルとして理解しておく必要があ る。

この検索機能が、作成中のテクストそのものに「自己言及」されることで、校正・

²⁹⁾ アカデミックな文章というものは、論点を立てて、論理で結んで書くものだ、という基礎的なことを改めて理解して身につけることが、多くの学生にとってはかなり難しい課題になるということを、リテラシーの授業を行いながら筆者も経験した。高校でも大学でも小論文やライティングの授業において箇条書きの答案を禁ずる指導があるということを、リテラシーの授業を担当するようになってから初めて知った。そのように徹底しておかないと、箇条書きしか書けない・書かない生徒学生が大量に出る、という理由らしい。しかし最初から何と何を書こうとしているのか自分の論点が立っていないようではライティングが始まらない。論理の流れを構築するためにいったん箇条書きにしてみるという、高等教養教育における新たな箇条書きの意味を指導する必要を感じている。

³⁰⁾ このような「目次」を構成できるアウトライナーというアプリも存在するほどである。Tak.著『アウトライナー実践入門―「書く・考える・生活する」創造的アウトライン・プロセッシングの技術』技術評論社、2016 年。
31) 注というものは、第 1 に必要や興味のある読者が誰でも確認できるように出典・論拠を示すために(科学としての再現性を確保するために)付けるものであり、第 2 に本文の論点が錯綜してきた場合に、本文には主要な論点を残し、重要な関連テーマだが、本文の流れからは副次的な論点を簡略に扱うために(科学としての論理性・明晰性を確保するために)付けるものだ、と筆者は学生時代に教わり、今も学生にそのように指導している。

校閲機能へと発展していく。

- ⑥ 校正 Proofreading
- ⑥ 校閱 Review and Emendation
- ⑦ ゲラ刷りとモニタリング Proof Print/Monitoring

組版されたテクストが本当に原稿通りかどうか、DTPが普及するまでは、ゲラ刷り(見本刷り・試し刷り)というもので実際に印刷してみなければ確認できなかった。今日でも紙媒体で出版する場合には、このゲラ刷りによる校正プロセスは十分に生きている。ワードプロセスの発達と、画面によるモニタリングの技術(WYSIWYG)が進んだから、コンピュータによる文書作成では、文書作成のプロセスがリアルタイムで確認できるようになった。その場で筆者自ら校正・校閲可能である。集団的な校正・校閲が可能になったのも大きな発展である。

これも当たり前のようでありながら、まだ使いこなせていない事例を目の当たりにする。特に誰かの文案を元に(俗に言う「たたき台」)、委員会など集団で文案を練っていく場合である。ヴァージョンの違いを明示しない草案が何度も印刷され、配布され、作業が混乱し、生産性が低下する。残された資料も整理が難しく、信頼度が低くなるため、仕事の継続性も確保できない。本来清書用であるプリンタによるハードコピー出力を乱用すると、作業効率がかえって下がる場合がある。モニタ画面ではどうしてもテクスト確認作業ができず、紙媒体に印刷した「ゲラ刷り」でないと校正が進まない、という紙媒体依存体質は世代を追うごとに「改善」されているだろうか32。実際に校正というのは高度な専門性を要する作業で、日本語の表記・表現と起こりやすい間違いに通暁しておらねばならず、現在でも新聞社や出版社において頼りになるうるさ型の校正のベテランが縁の下の力持ちを務めている。かれらがその該博な知識のほんの一部をたまにエッセイなどで披露してくれるのは、筆者のような言語学に携わる者にはまことに有り難い宝物である。

本来校正と校閲は別のものである。原稿通り(もしくは執筆者の意図通り)組版ができているか確認し、必要とあれば修正していくのが校正で、原稿の中に言語表現の

³²⁾ DTP がなかった時代に、会議の資料を作成し、あるいは会議で文案を決定するのはどうしていたか。戦後もしばらくは、秘書課や総務課に「清書係」がいて(ゴーゴリの『外套』と同じ世界である)、会議途中でも指示を受けると直ちに人数分資料を手書きで作り、議場に届けたそうである。こういう会議運営・意思決定の技術がどのように発達してきたかを知ることは、迂遠なようでも、今後の社会をになう学生にとって、データサイエンス教育に留まらない重要な知見であるかと思う。これは危機管理の問題でもあって、便利な道具がない(災害などで消失した)状況で合議の上意思決定する必要に遭遇するかもしれない。マッチで火が付けられない(使ったことがない)ために災害時に苦労をしたという若い世代の人々の話を耳にすると、デジタル技術をアナログ技術の発展から理解し、場合によってはアナログ技術にも通じておく必要は今も存在するように思われる。たとえば日本現代史を考える上で重要な意味がある「終戦の詔勅」(玉音放送)と呼ばれる文章は、1945年8月14日の午後から深夜にわたる長い閣議で決定されたが、成立過程が未だに明らかではないようだ。国立公文書館の刊行物に発表されている次の論文を参照。石渡隆之「終戦の詔書成立過程」『北の丸』第28号(平成8年)

https://www.digital.archives.go.jp/support/pdf/kaiteiban kitanomaru28gou.pdf。これほど複雑な改稿過程がどのように進められたのか。会議運営の技術面から考えることもできるように思う。

誤りや事実誤認がないか確認し、必要とあれば修正していくのが校閲である。これら を効率よく行う技術だけが普及したが、作業内容の難しさや専門性は広く理解された わけでもなく、まして普及したわけでもない。

言語表現の誤りや事実誤認については、現在ネット上で直ちに簡便に検索できる。 簡単なことであればまことに便利であり、翻訳などの際に必要だったその昔の手間を 考えると隔世の感がある。しかし少しでも必要とする情報が複雑なものになると、ネ ット上の知識が真正なものかどうか確認する(ウラを取る)必要がある。つまり、情 報リテラシーの技術が不可欠になる。

- ⑧ 印刷 Printing
- 製本 Bookbinding

印刷や製本までがワードプロセスに係わってくるということに疑問を持たれる向きがあるかもしれない。日本ほどの印刷・製本技術をもってしても、色塗りや裁断が全て完璧に行われることはあり得ない。そして製品上の 0.1 ミリのずれであっても、不快な不手際として知覚されるであろう。そこでトンボという基準点を設定し、塗り足し部分を作っておいて、万が一のずれでも吸収できるようにするのである³³⁾。ページレイアウトに「基準点」が置かれていて、普通デフォルトで紙型の周囲 1 インチにはいかなる文字も入れさせない仕組みになっている(もちろん強引に設定を変えることはできるが)のは、印刷・製本のことを視野に入れるからである。

