

日语语法标注规则刍议*

澳门大学 杨文江

提 要: 日语语法标注规则在文献中讨论的不多, 在国内的日语语言学论著中尚未得到广泛应用, 在国外的英文出版物中也存在着使用不规范的现象。本文介绍莱比锡标注规则在日语文本标注中的应用, 并通过分析以往研究中语法标注和语素切分相关的问题和争议, 提出初步的日语语法标注规则。

关键词: 语法标注、日语、语素切分、莱比锡标注规则

1. 引言

语言学论著中, 如果作为研究对象的语言不同于书写所用的元语言, 按照国际惯例, 需要对所有例句进行逐行标注 (interlinear glossing)¹。莱比锡标注规则 (The Leipzig Glossing Rules, 以下简称LGR) 是目前国际上使用最广泛的一套标注规则, 很多出版社和杂志社以此为规范, 或者在此基础上补充更详细的规定。LGR是由德国马克斯·普朗克演化人类学研究所 (Max Planck Institute for Evolutionary Anthropology) 的Bernard Comrie和Martin Haspelmath以及莱比锡大学语言学系的Balthasar Bickel合作编写, 它包括句法和语义逐行标注的十条规则与一个附录 (语法范畴标签和其缩略形式)。该标注系统的第一版于2004年公布, 2008年和2015年进行了两次修订。正如LGR的作者们所言, 基于不同的研究目标和研究对象, 每篇文章标注的精细度可以不同, 精细的标注能够让读者获得作者所讨论课题之外的有关目标语言的许多其他信息。同时, 特定语言的语素切分和语法标注方式经常存在分歧, 因此论著中实际采用的具体方式反映了作者的理论追求或曰理论假设。

在日语语言学的英文出版物中, 正如Dalla Chiesa (2020) 所述, 标注方式的混乱现象比较普遍, 一些著名语言学家的不同论著、甚至同一篇论文中也有标准不一或前后矛盾的情况。这一方面是由于日语形态学研究还存在很多争议问题, 作者们往往按照自己的理论假设进行语素切分和标注, 而有些作者制定的规则内部缺乏逻辑一致性。另一方面, 学界对逐行标注的规范问题讨论太少, 这导致不同作者对一些已有共识的问题采用不同的语素切分和标注方式。一个

* 本文的写作得到了澳门大学启动研究基金 (SRG2022-00026-FAH) 的资助, 谨致谢忱。

1 英语interlinear有两种含义。第一种是“插入在已写好或印好的两行之间的”, 汉语可译为“行间”, 如陈玉洁等 (2014), 但这不是interlinear glossing中该词的原义。第二种是“在不同的行使用不同语言或文本书写或印刷的”, 汉语难以用简洁的方式准确翻译, 本文暂时采用罗天华 (2023) 使用的译词“逐行”。

典型的案例是影山太郎虽然出版过很多分析日语形态的论著，Kageyama（2016）一文更是按照形态句法标准把日语的语素分为四类（本文4.1详述），但是他的很多论文中举出的例句仅使用词缀前的“-”和词间分写两种方式，Kageyama（2016）提倡的附缀和附属词的形态地位没有体现出来。虽然从另外的角度看，影山采用的是粗线条的标注方式，但是这种标注方法与他本人对日语形态学观点的差异不可避免地给读者带来误解或困惑。

在中文出版的日语语言学论著中，目前只有极少数文章对例句进行了逐行标注。我们知道，没有文本标注的文章预设了读者能够读懂日语的例句，因此实际上大多数国内出版物把读者限定在很小的范围，这在客观上阻碍了日语学界同其他语种研究者之间的学术交流。这一点可能也是国内日语语言学论文引用率较低的原因之一²。

据我们调查，在已经出版的讨论文本标注的文献中，除了以上提到的LGR，Lehmann（1982, 2004）也是广为引用的两篇文章，其中Lehmann（1982）是LGR的主要参考文献之一，Lehmann（2004）则在1982文章的基础上更详细地分析了文本标注的问题。国内讨论文本标注的文章有黄成龙（2005）、陈玉洁等（2014）和罗天华（2023）。黄成龙（2005）参考了LGR的第一版，重点分析了国内少数民族语言的标注问题。陈玉洁等（2014）详细介绍了LGR的第二版，并讨论了中文普通话和部分方言的英文标注问题。罗天华（2023）则重点介绍了Lehmann（2004）提出的标注规则，同时参考LGR，对汉语的文本标注提出了若干建议。在日语文本标注方面，Dalla Chiesa（2020）分析了日语语言学的英文出版物中格助词语素切分和语法标注的混乱现象，ナロック（2023）重点介绍了使用英文写作日语语言学论文的一些注意事项，随后简单提及了日语语素切分的问题，最后用很大的篇幅列出了日语功能语素的标注方式及其实例。

基于以上背景，又考虑到国内很多日语语言学研究不太了解日语的文本标注，本文将介绍并讨论一些基本的技术细节，期待为日语文本标注的规范化提供参考。此外，我们还将提及日语语素切分的争议问题。本文讨论的标注规则主要基于“莱比锡标注规则”2015年第三版和Lehmann（2004），日语的标注问题除了参考ナロック（2023）和Dalla Chiesa（2020），还借鉴了一些日语语言学的英文出版物，包括《日语语言学学报》（*Journal of Japanese Linguistics*）³和《日本语言与语言学手册》（*The Handbook of Japanese Language and Linguistics*）⁴丛书以及一些知名语言学家用英文出版的专著和论文。

2 国内日语语言学论著的另外一个问题是很多作者仍然在使用日本学校语法的术语，这些术语很多与普通语言学脱节，造成在中文语境下很多概念与国内其他语言学同行难以正常沟通。

