同濟大學
TONGJI UNIVERSITY

博士学位论文

中学化学教材中的知识构建研究：多符号视角

姓　　名：喻志刚

学　　号：1710369

所在院系：外国语学院

学科门类：外国语言文学

学科专业：外国语言学及应用语言学

指导教师：张德禄教授

副指导教师：Prof. Karl Maton, Dr. Yaegan Doran

二〇二一年五月

同濟大學

TONGJI UNIVERSITY

A dissertation submitted to

Tongji University in conformity with the requirements for

the degree of Doctor of Philosophy

Exploring the Knowledge-building of Chemistry in Secondary School Chemistry Textbooks: A Multisemiotic Perspective

Candidate: Zhigang Yu

Student Number: 1710369

School/Department: School of Foreign Languages

Discipline: Foreign Linguistics and Literature

Major: Foreign Linguistics and Applied Linguistics

Supervisor: Prof. Delu Zhang,

Prof. Karl Maton, Dr. Yaegan Doran

May, 2021

学位论文版权使用授权书

本人完全了解同济大学关于收集、保存、使用学位论文的规定，同意如下各项内容：按照学校要求提交学位论文的印刷本和电子版本；学校有权保存学位论文的印刷本和电子版，并采用影印、缩印、扫描、数字化或其它手段保存论文；学校有权提供目录检索以及提供本学位论文全文或者部分的阅览服务；学校有权按有关规定向国家有关部门或者机构送交论文的复印件和电子版；在不以赢利为目的的前提下，学校可以适当复制论文的部分或全部内容用于学术活动。

学位论文作者签名：

年　　月　　日

同济大学学位论文原创性声明

本人郑重声明：所呈交的学位论文，是本人在导师指导下，进行研究工作所取得的成果。除文中已经注明引用的内容外，本学位论文的研究成果不包含任何他人创作的、已公开发表或者没有公开发表的作品的内容。对本论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明。本学位论文原创性声明的法律责任由本人承担。

学位论文作者签名：

年　　月

摘要

本文从多符号视角研究中学化学教材中的知识构建，旨在探究包括化学符号（化学式、结构式和化学方程）、图像和语言在内的多种符号资源如何构建中学化学知识。为实现这一研究目标，本文探究以下研究问题：

（1）化学符号如何构建中学化学知识？

该问题由两个子问题组成：

（1a）化学符号如何组织建构意义？

（1b）化学符号在各年级中学化学教材中如何发展构建化学的等级知识结构？

（2）图像如何构建中学化学知识？

该问题也由两个子问题组成：

（2a）中学化学教材中的图像实现哪些意义以及如何实现这些意义？

（2b）图像在各年级中学化学教材中如何发展构建化学的等级知识结构？

（3）化学符号、图像和语言如何在中学化学教材中合作构建化学知识？

为解决上述研究问题，本文主要使用系统功能语言学的基本理论轴关系（系统和结构）与语场以及合法化符码理论的意义性符码分析澳大利亚新南威尔士州中学所使用的化学教材。依据新南威尔士州中学化学课程大纲，这些教材可分为三个阶段：阶段 4（7、8 年级）、阶段 5（9、10 年级）和阶段 6（11、12 年级）。基于上述理论工具和质性研究方法，本文对研究问题的发现如下：

（1）针对研究问题（1a），即化学符号如何组织建构意义，本文基于系统功能语言学的轴关系理论探究化学符号的语法。化学符号包括三大类：化学方程、化学式和结构式，三者皆为不同的符号系统。对化学方程的语法描写显示该语法包括四个级阶：方程（equation）、项（term）、式（formula）和（元素）符号（symbol）。这些级阶中的语法系统主要实现经验和逻辑元功能。其中，经验元功能遍布各个级阶，而逻辑元功能出现在除方程外的所有级阶。另外，化学方程的语法系统也实现语篇元功能，但仅限于项级阶。由于化学式是化学方程的组成部分，化学方程的式和符号级阶中的语法系统也属于化学式的语法。对结构式的语法描写显示，结构式的语法包括式和符号两个级阶。与化学式和化学方程类似，结构式的语法也主要实现经验和逻辑元功能。

为探究以上三种化学符号系统在建构意义上存在的异同，本文从语法组织及其所实现的意义两方面对这些符号进行功能符号类型学分析。研究发现，化学式和结构式主要呈现单变元结构，均允许化学元素符号复现。但与化学式不同，结

构式中的元素符号可以在二维空间中复现，形成空间排列结构。这导致二者实现的语场意义不同，前者只实现构成分类关系，而后者既实现构成分类关系又实现空间属性。与化学式和结构式不同，化学方程主要呈现多变元结构。从语场角度来看，这些多变元结构在方程级阶实现活动，即不同类别的化学反应，在项级阶实现多种属性和化学反应所涉及的量化关系。另外，与化学式不同，结构式和化学方程均使用概念化的韵律结构实现化学物质的属性和化学反应的量化关系。

针对研究问题（1b），即化学符号在中学各年级化学教材中如何发展构建化学的等级知识结构，本文首先依据化学符号的语场意义分别为“式”（化学式和结构式）和化学方程构建意义密度分析框架，然后利用该框架分析它们的意义密度在各年级教材中如何发展变化，从而构建有关化学物质和化学反应的等级知识结构。研究发现，中学化学教材中“式”的意义密度包含五个等级，其中最弱的为分子式，最强的为立体结构式，它们形成一个连续体，构成“式”的意义密度分析框架。利用该框架分析“式”在各年级教材中的发展显示，不同意义密度的“式”出现在不同年级。从横跨所有年级的角度来看，各年级“式”的意义密度范围不断拓宽，底层是表达化学物质构成的分子式，出现在所有年级阶段，而顶层的意义密度随着年级的增加不断增强，所表达的关于化学物质的语场意义越来越复杂，从而构建化学的等级知识结构。

中学化学教材中化学方程的意义密度包含六个等级，其中最弱的为非完整分子方程，最强的为可逆离子方程，它们形成一个连续体，构成化学方程的意义密度分析框架。利用该分析框架分析化学方程在各年级阶段教材中的发展显示，从阶段 4 到阶段 5 化学方程的意义密度呈现连续增强，在原来的化学物质构成分类关系和活动（正向反应）语场意义基础上增加有关这些物质的量化关系和物理状态这两项语场意义。基于阶段 5 中的完整分子方程，阶段 6 中的 11 和 12 年级分别引入离子方程和可逆方程，分别构建了有关这两类反应的化学知识。从横跨所有年级阶段的角度来看，化学方程的意义密度不断增强，实现的有关化学反应的语场意义越来越复杂，从而构建化学的等级知识结构。

（2）针对研究问题（2a），即中学化学教材中的图像实现什么意义以及如何实现这些意义，本文从语场角度构建了一个描述上述图像所实现的语场意义的系统，它同时呈现了实现这些语场意义的语法资源。该系统显示上述教材中的图像主要实现三大类语场意义：项目（item），活动（activity）和属性（property），它们各自又有不同的子类别。

针对研究问题（2b），即图像在各年级中学化学教材中如何发展构建化学的等级知识结构，本文首先为教材中的图像构建一个意义密度分析框架，然后利用该框架分析图像的意义密度如何发展构建化学的等级知识结构。研究发现，图像

的意义密度分析框架包括两个参数：内部组织和外部关联。内部组织关涉图像实现的语场意义的数量和类别，包括构成关系、分类关系、活动和属性。外部关联描述这些语场意义的技术性。利用该框架分析图像在各年级教材中的发展显示，当构建关于同一主题的化学知识时，图像的意义密度连续增强，即从相对较弱到相对较强到更强。这表明在构建特定的化学知识时，教材需要使用日常图像来呈现日常生活中的现象，同时使用意义密度越来越强的技术图像构建关于该现象越来越复杂的化学知识。与上述连续性增强的意义密度变化不同，从横跨各个年级阶段的角度来看，图像的意义密度呈现不断拓宽的变化趋势，每个年级阶段都识解日常现象，同时各阶段最强的意义密度变得越来越强，表达的化学知识越来越复杂，从而构建化学的等级知识结构。

（3）针对研究问题（3），即化学符号、图像和语言在中学化学教材中如何合作共同构建化学知识，本文首先从上述多种符号所实现的语场意义出发为化学知识构建一个“多符号版图”，然后探究它们在不同年级的化学语篇中如何互动构建知识。针对第一项研究内容，研究发现语言、化学符号和图像在表达化学知识方面具有各自的功能，它们彼此形成互补构成化学知识的完整版图。

为探究上述多种符号在不同年级的化学语篇中如何互动构建知识，本文利用合法化符码理论的意义性符码分析这些符号资源如何互动构建知识。研究发现，从意义性角度来看，语言和图像之间存在三种符间关系：（1）语言的意义密度强于图像但意义引力弱于图像；（2）语言的意义密度和意义引力与图像相似；（3）语言的意义密度弱于图像但意义引力相似。第一种图文关系倾向于出现在低年级，而第二、三种一般出现在高年级。从语场角度来看，通过与语言互动，图像既可使语言所表达的化学概念引力增强，从而与常识语场产生联系，又可从语言中压缩意义构建技术性越来越强的非常识语场。相较而言，语言和化学符号间的符间关系呈现较为有限的意义性关系类别，化学符号的意义密度一般与语言相似或强于语言。这种符间关系使化学符号从语言中压缩意义，并允许在低年级的化学符号基础上进一步压缩意义。这使化学语篇高效地构建概括性和融合性越来越高的化学理论和概念。上述多种符号资源彼此分工形成互补，共同构建化学的等级知识结构。

上述研究发现在理论和教学实践方面均具有一定的价值和启示意义。理论方面，本文基于系统功能语言学的基本理论轴关系和化学符号本身的结构特点，描写了化学符号的语法系统，拓宽了符号描写的版图，并尝试突破 Mitchell（1986）指出的在符号描写中存在的“语言帝国主义”问题，为未来希望以符号自身的结构特点来描写符号系统的研究提供启示。在化学符号语法分析方面，本研究发现了一类新的结构：“概念化的韵律结构”（ideationalized prosodic structure），丰富

了现有的功能语法结构。在图像分析方面，本研究提出了一个从语场角度描述图像所实现的意义的框架，弥补了当前研究在较语法更为抽象的层次描写图像这一问题中存在的空白。其次，在图像语法分析方面，本文发现了两类新的结构，一个是“转变过程”（transformation process），另一个是“多层级隐性分类结构”（multi-leveled covert taxonomy structure），丰富了 Kress 和 van Leeuwen（2006）的图像语法描写。另外，本研究还提出了一个分析图像意义密度的框架，该框架同时考虑了语场意义的数量和技术性，弥补了现有研究存在的不足。在符号间性方面，本研究从合法化符码理论的意义性角度提出了三类图文关系类别，丰富了当前的图文关系研究。

本研究在教学方面也具有一定价值和启示意义。对化学符号和图像的语法分析揭示了它们建构意义的方式，为发展针对化学知识的多元读写教学法提供一定启示。对化学符号和图像如何构建化学知识的研究发现可帮助教师在特定年级使用合适的化学符号或图像进行教学，这些符号资源的意义密度与该年级阶段匹配，可使学生更有效地理解这些符号资源，从而成功习得化学知识。

关键词：知识构建、中学化学教材、多符号、系统功能语言学、合法化符码理论

Abstract

This thesis explores the knowledge-building of chemistry in secondary school chemistry textbooks from a multisemiotic perspective. It aims to investigate how the multisemiotic resources, including chemical formalisms (chemical formulas, structural formulas, and chemical equations), images, and language build the knowledge of secondary school chemistry. To achieve this objective, this thesis explores the following research questions:

(1) How do chemical formalisms (chemical formulas, chemical equations, and structural formulas) build the knowledge of secondary school chemistry?

This question includes two sub-questions:

(1a) How are the chemical formalisms organized to make meaning?

(1b) How do the chemical formalisms develop across schooling levels to build the hierarchical knowledge structure of chemistry in the secondary school chemistry textbooks?

(2) How do images build the knowledge of secondary school chemistry?

This question also includes two sub-questions:

(2a) What meaning do images in the chemistry textbooks construe and how do they realize the meaning?

(2b) How do the images develop across schooling levels to build the hierarchical knowledge structure of chemistry in the textbooks?

(3) How do chemical formalisms, images, and language work together to build chemistry knowledge in the textbooks?

To address these research questions, this thesis uses primarily axial relations (system and structure) and field from Systemic Functional Linguistics (SFL) and Semantics from Legitimation Code Theory (LCT) to analyze chemistry textbooks used by secondary schools in New South Wales (NSW), Australia. The textbooks are categorized into three stages according to the NSW curriculum of secondary school chemistry: Stage 4 (Years 7 and 8), Stage 5 (Years 9 and 10), and Stage 6 (Years 11 and 12). Based on the above theoretical tools and a qualitative research method, the thesis obtains the following findings to the research questions.

(1) To address research question (1a), i.e., how are chemical formalisms organized to make meaning, this thesis explores the grammars of chemical formalisms based on axial relations in SFL. Chemical formalisms include three major types: chemical

equations, chemical formulas, and structural formulas, each of which is a distinct semiotic system. The description of chemical equations shows that the grammar includes four ranks: equation, term, formula, and symbol. The grammatical systems at these ranks organize primarily the experiential and logical metafunctions. The experiential metafunction pervades every rank of the grammar, and the logical metafunction occurs at all the ranks below the equation rank. The grammar also realizes a textual component but only at the term rank. As chemical formulas are a part of chemical equations, the formula and symbol ranks also constitute the grammar of chemical formulas. The description of structural formulas shows that the grammar includes two ranks: formula and symbol. Similar to chemical formulas and chemical equations, the grammatical systems of structural formulas also realize primarily the experiential and logical metafunctions.

To understand the similarities and differences between the three chemical formalisms in making meaning, this thesis develops a functional semiotic typology for the formalisms in terms of their grammatical organization and meaning construed. The finding shows that chemical formulas and structural formulas are similar in that both are dominated by univariate structures, each capable of iterating chemical symbols. However, in contrast to chemical formulas, chemical symbols in structural formulas can iterate in a two-dimensional space, which affords a reading of spatial layout. This leads to a distinction in the field-specific meanings they realize − chemical formulas realize only a compositional taxonomy, whereas structural formulas realize both compositional taxonomies and spatial properties. In contrast to chemical formulas and structural formulas, chemical equations are organized largely around multivariate structures. In terms of field, these multivariate structures construe activities (chemical reactions) at the equation rank and various properties and quantitative relations of the reactions at the term rank. In addition, in contrast to chemical formulas, both structural formulas and chemical equations employ ideationalized prosodic structure to construe qualitative properties of chemical matter and quantitative relations in chemical reactions.

To address research question (1b), i.e, how the chemical formalisms develop across schooling levels to build the hierarchical knowledge structure of chemistry in the secondary school chemistry textbooks, this thesis first establishes models for analyzing the semantic density of chemical formalisms based on field-specific meanings they realize and then enact the models to analyze how the formalisms develop

through schooling in the textbooks to build the hierarchical knowledge structure of chemistry. The finding shows that the semantic density of the various formulas used in the textbooks can be sequenced as five scales, with molecular formulas embodying the weakest strength and stereo formulas exhibiting the strongest strength. This continuum of scales constitutes the model for tracing the development of the formulas' semantic density across curriculum stages in the NSW secondary school chemistry textbooks. Enacting the model reveals that formulas with different degrees of semantic density are used in different schooling levels. Across the curriculum stages, they exhibit a widening range of semantic density, with the bottom of the range being the most basic molecular formulas construing the compositions of chemical matter and the top reaching increasingly strong semantic density that construes more and more complex knowledge of chemical matter. Such a development of the formulas' semantic density builds the hierarchical knowledge structure of chemistry.

The semantic density of the chemical equations can be sequenced as six scales along a continuum, with incomplete molecular equations embodying the weakest strength and reversible ionic equations exhibiting the strongest strength. This continuum of scales constitutes a model for tracing the development of the chemical equations' semantic density across curriculum stages in the NSW secondary school chemistry textbooks. Enacting the model shows that there is a consecutive increase in chemical equations' semantic density from Stage 4 to 5, expanding the field of chemical reactions that involves only compositional taxonomies (compositions of chemical species) and an activity (a forward reaction) to the field that includes two extra properties concerning chemical reactions (quantities of chemical species involved in the reactions and their physical states). Building upon the complete molecular equations used in Stage 5, Stage 6 introduces ionic equations in Year 11 and reversible equations in Year 12, building the knowledge of ionic reactions and reversible reactions. Across the curriculum stages, the semantic density of chemical equations keeps growing and the field meanings construed by the equations become increasingly complex, building the hierarchical knowledge structure of chemistry.

(2) To address research question (2a), i.e., what meaning do images in the chemistry textbooks construe and how do they realize the meaning, this thesis establishes a model of chemical images in terms of field that links with the grammatical resources realizing the field-specific meanings. The model shows that chemical images construe three broad types of field-specific meanings: item, activity, and property, each

of which includes further subtypes of meaning.

To address research question (2b), i.e., how do the images develop across schooling levels to build the hierarchical knowledge structure of chemistry in the textbooks, this thesis first establishes a model for the semantic density of images used in the textbooks and then enacts it to analyze how the images' semantic density develops through schooling. The finding shows that the model for the semantic density of images builds upon two parameters: internal organization and external relations. Internal organization concerns the number and types of field-specific meanings construed by an image, including composition, classification, activity, and property. External relations describe the technicality of these field-specific meanings. Enacting the model to trace the development of images shows that when building knowledge about the same subject, the images' semantic density exhibits a consecutive progression: from relatively weak through relatively strong to even stronger. This suggests that to build certain knowledge of chemistry, the textbook needs to use everyday images to present empirical phenomena and technical images with increasingly strong semantic density to build more and more complex technical understandings of the phenomena. In contrast to the consecutive progression of semantic density, the analysis of images' development throughout the curriculum stages shows that it exhibits a widening range of semantic density, with each stage maintaining connections with the everyday empirical world while also reaching toward increasingly complex chemical theories, which expresses increasingly complex chemical knowledge and builds the hierarchical knowledge structure of chemistry.

(3) To address research question (3), this thesis first establishes a multisemiotic picture of chemistry knowledge in terms of field-specific meanings realized by the multisemiotic resources and then explores how they interact in chemistry texts to build knowledge across schooling levels. As for the first issue, the finding shows that language, chemical formalisms, and images have their own functionalities in organizing the field of chemistry and complement each other to constitute the full picture of chemistry.

To investigate how the semiotic resources interact in chemistry texts across schooling levels to build chemistry knowledge, this study examines the interaction between the resources in terms of LCT's semantic density and semantic gravity. The finding shows that the intersemiotic relations between language and image involve three types: (1) language embodies stronger semantic density and weaker semantic gravity than images; (2) language embodies similar semantic density and semantic gravity as images; (3) language embodies weaker semantic density than and similar semantic gravity as images. The first type tends to prevail at earlier schooling levels while the second and third types typically occur at higher schooling levels. In terms of field, through the interaction with language, images can either gravitate chemistry concepts to link to everyday common-sense fields or condense meaning to build increasingly technical uncommon-sense fields. The intersemiotic relation between language and chemical formalisms, in contrast, exhibits relatively limited variations in Semantics − the formalisms typically show similar semantic density as or stronger semantic density than language. The interaction allows chemical formalisms to consolidate meaning from language and enables the meaning to be further condensed based on their previous forms at earlier schooling levels. With these intersemiotic relations, chemistry texts efficiently develop increasingly integrative and generalized chemistry concepts. The division of labor across the semiotic resources and their complementarity build the hierarchical knowledge structure of chemistry.

The above findings hold certain theoretical and pedagogical significance and implications. Theoretically, this thesis broadens the boundaries of modeling non-linguistic resources by describing the grammatical systems of chemical formalisms based on the theoretical primitive of axis in SFL and chemical formalisms' internal structures. It can be viewed as an attempt to step out of what Mitchell (1986) has referred to as 'linguistic imperialism' in modeling non-linguistic semiotic systems, which is thus heuristic for future studies hoping to model semiotic systems on their own terms. In terms of grammatical analysis of chemical formalisms, this study enriches previous functional structures accounted for in SFL by proposing a new structure − an ideationalized prosodic structure. In terms of image analysis, this study also makes several contributions. Firstly, it proposes a system network that describes the meaning construed by chemical images from the perspective of field, filling the gap that little research examines images from a more abstract level than grammar. Secondly, in terms of visual grammar, this study discovers two new structures – transformation processes and multi-leveled covert taxonomy structures, which enrich Kress and van Leeuwen's description of the grammar of image. Thirdly, this thesis establishes a more comprehensive model for the semantic density of images by considering both the number and types of field-specific meanings and their technicality, filling the gap in previous studies. In terms of intersemiosis, this study proposes three types of language-

image relations in terms of LCT's Semantics, enriching the literature on language-image relations.