「巻」という概念の理解も、章立てと製本の両方にまたがるものである。内容上の大分類として作者が指定する「巻」と、製本の技術的必要や出版の営業的要請から1まとまりの長いテクストをいくつかの部分に分けて別の本として作られる場合の「巻」とは、本来別のものである。一致していることもあるので話がややこしい。英語でbook/volume、フランス語で tome/volume、ドイツ語で Buch/Band と使い分ける。これはまた参考文献の書き方にも密接に関わってくる。

関連するが、文庫本や新書の元になったペーパーバックの歴史³⁴⁾について、書物の持つ意味と共に知っておく必要があるだろう。「ペーパーバック革命」と言われる印刷・製本革命によって、安価に大量に書物が出回り、庶民も利用可能になった。多くの人が古典作品に近づきやすくなった一方³⁵⁾、大量消費の新たなコンテンツが生み出されていった。無糸製本あるいは無線製本はこれと同時に普及した。書物はもう永続的に使えるものではなく、使っているうちにどんどん背が割れて、ばらばらになって

 $30 \underline{\text{https://web.archive.org/web/20120204191804/http://paperbarn.www1.50 megs.com/Paperbacks/index.htm}.$

³³⁾ Adobe の Illustrator のマニュアル等に詳しい説明がある。

³⁵⁾ 以前はどの本にもついていたが、岩波文庫や角川文庫の発刊の辞を読むと、その高邁な理念は感動的である。「読書子に寄す ――岩波文庫発刊に際して―― 岩波茂雄/真理は万人によって求められることを自ら欲し、芸術は万人によってされることを自ら望む…昭和二年七月」「角川文庫発刊に際して 角川源義/第二次世界大戦の敗北は、軍事力の敗北であった以上に、私たちの若い文化力の敗退であった…一九四九年五月三日」

当たり前の物になった³⁶⁾。いわゆる酸性紙問題³⁷⁾とならんで紙媒体による資料保存に大きな問題となったのである。

しかし電子媒体に使われている素材は、紙よりも保存が利かないと言われる。電子 媒体に人類の遺産を全て読み込んでも、新たな革命的な活用方法が開発される一方で、 保管保存に関して安全とは決して言えない。

かくて総合的で新しい数理・データサイエンス教育は、人類の知的遺産をどのよう にアーカイブ化し、データベース化し、これまで以上に活用する一方で、これまで以 上に大切に保存していかねばならないかという、古くて新しい総合的な使命と直面す るのである。

駆け足で概観することしかできないが、筆者としては新しい数理・データサイエンス教育にのぞみ、ワードプロセスに上のような諸問題を認識している。この中で特に組版の問題に積極的に答えてくれると同時に、本来美しい数式を書くための技術から、多言語の諸文字をも処理できる能力を発達させた tex を、筆者は人文諸科学を学ぶ学生にとって必要不可欠なツールの一つととらえ、必要最低限の入門的知識をここにまとめてみたいと考える。教材として(解説付き配付資料や、スライドの形に)成形するのはこれからである。

_

³⁶⁾ 学生の頃、安いからと言ってペーパーバック版の辞典(latin-english だの french-english だの)に喜んで手を出し、あっという間に使いにくくなったのは、筆者などの苦い想い出である。筆者の世代が辞典類の電子化を歓迎しているのはこのような世代的経験によるところが大きい。

³⁷⁾ 東京都立図書館のサイトに、簡にして要を得た解説がある。

https://www.library.metro.tokyo.jp/guide/about_us/collection_conservation/conservation/deoxidation/。酸性紙問題が大きく取り上げられていた 1980 年代にドイツに留学していた筆者は、ドイツの図書館で司書から受けた古写本の講義で、関連してこの問題の詳細を学ぶと共に、本の修復に果たす和紙の役割の重要性についても教えられた。

4. 「人文科学研究のための LaTex2e による多言語処理入門」について

以下のような概略を考えている:38)

(1) ファイルの大枠

テクスト本体は¥begin{document}と¥end{document}で挟む。コマンドは¥で始めるが、これは本来バックスラッシュである。日本語キーボードでは円マークにあててあるため、こういう字面にってしまうが、機能に影響はない。

テクスト本体の前に、documentclass (組版の種類の大枠)と、preamble (前書き)で特にそのテクストで使用する package (別機能のセット)を指定する: ¥usepackage{......}。

¥end{document}

(2) documentclass

documentclass には次の各種類を区別できる。

欧文: article, report, book

和文 (横書き): jarticle, jreport, jbook; jsarticle, jsbook

和文 (縦書き): tarticle, treport, tbook

予定しているテクストの長さと構成の大きさに対応する。後述するが、章立ての可能 性が違ってくる。

和文横書きの js-は、奥村晴彦先生による j-の改良版で、これを使うことをお薦めする。

この他に、slides(プレゼンテーション用スライド)、ams-クラス(American Mathematical Society 指定の組版)がある。

生田誠三『LATEX2E 文典』(朝倉書店、2000年)

³⁸⁾ 板東慶太『インストールいらずの LATEX 入門』(東京図書、2019年)。

奥村晴彦『LaTex2e 美文書作成入門』(技術評論社、改訂第7版・2017年、第5版・2010年)

吉永徹美『LATEX2E 辞典』(翔泳社、2009年)

(3) documentclass options

○ フォントはデフォルトが computer modern という、tex の創始者 Knuth の作った字体。好みがあるが、次のようなパッケージによって、よく使われる別のフォントに変更できる。

¥usepackage{txfonts}:

本文: Times New Roman、見出し: Helvetica

¥usepackage{pxfonts}:

本文: Palatino、見出し: Helvetica

- その他に指定できるフォントもある。
- その他の組版のオプションを、次のように[]に入れて指定できる。

\(\frac{4}{2}\)documentclass[12\)pt,twoside,b5\(\paper\)]\(\{\text{jsarticle}}\)