3 日语语言学学报（*Journal of Japanese Linguistics*）由德古意特出版社（Mouton de Gruyter）出版，每年两期。德古意特出版的语学期刊论文和专著均要求例句的逐行标注要使用莱比锡标注规则。此外，日本语学会会刊《日语研究》（『日本語の研究』）的姊妹英文刊《日本的语言》（*Language in Japan*）于2024年创刊，它的投稿要求也规定例句的文本标注使用莱比锡标注规则。

4 该丛书由柴谷方良和影山太郎（Masayoshi Shibatani & Taro Kageyama）主编，德古意特出版社2015年至2025年间出版。

2. 日语的语法标注

语法标注在英文论著中一般采用缩略词的方式，而且通常使用小型大写字母。在中文写作的论著中，如罗天华（2023）所建议，我们可以使用国际通用的英文缩略词，也可以使用相应的中文术语，但是中文术语需要使用与词汇标注不同的字体以示区分。本文的例句采用英文缩略词⁵。

由于每种语言语法范畴的种类和数量各不相同，因此需要标注的语法类型也不尽相同。表1按英文字母顺序列出了日语常见的语法标注，除了第一、第二和第三人称按照惯例使用阿拉伯数字外，其他语法标注均为小型大写字母的缩略形式。表中优先采用LGR附录中列出的英文术语及其缩略词，相应的术语不使用下划线，LGR没有收录的首先引自Lehmann（2004），使用下划波浪线，这两个文献均未收录的则采用ナロック（2023）列出的术语，用下划直实线标出。以上三个文献都没有收录的是本文建议的术语，使用下划虚线。

表1 日语的语法标注及其实例

语法标注	英文术语	中文术语	语素	举例
1	first person	第一人称	<i>watasi</i>	<i>watasi</i>
2	second person	第二人称	<i>anata</i>	<i>anata</i>
3	third person	第三人称	<i>kare</i>	<i>kare</i>
ABL	ablative	离格	<i>kara</i>	<i>koko=kara</i>
ACC	accusative	宾格	<i>o</i>	<i>ringo=o</i>
<u>ADN</u>	<u>adnominal</u>	定语	N/AN= <i>na</i>	<i>sizuka=na</i>
<u>ADVR</u>	<u>adverbializer</u>	副词化标记	A- <i>ku</i> ; N/AN= <i>ni</i>	<i>naga-ku</i> ; <i>sigoto/sizuka=ni</i>
ALL	allative	向格	<i>e</i>	<i>koko=e</i>
<u>AVS</u> ⁶	<u>adversative</u>	转折	<i>ga</i> ; <i>kedo</i>	<i>ik-u=ga</i> ; <i>ik-u=kedo</i>
BEN	benefactive	受益	- <i>te mora(w)</i> -/age-/kure-	<i>it-te mora-u/age-ru/kure-ru</i>
<u>CAL</u>	<u>causal</u>	原因	<i>kara</i> ; <i>node</i>	<i>ik-u=kara</i> ; <i>ik-u=node</i>
CAUS	causative	致使	-(s)ase-	<i>ik-ase-ru</i>
CLF	classifier	类别词	- <i>hon</i>	<i>ip-pon</i>
<u>CMP</u> ⁷	<u>comparative</u>	比较	<i>yor</i>	<i>kore=yor naga-i</i>
COM	comitative	伴随格	<i>to</i>	<i>ane=to issyo=ni</i>
COMPL	completive	完结体	- <i>te sima(w)</i> -	<i>it-te sima-u</i>
<u>CONC</u>	<u>concessive</u>	让步	<i>noni</i>	<i>ik-u=noni</i>
COND	conditional	条件	<i>to</i> ; - <i>tara</i> ; -(r)eba; <i>nara</i>	<i>ik-u=to</i> ; <i>it-tara</i> ; <i>ik-eba</i> ; <i>ik-u=nara</i>
<u>COO</u>	<u>coordinator</u>	并列	<i>to</i>	<i>ringo=to nasi</i>

5 英文缩略词的优势在于，很多单个英文语法术语对应多个中文译词，虽然英文术语也存在争议问题，但相较于中文术语分歧较小，因此更容易取得共识。另外，英文的缩略语占用空间较小，易于排版。

6 Lehmann（2004）把adversative缩略为ADRVs，该缩略词中R和V顺序颠倒，故此处采用ナロック（2023）使用的缩略词。

7 “CMP/comparative比较”指日语用于指示比较对象的「より」，而不是Lehmann（2004）列出的“比较级”。

(续表)