Pedagogically, this study also holds some significance and implications. The grammatical analysis of chemical formalisms and images unveils how they work to make meaning, which is helpful for the development of literacy pedagogy for teaching chemistry. The findings in the analysis of knowledge-building through chemical formalisms and images are also potentially helpful for teachers to use these semiotic resources with appropriate semantic density at certain learning stages for teaching chemistry, so that students can more effectively engage with and understand these resources and hence succeed in learning chemistry.

Key Words: knowledge-building, secondary school chemistry textbooks, multisemiosis, systemic functional linguistics, legitimation code theory

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction[1]

1.1 Research background

Learning chemistry is crucial for secondary school students. It not only helps them understand the nature of the physical world from a chemical viewpoint but also prepares them for other scientific disciplines in higher education, with chemistry being considered a core science permeating several areas of knowledge, such as physics, biology, engineering, and medicine (Quadros et al. 2011: 232). However, it has been widely reported that the teaching and learning of secondary school chemistry have faced significant challenges in many countries (Bennett 2003; Gabel 1998; Herrington and Daubenmire 2016; Treagust et al. 2000). In chemistry education, research on learning that has aimed to address these challenges has long been dominated by views focusing on the development of mental structures within learners' minds, often grouped together under the term 'constructivism' (Bodner 1986; Herron 1975; Shiland 1999; Taber 2000; Wink 2014). This approach to chemistry learning explores the state of consciousness in learners' mental models (what we could call 'knowing'). However, much of this research tradition elides a focus on **what** is being known – the knowledge itself. Maton (2014) argues that approaches such as this suffer from '**knowledge-blindness**' – knowledge as an object of study is obscured. Focusing on the form that knowledge takes is crucial for understanding how chemistry works because the structure of knowledge itself has effects – what is being taught and learned affects the way it is being taught and learned. By focusing only on the development of mental structures in learners, this crucial view of the knowledge itself is missed. It thus fails to show what forms of chemistry knowledge are taught to students, how the knowledge progresses as the learning proceeds, and how this may influence and impact the pedagogies being used. Therefore, to more effectively inform the teaching and learning of secondary school chemistry, we need to study the knowledge-building of chemistry

[1] This thesis document uses hyperlinks (cross references) for easier navigation. You can jump to the referred chapters, sections, figures, or tables by clicking the hyperlinks (the numbers). For example, if you click '2.1' in 'Section 2.1', you will jump to Section 2.1. (You can try with this example, which has been hyperlinked to the referred section) You can return to the previously displayed page by either pressing the shortcut keys 'Alt" and '⬅' (the left direction key on the bottom right of your keyboard) at the same time or right-clicking anywhere on the screen and selecting 'previous view' from the drop-down menu. You can do this as many times as you need to return to the original place.

with a focus on the forms of knowledge that make up the discipline of chemistry itself.

Chemistry knowledge, as it is construed through chemistry discourse, is inherently multisemiotic (Parodi 2012). It comprises language, images such as diagrams, graphs, and charts, and a wide range of chemical formalisms including chemical symbols, chemical formulas, chemical equations, and structural formulas. To understand the forms of chemistry knowledge requires an in-depth understanding of its range of semiotic resources and how they realize its technical meanings. Such a semiotic perspective to the knowledge-building of chemistry is also significant for engaging with literacy challenges faced by secondary school students in learning chemistry. It has been frequently reported that these students face difficulties in understanding the highly technical chemical language, formalisms, and images when studying chemistry (Canac and Kermen 2016; Cheng and Gilbert 2009; Davidowitz and Chittleborough 2009; Lee 2005; Markic and Childs 2016; Taskin and Bernholt 2014; Wellington and Osborne 2001). Since learning disciplinary knowledge is to learn the privileged discourse of that discipline (Moje 2008), difficulties in understanding chemistry discourse deny students access to highly valued chemistry knowledge.

In recent decades, explicit pedagogical programs have been developed to address literacy challenges across disciplines (e.g. Rose and Martin 2012). These pedagogies have long acknowledged the need for concepts of literacy to be extended into multisemiotic literacy practices (The New London Group 1996; Unsworth 2001). However, it has long been recognized that to develop a literacy pedagogy for a particular type of discourse, it is necessary to first understand how that discourse works (Rose and Martin 2012). In the case of chemistry, this means that it is crucial to understand how the wide range of chemical formalisms used through schooling work. The semiotic perspective of this thesis responds to this call by exploring how the range of different semiotic resources of chemistry construe chemistry knowledge. In particular, it examines in detail how the critical range of formalisms used in secondary school chemistry make meaning. It does so by developing in-depth and explicit grammatical descriptions of the three main formalisms used in secondary school chemistry – chemical formulas, chemical equations, and structural formulas.

This focus on formalisms is combined with a detailed exploration of the images used in secondary school chemistry, another essential semiotic resource constituting chemistry discourse. In particular, it examines the meaning construed by the images and the grammatical resources realizing the meaning. In the field of chemistry education, images tend to be studied in terms of the referents they involve. For example, images are often considered in terms of whether they show macroscopic phenomena that can be sensed, microscopic phenomena such as atoms, molecules, and ions, or symbolic phenomena such as those including symbols, formulas, and equations (Gilbert 2005; Johnstone 1991). This perspective provides a useful classification of the different types of chemistry knowledge. However, it does not differentiate the different forms and functionalities of the images and hence provides limited implications on how the images construe meanings. The semiotic perspective of this thesis fills this gap by making explicit how the images organize the knowledge of chemistry.

Like all disciplines, chemistry has its own way of organizing its knowledge. It is often characterized as a typical 'hard' science (Biglan 1973) organized through, what Bernstein (1999) calls, a **hierarchical knowledge structure**, where it 'creates very general propositions and theories, which integrate knowledge at lower levels, and in this way shows underlying uniformities across an expanding range of apparently different phenomena' (ibid.: 162). Thus, studying chemistry from a semiotic perspective also entails exploring how its meanings construe the knowledge structure of chemistry. To see this more clearly, it is useful to provide a **dynamic** perspective on knowledge. Such a focus is less concerned with the static structure of chemistry knowledge, but more concerned with how it develops, changes, and builds over time. Looking at knowledge dynamically as knowledge-**building** enables an understanding of shifts in chemistry knowledge comparable to what students experience as they move through schooling.

Thus, in addition to the description of chemistry discourse, this thesis explores how the discourse develops through schooling to build an intricate and integrated knowledge structure of chemistry. In particular, it focuses on how the various chemical formalisms and images develop through secondary schooling levels in a way that allows increasingly complex meanings to be presented and build the knowledge of chemistry.

Upon the above background, this study will examine the knowledge-building of secondary school chemistry from a semiotic perspective in terms of how the multisemiotic chemistry discourse construes meaning and develops across the schooling levels to build the hierarchical knowledge structure of chemistry. It considers chemical formalisms and images in relation to language as critical components of the discourse of chemistry and investigates their roles in building chemistry knowledge.

1.2 Research aims and questions

As discussed above, this study aims to investigate how multisemiotic chemistry discourse used at the secondary school level builds knowledge. Specifically, it examines semiotic resources including chemical formalisms, images, and language used in English secondary school chemistry textbooks.[2] We focus on chemistry textbooks because they are the primary resource for chemistry instruction by teachers and the core reading material for students to learn chemistry (Chiappetta and Koballa 2002; Gkitzia et al. 2011; Upahi and Jimoh 2015). As argued in Section 1.1, exploring the knowledge-building of chemistry from a semiotic perspective entails examining both the construal of meaning by the multisemiotic chemistry discourse and how they develop across schooling levels to build the hierarchical knowledge structure of chemistry. To achieve this objective, this thesis explores the following research questions:

1. How do chemical formalisms (chemical formulas, chemical equations, and structural formulas) build the knowledge of secondary school chemistry?
   This question includes two sub-questions:
   a. How are the chemical formalisms organized to make meaning?
   b. How do the chemical formalisms develop across schooling levels to build the hierarchical knowledge structure of chemistry in the secondary school chemistry textbooks?
2. How do images build the knowledge of secondary school chemistry?
   This question also includes two sub-questions:
   a. What meaning do images in the chemistry textbooks construe and how do they realize the meaning?
   b. How do the images develop across schooling levels to build the hierarchical knowledge structure of chemistry in the textbooks?
3. How do chemical formalisms, images, and language work together to build chemistry knowledge in the textbooks?

The first research question concerning the knowledge-building of chemistry through chemical formalisms is addressed by Chapter 4. It tackles the first sub-question by describing the grammars of chemical formalisms and the second sub-question by analyzing how the meaning construed by the formalisms develop across schooling

[2] Note that the term 'English secondary school chemistry textbooks' does not necessarily mean that the textbooks use only one semiotic resource, i.e., the English language. The term only denotes that the language used in the textbooks is English.

levels in the textbooks to build the hierarchical knowledge structure of chemistry. The second research question concerning the knowledge-building of chemistry through images is addressed in Chapter 5. It addresses the first sub-question by establishing a model of the images in terms of field in systemic functional linguistics (SFL) (Doran and Martin 2021) that links with the grammatical resources used and the second sub-question by analyzing how the images develop across schooling levels to build the hierarchical knowledge structure of chemistry. The third research question brings together chemical formalisms, images, and language. It is addressed in Chapter 6 by first probing their respective disciplinary affordances and then analyzing how they interact to build chemistry knowledge in the secondary school chemistry textbooks.

1.3 Significance of the research

This research makes a number of contributions to the field. First, it significantly expands the modeling of semiotic systems in the current literature to include chemical formalisms (primarily Chapter 4). It provides comprehensive systemic functional descriptions of the formalisms' grammars that go considerably further than any previous description. Based on these descriptions, this study develops a functional semiotic typology of the formalisms in terms of their similarities and differences in making meaning. This functional semiotic typology builds upon Doran (2019) to expand functional linguistic typology (Caffarel et al. 2004) toward embracing semiotic systems other than language. This typology enables the formalisms to be compared and contrasted so as to understand why different formalisms may be used at different times. Through these descriptions, this thesis probes the complex meaning construed by the chemical formalisms and makes explicit their varying degrees of complexity of meaning. This offers an analysis of the knowledge-building of chemistry that reveals the developmental features of chemical formalisms across schooling levels.

To complement this focus on formalisms, the thesis also establishes a model of images in chemistry that makes explicit the highly technical meaning they construe (through the SFL modeling of field; Doran and Martin 2021) as well as their complexity of meaning (through the Legitimation Code Theory concept of semantic density; Maton 2014). Together, this enables the present study to zoom in on individual images to understand their specific meaning in detail and zoom out to see the development of images across multiple years of secondary schooling. Finally, this study maps out the disciplinary affordances of chemical formalisms, images, and language in chemistry in terms of field and reveals how they interact with each other in chemistry to build the knowledge of chemistry. This adds a knowledge-building perspective to the study on intersemiotic relations, which significantly extends the scope of multimodal discourse analysis. In all, this thesis produces the most detailed and comprehensive model of the multisemiotic discourse of chemistry to date.

1.4 Organization of the thesis

This thesis is organized into seven chapters. Chapter 2 builds upon the present chapter by examining relevant literature to the research questions and identifying gaps in the literature. Following this, Chapter 3 outlines the research design for this study. It introduces the theoretical framework and methodology and explains how they address the research questions. Chapters 4-6 address the research questions and are introduced in more detail as follows.

Chapter 4 first explores how chemical formalisms are organized to make meaning. It describes the grammars of chemical equations, chemical formulas, and structural formulas from the perspective of SFL. Through the description, it reveals the intrinsic functionality of the formalisms and their grammatical organizations. To understand the similarities and differences between the formalisms in making meaning, the chapter establishes a functional semiotic typology for these formalisms. By building grammatical descriptions on the fundamental theoretical primitive of axis – the interaction between the paradigmatic and syntagmatic relations (system and structure) – this study is able to show the intrinsic functionality of chemical formalisms and compare and contrast these formalisms in terms of their meaning-making. To understand how the meaning construed by the formalisms builds the hierarchical knowledge structure of chemistry, Chapter 4 then analyzes how the formalisms develop across schooling levels to build the knowledge of chemistry. In particular, it focuses on the knowledge-building of two critical components of chemistry knowledge – chemical matter and chemical reactions – by tracing the development of formulas and chemical equations in terms of their semantic density across schooling levels.

Chapter 5 explores the meaning construed by images in secondary school chemistry textbooks and how they develop across schooling levels to build the hierarchical knowledge structure of chemistry. It first models the images from the perspective of field within SFL that links with grammatical resources realizing the meaning and then analyzes how the images develop in the textbooks to build the hierarchical knowledge structure of chemistry. To trace the images' development, this chapter first establishes a model for their semantic density – their complexity of meaning – and then enacts it to analyze their developmental features across schooling levels.

Chapter 6 investigates how the range of semiotic resources used in chemistry, including chemical formalisms, images, and language, work together to build knowledge in secondary school chemistry textbooks. It will first establish a multisemiotic picture of the secondary school chemistry knowledge in terms of field and then explore how these resources interact in chemistry texts in terms of the Semantics dimension of Legitimation Code Theory (LCT) (Maton 2014). To understand the role of the interaction in building chemistry knowledge, the chapter will trace its development across schooling levels.

Chapter 7 concludes by summarizing the findings from the above substantive chapters and then discussing how this research contributes to existing knowledge. Finally, it outlines potential limitations to this study and directions for future studies.

Chapter 2 Literature Review

Chapter 1 has established a rationale for this study that probes the knowledge-building of secondary school chemistry from a semiotic perspective and outlined the research questions to achieve this goal. This chapter will extend this rationale by examining existing literature in the field to show what is known about the knowledge-building of chemistry and what is not yet known. As noted in Chapter 1, this study concerns two major issues: how the multisemiotic resources in secondary chemistry textbooks construe knowledge and how they develop across learning levels to build the knowledge. Accordingly, this chapter will review the literature on knowledge-building analysis (Section 2.1) and the meaning-making by multisemiotic resources in disciplinary discourse (Section 2.2). Following this, it will review chemists' conceptualization of chemistry knowledge and the knowledge-building of chemistry in the field of chemistry education (Section 2.3).

2.1 Different approaches to knowledge-building

Knowledge-building is one of the key concepts in this study. This section will review two major approaches to knowledge-building: systemic functional linguistics (SFL) (Section 2.1.1) and sociology of education (Berstein's theory of knowledge structure and legitimation code theory [LCT]) (Section 2.1.2). It will first review each approach's conception of knowledge and ways of analyzing knowledge-building and then situate the present study based on comparing the approaches' strengths and weaknesses (Section 2.1.3).

2.1.1 SFL to knowledge-building

Examining knowledge-building entails first addressing the key question - what is knowledge? As a linguistic theory, SFL does not distinguish between knowledge and language. It believes that it is illusory to see knowledge as something that exists independently from language and may then be coded or made manifest in language (Halliday and Matthiessen 1999: 3). This is further expounded in the following quote from Halliday (2007[1988]: 346–347):

> "[l]anguage is not the means of knowing; it is the form taken by knowledge itself. Language is not how we know something else, it is what we know; knowledge is not something that is encoded in language – knowledge is made of language."

For Halliday, knowledge and language are not two distinct phenomena. Learning knowledge is inherently a semiotic process (Halliday 1993b: 94). Knowledge-building can thus be seen as language development. In terms of an individual's language development, Halliday (2004[1998]: 27) proposes three crucial stages: (1) moving from protolanguage to language in the second year of life when a child develops grammar; (2) moving from everyday spoken grammar to the grammar of literacy around age five when he starts primary school; (3) moving from the grammar of written language to that of the language of the subject disciplines when he finishes primary school and starts secondary school. Halliday corresponds these critical stages of language development to three stages of knowledge development: common-sense knowledge (age 1-2), educational knowledge (age 4-6), and technical knowledge (age 9-13). This suggests that language development is arguably the development of knowledge itself.

In addition to examining knowledge-building in terms of individual's language development, or in SFL's terms, in terms of "ontogenesis" (Halliday 1993b), SFL also explores linguistic resources in text for building knowledge, which can be reviewed in terms of the different strata in Martin's (1992) model of language: lexicogrammar, register, and genre.

2.1.1.1 Lexicogrammar and knowledge-building

At the stratum of lexicogrammar, the core resources for building knowledge are grammatical metaphor and technicality. Grammatical metaphor is ubiquitous in disciplinary discourse (Hao 2020) and forms an essential knowledge builder for the disciplines. Among the three types of grammatical metaphor: ideational metaphor and interpersonal metaphor proposed by Halliday (1985, 1994) and textual metaphor identified by Martin (1992), ideational metaphor is most frequently used in disciplinary discourse for knowledge-building. Within ideational metaphor, nominalization has been the focus of previous studies on its role in building scientific knowledge. Numerous studies have examined nominalization in different science disciplines, for example, physics (Doran 2018c; Halliday 1993a; Zhao 2012), geography (Wignell et al. 1993), and biology (Hao 2020; Martin 2013a). These studies show that the power of nominalization in building science lies in that it allows a nominalized lexis to become a technical term that enters a technical taxonomy. For example, the nominalized ‘planetary motion’ from the congruent form ‘planets move’ is a technical term in astrophysics and relates to numerous other entities in the technical taxonomy, such as planet rotation and revolution. Martin (2013a) calls such a process of nominalization ‘distillation’, whereby technical language both compacts and changes the nature of everyday words. Therefore, nominalization as a grammatical metaphor provides a pivotal avenue for science to condense meaning from the common-sense knowledge. To highlight nominalization’s role in science apprenticeship, Martin (2013a) refers to it as ‘power grammar’, an essential code that students need to grasp to access scientific knowledge. All of these studies suggest that nominalization as a grammatical metaphor is an essential resource for building scientific knowledge.

Complementing grammatical metaphor, some scholars have also investigated the role of technicality in building scientific knowledge. Wignell et al. (1993: 161) define technicality as “the use of terms or expressions (but mostly nominal group constituents) with a specialized field-specific meaning”. For example, the term ‘salt’ in the context of chemistry refers to a compound produced by the reaction of an acid with a base, construing a specialized meaning in the field of chemistry. According to Wignell et al. (1993: 179), technicality has a field-creating function, allowing disciplines to develop technical terms to describe their own fields. Another significant function of technicality is that it enables the setting up of technical taxonomies, whereby different technical concepts are interrelated, though in an implicit way. This provides a critical means for disciplines to build increasingly complex domains of knowledge. Doran (2018c) demonstrates that through technicality, language enables physics to build a large network of field-specific meanings in a rather efficient manner.

In addition to technical taxonomies afforded by technicality, one particular grammar of technicality is also crucial for building knowledge, that is, definitions of technical terms. According to Martin (1993: 229), definitions are a special type of relational clause that translates common sense into specialized knowledge. For example, the following is a definition of ‘precipitation’:

> <u>*Precipitation*</u> *refers to all forms of water which fall (precipitate) from the sky*. (c.f.(Martin 1993: 229)

The above definition relates the known common sense, i.e., all kinds of water that fall from the sky, to a technical term - precipitation. Furthermore, a definition can also relate previously defined technical terms to new technical terms, building increasingly complex knowledge. The following definition of 'biosphere' is a case in point.

The biosphere is made up of many hundreds of different ecosystems.

(Martin 1993: 230)

The above definition involves two technical terms: 'biosphere' and 'ecosystems'. The relational clause relates the previously defined technical term 'ecosystems' to the new technical term 'biosphere', building increasingly complex technicality. This can be analogized to the snowball effect, gaining increasing mass as the snowball rolls along.

Technicality and nominalization are closely related. As reviewed above, nominalization as a grammatical metaphor can turn an everyday congruent figure into a technical entity that becomes part of a technical taxonomy. Wignell et al. (1993) thus take it as one of the crucial grammatical configurations for realizing technicality. Doran (2018c) also shows that technicality often interacts with grammatical metaphors to build large sets of uncommon-sense knowledge.