- ア) 字のサイズは 10pt がデフォルトで、11pt と 12pt が指定できる。
- イ) 左右ページのレイアウトを同じにする oneside がデフォルトで、左右対称にする twoside が指定できる。
- ウ) 紙の大きさは a4paper がデフォルトで、和文では a5paper・b5paper が指定できる。欧文ではその上に letterpaper(11×8.5 mm)・legalpaper(14×8.5 mm)・executivepaper(10.5×7.25)が指定できる。
- エ) なお landscape と指定すると、紙を横置きにする。
- オ) 段組は一段組 onecolumn がデフォルトで、twocolumn で二段組みにできる。
- カ) 二段組みの場合、¥setlength{¥columsep}{長さ}で、段と段の間の開きを、{長さ}分指定できる。
- キ) また二段組みの場合、¥setlength{¥columnseprule}{長さ}で、段と段の間に、 {長さ}によって指定した幅の罫線を引くことができる。

「長さ」の概略を以下にまとめておく。

(tex 独自の定義である)

1pt=0.35mm 1mm=2.835pt

1pc (パイカ) =12pt

1 cm = 28.3 pt

1zw=全角字幅

1in(inch) = 25.4mm = 72.27pt

(4) page layout 基本

○ 行の長さをデフォルトから変更したければ、プリアンブルの箇所に、

¥setlength{textwidth}{長さ}

と指定すれば、{長さ}の1行になる。

○ 行送りをデフォルトから変更したければ、¥begin{document}の後に、

¥setlength{¥baselineskip}{長さ}

と指定すれば、{長さ}で指定した行送りになる。行送り(line spacing/leading ちなみに後者は lead「鉛」から来ているので発音 leading[lédin]に注意)とは、ベースラインから次の行のベースラインまでの幅を指す。(ノートの横罫の幅に当たる)

○ 行送りをデフォルトから拡大・縮小させたければ、プリアンブルの箇所に、

¥linespread{倍率}

と指定すれば、 ${倍率}$ 分 (×0.8, 1.5 等) デフォルトの行送りが拡大・縮小される。

○ 左マージンをデフォルトから変更したければ、プリアンブルの箇所に、

¥setlength{¥oddsidemargin}{長さ}

と指定すれば、1インチ+{長さ}の左マージンとなる。 {長さ}にはマイナスも指定できる。

○ 上マージンをデフォルトから変更したければ、プリアンブルの箇所に、

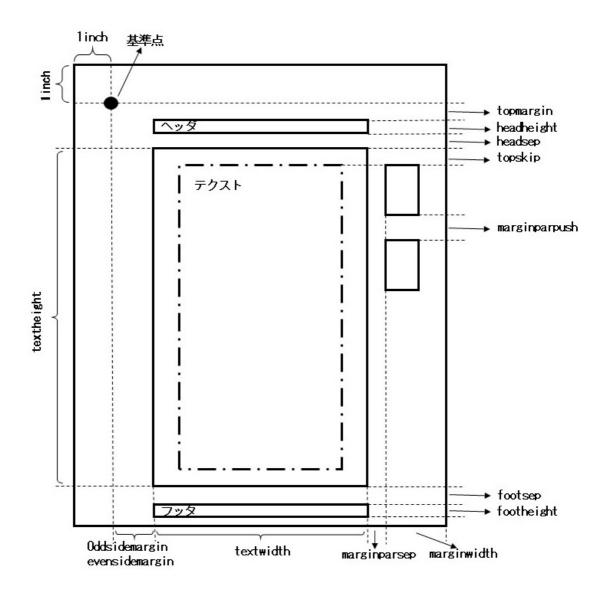
¥setlength{¥topmargin}{長さ}

と指定すれば、1インチ+{長さ}の上マージンとなる。

○ 本文領域は、行の長さと行送りで自動的に決まる。

(5) page layout 詳細

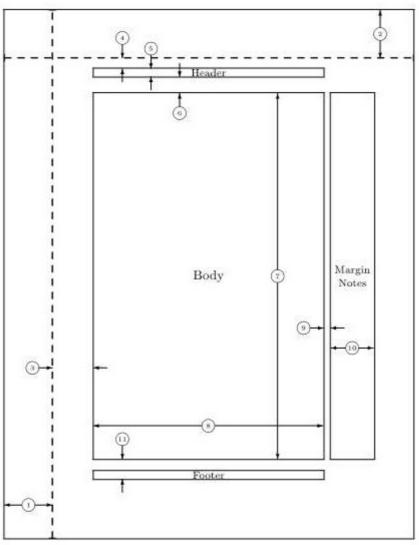
○ ページレイアウトの全体は以下の通り:



marginpar=marginparagraph とは傍注を書き込むスペースである。

小口と天にあたるところに最低必ず1インチの幅を取り、ここにページレイアウト全体の基準点を設けるのが基本。しかしこの基準点をずらす(オフセット offset と呼ぶ)ことも可能。その場合には、水平方向に(horizontal) ずらした距離を hoffset、垂直方向に(vertical) ずらした距離を voffset と呼んで指定するが、避けた方が良い。なお tex のデフォルトでは、文字の大きさ 10pt、用紙は A4 縦置き、左右ページのレイアウトを同一に(oneside)、一段組(onecolumn)として、以下の通り:

documentclass = article の場合

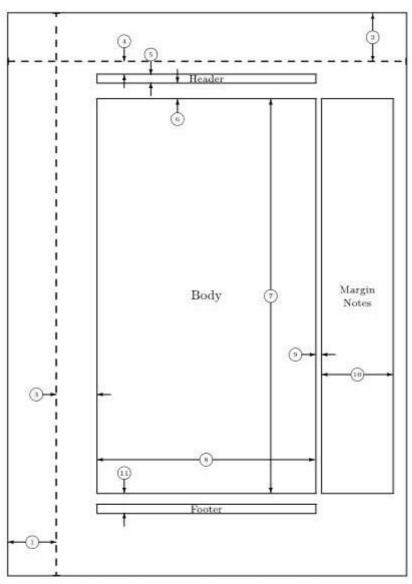


- one inch + \hoffset
- \oddsidemargin = 62pt
- \headheight = 12pt
- \textheight = 550pt
- \marginparsep = 11pt
- \footskip = 30pt \hoffset = 0pt 11 \paperwidth = 614pt
- 2 one inch + \voffset
- 4 \topmargin = 16pt
- 6 \headsep = 25pt 8 \textwidth = 345pt
- 10 \marginparwidth = 65pt \marginparpush = 5pt (not shown)

