

21世纪的田野调查*

——基于自然语料、归纳法和目的语母语人的认知范畴

罗仁地 (Randy J. LaPolla) 卢琳

提 要：这篇文章主要介绍怎样做21世纪的田野调查和语言记录描写，是基于作者的经验之谈，以及在田野调查中遇到的问题及其解决办法。本文所提出的建议包括：调查的时候应该尽量在使用目的语的社群中，选择沉浸式“参与观察者”（participant observer）的方式来收集自然语料（母语调查者同样要收集自然语料）；转写工作也最好在当地进行；最好先做社会语言学调查；调查和分析时要基于目的语母语者的认知范畴；应该使用归纳法来分析所搜集的自然语料。

关键词：语言学田野调查 语言类型学 人类语言学 功能语言学

○ 引言

语言是一种工具，跟其他工具一样，为了了解其形式的所以然，一定要了解其用法（罗仁地 2007）^①，了解语言用法需要对其进行细致观察，而观察语言用法的最佳方式是通过沉浸式田野调查。语言类型学是所有语言学的基础，目的是研究语言所呈现形式的多样性和共性，为了提高类型学的科学性和概括性，要加强对语言事实的了解，只能通过调查和描写现存的语言。除此之外，调查和描写一门陌生的语言也是学生学习语言学最好的训练方式，因为在撰写参考语法时需要学习和描写语言的各个方面。同时，调查时所收集的语料也会成为学生们取之不尽用之不竭的学术储备，可在将来用于撰写文章和进行更深入的分析。

本文讨论的是怎么做21世纪的田野调查和语言记录描写，是基于作者的经验之谈，很多意见和建议都来源于作者在田野调查过程中实际遇到的问题和棘手之处，以及因此所做的相关总结。20世纪的田野调查分析和21世纪的主要差别在于：（1）调查者应该在母语者社群里搜集自然语料，切勿使用翻译词表和句子的调查方式；（2）用归纳法来对自然语料进行分析，而不是先入为主地把别的语言中的范畴强行套加在目的语上；（3）调查和分析应基于母语者的认知范畴，尊重不同文化背景母语者认知范畴的多样性，不应断然地认为所有语言的词和结构所反映的认知范畴是一样的。

一 尽量在使用目的语的社群中调查

不同的研究目的和对象对应不同的调查方式，但如果调查的目的是记录和描写自然语言，那么在使用目

收稿日期：2023-12-06；定稿日期：2024-04-16。通讯作者：罗仁地。

* 在本文的撰写过程中，郑武曦老师提出了许多宝贵意见，在此深表谢意。

① 罗仁地（2007）更详细地讨论了记录语言的各个方面，如怎么选择调查目的语言，开始前所需要做的准备和具体的调查过程等。本文着重于讨论21世纪宏观视角下的田野调查。

的语的当地社群，选择沉浸式“参与观察者”（participant observer）的方式来收集自然语料是最科学的。母语研究者也应该选择这样的方式。^① 调查者在调查时不能只靠自己的想象，应该置身于目的语使用社群观察和收集自然语料，这样会发现很多意想不到的语言和其他文化、思维方式所呈现的现象。因为在实地调查时，可以随时听到自然的会话和其他语体的语料，还可以当场指着物品或在观察到某现象时跟母语者问其名询其用，这样调查者会对目的语言的用法有更为清晰和可靠的认识。在实地调查时，可以同时邀请多位合作者（其实整个社群成员都可以成为你的合作者）一起帮忙搜集语料，调查者应该设法与该社群建立良好的长期合作关系。

语言与其使用地点的诸多方面都具有相关性（Enfield 2002；De Busser and LaPolla 2015；Fiddler 2023）。如果调查者不在当地，则很难了解所使用语言及其与地理、气候、当地植被与动物、农耕文化、游牧文化、烹饪和饮食文化、建筑风格及相关文化、人体特征等各方面的关系。例如，地理环境对认知方向系统有决定性的影响，如果不了解当地的地理特征，则很难理解母语者语言中所表达的方向系统。再比如，农耕文化和当地的气候对建筑风格会有影响，也会在语言中有所反映。另外，如果调查者只熟悉或生活在农耕文化背景中，不实地置身于游牧环境，则很难收集和了解有关游牧文化的用语，等等。比如作者卢琳在采用沉浸式的方法调查四川省雅江县俄托藏语时，发现和记录了许多当地人只在游牧过程中和冬夏转换草场时的用语。