语法标注	英文术语	中文术语	语素	举例
COP	copula	系词	<i>da</i>	<i>koko=da</i>
CTR	<u>contrastive</u>	对比	<i>wa</i>	<i>ane=wa</i>
DAT	dative	与格	<i>ni</i>	<i>ane=ni</i>
DEM	demonstrative	指示词	<i>kore; sore; are</i>	<i>kore; sore; are</i>
DEO	<u>deontic</u>	道义情态	<i>beki</i>	<i>ik-u=beki</i>
DES	<u>desiderative</u>	意愿情态	<i>-ta-</i>	<i>iki-ta-i</i>
DIS	<u>disjunctive</u>	两者取一	<i>ka</i>	<i>ringo=ka nasi</i>
ENU	<u>enumerative</u>	列举	<i>ya; si</i>	<i>ringo=ya nasi=ya; ik-u=si</i>
EPI	<u>epistemic</u>	认识情态	<i>daroo; hazu</i>	<i>ik-u=daroo; ik-u=hazu</i>
ESS	<u>essive</u>	存格	<i>de</i>	<i>koko=de</i>
EXM	<u>exemplative</u>	例举	<i>tari</i>	<i>it-tari</i>
FOC	focus	焦点	<i>mo; sae</i>	<i>koko=mo; koko=sae</i>
GEN	genitive	属格	<i>no</i>	<i>ane=no</i>
GER	<u>gerund</u>	分词	<i>-te</i>	<i>it-te</i>
HON	<u>honorative</u>	尊敬	<i>o-; go-</i>	<i>o-denwa; go-ryoosin</i>
HORT	<u>hortative</u>	劝告	<i>-(y)oo</i>	<i>ik-oo</i>
HML	<u>humble</u> ⁸	自谦	<i>o/go-V-si-</i>	<i>o-yobi-si-; go-hookoku-si-</i>
IJL	<u>interjectional</u>	插入	<i>sa; ne</i>	<i>watasi=wa=sa; watasi=wa=ne</i>
IMP	imperative	命令	<i>-e, -ro</i>	<i>ik-e, tabe-ro</i>
INFR	<u>inferential</u>	推知	<i>mitai; yoo; -soo</i>	<i>it-ta=mitai=da; it-ta=yoo=da; iki-soo=da</i>
INS	instrumental	工具格	<i>de</i>	<i>basu=de</i>
INT	<u>interrogative</u>	疑问	<i>ka</i>	<i>ik-u=ka</i>
INTT ⁹	<u>intentional</u>	意向情态	<i>-(y)oo; tumori</i>	<i>ik-oo; ik-u=tumori</i>
JOI	<u>joining morpheme</u>	连接语素	<i>sa</i>	<i>yo-sa-soo=da</i>
LIM	<u>limitative</u>	限定	<i>dake; sika; nomi</i>	<i>koko=dake; koko=sika; koko=nomi</i>
LOC	locative	处所	<i>ni</i>	<i>koko=ni</i>
NCONJ	<u>negative conjectural</u>	否定推测	<i>mai</i>	<i>ik-u=mai</i>
NEG	negation	否定	<i>-(a)na-; -(a)zu</i>	<i>ik-ana-i; ik-azu</i>
NMLZ	nominalizer	名物化标记	<i>A/AN-sa; no/koto</i>	<i>naga/sizuka-sa; ik-u=no/=koto</i>
NOM	nominative	主格	<i>ga</i>	<i>koko=ga</i>
NPST	<u>non-past tense</u>	非过去时	<i>-(r)u</i>	<i>ik-u</i>
PASS	passive	被动	<i>-(r)are-</i>	<i>ik-are-ru</i>
PL	plural	复数	<i>-tati</i>	<i>watasi-tati</i>
PN	<u>proper name</u>	专有名称	<i>Taro; Kyuusyuu</i>	<i>Taro; Kyuusyuu</i>
POL	<u>politeness</u>	礼貌	<i>des-; -mas-</i>	<i>koko=des-; iki-mas-</i>

8 Narrog (2009) 和ナロック (2023) 写出的英文是humilitive, 当为humiliative之误, 也表示“自谦”。

9 “INTT/intentional意向情态”在ナロック (2023) 被缩写为INT, 此处与“interrogative疑问”重复, 故改为INTT。ナロック (2023) 将“interrogative疑问”缩略为QUE。

(续表)

语法标注	英文术语	中文术语	语素	举例
POT	<u>potential</u>	可能	-(r)e-	<i>ik-e-ru</i>
PRF	<u>perfect</u>	完成体	-te i-	<i>nakunat-te i-ru</i>
PROG	progressive	进行体	-te i-	<i>ason-de i-ru</i>
PROH	prohibitive	禁止	<i>na</i>	<i>ik-u=na</i>
PST	past tense	过去时	-ta	<i>it-ta</i>
PURP	purposive	目的	<i>yoo; tame</i>	<i>ik-u=yoo=ni; ik-u=tame=ni</i>
QUOT	quotative	引用	<i>to; tte</i>	<i>ik-u=to; ik-u=tte</i>
RECP	<u>reciprocal</u>	相互	-a(w)-	<i>naguri-a-u</i>
REFL	reflexive	反身代词	<i>zibun</i>	<i>zibun</i>
RES	resultative	结果体	-te i-	<i>sima-te i-ru</i>
RPT	<u>reportative</u>	闻知	<i>soo; rasi-</i>	<i>ik-u=soo=da; ik-u=rasi-i</i>
SEF	<u>sentence-final particle</u>	句末语气词	<i>ne; yo</i>	<i>ik-u=ne; ik-u=yo</i>
SG	singular	单数	<i>watasi</i>	<i>watasi</i>
SIM	<u>simultaneity</u>	同时	-nagara; -tutu	<i>yomi-nagara; yomi-tutu</i>
STAT	<u>stative</u>	状态	-te i-	<i>sugure-te i-ru</i>
TOA	<u>term of address</u>	称呼	-san; -kun	<i>Taroo-san; Taroo-kun</i>
TOP	topic	话题	<i>wa</i>	<i>koko=wa</i>
VBZ	<u>verbalizer</u>	动词化	-kar-	<i>naga-kat-ta</i>

以下对表1中使用的符号和部分语法标注做一下说明。语素一栏“-”表示该语素是前缀（如*o-*）或后缀（如*-te*），或者是后缀但仍然需要后接其他后缀（如*-ana-*），或者需要后接后缀（如*-te i-*中*i-*的后面），语素前未加“-”的除了少数独立词（如*watasi*），大多是附缀（clitic），它们在实际使用时前接“=”。语素一栏出现的大写字母表示词类，N为名词（noun），V为动词（verb）词基¹⁰，A为形容词（adjective）词干，AN为形容名词（adjectival noun），即学校语法所说的形容词。表1中有的语素有多种语法功能，已分别列出，典型如*-te i-*有“进行体”“结果体”“完成体”“状态”等多种功能，*wa*则兼有“话题”和“对比”两种功能，“劝告”和“意向”则都可用-(y)*oo*表达。有些语法标注列出的语素是完全列举，如VBZ“动词化标记”*-kar-*，有些语法标注则仅列出常见的语素，如ENU“列举”还包括*toka*和*dano*等，CAL“原因”还包括格助词*de*和*ni*等。此外，表1并未穷尽日语所有的语法功能，使用者可以参考Lehmann（2004）列出的其他语法标签并在标注时灵活处理。