\voffset = Opt

\paperheight = 794pt

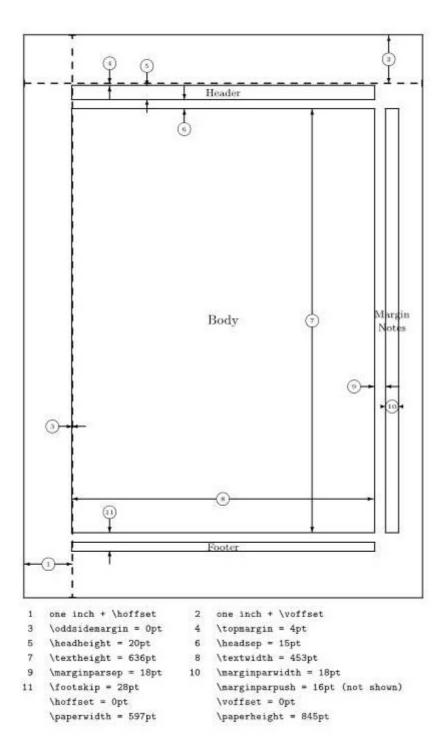
documentclass = jarticle の場合



- one inch + \hoffset 1
- \oddsidemargin = 62pt
- \headheight = 12pt
- \textheight = 592pt
- \marginparsep = 10pt
- 11
- \footskip = 30pt \hoffset = 0pt \paperwidth = 597pt
- one inch + \voffset
- 4 \topmargin = 20pt
- 6 \headsep = 25pt
- 8 \textwidth = 327pt
- \marginparwidth = 106pt 10 \marginparpush = 5pt (not shown) \voffset = 0pt

\paperheight = 845pt

○ documentclass = jsarticle の場合



○ なお上の 3 例では、article が letter paper、jarticle と jsarticle が A4 を用紙サイズとしているので、paperwidth と paperheight の数字が違っている。

○ これらの図は、以下のコマンドで呼び出すことができる:

¥documentclass{xxx}

¥usepackage{layout}

¥begin{document}

¥layout

¥end{document}

これで{xxx}で指定した documentclass のレイアウト図を返してくる。

(6) 柱とノンブル

- プリアンブルに¥pagestyle{xxx}で指定する。
- デフォルトは ¥pagestyle{plain}。ページ下部にノンブル、柱はなし。
- ¥pagestyle{empty}。ノンブルも柱もなし。
- ¥pagestyle{haedings}。 ノンブルと注がページ上部。 documentclass が article の 場合、柱に section 名が来る。 two side にすると、偶数 (左) ページに section 名、 奇数 (右) ページに subject 名が来る。
- ¥pagestyle{myheadings}。¥markright{柱の指定}、¥markboth{左の柱の指定}{右の柱の指定}となる。
- ¥thispagestyle{headings}。当該ページのみの指定ができる。
- jarticle/jreport/jbook のみ、次の指定が可能。¥pagestype{headnombre}上にノンブル、柱なし。¥pagestyle{footnombre}下にノンブル、柱なし。 ¥pagestyle{bothstyle}上に柱、下にノンブル。
- ¥maketitle/¥chapter においてはデフォルトで plain なので、タイトルページや 各章の最初のページにノンブルを必要としない場合には、empty を改めて指定す る必要がある。
- ¥pagenumbering{xxx}で、ページの付け方のオプション指定: arabic (アラビア数

- 字), roman (ローマ数字小文字), Roman (ローマ数字大文字), alph (アルファベット小文字), Alph アルファベット (大文字)、ができる。
- ノンブルの変更は、¥setcounter{page}{xxx}で、現行ページを xxx ページに変える。⇒(49)カウンター参照。
- ノンブルの形式変更は、例えば:

 $\label{thepage} $$ {\operatorname{Parabic}_{page}}^{-} \Rightarrow -1^{-}, -2^{-}, -3^{-} \dots $$ {\operatorname{Parabic}_{page}}^{-} \Rightarrow (a), (b), (c) \dots $$ $$$

(7) 段落と章立て

○ 章立ての構成と番号のデフォルトは次の通り(いわゆる「理科系方式」のナンバー リングとなる):

article		report, book		
¥part	I, II,	¥part	I, II,	
		¥chapter	1,2,	
¥section	1, 2,	¥section	1.1, 1.2,	
¥subsection	1.1, 1.2,	¥subsection	1.1.1, 1.1.2,	
¥subsubsection	1.1.1, 1.1.2,	¥subsubsection	1.1.1.1, 1.1.1.2	
¥paragraph	1.1.1.1, 1.1.1.2	¥paragraph	1.1.1.1.1, 1.1.1.1.2,	
¥subparagraph	1.1.1.1.1, 1.1.1.1.2,	¥subparagraph	1.1.1.1.1.1, 1.1.1.1.2,	

- ¥section{xxx}として、各章などのタイトル{xxx}を付ける。自動的に柱や目次に xxx が反映される。
- ¥section*{xxx}とアステリスクを付けると、番号のない「項目タイトル」のみが出力 される。(目次にも)
- 行頭行末の空白、改行は無視される。行末に¥¥をおけば、改行になる。空行は段落 改めになる。
- 改行後は一字下げなし。段落を改めると一字下げ。新たなパラグラフの冒頭に ¥indent をおくと一字下げ、¥noindent を置くと字下げなしとなる。

特別に字下げを行い、「引用」をしたい場合、quotation(頭一字下げあり)と quote(頭一字下げなし)の2つの命令がある:

¥begin{quotation}

(引用部分:頭下げ)

¥end{quotation}

¥begin{quote}

(引用部分:頭下げなし)

¥end{quote}

- (8) タイトルと目次
 - タイトル・作者・日付は、別立てに指定すると、ページ上部に大きい活字とセンタ リングで作成される。

¥title{xxx}

¥author{yyy}

\date{zzz}

¥maketitle

○ 目次は、¥begin{document}の中に、¥tableofcontents の命令を書き込めば章立てを自動的に反映して作成される。ただし、2 度コンパイルする必要がある。1 度目で組版が決定し、2 度目でそれを目次に反映させるからである。また¥clearpage を続けて命令しておかないと、タイトルページ・目次・本文テクストが改ページにならず、続けて組版されてしまう。

(9) text 入力・基本

○ Unicode をベースに打ち込んでいけば、ラテンアルファベットはもちろん、キリル

文字、ギリシャ文字、日本語、ハングル等も IME や ATOK で変換した通りに本文 テクストに反映される。

- \$...\$として\$で挟んだ部分は、数式モードの処理になる。文字はデフォルトでイタリックになる。もともと tex は数式を多用する著作を読みやすく美しく組版するために開発されたソフトなので、数式用の特殊文字や配置に豊富な機能が用意されている。この機能を利用することで、多言語処理が可能になったのである。
- %は注釈行に用いる。この記号から先は無視される。ソースコードにメモを書き込む場合等に利用する。%という文字そのものをテクストに出力したい場合には¥%と入力する。
- コマンドがアスキーベースなので、そのままでは本文テクストに使えない記号が 多い。極端な例が「円マーク」 "¥" と 「半角バックスラッシュ」である。 「円マーク」は、¥usepackage{textcomp}とプリアンブルに指定してから、 ¥textyenで呼び出せる。