二 社会语言学调查

调查者在搜集语料之前最好先做相关的社会语言学调查，包括但不限于语言的使用数据，调查地双语、三语的使用情况，教学及学生学校用语情况，地理特征，宗教或其他对当地语言或语言使用有影响的各个方面，以及母语者对他们本民族所使用语言的语言态度等。这些方面不能局限于某个人的主观认识或介绍，需要调查者进行实地调查。例如在了解双语、三语使用情况时，语言使用社群的比例、各语言的流利度、双语或三语产生和持续的时间等各方面相关内容，都只有通过调查者在当地的实地调查才能获得。

三 合作人的选择

调查者在寻找合作人时，如果目的语是调查者完全陌生的语言，可以先试着找双语者做个引介，逐渐熟悉和了解目的语，然后尽快开始找单语母语者。如果要调查少数民族语言，最好能找到不懂汉语或不怎么接触和使用汉语的母语者。在调查汉语方言时也尽量要找与普通话接触较少或平时不常使用普通话的母语者。这样的母语者才是比较理想的合作者，因为他们还保留本民族的认知思维方式以及目的语本身的语言范畴。汉语方言和少数民族语言双语或三语者、或者普通话和方言的双语者时常会因为汉语方言或普通话的影响而给出以汉语主导的认知范畴和语言形式，因而影响调查者对目的语的认识。

此外，调查者一定要尽可能地融入目的语的使用社群，努力与当地建立起信任关系，最理想的是能够与合作者像家人一般相处，互帮互助。调查者跟合作者之间应该是合作关系，而不是雇佣关系。当然，如果合作者为了帮助调查者付出了时间，那就应该要以合作者认为合适的方式来回馈他们的付出，具体形式可以是有形或者无形的各种物质。如何选择需要看调查者对当地社群文化的观察和了解程度，及其与合作者之间的关系来决定。但无论选择哪种方式，双方都应该处于平等的合作关系。

^① 参看本期《语言学论丛》中郑武曦（2024）一文。

四 调查记录方式

调查者观察和记录的对象必须是自然的语言材料,且数量越多越好,同时应该包括各种语体,如双人(或多人)对话、日常会话、传统故事(如创世纪、迁徙故事)和现代故事(程序类的,如怎么制造各种工具和其他日用品、其用法是什么等;叙述类的,如觅食、打猎、烹调等过程和期间的故事),还有各种节日、习俗(婚丧喜庆)用语、格言、诗歌、音乐及歌唱方式、人和事物的命名方法,等等。这不仅是为了搜集有关当地的文化知识,更重要的是,不同的语体可能呈现不同的语法现象,例如,在独龙语中,某种成段的语法结构只出现在程序性叙述语料里(LaPolla 2001),而不见于其他语体。

如果目的语对调查者来说是完全陌生的,在开始时可以使用一些无语言的“启发材料”(如小视频和漫画书等),请合作者根据材料里的故事或行为来进行讲述。此时调查者已经熟悉启发材料里的故事或情节,因而比较容易理解合作者所讲的话[这里提供一些有用的启发材料和有关田野调查、语言记录、语法写作、字典编纂、语言存档等方面的信息^①,此外可参考戴庆厦等(2008)]。

另外,调查者一定要尽可能地学习目的语,在学习目的语的同时可以通过使用目的语来检验你对目的语的了解程度。如果按照你的理解来遣词造句,但是当地母语者听不懂,或者误会你的意思,那么你的理解可能有误。