表1中有三个“~化标记”，即adverbializer（副词化标记）、nominalizer（名物化标记）和verbalizer（动词化标记），ナロック（2023）提供的英文名称分别是adverbial、nominalization和verbalization，但是该文列出的语法标记中还有intensifier（强调标记，如*hik-par-*和*syop-pik-*）和complementizer（补语化标记，标句词），我们认为最好统一使用英语-er后缀的术语。另外，ナロック（2023）把日语传统语法所说的形式名词*no*分列在两处，一处归入接续助词或接续后缀（接続助詞、接続接尾辞），写作*no*，标为complementizer（补语化标记），另

10 动词词基相当于学校语法所说的动词连用形。本文4.2节会提及。

一处归入句法功能的后缀（統語機能の接尾辞），写作=*no*，标为nominalization（名物化），但两处举的例子都是*yom-u=no*。一部分迂说形式（迂言的形式（一部））又收入了complementizer，举出的例子是*koto*和*mono*，可见他的处理标准比较混乱。本文建议，除了生成语法背景的论文可以把*no*和*koto*标注为complementizer，一般的论文可以统一标注为nominalizer（名物化标记）。

关于语法标注的原则，Lehmann（2004：1841）主张应当尽量精确，例如我们不能把所有的体标记都标注为ASP，因为这样不能提供足够的信息，除非这种语言只有一种体。按此原则，日语的时范畴有过去时（past tense）与非过去时（non-past tense）的对立，示证范畴有推知（inferential）和闻知（reportative）的对立，实际标注就不需要时（tense）和示证（evidentiality）的标签，因此表1未列出这种大类的语法标注。句末语气词如「ね」「よ」等按照这个原则同样应该细分为很多不同的功能标签，但是很多句末语气词功能复杂，学界对于这些功能的分类还缺乏统一的认识，在目前的文献中还很少有精细标注，因此我们只是把表示疑问的句末语气词单独标为INT，其他均使用SFP。

3. 莱比锡标注规则与日语的转写和标注

在英文出版物中，日语的拉丁化转写一般使用训令式，但有少数作者习惯用黑本式。训令式是1937年公布的内阁训令第1号中规定的日语拉丁字母转写法（ローマ字のつづり方）。关于拨音韵素的转写方式，比较常见的是按照训令式的规定，全部使用小写字母*n*，仅在后接元音或*y*时，为防止误读，加隔音符“'”，如*gen'eki*（現役）¹¹。促音韵素与后续辅音构成长辅音，所以转写一般用双写辅音的形式，如*gakkai*（学会）。长元音的转写在训令式中规定在元音上方加“^”，但是日语语言论著为了打字方便，实际一般双写元音，如*okaasan*（お母さん）、*kiiboodo*（キーボード）、*kuuki*（空気）、*keeki*（ケーキ）、*tooru*（通る）。才段假名后接ウ实际发音是长元音*oo*，如*soodan*（相談），エ段假名后接イ虽然一般读作长元音*ee*，但基于音系原因的考虑（参看Vance 2008：63-67），按惯例写作*ei*，如*sensei*（先生）。

按照国际惯例，逐行标注一般采用三行的格式，第一行是目标语言的拉丁化转写或使用国际音标的语音转写，并同时标明语素的切分方式，第二行是逐个语素的词义或语法标注，第三行是元语言的翻译，通常放在单引号内。第二行的标注中，目标语言的词汇性成分使用元语言中语义相同或近似的词或短语进行翻译，语法性成分（如日语的屈折词缀和附缀）使用语法标注。第三行的元语言翻译使用文章的书写语言，本文统一使用中文，以下引自英文文献的例句也一律翻译为中文。此外，有些论著的例句会在最上面增加一行目标语言的原文书写形式，例如日语是汉字假名混写。为了方便熟悉日语的读者，本文每个例句增加第一行日语原文，即实际使用四行格式。以下我们以日语的语料为例介绍莱比锡标注的十条规则，语法标注使用表1所列的英文缩略词。

11 Heiko Narrog的英文论著（如Narrog 2009）对拨音韵素一般采用精细化的转写，即全部使用小型大写字母*N*，如*geNeKi*（現役），但ナロック（2023）一文只是介绍一般情况用*n*，并未提及*N*。

LGR规则1: 逐词对齐。每个词的原文与标注左侧对齐。如:

(1) 通りに出る前に、ちらりと自転車置き場に目を向けた。

Toori=ni de-ru mae=ni,

大街=ALL 出-NPST 前=LOC

tirari=to zityensya-okiba=ni me=o muke-ta

看一眼=ADVR 自行车-停放处=ALL 目光=ACC 投向-PST

‘(他)来到大街之前,向自行车停车场望了一眼。’(東野圭吾2005『容疑者Xの献身』)

LGR规则2: 语素逐个对应。可切分的语素之间用短横线“-”隔开, 附缀之间用等号“=”隔开, 例句和各语素的标注必须一一对应, 例句和标注中的语素切分符号也必须一致。日语的附缀以及语素切分的问题将在第4节介绍。

按照日语语言学英文出版物通行的做法, 很多复合词通常不做语素切分, 而是对整个词做语义标注。例如二字汉源词「教室」*kyoositu*虽然实际包括两个语素, 但一般不写做*kyoo-situ*, 英文文献一般把词义标为*classroom*, 而不是*teaching-room*。例(1)由两个复合词构成的复杂复合词「自転車置き場」则是切分为两个复合词*zityensya-okiba*, 并分别做了语义标注。日语的屈折词缀必须做出标注, 但是莱比锡标注规则没有说明派生词缀是否需要全部标注, 因此日语可做灵活处理。我们认为, 如果作者希望做精细标注, 那么可以把所有的派生词缀都切分出来。例如「子供らしい」*kodomo-rasi-i*(孩子-像-NPST ‘像孩子一样的’)中, *-rasi-*是派生后缀。一般情况下, 派生词缀不必切分, 用整词对译即可, 例如「運転手」*untensyu*(‘司机’)不必写作*unten-syu*。此外, 使用频率较低的派生词缀没有必要专门列出标签, 如ナロック(2023)举出的强调标记的例子*syop-pik-*‘强拉’, 前缀*syop-*仅用于这一个词, 这种情况下通常不需要专门切分出来。