「半角バックスラッシュ」は、特にプリアンブルで指定する必要はないが、数式でしか使えないので数式であることを示すために \$...\$ で挟む。\$¥backslash\$ または \$¥setminus\$とすれば呼び出せる。

ちなみに、日本語変換において "¥"と「半角バックスラッシュ」の文字コード 005C がかぶっているため、日本語の環境では「半角バックスラッシュ」の使用は通常困難である。

○ 本文テクストにそのまま使える文字は以下の通り:

$ ext{LAT}_{ ext{E}} ext{X}2_{arepsilon}$ $ ext{C}$	そのま	ます	使用で	きる	記号	号類
入力	Į.	,	()	*	+
出力	ļ ţ	,	()	*	+
入力	,	-	•	1	•	;
出力	,	_	. 8 € 2	1	:	•
入力	=	?	Q.	E]	É
出力	=	?	@	Γ	1	4

○ ¥の次に置かれた文字は直ちにコマンドだと解釈され、該当するコマンドがなければ無視されるか、エラーが出る。本文テクストの中で ¥XXX というコマンドを用いたら、その次に半角の空白スペースを置いてコマンドの終了を宣言する。その後にさらに空白が必要な場合は、別途空白指定をしなければならない。

¥LaTeX で latex のロゴが呼び出せるが、続けて文字を打つ場合には、¥LaTeX 」として空白を置く。この後にさらに空白を置きたければ、コマンド終了の宣言を置き、さらにテクスト上の空白を設定する。以下の三通りの方法がある:

¥LaTeX¥..., ¥LaTeX{}..., {¥LaTeX}...,

これで、それぞれのコンマの前の半角1文字分の空白が書ける。

- 半角空白は単語の切れ目と解釈され、改行を呼び起こす可能性がある。改行しない 空白を付けたいときには"~"を用いる。これは¥を必要としない特殊なコマンドで ある。
- 画面に書いたままの体裁で出力するコマンドもある:

¥begin{verbatim}

(テクスト)

¥end{verbatim}

または、¥verb|テクスト|でも可能。ただし ¥section{xxx}などの章タイトルの中では使えない。

(10) 行揃え

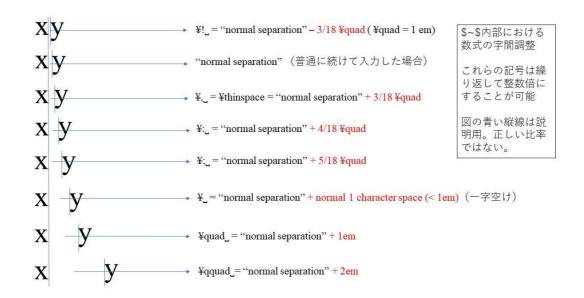
左寄せを flushleft、右寄せを flushright、中央揃えを center として、次のように指定する。

¥begin{flushleft}
(テクスト)
¥end{flashleft}

¥begin{flushright} (テクスト) ¥end{flashright} ¥begin{center} (テクスト) ¥end{center}

(11) 空白

- 水平方向(horizontal)の空白は以下のように指定する: ¥hspace{長さ}。¥hspace{¥fill}とすると行末までの空白となる。{長さ}にはマイナスも指定できる。行頭で効力を失うので注意。
- 垂直方向(vertical)の空白は以下のように指定する: ¥vspace{長さ}。¥vspace{¥fill}とするとページ末までの空白となる。{長さ}にはマイナスも指定できる。ページ頭に置かれると効力を失うので注意。
- 数式において語間と行間を微細に調整するコマンドが他にある。



○ 欧文活字は (タイプライター書体以外) 同じフォントであっても、横幅が文字 ごとに異なっている。それが文字種や文字サイズによってさらに異なる。あるフォントセットで代表的な小文字横幅をm活字の横幅ととし、それを em という単位で表す。 (小文字の高さは x 活字で代表させ、それがベースラインとミーンラインの距離になる。)

しかし、文字を間隔なしに並べても、必ず経験的に必要とされる一定の隙間を

LaTeX は自動的に空ける。上の表でこれを normal separation と呼んでおいた。 すべて臨時の字間調整は、この normal separation を基本に、これにプラスマイナスで取っていく。それが上の表である。(ただし縮尺は全く正しくなく、目安にならないので注意のこと) 1em の 18 分の 1 という刻みで調整する。「半角 1 字空け」にすると、標準の一文字分(しかし 1em よりやや短い長さで、これを上の表で mormal 1 character space と呼んでおいた)を加える。

(12) 特殊文字・基本

- Ångstrøm と出力させたい場合、Å が ¥AA であり、ø が ¥o というコマンドであるからといって、¥AAngstr¥om としても、エラーが出る。¥の次に切れ目なく置かれた文字は全てコマンドだと理解されてしまうが、¥AAngstr とか ¥om というコマンドが存在しないので、エラーを返してきたのである。¥AA_ngstr¥o_m として、¥AA と ¥o の部分がコマンドであり、続く文字はコマンドには含まれないことを明示してやる必要がある。
- 下の表に基礎的な特殊文字の出力方法をまとめる。アクセント類は項目を改めて一覧を示す。

千葉大学人文公共学府 研究プロジェクト報告書 第 350 集 pp. 1-42 2020 年 高等教養教育研究

	IATEX $2_{oldsymbol{arepsilon}}$	によ	る 特 殊 文 字	
入力	\#	\\$	\%	\&
出力	#	\$	%	&
入力	\	\{	\}	\s
出力		{	}	§
入力	\P	\dag	\ddag	\copyright
出力	P	†	‡	©
入力	\pounds	\oe	\0E	\ae
出力	£	œ	Œ	æ
入力	\AE	\aa	\AA	\0
出力	Æ	å	Å	Ø
入力	\0	\1	\L	\ss
出力	Ø	ł	Ł	В
入力	?'	1,	\i	\j
出力	i	i	1	J
入力	£ .	,		, ,
出力	6	,	44	"
入力	*	""	'''	_
出力	*	44 4	, ,,	- (hyphen)
入力	(==		\textregistered	\texttrademark
出力	- (dash)	— (stroke)	®	TM
入力	\textasciitilde	\TeX	\LaTeX	\LaTe X e
出力	~	T_EX	IAT _E X	IAT $_{ m E}$ X $2_{m arepsilon}$