除了语言学家以外,目的语使用者和他们的后代都可能是调查者所做语言记录的使用者。在考虑要记录什么、怎么记录之前,最好照顾到目的语使用者的需求,了解他们期望的成果形式,调查者应该尽量帮助他们完成。调查过程中务必注意保护合作者的隐私,询问他们是否允许你使用他们的真名和有关他们的其他信息进行写作或出版,若有顾忌则需要尊重当地人的意见进行匿名处理。此外,调查者应该主动告知合作者录制内容的用途,如果他们认为不能录音录像,或者录完之后不能让别人听到或看到,调查者也应该尊重当地人的意见。^②

罗仁地在调查缅甸的日旺语时,目的语母语者认为别人看不起他们的语言,希望罗仁地给他们编一部词典,来证明他们所说的话是有价值和地位的语言。虽然编词典不是当时调查的重点和目的,但还是按照当地人的意愿为他们编成了词典(LaPolla and Sangdong 2015),来帮助他们提高自己的语言自信。^③

在实地调查时,如何选择工作地点和很多其他的相关问题,也要尊重合作者的意愿,应该以他们的便利为主。切记,母语者是在帮助调查者,因为调查者的工作实际上是在占用当地人的时间,调查者应该时刻保持谦逊的姿态,全力融入当地人的日常生活,再酌情调整和计划自己的研究。

有关调查工具的选择方面,切记调查时不要只使用电脑记录语言,因为电脑有可能会出现故障,而且更新迭代很快,过一段时间电脑软件升级后,很有可能会打不开原来所记录的文档和信息。所以在调查中应同时使用纸质笔记本记录,最好是线圈笔记本,这样方便一手拿本子,一手写字,如图 1 所示。



图1 线圈笔记本

① 参看网址: <https://www.dropbox.com/sh/nqlz0w600j18zzw/AADVfXaqxHpWbQzm6LPT8la?dl=0>, 访问日期: 2023年10月16日。

② 现在不少学者为了更客观地了解说话者的来源地,还会搜集遗传基因测试资料,但采用这样的方式要特别注意保护样本提供者的隐私。

③ 如果用 Toolbox 或 FLEx 这样的软件输入所收集的长篇语料,很容易总结和整理出词典。

此外，随身携带纸质笔记本也方便随时记录所听到的口语或受口语启发的研究思路等。尽量不要总待在室内记录语言，应该试着多跟合作者去户外走一走，了解周围的环境、植物、动物等现象。在条件允许的情况下，除了纸质笔记本，调查者也可以同时使用手机或者别的录音、录像设备来捕捉自然口语。

在搜集语料后的标注方面：由于调查中常常不止跟一位母语合作者工作，每次记录的语料（无论是在专门的录制场景中搜集的还是在不同场景里听到的），或者其他相关的文化信息，都要注明是哪一位合作者所提供的何种类型的语料。这主要是由于调查者可能会在将来的研究中发现某一位合作者提供的语料和之后的了解有出入，或者想继续深入研究某种概念的相关问题，但只要做好标注便不会有问题。

此外，在记录长篇语料的时候最好同时使用录音设备和录像设备进行留档：一是由于长篇语料非常珍贵，如果录音出了问题，可以使用录像进行弥补，反之亦然；二是由于录像中可以观察到很多合作者在说话时的表情和动作，这些非语言的沟通形式对于调查者理解和进行后期的语言研究也非常重要，因为沟通的形式和意义的创造不只局限于有声语言之中。

另外还需要注意：调查者应该尽量在当地进行转写，大部分情况下，当调查者离开当地以后很难再有效地进行语料的转写工作。这主要是由于语料中常常会出现对于调查者来说很多种新的用法，需要在母语者的帮助下才能做到准确转写。此外，转写和分析语料工作需要投入大量的时间，而理想的语料提供者一般年龄偏大，他们可能没有精力来跟调查者一起进行这项工作。因此，留在当地进行转写，也方便调查者可以找到合适的、愿意花时间的母语者作为主要的转写合作者。