The above review suggests that grammatical metaphor and technicality are two significant linguistic resources for condensing meaning and building knowledge. However, the scholarship appears to have largely restricted to language, and few studies examined meaning condensation by semiotic resources other than language. O'Halloran (1999a, 2005) examined the metaphorical expansion of meaning among language, mathematical symbolisms, and images, which she terms 'semiotic metaphor'. Nevertheless, the expansion of meaning is between semiotic resources, and it remains unknown how the respective non-linguistic resources themselves build knowledge through meaning condensation. To fill this gap, the present study will investigate how the semiotic resources other than language, i.e., chemical formalisms and images, build knowledge in secondary school chemistry textbooks through the condensation of meaning. In specific, Chapters 4 and 5 will examine the different levels of density of meaning embodied by chemical formalisms and images and their roles in building chemistry knowledge.

2.1.1.2 Field and knowledge-building

At the stratum of register, previous studies have interpreted knowledge-building mainly in terms of the variable field. Field is defined by Martin (1992) as "sets of activity sequences oriented to some global institutional purpose". He models field as a system of activities as shown in Figure 2.1.

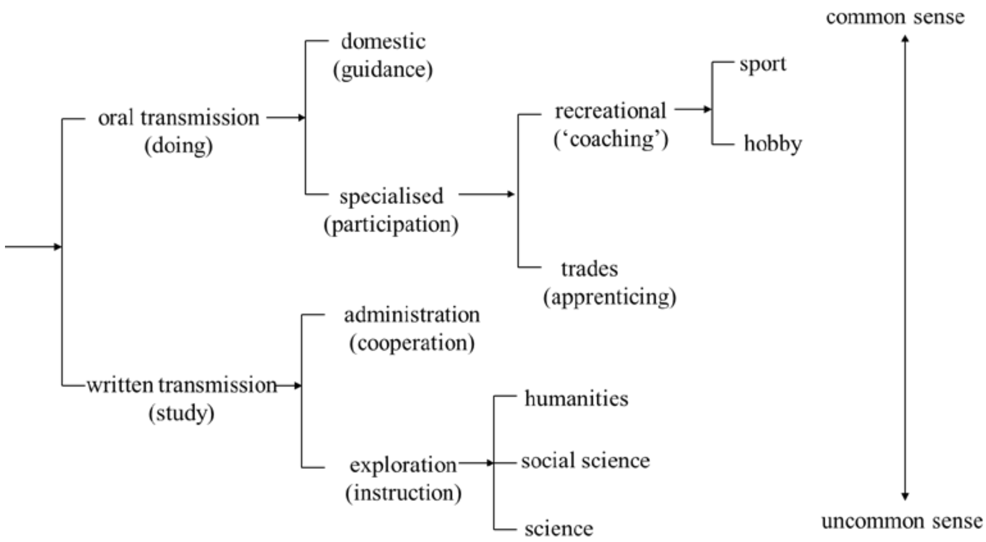


Figure 2.1 A provisional classification of fields (Martin 1992: 544)

Martin argues that the different fields are topologically categorized on a continuum from common-sense to uncommon-sense knowledge. The most common-sense fields are domestic activity sequences learned by doing by children under the guidance of caregivers, while the most uncommon-sense fields are activity sequences of exploration associated with academic disciplines. These topological categories of field from common sense to uncommon sense echo Halliday's (2004[1998]) view of knowledge development through language reviewed above. The domestic fields tend to relate to knowledge learned at home by children before the age of four, and the exploration fields are technical knowledge learned by students at schools. In this sense, knowledge development is in fact a process of developing an individual's capacity of construing increasingly specialized fields.

Martin (2017) revises his model of field that was previously treated as a set of activity sequences serving some global institutional purposes. The new model adds that each field additionally involves specialized taxonomies of people, places, and things involved in these activities, organized by classification and composition; these people, places, and things, whether abstract or concrete, may in addition to be graded in relation to one another in arrays. (ibid.: 114) The revised model takes into consideration taxonomic relations among entities and is more useful for understanding the differences between common-sense and uncommon-sense fields. Doran (2018c) applies this model to analyze knowledge-building of physics through language and finds that language builds physics knowledge through cumulatively construing large and deep taxonomies across multiple texts and multiple years of schooling. A similar finding is shown in Hao

(2020), which examines the development of taxonomies in undergraduate biology. It analyzes four biology experimental reports written by undergraduate students at different stages and reveals that both the diversity and depth of taxonomies develop across the stages. That is, as the knowledge is being built, the taxonomies become increasingly large and deep.

The studies reviewed above have provided an insightful approach to the knowledge-building of science in terms of the development of taxonomies construed by language. However, at the current stage, this approach has been largely limited to language. As science is inherently multisemoitic (Lemke 1998), it is equally significant to investigate how semiotic resources other than language, for example, chemical formalisms and images in this study, build scientific knowledge in terms of field. The present study will thus explore this issue. More specifically, Chapters 4 and 5 will examine how chemical formalisms and images build chemistry knowledge in secondary school chemistry textbooks from the perspective of field.

2.1.1.3 Genre and knowledge-building

Genre is critical for building disciplinary knowledge. Martin (2007: 56) factors disciplines as systems of genres. To access the disciplinary knowledge is thus to access genres. In science, there are four primary genres: reports that classify and describe, explanations of causes and effects, procedures for observing and experimenting, and procedural recounts for reporting on observations and experiments (Martin and Rose 2008: 141). These four genres are general genre families, which can be further distinguished into more delicate genres, as shown in Table 2.1 with their social purposes and generic stages presented as well.

Table 2.1 Genres in school science (Veel 1997: 172)

Genres	Social purposes	Stages
procedure	To enable scientific activity, such as experiments and observations, to occur	Aim^ Materials^ Steps
procedural recount	To recount in order and with accuracy the aim, steps, results and conclusion of a scientific activity	Aim ^ Record of Events ^ Conclusion
sequential explanation	To explain how something occurs or is produced – usually observable sequences of activities which take place on a regular basis	Phenomenon identification ^ Explanation sequence (consisting of a number of phases)
casual explanation	To explain why an abstract and/or not readily observable process occurs	Phenomenon identification ^ Explanation sequence (consisting of a number of phases)
factorial explanation	To explain events for which where are a number of simultaneously occurring causes	Phenomenon identification ^ Factor [1-n]
theoretical explanation	To introduce and illustrate a theoretical principle and/or to explain events which are counter-intuitive	Phenomenon identification/ Statement of theory ^ Elaboration [1-n]
consequential explanation	To explain events which have a number of simultaneously occurring effects	Phenomenon identification ^ Effects [1-n]
exploration	To account for events for which there are two or more viable explanations	Issue ^ Explanation 1 ^ Explanation [2-n]
descriptive report	To describe the attributes, properties, behaviour, etc. of a single class of object	General statement ^ Description
taxonomic report	To describe a number of classes of thing in a system of classification	General statement ^ Description
exposition	To persuade the reader to think or act in particular ways	Thesis ^ Arguments 1-n ^ Reinforcement of Thesis
discussion	To persuade the reader to accept particular position on an issue by considering more than one perspective	(for example) Issue ^ Dismissal of opponent's position ^ Arguments for own position ^ Recommendation

These genres can be arranged in a way that facilitates the development of learner pathways and helps students move smoothly from the control of one genre to another (Martin 1997: 16). According to the social purposes of the genres, Veel (1997) proposes an idealized knowledge path from procedures and recounts to theoretical explanations, as shown in Figure 2.2. In this pathway, the changing configurations of grammatical features lead students away from the kinds of meanings which are linked to the here-and-now towards the abstract and technical meaning (ibid.: 188). Through such development of genres, school science is cumulatively built throughout different stages.

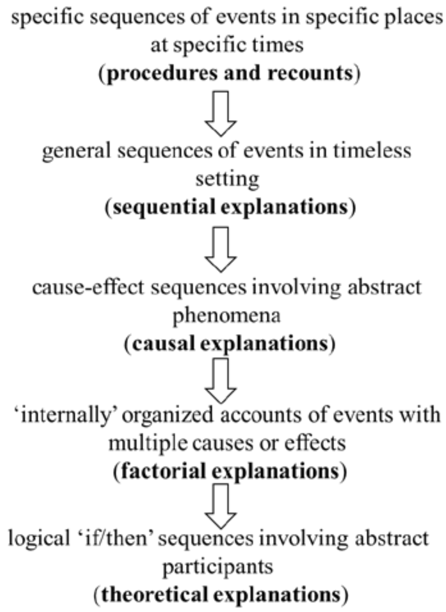


Figure 2.2 An idealized knowledge path in school science (Veel 1997: 189)

In addition to building school science in general, previous studies have also investigated knowledge-building through genres in specific disciplines. These studies mainly focus on physics. For example, Zhao (2012) analyzes experimental procedures in Australian primary and secondary physics textbooks and finds that generic stages vary as the schooling level progresses. At the lowest schooling level (level 1), four stages occur: Sum of experimental procedures, Theoretical warming-up, Experiment, and Theoretical summary; [3] the middle level (level 2) employs the same stages but uses 'Theoretical summary' much more frequently than level 1; the highest level (level 3) adds in a new stage – Theoretical exploration, but interestingly, it removes the 'Theoretical summary' stage. Doran (2018c) takes a step further than Zhao by examining genres in physics constituted by both language and mathematical

[3] According to SFL tradition of labelling generic stages (Martin and Rose 2008), the first letter of the stage name is capitalized.

symbolisms for building physics knowledge from primary school to undergraduate university. These genres are called mathematical genres. Doran develops two mathematical genres: ‘derivation’ and ‘quantification’. Derivations build new mathematical relations that have not yet been specified, while quantifications connect mathematical equations to empirical data. These genres do not occur until junior high school. Quantifications occur first in junior high school and derivations in senior high school. In contrast, undergraduate university employs large quantification and derivation complexes to a degree far surpassing that of senior high school. The knowledge-building of physics is embedded in the deployment of these genres across the schooling levels.

The above review suggests that genre plays a significant role in building scientific knowledge, and it shows certain variations as schooling progresses. However, most of the previous studies focus on genre constituted by language. It is not yet clear how genre composed of multisemiotic resources, for example, language, chemical formalisms, and images in the present study, builds knowledge. Doran’s work on mathematical genres constituted by language and mathematics can be a particularly good example for expanding the knowledge-building analysis to include genres realized by multisemiotic resources.

Overall, SFL approaches knowledge-building from a social semiotics perspective and views knowledge as meaning construed by semiotic resources. It sees knowledge-building as the development of meaning realized by these resources both in individuals (ontogenesis) and in texts (logogenesis). Previous studies drawing on this approach focus on knowledge-building in school science in terms of three strata of language – lexicogrammar, field, and genre. More studies are warranted to investigate the knowledge-building of science by semiotic resources other than language.

2.1.2 Bernstein’s knowledge structure and LCT to knowledge-building

Different from many sociologically informed approaches to education which share a subjectivist account of knowledge and ignore the key issue – what knowledge is (Maton 2014: 6), Berstein’s theory of knowledge structure and LCT take knowledge as an object of study on its own. That is, knowledge has its form and can be analyzed explicitly. This approach to knowledge is allied to a branch of sociology called “social realism” (Moore and Young 2001). Taking the forms of knowledge as an object of study shares some similarities with SFL, which studies knowledge in terms of semiotic resources that realize it. However, in contrast to SFL focusing on how the resources construe knowledge, Berstein's theory of knowledge structure and LCT are more concerned with organizing principles of knowledge in their knowledge-building analysis. The following will review the two theories' theorization of knowledge and empirical studies on knowledge-building drawing on these theories.

Berstein's theory of knowledge structure distinguishes between two types of discourse: 'horizontal discourse' and 'vertical discourse', which are in fact two forms of knowledge. Horizontal discourse refers to everyday knowledge and is characterized by being 'oral, local, context-dependent and specific' (Bernstein 1999: 159). In terms of SFL's conception of field, horizontal discourse is arguably equivalent to uncommon-sense field discussed above. Vertical discourse, in contrast, relates to disciplinary knowledge learned at schools and embodies "a coherent, explicit, and systematically principled structure, hierarchically organized, as in sciences, or it takes the form of a series of specialized languages with specialized modes of interrogation and specialized criteria for the production and circulation of texts, as in the social sciences and humanities" (ibid.: 159). Correspondingly, this type of discourse is associated with uncommon-sense field as described in SFL. The two types of discourse tend to be used in different places, with horizontal discourse more likely occurring in domestic settings, such as dialogue between parents and children, and vertical discourse in educational institutions, such as the discourse for learning knowledge at schools. Therefore, vertical discourse is more closely related to knowledge-building, which might be part of the reason why vertical discourse has been the focus of studies in either sociology of education or educational linguistics.

Alongside the distinction between horizontal discourse and vertical discourse, Bernstein distinguishes between two types of knowledge structure within vertical discourse – horizontal knowledge structures and hierarchical knowledge structures. Horizontal knowledge structures consist of "a series of specialized languages with specialized modes of interrogation and criteria for the construction and circulation of texts" and tend to be found within humanities and social sciences (ibid.: 162), which can be visualized as:

$$L^1 \quad L^2 \quad L^3 \quad L^4 \quad L^5 \quad L^6 \quad L^7 \quad \ldots \quad L^n$$

Hierarchical knowledge structures, on the other hand, "create very general propositions and theories, which integrate knowledge at lower levels, and in this way shows underlying uniformities across an expanding range of apparently different phenomena" (Bernstein 1999: 162). They are the way how natural sciences organize their knowledge. Bernstein symbolizes these structures as a triangle shown in Figure 2.3. The closer it approaches the top, the more integrated the knowledge.

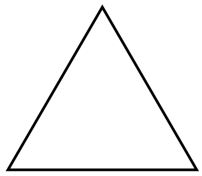


Figure 2.3 Hierarchical knowledge structures

The two knowledge structures exhibit different ways of organizing knowledge. A key difference is their mode of knowledge development: extension and integration of knowledge in hierarchical knowledge structures and accumulation and segmentation of knowledge in horizontal knowledge structures (Maton 2014: 108). Bernstein's model of knowledge structures offers a way of understanding knowledge-building over time in the two general disciplines. However, as Maton (ibid.: 109) has pointed out, the dichotomous types of knowledge structures provide a first step for theorizing knowledge practices. A key issue it has not addressed is that it does not offer organizing principles for analyzing how these knowledge structures arise from the discourse itself. Based on this, Maton (2014) develops Bernstein's code theory and establishes LCT, a theoretical toolbox that offers a series of organizing principles for analyzing knowledge-building practices. Among these principles, Semantics is the most pertinent dimension to this study that allows for 'seeing' knowledge-building in the discourse.[4]

Semantics is concerned with context-dependence and condensation of meaning in knowledge practices. It comprises two critical concepts: semantic gravity and semantic density. Semantic gravity refers to the degree to which meaning relates to its context, and semantic density refers to the degree to which meaning is condensed within symbols (terms, concepts, phrases, expressions, gestures, etc) (Maton 2011: 65-66). Semantic gravity and semantic density can be relatively stronger or weaker: the stronger the semantic gravity, the more context-dependent the meaning; the stronger the semantic density, the more condensed the meaning. In the knowledge-building process, semantic gravity tends to become increasingly weaker while semantic density grows stronger. The movement along the strength of Semantics over time forms a 'semantic wave', where knowledge is transformed between relatively context-dependent, simplified meanings and decontextualized, condensed meanings (Maton 2013).

[4] Note that 'Semantics' in initial capital refers to LCT's Semantics code. This is to distinguish it from 'semantics' in linguistics.

Semantics code provides a useful way of ‘seeing’ cumulative knowledge-building through discourse over time, which advances beyond Bernstein’s theory of knowledge structures.

Since the advent of Semantics, it has been applied to the knowledge-building analysis of numerous educational practices and has been proven rather fruitful. These studies examine diverse knowledge practices in various disciplines, covering classroom teaching of English for academic purposes (Brooke 2020; Kirk 2018; Zhao 2020), secondary school history (Matruglio et al. 2013), law (Clarence 2017b), high school biology (Maton 2013; Mouton and Archer 2019), university physics (Georgiou 2016), Jazz music (Richardson 2020), ballet dance (Lambrinos 2020) as well as knowledge-building in text such as story genres in humanities (Hood 2016). Most of these studies follow the analytical pattern - first analyze semantic gravity and semantic density of certain stages of the practices and then capture semantic waves that describe the cumulative knowledge-building over time. This can be illustrated through the analysis of classroom teaching of high school biology by Maton (2013) shown as follows.

The following is a transcription of the Year 11 biology classroom which teaches the concept ‘cilia’. (ibid.: 15)

> Teacher: Okay [student’s name] what are the ‘cilia’. What was it? No? [Student’s name] do you know what cilia is? No? Someone must know what they are…
>
> Student: Hairs
>
> Student: The little hairs?
>
> Teacher: The little hairs. And basically, they beat in an upward motion from inside your body out through to your nose. [Teacher is waving arms upwards]. So, they beat up and they take the pathogens away with them. And, guys, I don’t know if I’ve ever told you this, but when you smoke cigarettes, the tar actually causes your cilia to, because it’s so heavy, to drop, and so your cilia don’t work properly after that because they’re too heavy, they’ve dropped, so they can’t beat the pathogens out of your body! So that’s one reason that smoking’s bad as well. Okay! Alright, write this down under description!

(After telling the students to write this down under description, the teacher writes on the board: 'cilia', a brief definition, and a description of a function they serve in the body, see the following table.)

cilia	Hair-like projections from cells lining the air passages.	Move with a wavelike motion to move pathogens from the lungs until it can be swallowed into the acid of the stomach.

The teacher first introduces the abstract scientific term 'cilia' condensed with meaning and then unpacks it through a concrete example in everyday life, that smoking stops cilia from performing a function integral to their definition, and finally repacks the meaning in the technical term and its definition shown in the table. From introducing the scientific term through explaining it with everyday concrete examples to summarising the knowledge with the term and its definition, semantic gravity first increases and then decreases, whereas semantic density shows opposite shifts: decrease first and then increase. The knowledge-building process can be analyzed as the semantic wave depicted in Figure 2.4.

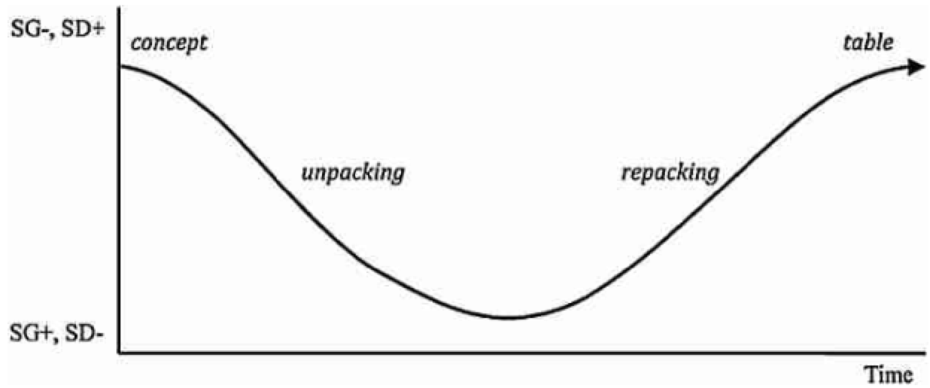


Figure 2.4 Example of a semantic wave in biology teaching (Maton 2013: 15)

As we can see, Semantics is rather helpful for unraveling how educational practices build knowledge and is potentially very useful for providing valuable feedback for teaching. However, it should be noted that the analyses of semantic gravity and semantic density are typically broad-brush as LCT emphasizes organizing principles of knowledge rather than how the knowledge practices make meaning, which is what SFL is concerned with. For example, the analysis of the above example is not built upon linguistic analysis, which may beg the question of why technical terms have weaker context-dependence and stronger condensation of meaning. The ongoing development of Semantics seems aware of this issue and begins to develop models for knowledge practices that more robustly explain the relation between Semantics and forms of knowledge. These models are called ‘translation devices’ (Maton and Chen 2016). For example, Maton and Doran (2017a, 2017b) establish a series of translation devices for relating semantic density and English discourse. The devices can serve for not only knowledge-building analysis but also text analysis. Such studies are critical as they provide a more solid basis for analyzing Semantics of knowledge practices. However, these studies are still in their infancy at the current stage. In addition to language, knowledge has many other forms, necessitating more studies exploring how they relate to Semantics. For the present study, we aim to contribute to this field by examining how Semantics relates to chemical formalisms and images used for building secondary school chemistry. In specific, Chapters 4 and 5 will examine the semantic density of chemical formalisms and images and show how their semantic density develops through schooling to build chemistry knowledge in secondary school chemistry textbooks.