LGR规则3: 语法标注。LGR的附录列出了常见的语法范畴名称及其缩略词, 本文表1已经列出日语主要的语法标注。在日语语言学英文出版物中, 语法标注通常使用小型大写字母, 但也有些论著使用普通的大写字母, 比如《日本语言与语言学手册》。LGR提到, 缩略词可以使用不同的形式, 如果一个范畴在某一语言中高频使用, 更短的缩写词会更方便, 例如“完结体”*COMPL*也可以写作*CPL*。由于莱比锡规则只列出了跨语言常见的语法范畴, 特定语言的很多范畴没有包括在内, 此时作者可以自己设计缩略形式。例如LGR没有收录*NONPAST*“非过去时”, 它在日语中是一个常用的屈折词缀, 《日本语言与语言学手册》缩写为*NONPST*, 而ナロック(2023)缩写为*NPS*。因此, 只要不会与其他标签相混, 原则上缩略形式比较自由。LGR还提到, 在很多情况下, 根据标注的需要, 元语言既可以采用语法标注, 也可以使用对应的词来标注, 并举出了俄语文本中俄语的介词用英语介词标注的例子。然而, 正如Lehmann(2004: 1840-1841)所指出的那样, 我们不能用元语言的自由语素标注目标语言的黏着性语法语素或派生语素。由于日语中语法性强的范畴基本采用黏着语素, 所以例(2)把*kara*标注为“因为”、把*to*标注为“和”的标注方式不妥。虽然这样的标注方式能够反映目标语言中对应语素的语义, 但是容易被读者误以为它们与元语言的词汇形式有同样的语法地位, 因此我们

建议使用例(2)第一个标注行的语法标签模式进行标注。

(2) 天気がいいから、娘と散歩しましょう。

<i>Tenki=ga</i>	<i>i-i=kara,</i>	<i>musume=to</i>	<i>sanpo-si-mas-yoo.</i>
天气 =NOM	好 -NPST=CAL	女儿 =COM	散步 - 做 -POL-HORT
天气 =NOM	好 -NPST= 因为	女儿 = 和	散步 - 做 -POL-HORT

‘天气很好，咱们和女儿去散步吧。’

LGR规则4：一对多的对应关系。目标语言的一个成分对应元语言的多个成分时，元语言用英语的句号“.”隔开。例如，日语的系词「です」兼有“礼貌”的特征，应该标注为 *des=u* (COP.POL=NPST)。再如，日语口语中由于语音省略或语音融合造成语素界限模糊时，需要用一对多来标注，例如「そうじゃない」*Soo=zya na-i.* (那样=COP.TOP NEG-NPST ‘不是那样。’)。

除了上述一对多，还有多对一的现象。例如 *kaomisiri* (顔見知り) 按语素切分应该是 *kao-mi-siri* ‘脸-看-知道’，但是除了研究构词的论文中需要这样切分和标注外，一般论文中对于此类词不需要做语素切分，把整个词一起标注反而读者更容易理解，即 *kaomisiri* ‘相识’。日语中有大量表示情态的迂说结构，它们一般也不需要按语素切分，如 *kamosirena-i* (EPI-NPST)，其中 *kamosirena*-含有五个语素 (*ka-mo-sir-e-na-*)，但是整体作为认识情态的标记不需要切分，只是由于最后一个语素 *-i* 有屈折变化 (可以出现过去时)，所以此处单独切分出来。

LGR规则5：人称与数的标签。动词有人称和数范畴屈折变化的语言中，人称与数是动词标注的必要内容。日语动词没有这两个语法范畴，因此人称和数只用于标注人称代词，这种情况下也可以用汉语相应的人称代词来标注，复数人称代词也可以把复数后缀单独标注，如例(3)的 *watasitati* 还可以标注为 *watasi-tati* (我-PL)。为了简洁，数与人称之间通常不加“.”。

(3) 彼は私たちの友達です。

<i>Kare=wa</i>	<i>watasitati=no</i>	<i>tomodati=des-u.</i>
3SG=TOP	1PL=GEN	朋友 =COP.POL-NPST
他 =TOP	我们 =GEN	朋友 =COP.POL-NPST

‘他是我们的朋友。’

LGR规则6：非显性成分。如果标注对象在目标语言里是非显性成分，可以用中括号把其对应的语法标签标记出来。LGR举出了拉丁语主格单数名词的例子。日语中通常不标注非显性成分，如果确有必要标出，例如在例(4)宾格助词省略的情况下，可如下标注。

(4) みんなでお昼ご飯食べた。

<i>Minna=de</i>	<i>o-hirugohan</i>	<i>tabe-ta.</i>
大家 =INS	HON- 午饭 [=ACC]	吃 -PST

‘大家一起吃了午饭。’

LGR规则7: 固有范畴。有些语言的固有非显性范畴, 如性范畴, 可以体现在标注中。日语没有相关范畴。

LGR规则8: 二分成分。如果一个成分的两部分被其他成分分离, 两个部分可以用同样的词语或缩写标注, 也可以对其中一个使用特殊的标签。日语动词的尊敬表达「お~になる」和自谦表达「お~する」可以看作二分成分, 如下例的自谦表达中*o-*和*-si-*都标记为HML。

(5) お話しします
o-hanasi-si-mas-u
 HML-说-HML-POL-NPST
 ‘(我)说’