如果因为战争、政治、自然灾害等不可控因素而无法到目的语的主要使用社群进行调查，那么调查者也可以设法在自己所住的城市，或别的地方找到使用目的语的社群或个人来进行合作。比如 2010 年，因为战争，罗仁地没办法到缅甸北部去调查日旺语，但了解到在泰国西北部设有日旺人难民村，于是就选择到泰国进行调查。

如果调查者在异地调查过程中只能找到一位母语合作者，这种情况实际上并不理想，因此要酌情进行调查和语料的收集。因为如果合作者长期不在社群里生活，会很大程度上受到别的语言的影响，而居住在异地则意味着没有其他母语者可以证实该合作者所提供的语料。不过调查者还是可以了解和搜集一些有关自己调查语言的基本语料和其他的相关信息，但若要使用这种情况下搜集的语料撰写或发表文章，必须对母语合作者当时的语言背景情况进行说明。

五 调查和分析要基于目的语使用社群的认知范畴

每种语言都是独特的，我们进行语言研究是为了了解每种语言所呈现出的母语者特有的思维方式。20 世纪的语言调查中，使用词汇表或定式的句子请合作者进行翻译是很常见的调查方法，但这样做会导致母语者（特别是双语或多语母语者）使用调查者工作语言的范畴和形式来进行回答，因而不能体现出目的语的真实面貌和其反映的思维方式。在 21 世纪的语言调查中，尤其是现今人工智能极速发展的时代背景下，搜集基于用法的自然语料，无论是对认识真正的语言多样性，还是对提高人工智能所需数据的科学性，都大有裨益。

进行语言分析的时候要用归纳法，即按照自然语料所呈现的现象描写该语言，避免先入为主地以别的语言的范畴概念套用在目的语上。例如在搜集词汇时可以邀请多位合作者同时跟你合作，调查者可以提出不同的语义领域，如“动物”“自然现象”“宗教活动”“民族特点”“食物”等，请合作者就着所提出领域里的词汇、用法和相关的文化知识进行介绍。

在记录词汇的时候，要注意词汇标签所反映的核心现象、使用范围和跨度都可能跟汉语（或工作时所使用的其他语言）不一样，一定要加问母语者“这个词可以用在什么样的情况下？最典型的例子是什么”。同时调查者一定要注意自己的认知范畴和母语者的认知范畴是不一样的，要按照母语者的范畴来进行语言

记录。比如说，如果用汉语来记录英语，问图 2 里的植物是什么，英语母语者会回答“pumpkin”，汉语者很可能会总结出 pumpkin 即是指“南瓜”，因此看到图 3 里的植物时也会记录为“南瓜”^①，因为汉语母语者思维范畴中对这两种植物并不做区分。但对于英语母语者来说这两种植物完全不同，图 3 中的植物称作 squash，而不是 pumpkin。此时若只记录了“南瓜”，那么就忽略了英语母语者在这个事物概念认知范畴上的区别。有关这个概念的误用随处可见。如图 4 中的实物是 squash，但因为是华人公司的广告，是以汉语思维为主导，便命名为 Pumpkin Crackers。再看图 5 中的植物，汉语母语者可能会认为与图 3 的植物不一样，但对英语母语者来说，这也是一种 squash。



图2 pumpkin



图3 squash



图4 Pumpkin Crackers



图5 squash

再如，英语的 soup 和汉语的“汤”也无法直接对应，如果调查者只记录 soup 为“汤”的话，那么就又忽略了两种语言认知范畴的差别，而会导致错误地描写和其他学者的误用。英文中 soup 的核心概念（即最典型的例子）通常指很稠的汤羹（类似汉语中的“羹”，但也不能完全等同），如图 6 里的 mushroom soup。此外，英语中 soup 的使用范围和汉语的“汤”也不一样。比如英语母语者会说 eat soup，而汉语母语者则说“喝汤”。罗仁地在某篇讨论中国茶道的英文文章中可以明显看出译者忽略了英语母语者的认知范畴和英语的实际用法。在谈到茶的汤色时译者将汉语的表达直接译为 the soup color of the tea，但对于英文母语者来说这个表达实际上没有任何意义，如果没有一定的汉语知识，则不知所云。