Overall, Bernstein’s theory of knowledge structure and LCT take the forms of knowledge as their object of study, but they are more concerned with organizing principles of knowledge practices. LCT takes one step further than Bernstein’s knowledge structure theory by considering how concepts relate to their context and how lower-order concepts are integrated into higher-order concepts within the theory in terms of the density of the meaning. The Semantics code makes knowledge-building analysis more operable. Nevertheless, since most LCT studies do not focus on how knowledge practices make meaning, the relation between Semantics and forms of knowledge is somewhat broad-brush, necessitating further research that combines LCT and SFL.

2.1.3 Comments on the two approaches

The above review suggests that the two approaches to knowledge-building share similarities in their view of knowledge and differ in ways of analyzing knowledge-building. They are similar in that they both take forms of knowledge as the object of study, which contrasts with constructivism that perceives knowledge as mental processes and states of consciousness that reside within learners’ brains (Maton 2013: 9). However, the two approaches give prominence to different aspects of knowledge-building analysis. As a linguistic theory, SFL emphasizes the way knowledge is construed by language and other semiotic resources and studies knowledge-building in terms of the development of meaning construed by these resources. Berstein's theory of knowledge structure and LCT, in contrast, offer a theorization of knowledge that focuses on organizing principles of knowledge practices. In some sense, the two approaches are **complementary**. SFL analyses what meanings knowledge practices make, but it does not concern organizing principles of these meanings underpinning knowledge-building. Berstein's knowledge structure and LCT, on the other hand, provide such organizing principles, but they do not investigate meaning-making by the forms taken by knowledge practices. Just as Veel (1997: 170) points out, one could not comprehend the use of language in a specific discipline without also having a sense of the way that discipline organizes knowledge. Knowing what meaning the use of language construes in a discipline and how this discipline organizes knowledge are two components of a jigsaw. Missing either component would make it incomplete.

The complementarity between the two approaches in knowledge-building analysis resonates with the fruitful ongoing dialogue between SFL and LCT, which has been sparking advances in both approaches and providing ground-breaking insights on mutual concerns (for the history of the dialogue see Maton and Doran 2017d). The most recent outcome of the dialogue is the establishment of SFL's concepts of mass and presence (Martin 2017; Martin and Matruglio 2019) in response to LCT's semantic density and semantic gravity (Maton 2014). The complementarity between LCT's Semantics and SFL's certain concepts has also inspired a large number of empirical studies examing knowledge practices, for example, building knowledge through language in primary and secondary physics textbooks (Zhao 2012), through language and mathematics in physics textbooks from primary school to undergraduate university (Doran 2018c), and through spoken discourse in face-to-face lectures in health science (Hood 2017). These studies have shown that employing both SFL and LCT is rather useful for knowledge-building analysis. Following this, the present study will take advantage of this complementarity between the two theories and expand the object of study to knowledge-building of chemistry through the multismemiotic resources used in secondary school textbooks. Situating the present study in the complementarity between the two theories means that we take SFL's conception of knowledge as **meaning** and analyze how chemistry is contrued by semiotic resources, namely language, chemical formalisms and image, and employ LCT's organizing principles of knowledge to examine how the resources develop across learning levels to build the hierachical knowledge structure of chemistry.

This section has reviewed two approaches to knowledge-building and positioned the present study based on their strengths and weakness. As discussed above, we follow SFL and adopt a semiotic perspective of knowledge, and employ LCT to analyze how the meaning construed by semiotic resources builds chemistry knowledge. A key issue to this, however, is that chemistry as a scientific discipline is inherently multisemiotic. To analyze the knowledge-building of chemistry, it is thus crucial to have a sense of multisemiotic resources used in scientific disciplines at large, which will be reviewed in the next section.

2.2 Multisemiosis in disciplinary discourse

In Section 2.1, we have positioned ourselves in the conceptualization of knowledge and ways of analyzing knowledge-building based on a critical review of SFL and LCT's approaches to knowledge-building. This has clarified the key term of 'knowledge-building' in our research questions. As discussed above, we approach knowledge from a semiotic perspective, which raises another critical issue – multisemiotic resources used for building chemistry knowledge. This issue involves the way different semiotic resources construe meaning and build disciplinary knowledge. To situate this in a broader context, this section will review multisemiosis in disciplinary discourse with a focus on how they make meaning and build knowledge, which will help pinpoint key issues under debate in the scholarship where the present study is situated and identify gaps to be filled.

As O'Halloran (2011: 124) usefully remarks, theoretical and analytical issues in multimodal discourse analysis include modeling semiotic resources fundamentally different from language and analyzing intersemiotic expansions of meaning as semiotic choices integrate in multimodal phenomena. Accordingly, we will review multisemiosis in disciplinary discourse in two aspects: semiotic resources other than language (Section 2.2.1) and intersemiotic relations in multisemiotic texts (Section 2.2.2).

2.2.1 Semiotic resources other than language

This section will review meaning-making and knowledge-building by semiotic resources other than language. Following an inverted pyramid structure, we will first focus on the general background theories and then review empirical studies relevant to the present study. To begin with, we will focus on descriptions of meaning-making by non-linguistic semiotic resources.

2.2.1.1 Semiotic description of non-linguistic resources

This section will review first relevant theoretical issues (Section 2.2.1.1.1) and then empirical studies on non-linguistic semiotic resources in disciplinary discourse (Section 2.2.1.1.2).

2.2.1.1.1 Theoretical issues

Modeling meaning-making of semiotic resources other than language has been a key issue in Social Semiotics (Hodge and Kress 1988; Kress and van Leeuwen 2001; van Leeuwen 2005) and Systemic Functional tradition to multimodal discourse analysis (Baldry and Thibault 2006; O'Halloran 1999b, 2005; O'Toole 1990). Both approaches derive from Halliday's social semiotics and systemic functional theories (Jewitt 2017: 32–34). However, they exhibit considerable differences in certain aspects.

A key difference between the two traditions is that the social semiotics approach emphasizes sign-makers and their uses of semiotic resources in a given socio-cultural context, whereas systemic functional multimodal discourse analysis (SF-MDA) focuses on semiotic phenomena. With a strong emphasis on the agency of sign-makers, the social semiotics approach foregrounds individuals' 'interest' (Kress 1993) in a particular social context and focuses on how it motivates his choice of one semiotic resource over another. Put simply, the social semiotics approach underscores the process of meaning-making. SF-MDA, on the other hand, focuses on semiotic resources themselves and models their meaning potential in general that constitutes culture, contrasting with the social semiotics approach that examines individuals' uses of these semiotic resources. Overall, SF-MDA can be viewed as giving prominence to 'product', in contrast to the social semiotics approach's emphasis on 'process'.

In addition to the different focuses between the two approaches, they also differ in the degrees of borrowing theoretical categories from SFL. Generally, SF-MDA is arguably more SFL-loaded than the social semiotics approach as the former employs much more SFL categories than the latter. For example, O'Toole's (1990) model of visual art employs SFL's concept of 'rank', whereas Kress and van Leeuwen's (2006) description of image grammar is rank-free (for a detailed comparison of the two models in terms of rank see Zhao 2010b). As a representative work in SF-MDA, O'Halloran's (2005) description of mathematical symbolism pushes one step further than O'Toole by assuming almost all theoretical categories from Martin's (1992) model of language:

metafunction and rank as well the three strata – discourse semantics, grammar and an expression stratum termed ‘display’ (graphology and typography). Such a full-scale assumption of SFL theoretical categories is not seen in the social semiotics approach.

Despite the differences, the two approaches share some similarities. A key feature they share is that they tend to model non-linguistic semiotic resources in terms of the three metafunctions developed from Halliday’s (1967a, 1967b, 1968, 1969, 1970) description of English. Examples of such semiotic descriptions include mathematical symbolisms (O'Halloran 1999b, 2005), images (Kress and van Leeuwen 2006; O'Toole 1990; Painter et al. 2013), animation in science education (He 2020), bodily action (Martinec 1998, 2000, 2001), three-dimensional space (Ravelli and McMurtrie 2016; Stenglin 2009). These models tend to take the three types of meaning, ideational, interpersonal, and textual, as the basis and look for grammatical resources realizing these meanings in the semiotic resources under study. Many terms in the functional grammar of English, for example, ‘process’, ‘participant’, and ‘circumstance’, are incorporated in these models. Being closely related to language’s metafunctional organization, the semiotic descriptions listed above are often referred to as a ‘language-based model’ (Machin 2009: 182). The above numerous semiotic descriptions suggest that this approach is remarkably productive in modeling meaning-making by semiotic resources other than language.

However, productive as the language-based model is, it suffers from several problems. A key issue is that unproblematically assuming the three metafunctions developed from the description of English to other semiotic systems may homogenize semiotic description. This point is made patently clear by the following quote from Doran (2018b: 460):

> “…metafunctions were initially derived from descriptions of a small handful of languages (primarily English, Halliday 1969) and have tended to be carried over to other semiotic resources unquestioningly. This is problematic if we wish to build descriptions that bring out the specific functionality of each semiotic resource. If we take metafunctionality to be one of the broadest means by which these traditions conceptualize the intrinsic functionality of semiotic resources, by simply assuming metafunctions across semiosis, we run the risk of homogenizing descriptions and making everything look

> like the first resource to be comprehensively described (i.e., English). That is, we risk watering down the specific functionality of each resource."

Doran's questioning of the assumption of the three metafunctions across semiotic resources is echoed by numerous multimodal studies that have critiqued the applicability of the metafunctions in the analysis of non-linguistic semiotic resources, for example, van Leeuwen (1999, 2009) on sound, Djonov and van Leeuwen (2011) on texture, Zhao (2010a) on web-based multimodal information texts, and Doran (2018c) on mathematics. These studies have shown that not all of the three metafunctions are applicable to the semiotic resources under investigation. A growing number of scholars believe that the three metafunctions developed from the description of English cannot be taken for granted and unproblematically assumed to all semiotic resources, but instead, they should be motivated from the uses of the resources or, more specifically, their internal grammatical organization. van Leeuwen (1999: 190) makes this point patently clear in the following quote:

> "…different semiotic modes have different *metafunctional configurations*, and that these metafunctional configurations are neither universal, nor function of the intrinsic nature of the medium, but cultural, a result of the uses to which the semiotic modes have been put and the values that have been attached to them." (italics are original emphasis)

For van Leeuwen, metafunctions are **not universal** but specific to each semiotic resource and motivated by the uses of that resource. Bateman (2019: 534) also points out that motivation for metafunctions should be "drawn from considerations of grammatical, or more generally formal, organizations of the realizations of semiotic resources". Pushing a step further, Martin (2015b: 57) argues that metafunctions should be justified by system/structure cycles shown in the grammatical description of semiotic resources under investigation. These opinions are consistent with how Halliday develops the three metafunctions. Halliday's early papers, including "notes on transitivity and theme in English (parts 1, 2, and 3)" (1967a, 1967b, 1968), "function diversity of language as seen from a consideration of modality and mood in English"

(1970), and "options and functions in the English clause" (1969), show that his proposal of the three metafunctions is based on the grammatical systems of English, which is articulated more clearly in the following quote from Halliday (2002[1970]: 174):

> It is fairly obvious that language is used to serve a variety of different needs, **but until we examine its grammar there is no clear reason for classifying its uses in any particular way**. However, when we examine the meaning potential of language itself, we find that the vast numbers of options embodied in it combine into a very few relatively independent "networks"; and **these networks of options correspond to certain basic functions of language**. This enables us to give an account of the different functions of language that is relevant to the general understanding of linguistic structure rather than to any particular psychological or sociological investigation. (Bold fonts are my emphasis)

The above quote indicates that metafunctions are motivated by Halliday's grammatical description of English. The recognition of the three metafunctions is based on the systems of grammatical structures of English that are relatively bundled together. The systems of TRANSITIVITY, MOOD, and THEME are relatively independent of each other, suggesting that they realize three distinct metafunctional components, termed ideational, interpersonal, and textual, as shown in Figure 2.5.

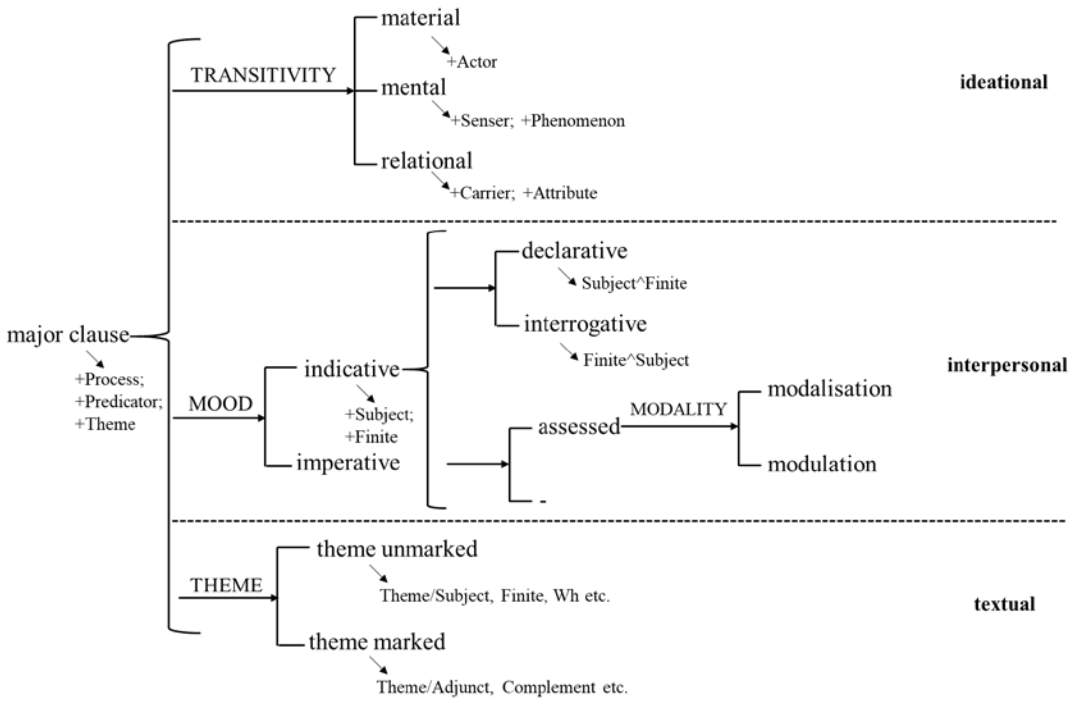


Figure 2.5 The three independent grammatical systems (simplified) and the three metafunctions (adapted from Halliday 2003[1973]: 315)

Figure 2.5 shows that metafunctions are in nature **clusters of systems** relatively interdependent of each other. It is this interdependency of systems in the description of English that gives rise to metafunctions, as attested in the following quote from Halliday (1979: 301):

> "The categories of ideational, interpersonal and textual appear clearly in the semantic system itself, as system networks each having a high degree of internal dependence but a very low degree of external dependence. Choices made within one component have a great deal of effect on other choices within the same component but hardly any effect on choices in the other components."

Martin (2015a: 280) also usefully summarizes Halliday's generation of the three metafunctions from the grammatical systems of English:

> "Interdependency of this kind is the inspiration for Halliday's concept of metafunctions, which he uses to characterize the intrinsic functional organization of language. The relatively interdependent

> systems we are focusing on here cluster together as interpersonal meaning from a metafunctional perspective, complemented by additional clusters of experiential and textual systems. Critically, it is thus **axis** (i.e., system privileged over structure and formalized in system networks) that give rise to and ultimately underpins this well-known dimension of SFL theory – the interpersonal, experiential and textual metafunctions.” (Bold fonts are my emphasis)

The fact that metafunctions derive from systems suggests that the axial relations between system and structure are the more fundamental theoretical categories. Therefore, to bring out the intrinsic functionalities of non-linguistic semiotic resources, we should follow the same path that Halliday has taken for the description of English. That is, semiotic descriptions should begin with describing the grammatical systems of semiotic resources based on system and structure and then generate the metafunctions from the axial principle.

A huge advantage of taking system and structure as the point of departure for semiotic description is that it can avoid watering down the intrinsic functionalities of non-linguistic semiotic resources. This can be illustrated by a brief comparison of Doran’s (2018c) and O’Halloran’s (2005) descriptions of mathematical symbolism. As mentioned previously, O’Halloran follows a language-based approach to model mathematical symbolism, assuming metafunctions and ranks developed from the description of English. Basically, her model takes the four metafunctional components, i.e., experiential, logical, interpersonal, and textual, as the point of departure and looks for grammatical resources realizing these meanings at four ranks: statement, clause, expression, and component, parallelling with the ranks of clause complex, clause, groups/phrase and word in the grammar of English. This model focuses more on the structure and does not involve system/structure cycles. Martin (2015b: 45) takes such description as “based more on Halliday’s earlier scale and category modeling”.

In consistent with Halliday’s approach to the description of English, Doran also takes the axial principle as the descriptive foundation and develops exhaustive grammatical systems for mathematical symbolism based on comprehensive data from primary school to undergraduate university. With these systems, he generates three metafunctions: operational, logical, and textual. In addition, the system and structure circles in his grammatical description suggest two types of hierarchy between grammatical units in mathematical symbolism: a nesting hierarchy involving statements and symbols and a rank-scale involving symbols and elements.[5]

As we can see, instead of simply assuming the metafunctions and ranks developed from the description of English, Doran takes the axial principle as the point of departure and discovers mathematical symbolism's intrinsic functionalities and hierarchical relations between its grammatical units. This allows for comparing mathematical symbolism and language in a more principled nature and revealing their similarities and differences in their functionalities and organization. The language-based model, in contrast, misses this. This once again suggests that to bring out the intrinsic functionalities of non-linguistic semiotic resources, it is best to take the axial principle as the descriptive foundation.

The above has reviewed the two approaches to the modeling of meaning-making by semiotic resources other than language. The Social Semiotics approach tends to emphasize sign-maker and process of meaning-making, while SF-MDA typically focuses on multimodal phenomena and models the meaning potential of these semiotic systems. Since the present study aims to investigate how chemical formalisms make meaning, our approach is relatively more associated with SF-MDA. However, as discussed above, we intend to **move away** from the language-based model and do not assume secondary theoretical categories such as metafunctions developed from the description of English. Instead, following Halliday's method of modeling English, we take the more fundamental theoretical primitive of axial principle as the point of departure and describe chemical formalisms in terms of their own grammatical organization. Taking the axial principle as the descriptive foundation does not mean that the three metafunctions would not occur in semiotic resources other than language, but that they need to be justified in terms of their internal grammatical systems.

2.2.1.1.2 Empirical studies

Having reviewed the background theories underpinning semiotic description, we will now zoom in to focus on studies on meaning-making by semiotic resources other than language in disciplinary discourse.

The most frequently examined semiotic resource other than language in disciplinary discourse is mathematical symbolism, which is widely used in a range of disciplines. As reviewed above, O'Halloran (2005) and Doran (2018c) have studied

[5] The term 'symbol' here refers to mathematical symbols. It is not the theoretical term 'symbol' in Perice's classifications of signs, i.e., symbol, icon, and index.

mathematical symbolism in mathematics and physics, respectively. They approach the grammar of mathematical symbolism with different theoretical assumptions. O'Halloran follows a language-based approach, while Doran takes the axial principle as the point of departure. It has been argued above that Doran's approach is more advantageous as it captures the intrinsic functionalities of mathematical symbolism and the hierarchical relations between the grammatical units. The present study will thus take system and structure as the point of departure for the description of chemical formalisms in Chapter 4.