LGR规则9: 中缀。中缀在目标语言和元语言中都需要加尖括号。日语中没有该现象。

LGR规则10: 重叠。重叠的处理方式与加缀相似, 但是用“~”连接词干与重叠的部分¹²。日语中没有LGR举出的如希伯来语和他加禄语那样部分重叠词干的现象, 但是有大量由词基完全重叠构成的叠词, 如きらきら*kirakira*。因为此类重叠大多是构词性形态, 不涉及语法功能, 因此可以不做语素切分, 用词汇义标注即可, 也可以用“-”隔开(如*kira-kira*), 两个语素均标注同样的词汇义。

4 日语语素切分的争议问题

4.1 助词的形态地位

日语学校语法所说的“助词”从普通语言学角度看通常不被认定为独立词, 不过它们的形态地位一直存在争议, 在实际的文本标注中, 不同的作者或同一作者的不同论著经常使用不同的切分方式, 有的采用与独立词相同的书写形式, 即在助词前加空格, 有的标为后缀(前加“-”), 还有的标为附缀(前加“=”)。格助词的这三种切分方式在柴谷方良的论著中都曾使用过(Dalla Chiesa 2020: 517-518)。例(6)展示的是柴谷方良和影山太郎主编的《日本语言与语言学手册》丛书中的处理方式, 即把助词按照独立词的方式切分, 但是该丛书收录的Kageyama(2016)一文把它们看作附属词, 区别于独立词。我们认为, 《日本语言与语言学手册》这样处理不过是为了统一丛书的书写体例, 例(6)的书写方式未必表示各章作者或丛书编者都认为助词是独立词。英文期刊《日语语言学报》最近几年刊载的论文中, 日语助词主要采用两种书写形式, 即前加空格或前加“-”。由此可见, 对于助词的形态地位及其标注方式, 学界并未达成一致的观点, 而且由于讨论日语附缀的文献还不多, 日语语言学研究者们往往回避使用“=”, 在例句中仅采用粗略的形态分析。

12 Narrog(2009)等文献在日语复合词和叠词的词基之间用波浪号“~”连接。Lehmann(2004)的第十四条规则称, 复合和派生词干的界线用加号“+”表示, 第十九条规则与LGR第十条规则相同, 即用“~”连接词干与重叠的部分。鉴于日语的重叠模式比较简单, 重叠的语素之间可以与构成更大复合词的两个复合词之间采用相同的符号, 即或者采用Narrog(2009)使用的“~”, 或者采用简化的方式统一使用“-”, 本文暂采用后者。

(6) 太郎は次郎と東京へ行って靴下を買った。

Taroo wa Ziroo to Tookyoo e it-te kutusita o kat-ta.
太郎 TOP 次郎 COM 东京 ALL 去 -GER 袜子 ACC 买 -PST
‘太郎跟次郎去了东京，买了袜子。’ (Shibatani & Kageyama 2016: xviii)

在传统的形态学研究中，非自由语素包括黏着词干、词缀和附缀 (Haspelmath & Sims 2010: 196)。日语的黏着词干争议不大，但附缀与词缀的界定问题在文献中讨论不多，还存在一些模糊之处。附缀是句法上独立而音系上依附于宿主的一类形态单位，它可以构成独立的语法词，但不能成为音系词，即不具备音系词的音段和韵律特征。普通语言学文献通常把黏着性介于词缀和独立词之间的非实义形态成分归入附缀，这也是一些日语研究文献 (如 Jarkey 2020) 的观点，即把格助词、小句接续助词、句末语气词和系词等都看作附缀。

Vance (1993) 利用 Zwicky & Pullum (1983) 和 Zwicky (1985) 提出的附缀测试标准对日语格助词、并列助词和焦点助词 (即传统语法所说的「副助詞」或「とりたて助詞」) 进行了测试，该文的结论是很多焦点助词是独立词，而其他助词比英语的介词更像是独立词。由于 Zwicky (1985) 把英语的介词看作附缀，Vance (1993) 认为日语的助词不是附缀，但并未说明它们是否属于独立词。Kageyama (2016) 进一步提出日语助词是介于附缀和独立词之间的“附属词 (non-independent word)”。附属词的概念由服部 (1950) 首次提出，他使用三条标准对多种语言的附属词进行了界定。Kageyama (2016) 在此基础上使用了新的测试手段论证了日语附属词与附缀的区别，表2是他对日语四类形态单位特征的归纳。虽然表中显示附属词和附缀的差异在于形态独立性的有无，但 Kageyama 在文中并没有明确指出应该如何判定形态独立性，也就是说附属词和附缀如何划界还有待商榷。此外，Kageyama (2016) 从句法辖域的角度把一些传统上视作屈折词缀和派生词缀的语素归入附缀，但是文中仅举出数例，并未进一步说明日语附缀还包括哪些语素。按照他的判定标准，动词后接的所有屈折词缀似乎都可以看作附缀，但这样就会产生一个新的问题，即日语和其他语言中动词的屈折词缀是否都是附缀，或者说附缀与屈折词缀应该如何区分。Kageyama (2016) 对于日语附属词和附缀的重新界定为日语黏着语素的分类提供了新的研究思路，但是如上所述，它们的判定还缺乏清晰的标准，而且目前我们还没有看到其他学者支持这种新分类，因此本文对表2的分类持保留意见。在上文表1以及第3节的例句中涉及语素切分的地方，我们仍然支持ナロック (2023) 所持的传统观点，把既非独立词又非词缀的成分暂时认定为附缀¹³，使用“=”与它们的前接成分分隔开。

13 铃木 (1972: 第5章) 把学校语法所说的“助词”称作「くつつき」，包括「格のくつつき」「とりたてのくつつき」「ならべのくつつき」。「格のくつつき」是名词的词形变化，它不同于动词的屈折后缀，但又是黏着在名词上 (铃木1972: 204)。这种观点与ナロック (2023) 的观点类似。

表2 独立词、附属词、附缀和词缀的特征 (Kageyama 2016: 519)