○ 関連して、ワープロソフト一般に、ダッシュの指定に要注意。文字コードは"2014"で同じだが、1 文字分の長さで両端に小さなスリットが入り、並べて入力しても破線状態になるものと、長さが 1em で、並べて入力すると切れ目のない直線になるものと区別のある場合がある。また hyphen、dash、stroke、minus、長音符の区別に注意のこと。横書きで、全角ダッシュの代わりに長音符が置いてあるのをしばしば見かけるが(起筆収筆がきちんと表されている)見苦しいだけでなく、場合によっては意味が変わる。

(13) アクセント

	IATEX $2_{\mathcal{E}}$	によるア	クセント数	領
入力	\'{o}	\'{o}	\^{o}	\"{o}
出力	δ	ó	ô	ö
入力	\~{o}	\={o}	\.{o}	\u{o}
出力	õ	ō	Ó	ŏ
入力	\v{o}	\H{o}	\t{oo}	\c{o}
出力	ŏ	ő	රව	Q
入力	\d{o}	\b{o}	\r{a}	\k{a}
出力	ó	ō	å	ą

使用例:

Schr¥"{o}dinger → Schrödigner

$\bar{i} \leftarrow Y=\{Yi\}$

まず¥i として、i の上の点を取ってから¥={}の中に入れ、上につく横棒(マクロン)を加える。

○ また数式においても次のようにアクセントを付けることができる。これらのコマンドは繰り返して重ねることができる(次表の最下段参照):

数式におけるアクセント	類
$\usepackage{amsmath}$	
$\Lambda cute\{x\}$	x
\$\Grave{x}\$	à
\$\Hat{x}\$	\hat{x}
\$\Bar{x}\$	\bar{x}
\$\Breve{x}\$	$ \ reve{x}$
$\c \c \$	\check{x}
$Tilda{x}$	\tilde{x}
Δx	\dot{x}
\$\Ddot{x}\$	\ddot{x}
\$\Grave{x}\$	à
<pre>\$\Acute{\Acute{x}}\$</pre>	ź

○ 同じく数式において、ダッシュあるいはプライム記号を次のようにして付けることができる:

数式におけるダッシュあるいはこ	プライム記号のつけ方2種類
\$a'\$	a'
\$a''\$	$a^{\prime\prime}$
\$a'''\$	$a^{\prime\prime\prime}$
\$a\prime\$	a/
<pre>\$a\prime\prime\$</pre>	$a\prime\prime$
<pre>\$a\prime\prime\$</pre>	a''''

○ 数式における上下の添字は、以下のように付けることができる:

\$x^2\$	$ x^2 $
\$x_2\$	x_2
\$x^{123}\$	x^{123}
\$C^3_6\$	C_6^3
\${}^2x\$	^{2}x
\${}_2x\$	2x
\${{}_2x}^2\$	$\frac{1}{2}x^2$

○ ついでながら、数式における分数・平方根・積分は次のように表すことができる:

分数・累乗根・積	分
\$1/2\$	1/2
\$\frac{1}{2}\$	$\frac{1}{2}$
\$\sqrt{7}\$	$\sqrt{7}$
\$\sqrt[12]{a}\$	$\sqrt[12]{a}$
\$\int_{-a}^{+a}f(x)dx\$	$\int_{-a}^{+a} f(x) dx$

○ 数式には以下のようなコマンドがあり、数式の上に数式を乗せることができる:

これらを転用し、和文組版における「ルビ」を作る。本論の趣旨を超えるので、参考までに紹介するに留める。

(14) フォント・基本

- 普通フォントと呼んでいるものには、5つの要素がある。
 - 1) encoding: 文字コードのことである。
 - 2) family: 様々な書体をそろえた同じデザインのワンセットのフォント。Times New Roman とか Century などといった名称はこれを表す。
 - 3) series: 文字線の太さによる区別。フェイスともいう。ボールドやメディアムなどがある。そのほかに活字の横幅を広げたり狭めたりするフェイスもあるが、一般には用いない。
- 4) shape: ローマンやイタリックなどの書体。LaTeX2e では 3)と 4) を合わせて 7 種類を区別している。
- 5) size: 文字の大きさ。アセンダーラインからディセンダーラインまでの長さを pt で表す。

○ 和文と欧文に、次のようなフォントの区別が可能である:

	千葉大学	
標準	千葉大学	千葉大学
ゴチック	\textgt{千葉大学}	千葉大学

Chiba University (Times New Roman)				
標準	Chiba University	Chiba University		
ローマン	\textrm{Chiba University}	Chiba University		
ボールド	\textbf{Chiba University}	Chiba University		
イタリック	\textit{Chiba University}	Chiba University		
斜体	\textsl{Chiba University}	Chiba University		
ゴチック	\textsf{Chiba University}	Chiba University		
タイプライター	\texttt{Chiba University}	Chiba University		
スモールキャプ	\textsc{Chiba University}	Chiba University		

「ローマン」というのは、「標準」と同じ。¥textrm というコマンドは、別の書体であったものを「標準」に戻す際に用いる。

「ボールド」と「ゴチック」がしばしば混同されるが、上の表の通り、本来別物である。

「ボールド」は、「標準」をそのまま太字にしたものである。これに対して「ゴチック」は、本来「サンセリフ(セリフなし)」と呼ばれ、縦横の線の端にある小さな出っ張り(=セリフ)をなくした形をいう。和文フォントでも欧文フォントでも始筆と収筆における筆(ペン)の力の入れ方を示す出っ張りがデザイン化されている。それを取り去ってつるんとした形になるのがサンセリフ=ゴチック/ゴシックである。また「斜体」と「イタリック」もしばしば混同されるが、上の表の通り、本来別物である。「斜体」は「標準」を斜めに倒しただけだが、「イタリック」は筆記体に寄せて華麗なデザインにしたもので、筆記体に似せたので傾いているのである。ただし、フォントセットによっては、「斜体」と「イタリック」のいずれかしか持たず、「斜体」と「イタリック」で同じものを指している場合がある。これが混乱に拍車をかけたと思われる。