In addition, Doran (2019) takes a step further by comparing the meaning-making of a series of academic formalisms used in academic discourse – system networks and tree diagrams in linguistics, and algebraic mathematical symbolism and nuclear equations used in physics – in terms of their types of structure (Halliday 1981[1965]) and field-specific meanings realized. Doran finds that from the perspective of field, they each construe significantly different meanings oriented to their disciplines. Nevertheless, they display a significant similarity in terms of their structural organization, i.e., they are primarily organized around iterative structures. By not assuming certain categories such as metafunction and rank from the description of English, Doran is able to compare the similarities and differences in meaning-making by these academic formalisms and develops a functional semiotic typology for them. This is rather heuristic for examining the similarities and differences among the various chemical formalisms to be investigated in the present study. For example, chemists use chemical formulas, chemical equations, and structural formulas to construe chemistry. How similar or different these formalisms organize meaning is critical for understanding how the chemistry discipline organizes its knowledge. Establishing a functional semiotic typology for these chemical formalisms would help address this issue. The present study will thus develop a functional semiotic typology for chemical formalisms in Chapter 4.

In addition to the above academic formalisms, chemical formalisms are also a key component. As one of the objects of study in the present work, they will be reviewed in more detail. Chemical formalisms have been studied in both chemistry education and multimodal discourse analysis. In the field of chemistry education, chemists tend to focus on **formal rules** governing chemical formalisms' formation (Goodwin 2008; Jacob 2001; Taber 2009). For example, Jacob's (2001: 35) 'chemical orthography' provides the rules that govern the combination of chemical symbols to chemical formulas (e.g., valency, oxidation state). For example, the chemical symbols ‘Na’ (sodium) and ‘Cl’(chlorine) can be combined to form ‘NaCl’ (sodium chloride) using the rule that one sodium atom can be combined with one chlorine atom. ‘Chemical grammar’, on the other hand, provides the rules that govern reaction equations, which determines the stoichiometric coefficients, the use of a unidirectional or an equilibrium arrow, and ‘rection conditions’ as long as they are part of a reaction equation (e.g., solvent, temperature) (Jacob 2001: 35), for example, $2Na(s) + Cl_2(g) \rightarrow 2NaCl(s)$. These formal rules may be useful for teaching students to write chemical formulas and equations. However, how they are systematically organized to make meaning is relatively underexplored.

In the field of multimodal discourse analysis, a key study is Liu (2011) that associates with SF-MDA. Similar to O’Halloran’s (2005) description of mathematical symbolism, Liu follows a language-based approach to model chemical equations (chemical formulas are a part of chemical equations). He starts with the four types of meaning: experiential, logical, interpersonal, and textual, and three ranks: clause, expression, and component, paralleling with the three ranks of clause, group, and word in the grammar of English, and then looks for grammatical resources that notionally realize the meanings at these ranks. No grammatical systems are developed in the analysis.

There are several issues worthy of discussion in Liu’s analysis of chemical equations. Firstly, Liu’s language-based analysis of chemical equations is characterized by assuming the four types of meaning developed from the description of English and then looking for scattered grammatical resources realizing these meanings. Although this approach is rather productive for analysis, it appears not consistent with Halliday’s approach to the description of English that begins with describing grammatical systems based on system and structure and then see what metafunctional organization the description reveals (see discussion in Section 2.2.1.1.1). As argued above, assuming the four functional components developed from English would inevitably homogenize the description of chemical equations since it comes to this non-linguistic resource with the premise that it share the same functionalities as language without first investigating how that resource is used for making meaning. This is evidenced by a contradiction in Liu’s analysis that the interpersonal metafunction is assumed, but no grammatical systems are found realizing this meaning. Therefore, to bring about the intrinsic functionalities of chemical equations, it is best to follow Halliday’s approach by first describing the grammatical systems of chemical equations based on system and structure and then see what metafunctional organization the description shows. Based on Halliday's axial principle, the present study will thus provide an alternative description of chemical formalisms that aims to capture their intrinsic functionalities in Chapter 4.

In addition to the theoretical issue underpinning description, some specific descriptive issues in Liu's analysis are also worthy of discussion. One of these issues is that it draws heavily on the functional grammar of English (Halliday 1994), and some of the grammatical analyses are based on translating chemical equations into English. For example, when looking for grammatical resources realizing the logical metafunction, Liu translates a chemical equation into English and analyzes the equation in terms of the translation. For example, the chemical equation $Mg + O_2 \rightarrow MgO$ is translated into the English clause complexes *when magnesium burns in air, it combines with the oxygen to form magnesium oxide*, of which the analysis of the logical relations is shown in Table 2.3 (Liu 2011: 167).

Table 2.2 Analysis of logical relations in a linguistic representation of $Mg + O_2 \rightarrow MgO$

Clause 1	*Clause 2*	*Clause 3*
When magnesium burns in air,	it combines with the oxygen	to form magnesium oxide.
Modifier	Head	Modifier
×β	α	×β
Temporal		Causal-conditional

Liu argues that the chemical equation construes the same logical relations as the clause complexes, which are implicitly realized by the arrow sign in the equation. However, we argue that by translating chemical equations into English clauses and analyzing their structure in terms of English's model, the specific structural patterns of chemical equations will inevitably be neutralized, and chemical equations will become just like English. This can be illustrated through a comparison between the internal structures of the English clause complexes and the chemical equation. The clause complexes involve the complexing of three clauses, whereby the logico-semantic relations arise (Halliday and Matthiessen 2014: 428). The chemical equation, however, does not involve such complexes. There is no evidence from the equation's internal structure to suggest the logical relations. To genuinely captures chemical equations' grammatical organization, the description should focus on their internal structures and analyze their own structural patterns. The present study will thus describe chemical formalisms in terms of their internal grammatical organization in Chapter 4.

Another critical issue is about data. There are two points worthy of discussion here. One is that for systemic functional description, **exhaustiveness** is a key feature. Halliday and Matthiessen's (2014) description of English, for example, includes almost all types of clauses. Being exhaustive is critical because it allows for capturing the full meaning potential of semiotic resources under study. Liu's description, however, seems to be depending on only a handful of chemical equations, or more specifically, around four in total, as found in his study (Liu 2011). Moreover, these four chemical equations are of the same type, i.e., non-reversible equations. The data thus cover only part of chemical equations and misses some critical types, for example, reversible equations such as '$N_2(g) + 3H_2(g) \rightleftharpoons 2NH_3(l)$' and ionic equations such as '$NaCl(s) \rightarrow Na^+(g) + Cl^-(g)$'. Without exhaustive description, the meaning potential of chemical equations mapped is somewhat restricted. The present study will thus cover all types of chemical equations used in secondary school chemistry in Chapter 4.

Another issue about the data is that most of the chemical equations in Liu's analysis are incomplete equations. For example, the equation mentioned above, $Mg + O_2 \rightarrow MgO$ (Liu's original example), is an incomplete equation. A complete one should include coefficients (numbers indicating units of chemical species) and state symbols (signs denoting physical states of chemical species), for example, $2Mg(s) + O_2(g) \rightarrow 2MgO(s)$. The number '2' attached to 'Mg' and 'MgO' are coefficients, indicating that there are two units of magnesium and magnesium oxide. The symbols '(s)' and '(g)' denote two respective physical states: solid and gas. That is, magnesium and magnesium oxide are in solid state, and oxygen gas is in gas state.

Incomplete equations are often used in lower schooling levels and gradually develop into their complete forms in higher levels (Taber 2009). This then raises an issue as to which type of equation should be taken as objects for semiotic description. In contrast to Liu, we argue that complete chemical equations should be taken as the descriptive objects. We can analogize the two equations to the linguistic systems of children and adults. To describe the grammatical systems of English, adults' usage of English is preferred as the object for description because they include the full meaning potential of the language (Halliday 1985; Martin 1992). This is no different for chemical equations. To describe the full meaning potential of chemical equations, we must take the fully developed equations as the object for description as they encode all the meaning of the writing system. The present study will thus take the complete chemical equations as the descriptive objects in Chapter 4.

In a nutshell, the above has reviewed both theoretical and empirical studies on meaning-making of semiotic resources other than language in disciplinary discourse, through which we have pinpointed key theoretical issues under debate in the scholarship and argued for the present study's position. Through a relatively detailed review of Liu's work on chemical formalisms, we have identified gaps to be filled in the description of chemical formalisms. However, examining the meaning-making of semiotic resources is only a first step as it informs how they construe knowledge. Another major issue is how they build knowledge cumulatively, which will be reviewed in the next section.

2.2.1.2 Knowledge-building

As has been reviewed in Section 2.1.1, language builds disciplinary knowledge primarily through grammatical metaphor and technicality at lexicogrammatical stratum and genre development at genre stratum. A question arising from this is how semiotic resources other than language build disciplinary knowledge. This section will review previous studies on this issue.

As noted above, mathematical symbolism is one of the most frequently studied semiotic resources other than language in disciplinary discourse. Previous studies include, for example, O'Halloran (2005) and Doran (2018c). However, among these studies, only Doran examines how mathematical symbolism builds knowledge as learning develops. To understand the role of mathematical symbolism in building physics knowledge, Doran investigates its development in physics textbooks across schooling stages from primary school to undergraduate university. He finds that mathematics rarely occurs in primary school. Junior high school introduces a mathematical genre called 'quantification', which connects theories in physics with empirical data. In LCT's terms, this strengthens semantic gravity and describes the empirical world. Senior high school brings in a new mathematical genre called 'derivation', which develops and specifies new sets of relations that become part of the technical field. It strengthens semantic density and builds new knowledge. Undergraduate university continues employing large sets of derivations, building increasingly complex sets of relations and thus even stronger semantic density. The development of these mathematical genres reveals mathematical symbolisms' role in building physics knowledge: building knowledge and describing the empirical world

(Doran 2017). Doran's study is rather heuristic for investigating the knowledge-building of chemistry through chemical formalisms. For example, incomplete chemical equations are often used in lower schooling levels and gradually develop into complete ones, experiencing a process of strengthening semantic density and building new knowledge in chemistry. However, thus far, no attempt has been made to examine knowledge-building through chemical formalisms using Semantics from LCT. The present study will thus attempt to fill this gap by exploring how the semantic density of chemical formalisms develops across schooling levels and builds chemistry knowledge in Chapter 4.

Images are also one of the most researched semiotic resources. However, most studies mainly focus on what meaning images construe, and few investigate how they build knowledge. To the best of our knowledge, two works are relatively relevant here. One is Dimopoulos et al. (2003) that analyze the content specialization of images used in school science textbooks. They distinguish three types of images in science - realistic (images that represent reality according to human optical perception), conventional (graphs, maps, flowcharts, molecular structures constructed according to the techno-scientific conventions), and hybrids (images that include elements from both the other two types). They argue that conventional images correspond to strong, hybrids to moderate, and realistic to weak levels of 'specialization'. That is, the conventional images express the most techno-scientific knowledge, whereas the realistic ones convey 'everyday' knowledge. As the education level of the school science textbooks rises, more conventional images tend to be used, and the content specialization projected by the visual images increases. The above study provides a useful perspective of capturing knowledge development through analyzing types of images used across schooling levels. However, there appears to be a lack of robust connection between the three categories of image and their degrees of specialization. A problematic issue is that realistic images can also embody a relatively high level of specialization. For example, an image of a chemical apparatus, such as a distillation apparatus, could be 'realistic' but expresses relatively complex technical knowledge. This suggests that a more robust model is needed for capturing the complexity of knowledge expressed by images.

The other pertinent work is Doran (2018a), which examines diagrams and graphs in physics. One crucial point of this study is that it uses the SFL dimension of field (Doran and Martin 2021; Martin 2017) to analyze the field-specific meanings represented in the images and Semantics from LCT (Maton 2014) to study their roles in knowledge-building. It shows that a special feature of a diagram is that it can contain multiple structures overlaid on each other, realizing multiple field-specific meanings. In terms of semantic density, these diagrams display relatively strong semantic density. Such types of diagrams allow the field of physics to expand and build increasingly technical knowledge. For graphs, they realize a distinct dimension of field, i.e., array, which refers to properties that can be graded, such as color, size, and shape (Doran and Martin 2021). For graphs in physics, the field-specific arrays organize technical meanings in a field along a particular dimension. In terms of semantic gravity, through these arrays, graphs strengthen semantic gravity by connecting physics theories to their empirical objects of study and weaken semantic gravity by abstracting empirical measurements to theories. In comparison to the study by Dimopoulos et al, Doran establishes more robust reasoning for why some images have stronger semantic density than others in terms of field in SFL. Through field analysis of images, different field-specific meanings are shown to be layered over one another, directly indicating the density of meaning. This offers a scaffold for analyzing knowledge-building through images. However, it should be noted that comparing the semantic density of images based on how many structures they include seems impracticable for distinguishing the semantic density of everyday and technical images that embody the same number and types of structures. For example, a diagram of a tree and a model diagram of a water molecule, shown in Figures 2.6 and 2.7, both embody an analytical structure (Kress and van Leeuwen 2006) and realize one level of compositional taxonomy in terms of field (a tree is composed of trunks, branches, and leaves; a water molecule is composed of one oxygen atom and two hydrogen atoms).

Figure 2.6 A diagram of a tree Figure

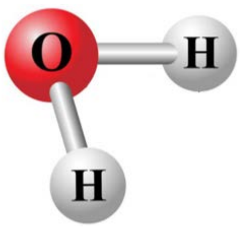


Figure 2.7 A diagram of a water molecule

However, assessing them as having the same level of semantic density clearly contradicts with the common sense that the concept of water molecules is much more complex than trees as the former involves specialized chemistry knowledge and sits in a complex constellation of meanings, while the latter is an everyday entity that involves rather simple meanings. This suggests that Doran's analysis of the semantic density of images needs to include the consideration of the complexity of concepts themselves represented in images. To fill this gap, the present study will develop a model for the semantic density of images that considers both the type and numbers of field meanings and their **technicality** in Chapter 5.

Another issue that Doran's study has not touched upon is the development of images across schooling levels. Tracing the development of images could reveal expansion or distinct shifts in meaning realized by images across schooling levels, which would help unveil the nature of cumulative knowledge-building through images. To fill this gap, this study will analyze the knowledge of chemistry by tracing the development of images across schooling levels in secondary school chemistry using the model for the semantic density of images developed in Chapter 5.

In summary, previous studies mainly focus on knowledge-building through mathematical symbolism and images. Doran's study on knowledge-building through mathematical symbolism is rather heuristic for analyzing how chemical formalisms build chemistry knowledge. Previous work on knowledge-building through images provides a useful first step for seeing the semantic density of images and a broad sense of how knowledge can be built through images. However, there appears to be a gap in the analysis of images' complexity – a more robust model for the semantic density of images is needed that can be used for analyzing knowledge-building of chemistry across learning stages. The present study seeks to fill this gap in Chapter 5 by establishing such a model for the semantic density of images and using it to trace the development of images across schooling levels in secondary school chemistry.

This section has focused on meaning-making and knowledge-building by semiotic resources other than language. However, in multisemiotic texts, they do not make meaning and build knowledge on their own but instead through interaction. The next section will thus review previous studies on this issue.

2.2.2 Intersemiotic relations in multisemiotic texts

Intersemoitic relations in multisemiotic texts as an object of study have been examined by numerous scholars from various perspectives. This section will tease out these studies and attempt to identify gaps to be filled by the present study. It will first review studies on intersemiotic relations in general and then zoom in to studies on disciplinary discourse relevant to this study. To start with, we will review frameworks that describe intersemiotic relations in multisemiotic texts in general.

2.2.2.1 Studies on intersemiotic relations in multisemiotic texts in general

Relations between language and image have been most frequently studied by previous research. The earliest work, to the best of our knowledge, is Barthes (1977[1964]) which proposes three types of language-image relations: anchorage (language fixes interpretation of the image), illustration (image determines the interpretation of language), and relay (language and image are at the equal status and are interdependent parts of a multisemiotic text). In contrast to relay, anchorage and illustration capture the unequal status between language and image, and the two categories are in reverse relation. This model is ground-breaking and stands as the starting point for all those concerned with language-image relations (Bateman 2014: 31). Subsequent studies more or less relate to Barthes' model and can be categorized as 'grammar-based' (Section 2.2.2.1.1) or 'discourse-based' (Section 2.2.2.1.2). The following will review these studies in terms of the two categories.

2.2.2.1.1 The grammar-based approach

The most representative work in the 'grammar-based' approach to the intersemiotic relation between language and image is Martinec and Salway (2005), which draws on SFL's theorization of logico-semantic relations (Halliday and Matthiessen 2014), the relations that combine clauses into clause complex. It is taken as grammar-based because it strictly follows the clause-combining relations in the grammar of English clause complex. Their model of image-text relations is shown in Figure 2.8.

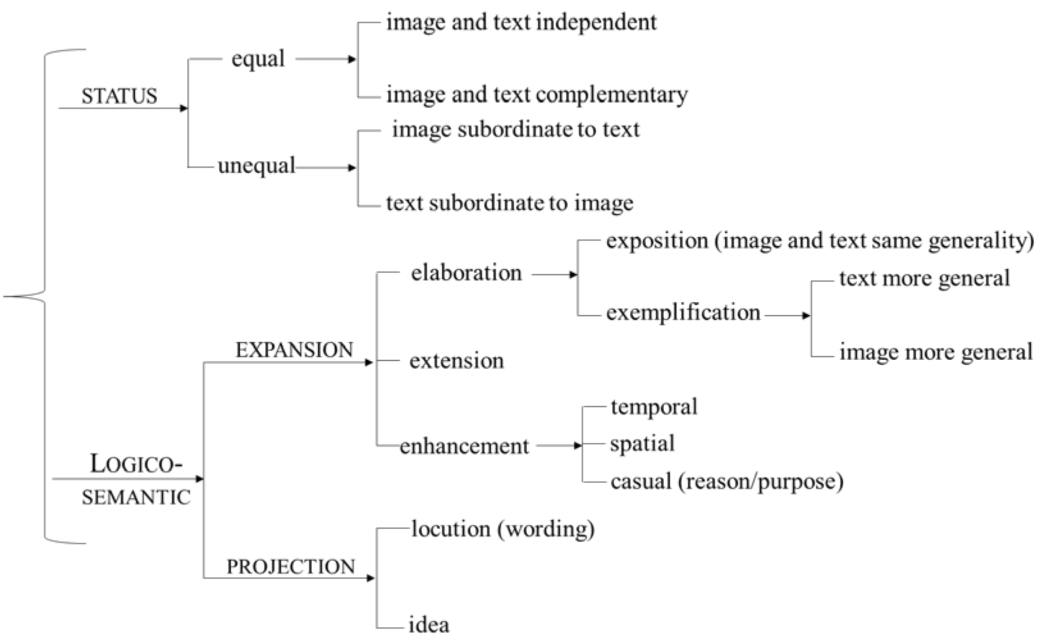


Figure 2.8 The network of image-text relations (Martinec and Salway 2005: 358)[6]

The categories 'expansion' and 'projection' in the LOGICOSEMANTIC system are the same relations through which English clauses are combined. What is different is that 'elaboration' is further specified as 'exposition' and 'exemplification'. 'Exposition' means that image and language text interact at the same level of generality, whereas 'exemplification' indicates their interaction at different levels of generality - either language text more general than image or vice versa. The system of STATUS draws on another type of relation between clauses – status (Halliday 1985, 1994). Like clauses can be of equal or unequal status, language text and image share similar relations. The equal status between language text and image proposed by Martinec and Salway resonates with Barthes' 'relay', capturing the equal importance of the two semiotic resources in a multisemiotic text. Martinec and Salway, however, take a step further and introduce two subtypes: independent and complementary. Barthes' 'relay' is equivalent to 'complementary', both of which describe that language text and image are equally important and contribute to the whole meaning of a multisemiotic text. The two subcategories within 'unequal status' share some similarities with but also differ from Barthes' anchorage and illustration. They are similar in that they all describe unequal relations between language text and image, but distinct in that Martinec and Salway describe the unequal relations in terms of subordination between language text and image, in contrast to Barthes' modeling in terms of whether image determines the

[6] 'Text' is often used interchangeably with 'language'. To avoid confusion over different senses of 'text' arising from differing contexts where it is used, for example, 'text' in 'image-text relations' and 'multisemiotic text', we use language text when 'text' refers to monomodal text composed of language.

meaning of language text or vice versa.