	独立词	附属词	附缀	词缀
句法地位	有	有	有	无
形态独立性	有	有	无	无
单独使用	可	否	否	否

4.2 动词后接的屈折词缀

莱比锡标注规则的第二条中有一项补充说明,即语素界线的增音应该分派给前面的语素或后面的语素,具体分派给哪个语素会由多种原则决定,很难给出统一的原则,因此该规则中没有提出具体的标准。日语的形态分析中有一些增音现象,如表1中的yo-sa-soo(好-JOI-INFR),按照ナロック(2023)的分析,-sa-标注为JOI即“连接语素”,鉴于-soo前接形容词时以词干为主,这一处理方式非常合理。其他一些与动词词尾和屈折后缀相关的增音现象则不容易处理。例如动词ik-‘去’后接否定义屈折后缀时有三种可能的分析方式:ika-na-i、ik-a-na-i和ik-ana-i(西山2016: 102)。第一种即增音a与词干ik-组合是学校语法的分析方式,由于ik-之类的动词分析为辅音词干动词已经是日语形态学的通识,ika-作为一个语素是不成立的。第二种-a-被分析为如同yo-sa-soo中的-sa-同样的增音,表面上讲得通,但是-sa-仅出现与yo-和na-两个形容词之后,它作为连接语素可以看作例外现象,而ik-a-na-i中的-a-是规则性的插入,即所有辅音词干动词后都要插入无实际语义的-a-才能后接-na-,这样的话就增加了分析的复杂性。相比之下,第三种分析最为俭约,-ana-被分析为-na-的语素变体,它们分别出现在辅音词干动词和元音词干动词之后,以满足日语辅音和元音交替出现的音位配列规则,而且由于-a-的插入无法预测(即为什么不插入其他元音,没有语音学和音系学的理据支持),-na-被分析为-ana-脱落-a-的结果则非常自然。

与-(a)na-相似的分析方法还适用于非过去时后缀-(r)u、致使后缀-(s)ase-、被动后缀-(r)are-、条件后缀-(r)eba以及劝告/意向后缀-(y)oo等。然而,同样的分析在应用于增音-i-时会遇到问题。如果按照以上的分析,ik-imas-u(去-POL-NPST)、ik-ita-i(去-DES-NPST)、ik-inagara(去-SIM)和ik-itutu(去-SIM)等动词后接的屈折后缀时,以礼貌后缀-imas-为例,它与-mas-构成语素变体,前者接辅音词干动词后,后者接元音词干动词后。另一方面,不同于上述的ika,iki可以单独使用,即学校语法所说的动词连用形的中顿用法,此时的-i可以按照Shibatani(2009: 749)的分析方式标注为INF(infinitive,不定式),按此思路,ik-i-mas-u(去-INF-POL-NPST)也是一种可能的分析方式。更进一步看,iki-还可以后接多个派生后缀,如-yasu-i(-容易-NPST)、-niku-i(-难-NPST)、-kane-ru(-难-NPST)等等,此时又出现了三种可能:-i-与动词词干合并(如iki-yasu-i)、-i-单独分析为一个不定式语素INF或无语义的连接语素JOI(如ik-i-yasu-i)、-i-与派生后缀合并(如ik-iyasu-i)。虽然第一种分析方式在上文谈及ika-na-i时已经被否定,但是由于iki与ika不同,它可以单独使用,因此这种分析方式成为了一个可选项,实际上Narrog(2009: 74)把ik-称作stem(词干),而把iki称作base(词基),该书中出现的所

有词基（如*iki-*）后接词缀的组合全部把词基分析为一个语素，如*iki-mas-u*等。由于*iki-*这样的词基后接词缀众多，Narrog（2009）的分析方式不失为一种非常俭约的做法，本文以上的分析均采用这种方式。第二种和第三种分析方式虽然在理论上都讲得通，但是与第一种相比标注稍显繁琐，因此从经济性角度看第一种方式更为理想。

5. 结语

莱比锡标注规则指出，一个语言的形态模式经常有多种分析方式，因此逐行标注不过是提供了一种可能的分析方法，也就是为文本或例句的结构提供了一些仅凭元语言翻译无法体现的语言信息。规则还指出，标注是分析的一部分，而不是数据的一部分，因此从其他已出版的文献引用例句时，如果作者希望采用不同的术语、风格或分析方式，标注形式可以相应做出改变。我们认为，虽然逐行标注不是严格的形态分析，但是应该力求内部逻辑的一致性，而且要尽量反映客观的语素切分方式，因此我们应该避免出现Dalla Chiesa（2020）指出的同一论著内的混乱现象。

本文主要基于莱比锡标注规则介绍并分析了以往文献中日语语法标注和语素切分的若干问题，并提出了我们的观点和建议。由于日语附缀与词缀的划界问题以及服部四郎和影山太郎论证的附属词的概念是否具有理论意义等问题还没有明确的答案，本文建议使用的语素切分和语法标注等方案并非定论，随着今后相关研究的进展，我们期待学界制定出更加严谨的标注规则。另一方面，出于国内不同语种之间语言学学术交流的需要，我们提议日语语言学的中文出版物中加入逐行标注，并期待更多的研究者参与讨论语法标注的规范化问题。

参考文献

- 陈玉洁、Hilario de Sousa、王健、倪星星、李旭平、陈伟蓉、Hilary Chappell, 2014, 莱比锡标注系统及其在汉语语法研究中的应用[J], 《方言》(1): 1-13。
- 黄成龙, 2005, 语法描写框架及术语的标记[J], 《民族语文》(3): 23-33。
- 罗天华, 2023, 莱曼标注和语法标注的规范问题[J], 《当代语言学》25(4): 577-594。
- 鈴木重幸, 1972, 日本語文法・形態論[M]。東京：むぎ書房。
- ナロック・ハイコ, 2023, 日本語をテーマとした英語論文の書き方についての覚書[J], 『日本語の研究』19(2): 89-105。
- 西山國雄, 2016, 第4章 屈折形態論：日本語動詞の活用と英語の不規則動詞[A], 漆原朗子（編），『形態論』[C]。東京：くろしお出版，84-112。
- 服部四郎, 1950, 付属語と付属形式[J], 『言語研究』15: 1-26。
- Comrie, Bernard, Martin Haspelmath & Balthasar Bickel. 2015. The Leipzig glossing rules: Conventions for interlinear morpheme-by-morpheme glosses [A]. <https://www.eva.mpg.de/lingua/resources/glossing-rules.php>
- Dalla Chiesa, Simone. 2020. Troubles on the border: Marking Japanese case-particle boundaries in grammatical annotation [J]. *Annali di Ca' Foscari. Serie orientale* 56: 501-562.