「タイプライター」は、もう失われた書字機材「タイプラター」の書体をまねて作

られたフォント。洗練されない無骨で古臭い活字が、全て同じ横幅で作られていると ころに特徴がある。古い公文書がタイプライターで書かれているため、セピア色に焼 けた写真と同じ、懐古的な雰囲気を持っていた。現在では、字間を空けた強調表示に 用いられる。

「スモールキャプ」は、「小さな大文字」という意味で、大文字を小文字の大きさ (x-height) に縮小した字体。

(15) フォントサイズ

○ 以下の 10 種類の文字サイズを指定できる。下の表は、documentclass のオプションで、標準文字サイズをデフォルトの 10pt とした場合の例である:

	ľ	
{\tiny Chiba University}	Chiba University	5pt
{\scriptsize Chiba University}	Chiba University	7pt
{\footnotesize Chiba University}	Chiba University	8pt
{\small Chiba University}	Chiba University	9pt
{\mormalsize Chiba University}	Chiba University	10pt
	Chiba University	2.00
{\large Chiba University}		12pt
{\Large Chiba University}	Chiba University	14.4pt
{\LARGE Chiba University}	Chiba University	17.28pt
{\huge Chiba University}	Chiba University	20.74pt
{\Huge Chiba University}	Chiba University	24.88pt

- これ以上大きな文字は、図形の拡大の機能を転用して作ることができる。
- 同じ文字列に対し、フォントの種類とサイズを同時に指定する際に、順番を入れ替えても同じ結果になる。¥textgt{¥large 千葉大学} と{¥large ¥textgt{千葉大学}}とでは、共に同じ「千葉大学」(12pt)という出力を得る。

(16) babel·基本

○ babel パッケージを使い、多くの欧文を簡便に処理することが可能である。概略を 以下に示す:

babelによる多言語処理例

新約聖書「ルカによる福音書」6.33

「汝等 おのれに 善をなす者 に 善を 為 すとも、何 の 嘉 すべき 事あらん、罪人 にても 然 するなり。」

入力 出力

古代ギリシャ語

ka'i g'ar >e'an >agajopoi~hte >agajopoio~untas <um~as, po'ia <um~in q'aris >est'in; ka'i o<i <amartwlo'i t'o a>ut'o poio~usin. καὶ γὰρ ἐάν ἀγαθοποιῆτε ἀγαθοποιοῦντας ὑμᾶς, ποία ὑμῖν χάρις ἐστίν· καί οἱ ἁμαρτωλοί τό αὐτό ποιοῦσιν.

ラテン語

Et si bene feceritis his, qui vobis bene faciunt, quae vobis est gratia? Siquidem et peccatores idem faciunt. Et si bene feceritis his, qui vobis bene faciunt, quae vobis est gratia? Siquidem et peccatores idem faciunt.

ドイツ語

Und wenn ihr euren Wohlt"atern wohltut, was f"ur Dank habt ihr davon? Denn die S"under tun dasselbe auch.

Und wenn ihr euren Wohltätern wohltut, was für Dank habt ihr davon? Denn die Sünder tun dasselbe auch.

フランス語

Si vous faites du bien \'a ceux qui vous font du bien, quel gr\'e vous en saura-t-on? Les p\'echeurs aussi agissent de m\^eme.

Si vous faites du bien à ceux qui vous font du bien, quel gré vous en saura-t-on? Les péheurs aussi agissent de même.

ラテン語と英語については、基本的に特殊文字・アクセント類を必要としないので 違いが見えにくいが、ハイフネーションが違ってくる。

Latex2eで本来可能な特殊文字・アクセント類の出力よりいっそう簡単な手続きで、各言語の特殊文字・アクセント類を出力することができる。ハイフネーションもそれぞれの言語の正書法に準拠して行われる。(ついでに、文語聖書の引用に、Latex2eで行ったルビの処理を例示しておいた)

○ プリアンブルで babel パッケージの使用を宣言すると同時に、オプションとして []の中に、使用する予定の言語名を並べておく。

各言語のテクストを入力する際に、Yselectlanguage $\{$ 言語名 $\}$ を書き込んで、使用する言語を宣言する。下の図が babel 使用の基礎である:

(17) babel, german/french

○ 具体的に、一番多言語処理の要求が高いであろう諸言語から入力方法を一覧にして示す。先ずフランス語とドイツ語から:

¥usepackage[english, german, french, latin, polutonikogreek, russian]{babel}

¥begin{document}

¥selectlanguage{french}

<フランス語のテクスト>

¥selectlanguage{english}

<英語のテクスト>

¥end{document}

babel: フラン	ノス語
∖ 'a	à
/ · A	À
\^i	î
/^I	Î
\'e	é
\'E	É
\c{c}	ç
\c{C}	Ç
	œ
	Œ
$\log{{}}xxxxx\fg$	« xxxxx »

babel: ドイツ語		
"a	ä	
"A	Ä	
"s, "z	В, В	
"S, "Z	SS. SZ	
xxxxx\grqq	"xxxxx"	
xxxxx\flqq	»xxxxx«	
">xxxxxx"<	»xxxxx«	
>>xxxxx<<	»xxxxx«	
xxxxx\grq	,xxxxx'	
xxxxx\flq	>XXXXXX<	

○ 特に注意して貰いたいのは、引用符 quotation marks である。

各国語で引用符が微妙に異なっている。別に英語でよく使われる引用符を代用で使って差し支えないし、理解可能だが、それぞれの言語の正書法や習慣には反するので、 違和感が残る。

引用符は性格上2つ一対で使われる。和文ではこれを「始め」「終わり」と区別するが、欧文では「左」「右」で区別する。

フランス語では引用符にギュメ(仏 guillemet/英 angle quotation mark, «:

U+00AB,»: U+00BB <: U+2039, >: U+203A)という括弧が用いられる。これはロシア語でも引用符として使う。山括弧(angle bracket, 《: U+300A, 》: U+300B, 〈: U+3008, 〉: U+3009)とは別のものである。

またこのギュメは、デンマーク語でも引用符として使い、ドイツ語でも使うことがあるが、こちらの2言語の場合には、フランス語やロシア語と左右が逆になる。

LaTeX2e で用いる引用符のコマンド¥glq などでは、最初の g がドイツ語を、f がフランス語を、次の l が左を、r が右を、最後の q が single quotation marks を、qq が double quotation marks を表している。図表の中で示した通り、 r と l に注意して左右逆に使えば、言語の間の微妙な習慣の違いを処理できる。