Overall, Martinec and Salway’s model develops Barthes’ framework and offers useful guidelines for identifying intersemiotic relations between language text and image. However, it is not without problems. One of the key issues is that Martinec and Salway’s model relies heavily on which parts of a language text being related to which parts of an image. For example, the relative status of an image and language text is considered independent when the whole image is related to the whole language text; an image is considered subordinate to a language text when it relates to only a part of the language text. (Martinec and Salway 2005: 346) This indicates that which parts of an image being related to which parts of a language text are key for their identification of language-image relations. However, as pointed out by Bateman (2014: 197–198), it is often difficult to know precisely which bits of a language text are being related to which bits of an image and on what basis this decision is made. For example, when analyzing the language-image relation in Figure 2.9, Martinec and Salway take the tittle ‘stars of the sea’ as a text separate from the subsequent paragraph and analyze the relation between it and the image. It is, however, not clear why the title and language text bodies should be taken as different texts. Without an explicit account of what counts as a language text unit, the analysis of language-image relation may not stand up to scrutiny.

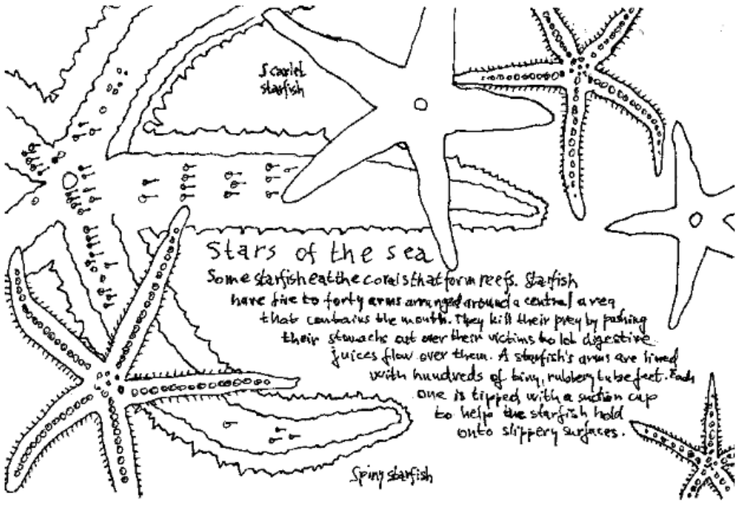


Figure 2.9 Martinec and Salway’s (2005: 346) example of image-subordinate-to-text relation

Another key study drawing on logico-semantic relations is Unsworth (2007), which extends Martinec and Salway’s framework. As shown in Figure 2.10, Unsworth extends both projection and expansion.

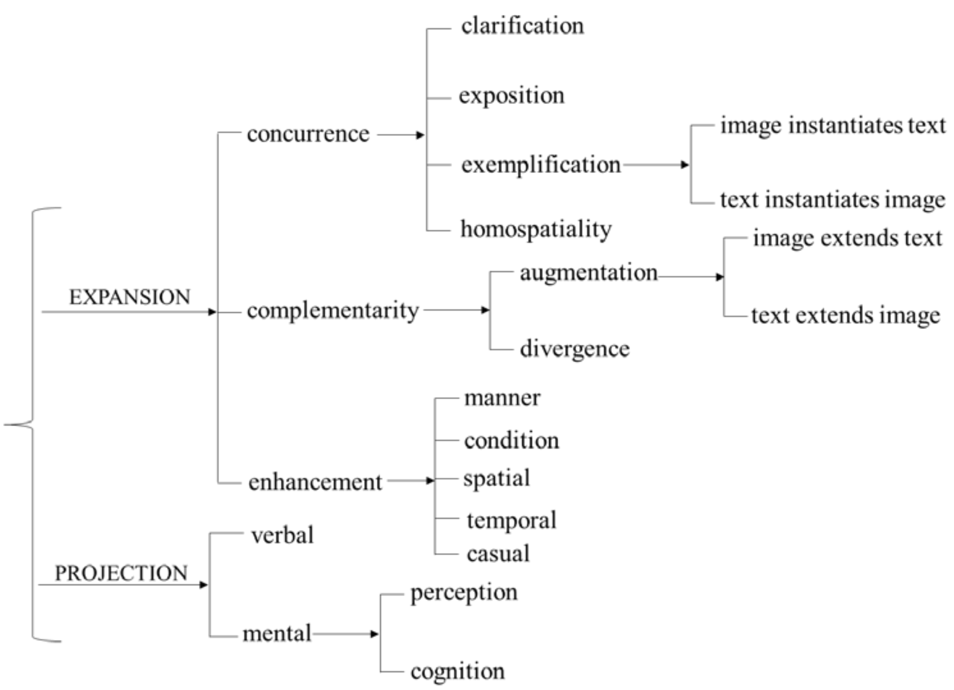


Figure 2.10 Network of image-language relations (Unsworth 2007: 1175)

Within projection, a further classification of 'mental' is specified: 'perception' and 'cognition'. Within expansion, elaboration and extension are renamed as 'concurrence' and 'complementarity'. Enhancement remains the same, but two subcategories are added, i.e., 'manner' and 'condition'. In 'concurrence' (elaboration), two categories, 'clarification' and 'homospatiality', are added. 'Clarification' refers to the relation that image clarifies or explains language text or vice versa. It can be glossed as 'to be precise'. 'Homospatiality' refers to multisemiotic texts where two different semiotic modes cooccur in one spatially bonded homogenous entity. (Unsworth 2007: 1175–1176). These semiotic artifacts can be construed simultaneously as words and as images. For example, Unsworth illustrates this relation with Figure 2.11. It shows a pile of wood burning with heat arising denoted by the curved lines, which at the same time can be read as the word 'HOT'. The typographical form of the word is integrated with the image for visual effect.

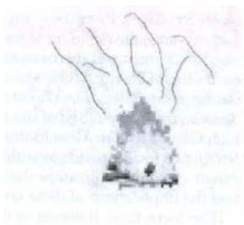


Figure 2.11 Homospatiality: 'hot' (Unsworth 2007: 1184)

Within 'complementarity' (extension), two further categories are included:

'augmentation' and 'divergence'. 'Augmentation' involves an image extending the meanings to those realized by language text or vice versa. Divergence, on the other hand, refers to the relation that the meanings of image and language text are discordant.

Overall, Unsworth's work on intersemiotic relation between image and language can be viewed as an extension of the LOGICOSEMANTIC system in Martinec and Salway's model. Some of the extensions are predictable from the grammar of language, for example, the further division of categories under projection, and some are more detailed categories generalized from Unsworth's data. It remains to be seen whether these finer categories apply to other data, or whether the framework can be further developed.

2.2.2.1.2 The discourse-based approach

In addition to the grammar-based studies, some scholars follow a discourse-based approach. These studies do not take the lexicogrammatical formulation of interdependency between clauses as their basis but draw on linguistic theories at the stratum of discourse for modeling intersemiotic relations between language and image. These studies will be reviewed in the following.

Based on conjunctive relations in English (Halliday 1985; Martin 1992), van Leeuwen (2005) models the links between language and image as Table 2.3.

Table 2.3 van Leeuwen's model of language-image relations (van Leeuwen 2005: 230)

Elaboration	Specification	The image makes the text more specific (illustration). The text makes the image more specific (anchorage).
	Explanation	The text paraphrases the image or vice versa.
Extension	Similarity	The content of the text is similar to that of the image.
	Contrast	The content of the text contrasts with that of the image.
	Complement	The content of the image adds further information to that of the text, and vice versa ('relay').

'Elaboration' includes two subcategories: 'specification' and 'explanation', and 'extension' consists of three subcategories: 'similarity', 'contrast', and 'complement'. The two further subcategories within 'specification' correspond to Barthes' 'illustration' and 'anchorage', and the category 'complement' corresponds to 'relay'. van Leeuwen's framework also shares some similarities with Unsworth's model discussed above. 'Explanation' and 'contrast' are similar to Unsworth's 'clarification' and 'divergence'. Overall, although van Leeuwen's work is based on conjunctive relations in English, it resonates with studies in the grammar-based approach.

Another study based on conjunctive relations in English is Liu and O'Halloran

(2009). Specifically, they draw on the four basic options in Martin's system of external conjunctions (Martin 1992; Martin and Rose 2003) to model intersemiotic logical relations between language and image. Their framework is shown in Table 2.4 (Liu and O'Halloran 2009: 384).

Table 2.4 Intersemiotic logical relations between language and image

Logical relations		**Meaning**
Comparative	Generality	Similarity
	Abstraction	
Additive		Addition
Consequential	Consequence	Cause
	Contingency	Purpose
Temporal		Successive

Most categories presented in Table 2.5 are included in Martin's classification of external conjunctions (Martin and Rose 2003: 133), except for 'generality' and 'abstraction'. As shown in Liu and O'Halloran's (2009: 379) analysis, language is usually more general and abstract than images, which could be explained by 'epistemological commitment' of modes (Kress 2003) – the visual mode depicts entities or activities while language symbolizes. Images thus tend to be more specific and concrete.

The work done by Liu and O'Halloran is a valuable attempt towards modeling intersemiotic relations between language and image based on conjunctive relations in English. However, their model seems relatively simple – subcategories are developed only in 'comparative' and 'consequential', much simpler than Martin's system of external conjunctions. It remains to be tested whether the model can adequately explain a wider range of multisemiotic texts composed of language and image.

In addition to studies drawing on SFL theories of conjunctive relations in English, some other studies approach language-image relations based on cohesion theories in SFL. The most influential work in this regard is Royce (2002, 2007). Drawing on Halliday and Hasan's (1976) cohesion theories, Royce models relations between language and image as 'intersemiotic complementarity'. A key concept for analyzing this is 'visual message elements' - visual features which carry semantic properties. The cohesion analysis between language and image is based on relations between these visual elements and their counterparts in language. With these, Royce establishes a framework for intersemiotic complementarity between language and images from a metafunctional perspective. In terms of ideational meaning, language and image are related lexico-semantically through intersemiotic sense relations - repetition, synonymy, antonymy, hyponymy, meronymy, and collocation; interpersonally, the two semiotic resources are related through intersemiotic reinforcement of address and intersemiotic attitudinal congruence and attitudinal dissonance relations; textually, they are related by the compositional relations of information value, salience, visual framing, visual synonymy, and potential reading paths.

In contrast to the studies on intersemiotic relations between language and image reviewed above, Royce provides a relatively fine-grained analysis of language-image relations. This level of granularity is afforded by decomposing an image into visual elements from which cohesive relations with language arise. Nevertheless, this raises some questions concerning the recognition of cohesive ties between language and image. A key issue is that it is rather tricky to independently identify which ‘parts’ of an image are visual message elements. As Bateman (2014: 173) points out, visual message elements are posited only after cohesive analysis is done. That is, we can hardly identify visual message elements independently of language, which makes the analysis of the cohesive ties less compelling.

In this section, we have reviewed frameworks of intersemiotic relations between language and image in multisemiotic texts in general that draw on SFL theories at the strata of lexicogrammar and discourse. From the above review, we can see that these frameworks provide a rich understanding of intersemiotic relations between language and image. However, none of them investigate how the semiotic resources interact to build disciplinary knowledge. To have a sense of how this issue has been researched, we will now zoom in to studies on the interaction between semiotic resources in disciplinary multisemiotic texts.

2.2.2.2 Studies on intersemiotic relations in disciplinary multisemiotic texts

This section will review intersemiotic relations in disciplinary multisemiotic texts, with a focus on how different semiotic resources interact to build knowledge. To begin with, we will review O’Halloran’s study on ‘semiotic metaphor’ in mathematic texts.

O’Halloran (1999a, 2005) proposes the concept of semiotic metaphor based on Halliday’s (1985) grammatical metaphor and argues that they are similar in that both involve a shift in the function of elements and an introduction of new entities. However, for semiotic metaphor, the semantic shifts take place as a result of movements between semiotic resources (O'Halloran 1999a: 321). For example, O’Halloran illustrates the semiotic metaphor phenomenon in a mathematic text with Figure 2.12. The circumstantial elements 'how high' and 'how wide' realizing Extent (spatial distance) in language are transformed into participants 'h' and 'h-10' in the diagram. The status of the functional element as circumstance undergoes a transformation to participants in the semiotic metaphor.

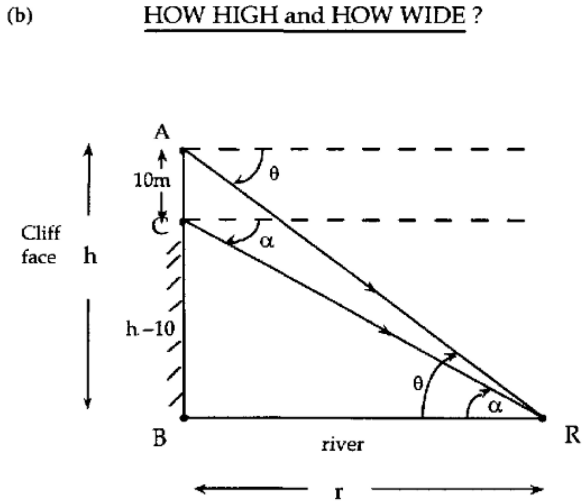


Figure 2.12 Visualization of the problem (O'Halloran 2005: 181)

O'Halloran argues that semiotic metaphor always involves a metaphorical expansion of meaning. According to the types of meaning expansion involved, she distinguishes between two types of semiotic metaphor: 'parallel semiotic metaphor' and 'divergent semiotic metaphor'. The former involves an expanded semantic field, but also one which is situated within the old, while the latter involves a functional element reconstrued into a new semantic field (O'Halloran 1999a: 348). O'Halloran illustrates these semiotic metaphors through a mathematic text from a research article shown in Figure 2.13.

> When the fluid is withdrawn from layers with different density, the withdrawn fluid will come from the layer adjacent to the point of removal until some critical value in flow rate is reached. At subcritical flow rates, the interface rises smoothly to an arch shape. At the critical flow rate, the interface is drawn upwards into a cusp shape … Above the critical rate, the fluid from the other layer will break through into the sink. That will influence the quality of the oil or fresh water. Therefore, in these cases, it is of significance to understand the behaviour of the flow induced when fluid is withdrawn.

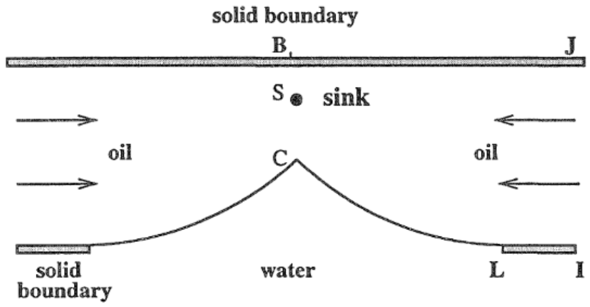


Figure 2.13 Part of a mathematics text from a journal article (Zhang and Hocking 1996: 240–242)

The diagram provides a visual illustration of withdrawing fluids such as oil through a sink line without drawing upon other layers such as water. O'Halloran takes the semiotic transition from the language text to the diagram as a semiotic metaphor. Specifically, it involves both parallel semiotic metaphor and divergent semiotic metaphor. For example, O'Halloran sees the phenomenon that the major linguistic participants 'sink', 'water', and 'oil' in the language text are displayed visually in the diagram as a parallel semiotic metaphor. The visual representation of these participants in the diagram adds meanings not included in the language text: their location and relative positioning, size, and shape. That is, this semiotic metaphor involves the expansion of meaning situated in the semantic field as the language text.

On the other hand, O'Hallloran takes the introduction of a new functional element, the 'interface' in the diagram, as a divergent semiotic metaphor. The corresponding congruent functional elements in the linguistic text are the 'oil', 'water', 'withdrawing' (material process), 'the fluid' (participant), and 'from layers with different density' (circumstantial adjunct). In the language text, the situation is described as *above the critical flow rate, the fluid from the other layer will break through the sink*. The diagram portrays the processes 'withdrawing the oil' or 'the flowing oil' visually as the arrows, and the 'sink' is made prominent by size and caption. The major participant is the curve representing the shape of the 'interface', which does not exist in the linguistic text but is introduced metaphorically in the visual diagram. This expanded meaning is not situated in the semantic field as the language text.

These expansions of meaning brought about by semiotic metaphor share some similarities with the two subcategories ‘image extends text’ or ‘text extends image’ in Unsworth’s model of language-image relations we have reviewed above. What is distinct in O’Halloran’s model is that she identifies specific types of meaning expansion, which is potentially helpful for analyzing how multisemiotic texts aggregate meaning and build knowledge. Nevertheless, there remains some difficult issues and challenges in applying semiotic metaphor. In the first place, as a ‘bottom-up’ approach (from grammar to semantics), analyzing semiotic metaphor entails first examining grammatical systems of the semiotic resources involved. For example, to analyze semiotic metaphor in mathematics texts, O’Halloran needs first to investigate the grammar of mathematical symbolism and mathematical images so that she can describe shifts of functional elements in the movement of semiotic resources. This is rather challenging for analyzing semiotic metaphors involving other semiotic resources since analysts must examine their grammatical systems first. Another issue is that it is difficult to tell which semiotic resource is the congruent form and which is metaphorical. For instance, in the example presented in Figure 2.13, it is not clear why the language text is the congruent form and the diagram is metaphorical. These issues suggest that more practical models are needed for examining meaning aggregation and knowledge-building through interaction between semiotic resources in multisemiotic disciplinary texts.

Another study more closely related to knowledge-building through interaction between semiotic resources in disciplinary multisemiotic texts is Unsworth (2020), which examines the intermodal construal of field (Doran and Martin 2021; Martin 2017) through language and image in biology texts. Specifically, the study focuses on how the knowledge of mitosis (a process where a single cell divides into two identical daughter cells) is construed by infographics in four different year-10 science textbooks used by Australian high schools. To compare the knowledge-building of miosis by the four textbooks, Unsworth draws on Martin’s (2017) conceptualization of ‘mass’ and ‘presence’, two concepts equivalent to semantic density and semantic gravity in LCT, to analyze intermodal construal of compositional taxonomies and activities by language and images in the four infographics. It shows that infographics with stronger mass are characterized by greater concurrence in the construal of compositional taxonomies and activities in the verbiage and images. Furthermore, stronger presence of images due to greater explicitness of depiction would facilitate intermodal construal of deeper compositional relations. These findings are potentially helpful for the intermodal design of more effective infographic explanation genres. However, Unsworth seems to focus more on the overall construal of field by language and images and gives much less attention to how they interact to build knowledge. For example, Unsworth discussed only the overall mass of the infographics, but it is not clear how this is obtained from language and image and whether there are shifts of mass between them. Similar issues occur in the analysis of presence. Unsworth only briefly discussed how the presence of images would influence the intermodal construal of compositional taxonomies. It remains unclear how language and image interact in terms of presence and if there are any shifts of presence between them. Therefore, more detailed studies are warranted to examine how different semiotic resources interact in disciplinary multisemiotic texts to build knowledge in terms of mass and presence, or in terms of semantic density and semantic gravity.

In this section, we have reviewed studies on intersemiotic relations in multisemiotic texts. The review suggests that previous studies show a relatively rich understanding of how semiotic resources interact in making meaning. However, there appears to be a paucity of exploration in how they interact to build knowledge. Unsworth's study provides a good first step for approaching this issue. The present study attempts to extend this and probe how the interaction among chemical formalisms, images, and language in texts builds secondary school chemistry across schooling levels in terms of their semantic density and semantic gravity in Chapter 6.

Thus far, we have reviewed multisemiosis in disciplinary discourse and knowledge-building through multisemiotic resources. Besides the semiotic perspective, the characteristics of chemistry knowledge conceptualized by chemists are also crucial to the knowledge-building analysis of chemistry. The next section will thus review chemists' conceptualization of chemistry knowledge in the field of chemistry education and their understanding of the knowledge-building of chemistry.

2.3 Chemistry knowledge and knowledge-building

Since chemistry knowledge is the key object in this study, this section reviews studies on chemistry knowledge and issues in relation to the knowledge-building of chemistry. It will first review studies on the classifications of chemistry knowledge (Section 2.3.1) and then literature on knowledge-building of chemistry in terms of these classifications (Section 2.3.2).

2.3.1 The three levels of chemistry knowledge

Chemistry knowledge has its own distinctive features. In chemistry education, it is modeled as consisting of three inter-related levels - ‘macrochemistry’, ‘submicrochemistry’, and ‘representational chemistry’ (Johnstone 1982). This seminal model is usually called the ‘chemical triplet’ (Talanquer 2011), shown in Figure 2.14.