- Haspelmath, Martin & Andrea D. Sims. 2010. *Understanding Morphology* (2nd edn) [M]. London: Hodder Education.
- Jarkey, Nerida. 2020. Words in Japanese [A]. In Alexandra Y. Aikhenvald, R. M. W. Dixon & Nathan M. White (eds.), *Phonological Word and Grammatical Word: A Cross-Linguistic Typology* [C], 39-77. Oxford: Oxford University Press.
- Kageyama, Taro. 2016. Lexical integrity and the morphology-syntax interface [A]. In Taro Kageyama & Hideki Kishimoto (eds.), *Handbook of Japanese Lexicon and Word Formation* [C], 489-528. Berlin: Mouton de Gruyter.
- Lehmann, Christian. 1982. Directions for interlinear morphemic translations [J]. *Folia Linguistica* 16: 199-224.
- Lehmann, Christian. 2004. Interlinear morphemic glossing [A]. In Geert Booij et al. (eds), *Morphology: An International Handbook on Inflection and Word-Formation* [C], vol. 2, 1834-1857. Berlin: Mouton de Gruyter.
- Narrog, Heiko. 2009. *Modality in Japanese: The Layered Structure of the Clause and Hierarchies of Functional Categories* [M]. Amsterdam: John Benjamins.
- Shibatani, Masayoshi. 2009. Japanese [A]. In Bernard Comrie (ed.), *The World's Major Languages* (2nd edn) [C], 741-763. London: Routledge.
- Shibatani, Masatoshi & Taro Kageyama. 2016. Introduction to the *Handbooks of Japanese Language and Linguistics* [A]. In Taro Kageyama & Hideki Kishimoto (eds.), *Handbook of Japanese Lexicon and Word Formation* [C], vii-xxxiii. Berlin: Mouton de Gruyter.
- Vance, Timothy J. 1993. Are Japanese particles clitics? [J] *Journal of the Association of Teachers of Japanese* 27(1): 3-33.
- Vance, Timothy J. 2008. *The Sounds of Japanese* [M]. Cambridge: Cambridge University Press.
- Zwicky, Arnold M. 1985. Clitics and particles [J]. *Language* 61: 283-305.
- Zwicky, Arnold M. & Geoffrey K. Pullum. 1983. Cliticization vs. inflection: English *n't* [J]. *Language* 59: 502-551.

通信地址：中国澳门氹仔大学大马路澳门大学人文学院（E21A）日文系3111室
电子邮箱：wenjiangyang@um.edu.mo

a notable decline in the number of published papers is observed, both in Japanese language and literature journals as well as comprehensive foreign language journals. One of the significant decreases is observed in the quantity of vocabulary-related research papers. In terms of linguistic theory, cognitive linguistics maintains its prominent position, with a pronounced focus on concepts such as metaphor, spatial representation, and subjectivity. Furthermore, Japanese media language research emerges as the most prominent growth area, indicating a rising interest and visibility in regional studies within the linguistic domain and suggesting the progress of applied research in fundamental disciplines.

Keywords: Chinese National Knowledge Infrastructure (CNKI); Japanese language; the number of published papers; cognitive linguistics; regional studies

Morphemic glossing rules for the Japanese language, by Yang Wenjiang(p.70)

The rules for Japanese morphemic glossing are not extensively discussed in literature, and have not been widely applied in Chinese publications of Japanese linguistics. Moreover, in English publications abroad, there exist non-standard practices. This article introduces the application of the Leipzig Glossing Rules in Japanese texts, and through analyzing issues and controversies related to morphemic glossing and morpheme segmentation in previous research, presents preliminary rules for Japanese morphemic glossing.

Keywords: morphemic glossing; the Japanese language; morpheme segmentation; the Leipzig Glossing Rules

A study of patterns and cognitive factors of encoding non-intentional resultative events of Japanese, by Song Jiaxin & Yao Yanling(p.83)

Non-intentional resultative events can be encoded through various linguistic constructions, including intransitive and transitive forms. The goal of this paper is to investigate the cognitive processes and construction selection of Japanese, and to figure out the influence factors of encoding non-intentional resultative events.

The results indicates that the cause of non-intentional resultative events has influence on cognition and encoding of Japanese. Consciousness of responsibility, availability and the selection ratio of transitive construction shows positive correlations. Furthermore the causal conditions of non-intentional resultative events have influence on event cognition stage encoded. With diminishing of clarity or controllability of cause, the cognition stage of Japanese shifts from cause to effect. Lastly cognition is closely related to the selection of constructions. The higher the consciousness of responsibility is, the higher ratio of transitive constructions will be, on the contrary, the lower the consciousness of responsibility is, the higher ratio of intransitive constructions will be.

Keywords: Non-intention; resultative events; construction selection; cognitive factors

A study on the formation of transitive sentences with inanimate subjects in Japanese, by Che Luming & Hu Jun (p.98)

Based on a corpus of written Japanese, this study investigates the usage of inanimate subject transitive verb constructions in Japanese and explores the factors influencing their formation. The results