图6 mushroom soup

六 总结

21 世纪是语言学发展的重要时期（罗仁地 2017），语言田野调查方法也不例外。在语言技术极速发展的今天，囿于 20 世纪盛行的语料搜集方法而产生的问题与日俱增，因此，基于语言类型学及功能、认知语言学的田野调查和记录在当下显得尤为重要。具体来说，在田野调查时调查者应该摒弃对词表和句子进行直接翻译的传统，转而采用沉浸式“参与观察者”的方式，到目的语社群所居住的地方搜集自然语料。为提高搜集语料的科学性，调查者在调查之前应该先做与目的语相关的社会语言学调查，尽量选择目的语母语者作

^① 图 3 是我一位华人朋友想做 pumpkin soup 的时候买的 pumpkin，做出来的是 squash soup，不是 pumpkin soup。

为合作者，以了解目的语的形式所呈现出的所以然。在调查过程中，应尽可能搜集不同语体和用法的自然语料，调查后尽量在当地进行转写工作，以便调查者更清晰地理解目的语。无论是搜集语料还是进行转写，都应该尊重当地人的文化与合作者的意愿，要注意保护他们的隐私。综上所述，21 世纪的语言田野调查应该侧重于搜集自然语料、了解目的语使用社群独有的认知范畴，同时在分析时采用归纳法，才能保证成果的科学性和客观性。

参考文献

- 戴庆厦、罗仁地、汪锋主编（2008）《到田野去——语言学田野调查的方法与实践》，民族出版社，北京。
- 罗仁地（2007）语言类型学 / 功能语言学派视野下的语言学田野调查，《语言学论丛》第三十六辑，商务印书馆，北京；又载戴庆厦、罗仁地、汪锋主编《到田野去——语言学田野调查的方法与实践》，民族出版社，北京。
- 罗仁地（2017）尊重语言事实 提倡科学方法——21 世纪语言学刍议，《光明日报》12 月 3 日，北京。
- 郑武曦（2024）重视自然语料：以羌语田野调查为例，《语言学论丛》第 4 期。
- De Busser, Rik and Randy J. LaPolla (eds.) (2015) *Language Structure and Environment: Social, cultural, and natural factors*. Amsterdam and Philadelphia: John Benjamins Publishing Company.
- Enfield, Nick (2002) *Ethnosyntax: Explorations in grammar and culture*. Oxford: Oxford University Press.
- Fiddler, Michael (2023) Conceptual metaphor in areal perspective: Time, space, and contact in the Sinosphere. *Language and Cognition, First View* (published online), 1-29. <https://doi.org/10.1017/langcog.2023.41>.
- LaPolla, Randy J. (2001) Dulong texts: Seven fully analyzed narrative and procedural texts1. *Linguistics of the Tibeto-Burman Area*, 24(2): 1-39.
- LaPolla, Randy J. and David Sangdong (2015) *Rawang-English-Burmese Dictionary*. <https://rawang.webonary.org>.

Fieldwork in the 21st Century——Based on Natural Language Data, Induction and Cognitive Categories of the Native Speakers

Randy J. LAPOLLA LU Lin

Abstract: This article discusses how to conduct fieldwork and language documentation in the 21st century. It is mainly based on the authors' experiences, as well as problems they encountered during fieldwork and their solutions. The suggestions put forward in this article include: when conducting surveys, we should try our best to collect natural language data in the community where the target language is spoken, and choose the immersive “participant observer” method to collect natural language data (native-language investigators should also collect natural language data). It is also best to transcribe the data locally; it is advisable to do a sociolinguistic survey first; the survey and analysis should be based on the cognitive categories of native speakers of the target language; the inductive method should be used to analyze the collected natural language data.

Keywords: linguistic fieldwork, language typology, anthropological linguistics, functional linguistics

罗仁地: 639798 新加坡, 南洋理工大学人文及社会科学研究中心 randylapolla@protonmail.com

卢琳: 160-0023 日本, 东京都株式会社 MUZEN lulinlin5260@gmail.com