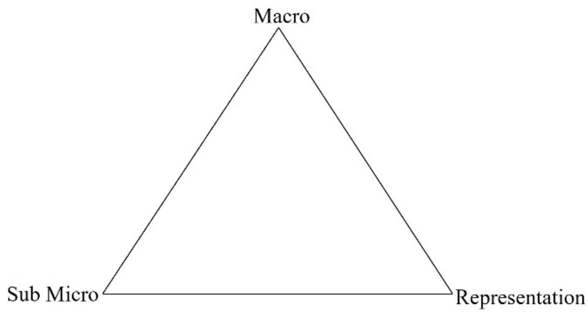


Figure 2.14 The chemical triplet (Johnstone 1993: 703)

Macro refers to tangible things or phenomena that can be seen, touched, and smelt; sub-micro refers to atoms, molecules, ions, structures, etc.; representational refers to symbols, formulas, equations, molarity, mathematical manipulation, and graphs. (Johnstone 2000: 11) The macro level is real and can be observed; the sub-micro level is based on real observations but still needs theories to explain what is occurring at the microscopic world; the representational level, on the other hand, is a representation of the physical reality. (Chittleborough 2004: 20–21) Johnstone argues that to successfully learn chemistry students are expected to cope with all three levels at once.

Chittleborough (2004) adapts this chemical triplet as three levels of representations: macroscopic, sub-microscopic, and symbolic. He calls all of the three levels of chemistry knowledge ‘representations’ and renames ‘representation’ in Johnstone’s model as ‘symbolic’. His model is illustrated in Table 2.5 using the example of the rusting of iron (Chittleborough 2004: 22).

Table 2.5 Description of 'the rusting of iron' at each level of chemical representation of matter

The rusting of iron	Levels of representation		
	Macroscopic	Sub-microscopic	Symbolic
Observations	Solid iron nail has a brown flaky coating on it that comes off easily when touched	Iron metal has iron atoms all closely packed together to form the solid nail. Some of the iron atoms next to the surface have reacted with the oxygen molecules forming a bond between an iron atom and an oxygen atom according to the formula Fe_2O_3.	The chemical equation summarises the reaction showing the number of iron atoms and oxygen atoms involved in the reaction. A ball-and-stick model and a computer simulation can depict the solid iron atoms being attacked by the oxygen molecule.
Real or representation	Real	Real – but too small to be seen with the naked eye.	Representation
Description	Tangible; quantitative	The particulate or molecular level according to the atomic theory of matter.	A depiction which may or may not be accurate but helps to provide a mental image.
Perception	Visible	Can't be seen with the naked eye, so mental image is based on descriptions, diagrams, explanations.	The model is a tool to help understand the real entity.

From this table, we can see that Chittleborough seems to use the term 'representation' differently from Johnstone. In Johnstone's model, 'representation' refers to certain semiotic resources, usually symbols, formulas, and equations, which represent chemistry knowledge, whereas Chittleborough takes it as an overarching term for the three types of knowledge. That is, 'representations' in Chittleborough's sense are not what we usually understand as forms of knowledge. However, it is interesting to note that Chittleborough appears to use 'representation' inconsistently, as evidenced by Table 2.5 where he takes 'symbolic' as representation but treats 'macroscopic' and 'microscopic' as physical reality. In contrast to 'representation' as a covering term for the three types of chemistry knowledge, it is interpreted from the semiotic perspective in the table.

More consistent with the term 'representation', Gilbert (2005) proposes a model of three representations based on Johnstone's chemical triplet. These representations are: macroscopic, the representation of observational experience in the laboratory and everyday life, for example, color change in a chemical reaction or in pictures of such situations; microscopic, the representation of the inferred nature of chemical entities (as atoms, ions, or molecules) and the relationships between them, for example, three-dimensional models of molecules; symbolic, the representation of the identities of entities (atoms, ions, or molecules), for example, chemical equations representing chemical reactions. (Gilbert 2005: 14) In contrast to Johnston (1982) and Chittleborough (2004) which take the macroscopic and microscopic levels as physical reality, Gilbert considers them as two types of representation of reality. These representations are distinguished based on scales of physical reality signified (macroscopic or microscopic world) and do not correspond to specific semiotic resources. To illustrate, macroscopic representations can be the verbal description of, for example, the rusting of nails or a photo of that phenomenon. In Hjelmslev's (1969) term, 'macroscopic' and 'microscopic' are meanings in the content plane which can be realized by various semiotic resources in the expression plane. However, in contrast to these two representations, the symbolic representation tends to be realized by certain semiotic resources, usually chemical formalisms including chemical symbols, formulas, equations, etc. This suggests that different from macroscopic and microscopic representations, there appears to be a correlation between symbolic representations and semiotic resources. Thus, Gilbert's model lacks a unitary criterion for classifying these representations in terms of semiotic resources they may take.

Gkitzia et al. (2011) notice the phenomena in school chemistry textbooks that different types of chemical representations can be put together and propose two more representations: multiple and hybrid. These added types derive from the way how the three original types are combined in a chemical representation. According to Gkitzia et al. (2011: 8), multiple representations are those depicting a chemical phenomenon simultaneously at two or three levels of chemistry knowledge. For example, Figure 2.15 is a multiple representation where macroscopic (the ice and water) and microscopic (the organizations of molecules forming ice and water) representations are integrated, each of which depicts the same phenomenon separately.

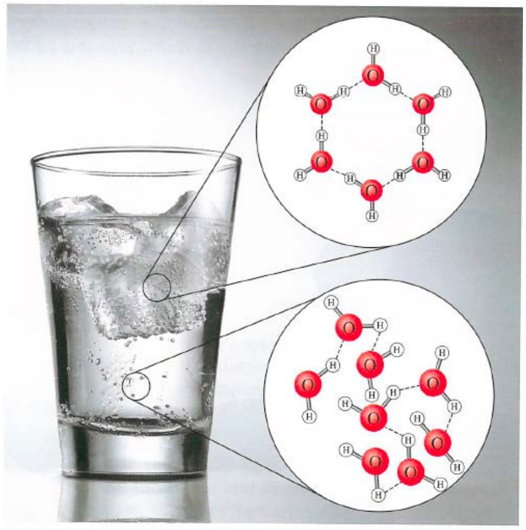


Figure 2.15 Ice is less dense than liquid water (Chan et al. 2018: 183)

Hybrid, on the other hand, refers to representations where two or three levels of chemistry coexist complementing each other forming one representation (Gkitzia et al 2011: 8). For example, Figure 2.16 is a hybrid representation of microscopic (ball-stick models representing molecules) and symbolic representations (structural formulas and mathematical expressions), which are integrated to describe one phenomenon – enthalpy change in the gas-phase reaction between methane and chlorine to produce methyl chloride.

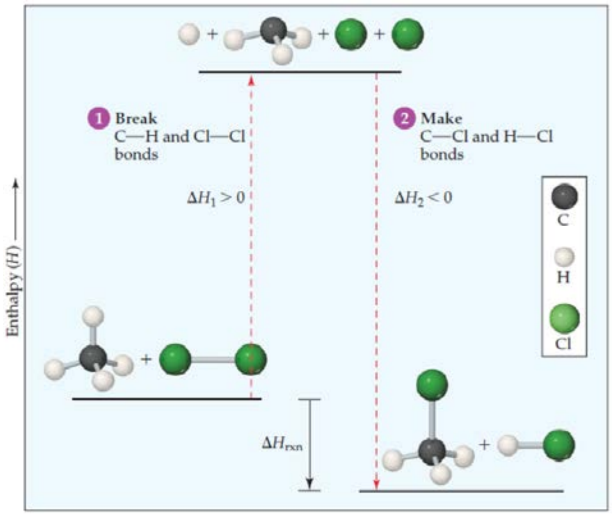


Figure 2.16 Using bond enthalpies to calculate$\triangle H_{rxn}$ (Brown et al. 2012: 317)

The above introduction shows that there is a high correlation between the two added representations and visual mode. This is because that the affordance of visual mode allows different elements to be presented in a single image as an eyeful, which is beyond the reach of language. Although the extension by Gkitzia et al. (2011) does not make substantial changes to the original model of chemistry representations, they have accounted for the possibilities of how different representations can be brought together in visual mode.

Overall, the adaptations of Johnstone's chemical triplet exhibit a shift from knowledge to representations. In Johnstone's original model, the representational is only one of the three levels of chemistry knowledge, whereas in Gilbert's model all the three levels are taken as representations. Put differently, such a shift gives prominence to the forms of chemistry knowledge, providing a semiotic perspective for accessing chemistry. In this sense, learning chemistry knowledge involves engaging with different semiotic resources representing the three types of knowledge. However, there seems to be a paucity of semiotic studies on the three levels of chemistry knowledge. The present study thus attempts to contribute to this field.

2.3.2 Transition among the three levels of chemistry knowledge

In chemistry education, transition among the three levels of chemistry knowledge is crucial for learning chemistry. As Johnstone (1991) points out, one reason why chemistry is difficult for students to learn is that it involves 'multilevel thought'. That is, students are expected to cope with all three levels of chemistry knowledge simultaneously. For example, in laboratories, students make observations of chemical reactions at the macroscopic level, but teachers expect them to interpret the phenomena at the microscopic level and show their mastery of the knowledge through the symbolic level. The significance of such a transition for building chemistry knowledge has been noted by numerous studies. Gabel (1993) examines teaching about the particulate nature of matter and finds that helping students make a connection between the three levels of chemistry would facilitate their learning of the particulate nature of matter. Kozma and Russell (1997) find that novices tend to use only one form of representation and rarely transform to other forms, whereas experts transform easily. Treagust et al. (2003) observed chemical representations used in chemistry lessons of introductory organic chemistry and finds that students' ability to transfer from macroscopic to submicroscopic and to symbolic representations is critical for mastering the knowledge of organic chemistry. All these studies corroborate the significance of the transition between the different levels of chemistry knowledge.

To understand how chemistry knowledge transfers between different levels and builds chemistry knowledge, a lens from the perspective of semiotics is critical. However, none of the above studies probe semiotic resources that the different levels of chemistry knowledge take. In the interdisciplinary field between chemistry education and semiotics, there are only a few studies examining the transition between different levels of chemistry with a focus on the forms of knowledge. Davidowitz and Chittleborough (2009) examine the role of diagrams in connecting different levels of chemistry knowledge. They find that chemical diagrams are featured by their capability of integrating elements that represent different levels of chemistry knowledge. For example, Figure 2.17 brings together three components which denote, from the top to the bottom, macroscopic, microscopic, and symbolic knowledges.

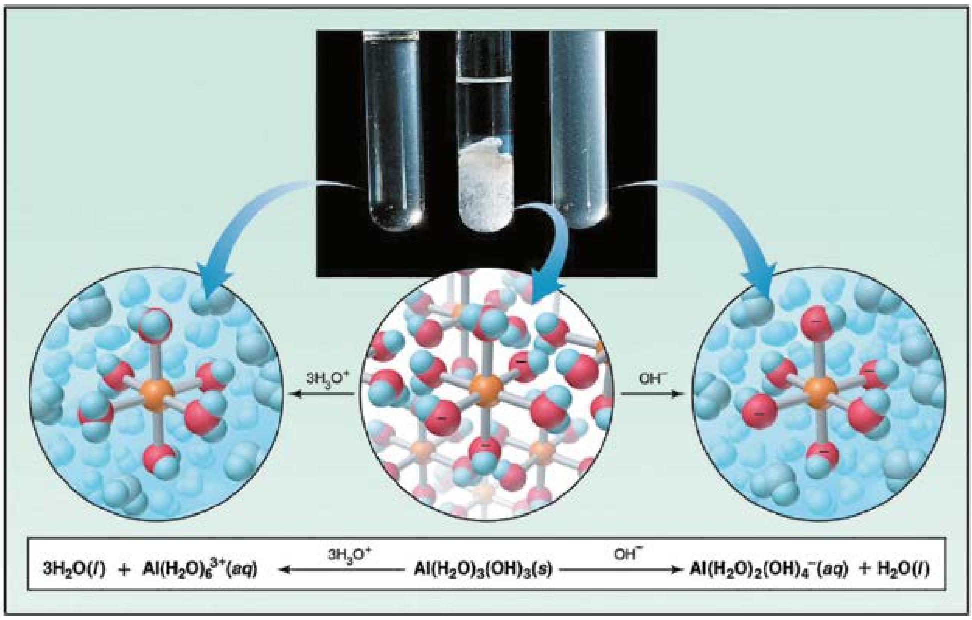


Figure 2.17 A three-level illustration (Davidowitz and Chittleborough 2009: 180)

The diagram makes explicit how the different levels of chemistry knowledge connect up through arrows and spatial arrangement of the elements. However, their study only points out this feature of chemical diagrams, but it does not address how meaning shifts to build knowledge as one level of chemistry knowledge transfers to another. To fill this gap, the present study attempts to approach this issue from a genuine semiotic perspective and focus on images that can represent all of the three levels of chemistry knowledge in Chapter 5. Specifically, it will probe how their meaning shifts to build chemistry knowledge along with the transition between different levels of chemistry.

2.4 Summary

This chapter began by reviewing the different approaches to knowledge-building analysis. It finds that the SFL and LCT approaches are rather useful and complement each other, with the former focusing on the meaning-making of semiotic resources and the latter emphasizing the organizing principles underpinning knowledge practices. However, there is a paucity of literature combing both to explore the knowledge-building of chemistry. To fill this gap, this study will take advantage of the two approaches' complementarity by using SFL to analyze how chemistry knowledge is construed by the multisemiotic resources, namely chemical formalisms, images, and language, and employing LCT to analyze how the resources develop across schooling levels to build the hierarchical knowledge structure of chemistry.

Following this, the second section reviewed multisemiosis in disciplinary discourse. It shows that existing literature follows a language-based approach to modeling chemical formalisms, which runs the risk of homogenizing description and watering down the specific functionality of the resource under study. To account for chemical formalisms on their own, Chapter 4 will move away from a language-based approach and take the fundamental theoretical primitive of axis as the descriptive foundation, which allows for revealing chemical formalisms' intrinsic functionality and grammatical organization. As for knowledge-building, scant studies have investigated how chemical formalisms develop across learning levels to build chemistry knowledge. To fill this gap, Chapter 4 will examine the knowledge-building of secondary school chemistry by chemical formalisms in terms of their development in semantic density across schooling levels. As for images, existing literature mainly focuses on the grammar of images, but few studies explore what types of meaning images in chemistry construe and how they develop across schooling levels to build chemistry knowledge. To fill this gap, Chapter 5 will model chemical images from the perspective of SFL's field and analyze the knowledge-building of secondary school chemistry through images in terms of the development in their semantic density. In terms of intersemiotic relations in texts, most of the previous studies probed the meaning-making of language-image relations. Few examine how the two resources interact to build knowledge. In addition, there is a paucity of work exploring the interaction between language and chemical formalisms in texts. To account for these issues, Chapter 6 will explore how the interaction among chemical formalisms, images, and language builds secondary school chemistry across schooling levels in terms of LCT's Semantics.

The last section reviewed chemists' classifications of chemistry knowledge in the field of chemistry education. It shows that chemists distinguish between three levels of chemistry knowledge: macroscopic, microscopic, and symbolic, and they believe that the transition among these levels is crucial for learning chemistry. However, without the emphasis on the forms of knowledge, it is rather difficult to see how the knowledge transfers from one level to another. To fill this gap, Chapter 5 will focus on images that can represent all three levels of knowledge and investigate how their meaning shifts to build chemistry knowledge along with the transition between the different levels of chemistry.

Having identified the gaps in previous literature, the following chapter will elaborate on the research design and explicate how the theoretical framework and methodology used in this study address these gaps.

Chapter 3 Research Design

The previous chapter has reviewed existing literature relevant to the research questions and identified primary issues to be addressed by this study. This chapter will outline the research design for the present study and explicate how it addresses the key research questions. It will begin by introducing the theoretical framework – systemic functional linguistics (SFL) and legitimation code theory (LCT), detailing key concepts and their complementarity for addressing issues related to the research questions (Section 3.1). Following this, it will introduce the methodology taken by this study to explore the research questions (Section 3.2). Section 3.3 then introduces the data used in this study to address the research questions.

3.1 Theoretical framework

As noted in Chapters 1 and 2, this study aims to explore the knowledge-building of chemistry through multisemiotic resources in secondary school chemistry textbooks. To achieve this objective, this thesis investigates both the construal of meaning by the multisemiotic resources and how they develop across schooling levels to build the hierarchical knowledge structure of chemistry. Addressing the two general issues necessitates theoretical tools that enable describing the meaning-making of the semiotic resources and analyzing how they develop through learning stages to build chemistry knowledge. For the first issue, this study will employ the basic theoretical concepts of SFL – a social semiotics theory accounting for meaning-making (Halliday 1978) (Section 3.1.1). We use SFL because it provides one of the most robust frameworks for modelling semiotic resources and analyzing how they make meaning (Martin 2013b). For the second issue, this thesis will use theoretical tools from LCT – a theory of sociology of education for understanding the organizing principles underpinning the development of knowledge (Maton 2014) (Section 3.1.2). LCT is used for addressing this issue because it provides the most practical tools that are capable of tracing the development of knowledge and revealing the patterns of the development (Maton 2013). The two theoretical tools complement each other to solve the research questions in this study, with the former addressing the issue of the construal of meaning by the multisemiotic resources and the latter underpinning the analysis of their development through schooling (Section 3.1.3).

3.1.1 Systemic Functional Linguistics

As noted in Chapter 2, systemic functional descriptions of semiotic systems other than language originate from the work on English by Halliday in the 1960s and 1970s, well known as Systemic Functional Linguistics. It is thus crucial to understand the basic theoretical concepts of SFL in order to account for chemical formalisms. This section will first introduce four SFL theoretical categories that are critical for semiotic description – axis (system and structure), metafunction, stratification, and rank (Section 3.1.1.1). Following this, it will introduce the theoretical concept of field in SFL, a critical tool for capturing types of content meaning realized by chemical formalisms, images, and language (Section 3.1.1.2).

3.1.1.1 Basic theoretical categories

This section will introduce four basic theoretical categories in SFL for modeling semiotic systems – axis, metafunction, stratification, and rank.

3.1.1.1.1 Axial relations: the theoretical primitive

The axial relations in SFL originate from Saussure’s (1959) concepts of syntagmatic relations and associative relations, which are adapted as syntagmatic and paradigmatic relations by Hjelmslev (1969) and as system and structure by Firth (1957). For Saussure, Hjelmslev, and Firth, neither the syntagmatic (structure) or paradigmatic (system) perspective is privileged. Halliday (1961), however, privileges system over structure and views structure as the realization of choices from a system. Thus, SFL inherits the complementarity between paradigmatic and syntagmatic relations from Saussure, Hjelmslev, and Firth but privileges paradigmatic relations. The axial relation between system and structure is the theoretical primitive in SFL since other macro-theoretical concepts such as stratification, metafunction, and rank can be derived from it (Martin 2013b, 2015a) (see Section 3.2.1 for an illustration of how the axial relations generate metafunctions and ranks in Halliday’s description of English). This section will introduce the two fundamental theoretical concepts.

3.1.1.1.1.1 System

SFL formalizes the paradigmatic axis as systems. A **system** is a representation of relations on the paradigmatic axis, a set of features contrastive in a given environment (Halliday 1966: 60). It is a choice relation among any set of alternatives. SFL represents systems through system networks, a type of representation comprising three components: entry conditions, system names, and features. Take the simplified MOOD network shown in Figure 3.1 as an example. 'major clause' is the entry condition that allows access to the system. 'MOOD' is the system name. The options [indicative] and [imperative] are the system's features.[7] These components have specific writing rules in SFL - entry conditions and features must be in lower case, and the system names in capital case (Halliday and Matthiessen 2014; Martin 2013b) (see Appendix B for the detailed writing conventions of system networks).

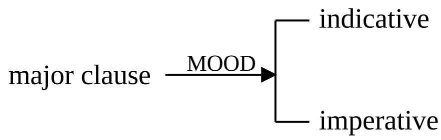


Figure 3.1 A simplified network of MOOD

The MOOD system in Figure 3.1 presents a systemic contrast between indicative and imperative clauses. However, this is only a partial system of MOOD. Indicative clauses have a deeper set of choices: declarative and interrogative clauses. SFL represents these more delicate choices in system networks through the cline of **delicacy** (Halliday 1961). As the delicacy increases, the primary class breaks into secondary ones, as shown in Figure 3.2. The feature [indicative] becomes the entry condition to the subsystem made up of two more delicate features - [declarative] and [interrogative]. The less delicate systems are on the left, while the more delicate ones are on the right.