ドイツ語で用いる引用符は、英語の引用符 "": U+0022 ではなく、また": U+201C,": U+201D でもない。後者を 180 度回転させ、しかも初めを下付に、終わりを上付きにする。結果として、ドイツ語引用符右側が英語引用符左側と同じ形になり、ドイツ語引用符左側は特別の形を取る: ": U+201E (Low-9 quotation mark)。 babel フランス語で用いるYog-Yfg と、babel ドイツ語で用いるYflqq-Yfrqq の違いは、前者の方がフランス語の組版の習慣に従っていて、記号の後にわずかなスペースを空ける仕様になっていることである。

 \bigcirc なおドイツ語の新正書法(1996 年 \bigcirc)に対応した ngerman のオブションもある。

(18) babel, ancient greek, Russian

○ アルファベットがそもそもラテンアルファベットと異なる言語としては、古代ギリシャ語とロシア語の要求が高そうでもあり、また babel から入りやすいので、次にその出力方法を紹介する。

現代ギリシャ語と古代ギリシャ語では少しだけ扱いが異なる。キリル文字を使う 言語はこれでほとんどカバーできる。

ただしキリル文字は、プリアンブルに ¥usepackage[OT2, T1]{fontenc} と OT2 をオプションで指定しておかなくてはならない。

babel:古代	代ギリシャ語
A, a	Α, α
B, b	Β, β
G, g	Γ, γ
D, d	Δ, δ
E, e	Ε, ε
Z, z	Ζ, ζ
H, h	Η, η
J, j	Θ, θ
I, i	Ι, ι
K, k	Κ, κ
L, 1	Λ, λ
M, m	Μ, μ
N, n	Ν, ν
X, x	至, ξ
Оо	O. o
P, p	Π, π
R, r	Ρ, ρ
S, s(c)	Σ , $\sigma(\zeta)$
T, t	Τ, τ
U, u	Υ, υ
F, f	Φ, φ
Qq	Χ, χ
Y, y	Ψ, ψ
W, w	Ω, ω
o,	ó
ʻi	ί
W	ũ
<'0	ő
>'0	ő
h	ñ
>H	'H

babel: ì	コシア語
A, a	A, a
B, b	Б, б
V, v	В, в
G, g	Г, г
D, d	Д, д
E, e	E , e
EO, eo	EO, eo
Zh, zh ; Z1,z1	Ж, ж ; Ж,ж
Z, z	3, 3
I, i	И, и
I0, i0	Й, й
K, k	К, к
L, 1	Л, л
M, m	М, м
N, n	Н, н
X, x	Ш, ш
Оо	O. o
P, p	П, п
R, r	P, p
S, s	C, c
T, t	Т, т
U, u	У, у
F, f	Φ, Φ
H,Kh; h, kh	Ч, ч
C, c; Ts, ts	Ц, ц; Ц, ц
Q, q; Ch, ch	Ч, ч ; Ч, ч
X, x; Sh, sh	Ш, ш ; Ш, ш
W, w; Xq, xq	Ш, щ; Ш, щ
p2	ъ
у	ы
p1	Ь
E1, e1	Э, э
Yu, yu; J2, j2	Ю, ю ; Ю, ю
Ya, ya; J1, j1	Я, я; Я, я

(19) TIPA (国際音声記号・発音記号)

- プリアンブルに tipa パッケージの使用を宣言すると、IPA(International Phonetic Alphabet)「国際音声記号」、俗に言う「発音記号」が出力できる。TIPA は東京大学の福井玲教授が開発したもので、以下のサイトにマニュアルが公開されている: http://ctan.math.utah.edu/ctan/tex-archive/fonts/tipa/tipa/doc/tipaman.pdf
- 2種類の入力方法がある。IPA の一つ一つに付けられた長い名前で一つ一つ呼び 出す方法(入力方法 1)と、YtextipaX として、一部の IPA に付けられた簡略な 記号を入れ、一括して変換可能な方法(入力方法 2)である。続けて IPA を入力し なければならない場合に、入力方法 2 は便利である。

全てのIPAに入力2のための簡略な記号名称が付けられているわけではないので、 入力方法2が常に使えるわけではない。

university の発音記号を出力する:

入力2の場合

[\textipa{""ju:n@"v@:s@ti}] [ju:nə'və:səti]

入力1の場合

[\textsecstress ju\textlengthmark n\textschwa
 \textprimstress v\textschwa\textlengthmark s\textschwa ti]

[ju:nəˈvə:səti]

○ 主要な発音記号の出力方法について、以下に紹介しておく:

MFX 2εによる IPA 出力(tipa)				
入力方法1	入力方法 2 \texttipa{xxxx}	出力		
a	a	a		
\textscripta	A	α		
\textturna	5	ß		
\textturnv	2	Λ		
\textcrb		ъ		
\c{c}	\c{c}	ç		
\textcrd		đ		
\dh	∖dh	ð		
\textschwa	@	Э		
\textepsilon	Е	ε		
\textcrg		g		
\textgamma	G	¥		
\texthvlig		h		
\textbari	1	i		
	N	ŋ		
\textopeno	О	э		
\textthorn	\th	þ		
\textesh	S	ſ		
\texttheta	Т	θ		
\textupsilon	U	ប		
textyogh	Z	3		
\textchi		χ		
\textcommatailz		z		
\textprimstress	и	j.		
\textsecstress	3111	т		
\textlengthmark	1	I		
\textsubring{u}	\r*u	ų		
\textsubarch{u}		ů		
<pre>ph</pre>		\hat{p}^h		
kw		k ^w		
d\textraiseglotstop		d ²		

- (20) ラテンアルファベットを中心とした多言語処理について一応の概略は以上の通りである。アラビア語、ヘブライ語、中国語、韓国語、ヒンディー語といったアジア系諸言語の処理がさらに大きな課題だが、スペースの関係でここでは触れられず、別の教科書に回す。
- (21) また LaTeX2e そのものの機能についても、
 - a) 箇条書きの方法
 - b) 注の書き方
 - c) 図版の挿入
 - d) 表の作成
 - e) box 機能
 - f) 作図
 - g) カウンター操作

が最低限欠かせない。次の教科書で説明していく。

(22) TEX はフリーソフトであり、CTAN というサイトに https://www.ctan.org/、TEX 関連のファイルやソフトが世界中の有志によって続々と開発され、保管されている。 誰でもダウンロードすることが可能である。

本稿はとりあえずここで閉じる。