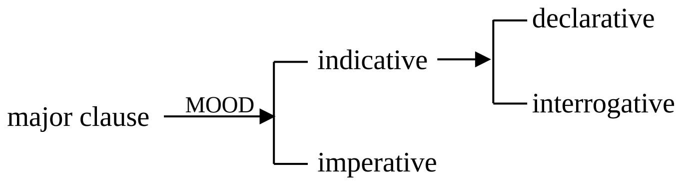


Figure 3.2 A more delicate network of MOOD

In SFL, grammatical systems represent the meaning potential of language (Halliday and Matthiessen 2014: 23). The instances of meaning are interpreted as the outward form taken by systemic choices. For example, MOOD represents the meaning potential of negotiating interpersonal relations. Choices from this system would generate specific clauses performing certain functions. For instance, choosing [imperative] leads to an imperative clause fulfilling the function of command. The choices in the system network would produce different clauses performing varying speech functions (Halliday 1985). These clauses are in fact different structures. This suggests that choices from system lead to structure, which is another crucial concept

7 Features are denoted by the notation '[ ]'.

we need to introduce.

3.1.1.1.1.2 Structure

In contrast to system that describes choice relations, **structure** is the syntagmatic ordering in language: patterns in what goes together with what (Halliday and Matthiessen 2014: 22). Halliday (1963a) also refers to it as “chain relations”. As the term implies, a structure comprises elements sequenced as a chain. These elements are called “**functions**” (Halliday 1966). A structure shows the relation between the functional elements. Example (3:1) involves a structural configuration of ‘Actor ^ Process ^ Goal’.[8] *The boy*, *kicked*, and *the ball* perform the three respective functions: Actor, Process, and Goal, which are relational terms representing the clause components’ grammatical functions with respect to one another.[9] Put simply, the structural configuration shows what role each element plays in the structure.

(3:1) *The boy kicked the ball*.

In addition to structure, there is another type of syntagmatic relations – **syntagm**, a linear succession of **classes** (Halliday 1966: 58). The relation between the two is ‘realization’. That is, structural configurations of functions are realized by syntagms of classes. For instance, in example (3:1), the structural configuration of ‘Actor ^ Process ^ Goal’ is realized by the syntagm ‘nominal group ^ verbal group ^ nominal group’. However, it must be noted that there is no one-to-one relation between the functions and classes. For example, a nominal group can realize both Actor and Goal.

Structural configurations and classes realizing functional elements within the configurations are represented in system networks through ‘**realization statements**’. Figure 3.3 illustrates the realization statements of the features in MOOD.

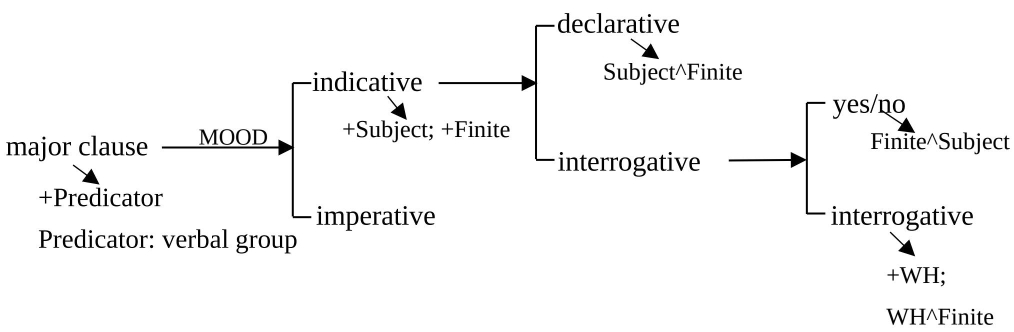


Figure 3.3 The simplified MOOD system of English clauses (adapted from Halliday and Matthiessen 2014: 162)

[8] Following SFL convention, function labels are written in initial capital.
[9] Italics mark examples.

The downward slating arrow '↘' indicates a realization statement, relating system to structure. Reading from left to right, we can see that [major clause] is realized by the insertion of a Predicator, indicated by '+Predicator' ('+' means 'insert'). Predicator is realized by verbal groups, denoted by the notation 'Predicator: verbal group' (':' means 'realized by'). [major clause] leads to the MOOD system. When [indicative] is chosen, the Mood elements of Subject and Finite are inserted, denoted by the statement '+Subject; +Finite'. These functions are not sequenced yet because there are further variations. When the more delicate feature [declarative] is selected, the Mood elements are configured as 'Subject ^ Finite'. When [interrogative] is chosen, there are two further variations: [yes/no] is realized by the configuration of 'Finite ^ Subject', and [WH-] is realized by the insertion of WH- element and the sequence 'WH ^ Finite'. The ongoing choices produce different configurations of Mood elements, determining the different types of moods.

The above has demonstrated how structure and syntagm are represented in systems. We conclude this section with a brief summary of the relations among system, structure, and syntagm. Axial relations in SFL include two axes: paradigmatic relations and syntagmatic relations. The paradigmatic axis is system, and the syntagmatic axis comprises structure and syntagm. Structure is configurations of functions, whereas syntagm is linear successions of classes. System is realized by structure, which is in turn realized by syntagm. These relations can be represented as Figure 3.4.

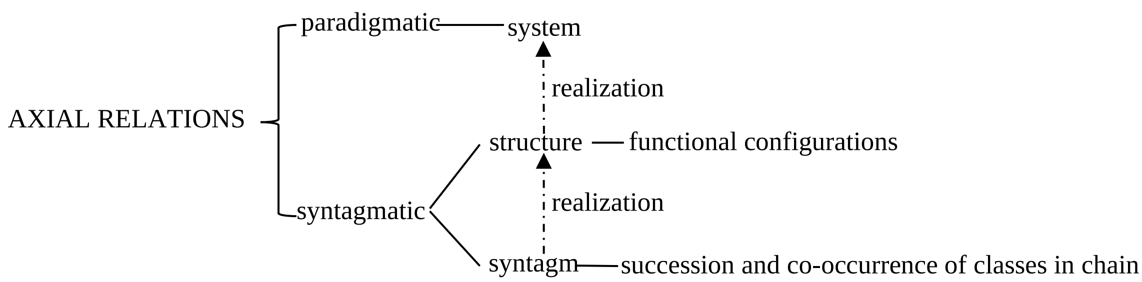


Figure 3.4 System, structure, and syntagm (adapted from Zhang 2020: 19)

Crucially for this study, the axial relations between system and structure will be the foundation for describing chemical formalisms' grammars, as will be shown in Chapter 4. It does not assume English's grammatical systems but investigates chemical formalisms' grammatical organization in its own right. The axial relations lay the foundation for achieving this objective.

3.1.1.1.2 Metafunction

SFL theorizes the meaning made by language as three broad types of **metafunctions**: ideational, interpersonal, and textual Halliday (1978, 1985). Ideational metafunction construes our experience of the world around us; interpersonal metafunction enacts relationships between people; textual metafunction organizes the ideational and interpersonal meanings as information in a text (Halliday 1985). The three metafunctions coordinate with the three register variables: field, tenor, and mode (Halliday 1978). Field concerns with what happens in social activities, tenor with relationships between interactants, and mode with the role of language in the activities. Figure 3.5 illustrates this metafunction-register hook-up.

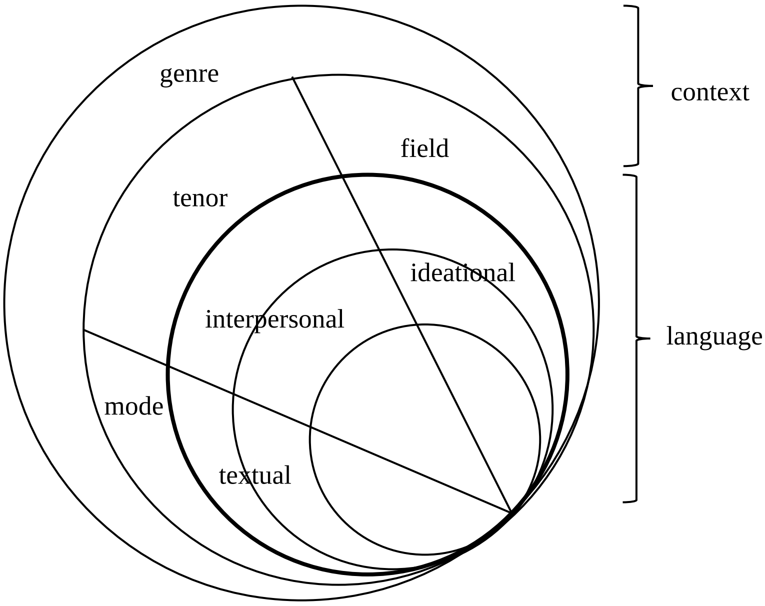


Figure 3.5 The metafunction-register hook-up (adapted from Martin 1992)

As discussed in Section 2.2.1.1.1, Halliday develops the notion of metafunction from English's grammatical systems. It is worth recapitulating how Halliday develops the three metafunctions from his description of English. Halliday's early papers in the 1960s and 1970s including 'notes on transitivity and theme in English, Part 1 (1967a), Part 2 (1967b), Part 3 (1968)', 'functional diversity of language in language as seen from a consideration of modality and mood in English' (1970), and 'options and functions in language clause' (1969) show that he first describes the grammatical systems of English based on system and structure and then generates the three metafunctions according to the systems bundled together. It must be stressed that it is the grammatical systems that give rise to the metafunctions, as articulated more clearly in the following quote from Halliday (2002[1970]: 174):

> "It is fairly obvious that language is used to serve a variety of different needs, **but until we examine its grammar there is no clear reason for classifying its uses in any particular way**. However, when we examine the meaning potential of language itself, we find that the vast numbers of options embodied in it combine into a very few relatively independent "networks"; and **these networks of options correspond to certain basic functions of language**. This enables us to give an account of the different functions of language that is relevant to the general understanding of linguistic structure rather than to any particular psychological or sociological investigation." (Bond fonts are my emphasis)

For Halliday, metafunctions are in fact **clusters of systems** bundled together. These systems are interdependent with the same functional components but independent of each other across the components, as explicated in the following quote from Halliday (1978: 187):

> "The semantic system is organized into a small number of components – three or four depending on how one looks at them – such that *within* one component there is a high degree of interdependence and mutual constraint, whereas *between* components there is very little: each one is relatively independent of the others." (original italics)

To illustrate this with the grammatical systems of English, the systems of TRANSITIVITY, MOOD, and THEME are relatively interdependent within themselves and independent between each other, suggesting that they realize three distinct metafunctional components, termed ideational, interpersonal, and textual, as shown in Figure 3.6.

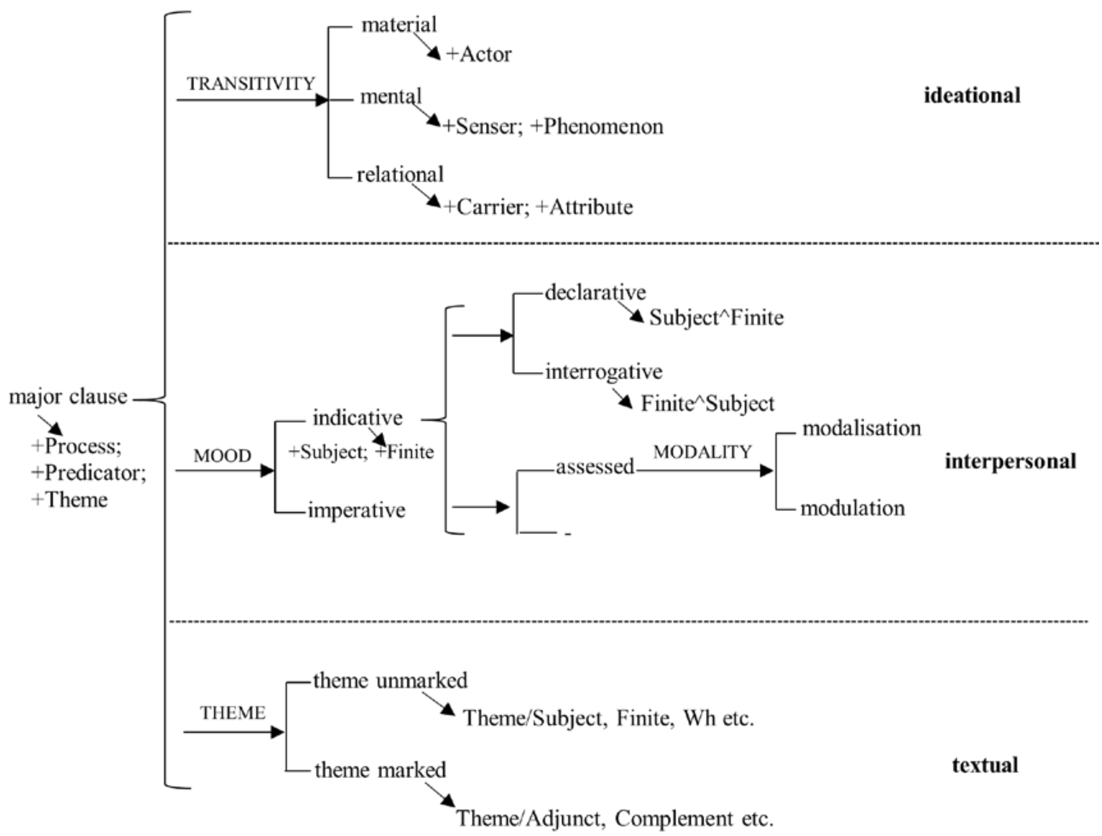


Figure 3.6 The three independent grammatical systems (simplified) and the three metafunctions (adapted from Halliday 2003[1973]: 315)

In addition to the paradigmatic independence, the metafunctions are associated with particular types of structure. Halliday (1979) proposes three types of structure: particulate, prosodic, and periodic, realizing ideational meaning, interpersonal meaning, and textual meaning, respectively. **Particulate structures** are segmental. Experientially, they divide bounded wholes into parts (constituency), and logically, they relate one part to another (interdependency). For example, in the transitivity structure of ‘Actor ^ Process ^ Goal’, a clause is divided into elements, between which the boundaries are clear-cut. This structure realizes experiential meaning. For another example, in Example (3:2), the two clauses are related through the interdependency of hypotactic relations. This structure realizes logical meaning.

(3:2) *After finishing homework, the boy starts playing football*.

The particulate structures then involve two subtypes, one based on constituency and the other on interdependency, which Halliday (1981[1965]) terms as multivariate and univariate structures, respectively. **Multivariate structures** involve more than one variable, with each occurring only once, whereas **univariate structures** involve only one variable that can repeat indefinitely. A typical example of multivariate structures is the transitive structure of a material clause ‘Actor ^ Process ^ Goal’. The clause consists of three distinct functional elements, each of which occurs only once. The clause complex shown in Example (3:2), on the other hand, is an example of a univariate structure involving iteration of the same element (the two clauses *after finishing homework* and *the boy starts playing football* are the same element).

Univariate structures consist of two subtypes in terms of the elements’ status: hypotactic and paratactic. Hypotactic structures involve one element depending on another, whereas paratactic structures comprise elements of equal status. For example, in the hypotactic clause complex shown in Example (3:2), the clause *after finishing homework* depends on the clause *the boy starts playing football*, because the former cannot occur by itself. In contrast, in the paratactic clause complex shown in Example (3:3), the two clauses are of equal status and can occur independently. The distinction between multivariate and univariate dimensions of particulate structures is rather crucial for describing chemical formalisms’ grammar in Chapter 4.

(3:3) *The playground closed and the boy went back home*.

Prosodic structures are suprasegmental and cut across a range of units. For instance, in Example (3:4), the negation realized by the Finite and indefinite deixis (underlined) maps over the clause, forming a prosody.

(3:4) *If you don’t get no publicity you don’t get no people at the fight*. (example cited from Martin 1996)

Periodic structures are wave-like and establish rhythmic peaks of prominence that bound units. For example, in English clauses’ thematic structure and information structure, the thematic prominence at the beginning of the clauses and the newness prominence typically occurring at the end of clauses exhibit a wave-like structure. Both prosodic and periodic structures are multivariate. Figure 3.7 summarizes the above different structures and the types of meaning they realize.

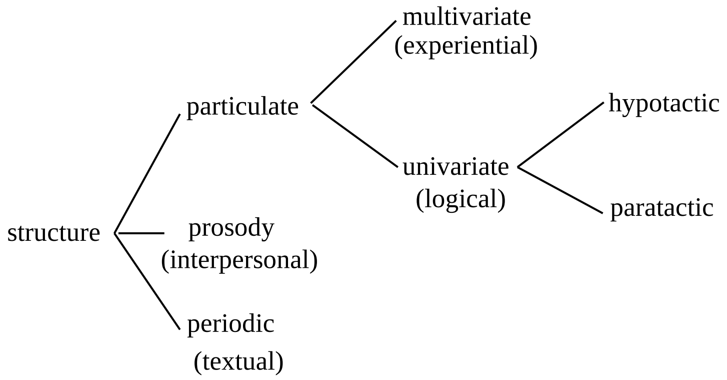


Figure 3.7 Types of structure and metafunctions realized (adapted from Doran 2018c: 26)

The paradigmatic independence of systems and the types of structure associated with the different functional components constitute the foundation for suggesting metafunctions of chemical formalisms in Chapter 4.

3.1.1.1.3 Stratification

SFL theorizes language as a stratified semiotic system involving three levels of abstraction. These levels of **abstraction** are called '**strata**' (Halliday 1985). The three strata are discourse semantics (semantics), lexicogrammar, and phonology/graphology, arranged on a cline of abstraction.[10] Discourse semantics is more abstract than lexicogrammar, which is in turn more abstract than phonology/graphology. The relation between these strata is **realization**, which entails a 'meta-redundancy', the notion of patterns at one level 'redounding' with patterns at the next level (Martin and Rose 2008: 10). That is, discourse semantics is realized as patterns of lexicogrammar, which is in turn realized as patterns of phonology/graphology. In terms of Hjelmslev's (1969) model of language as a stratified system, the strata of discourse semantics and lexicogrammar constitute the 'content plane' of language, and the stratum of phonology/graphology forms the 'expression plane' of language. Figure 3.8 illustrates the stratal organization of language.

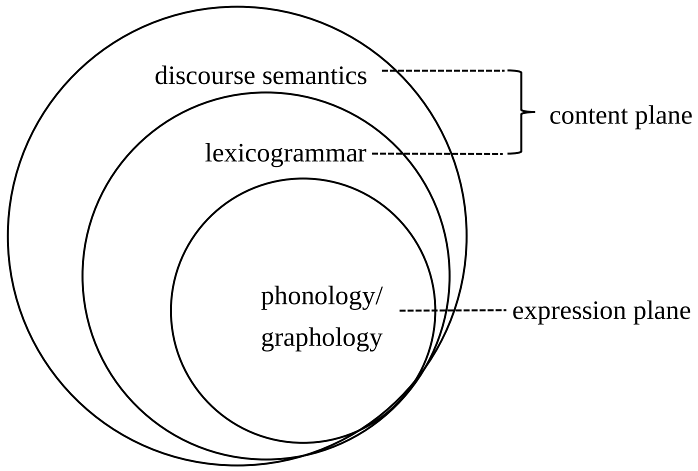


Figure 3.8 Strata of language (adapted from Martin 1992)

In addition to conceptualizing language as a stratified model, SFL also theorizes the context where language is used as a stratified connotative semiotic (Martin 1992).[11] Unlike Halliday's (1978) model of context, which positions context of culture and context of situation at one stratum on a continuum, Martin (1992) stratifies context as

[10] SFL scholars name the stratum more abstract than lexicogrammar differently – 'semantics' by Halliday (1985) and 'discourse semantics' by Martin (1992)

[11] Hjelmslev (1969) defines connotative semiotics as semiotic systems which have another semiotic system as their expression plane. The semiotic systems that realize the connotative semiotics are called denotative semiotics.

two strata: genre and register, corresponding to context of culture and context of situation. Genre is the highest stratum and describes a text's global social purposes, coordinating the three register variables at the lower stratum: field, tenor, and mode (Martin 1992). Field concerns with what happens in social activities, tenor with relationships between interactants, and mode with the role of language in the activities (Martin and Rose 2008: 11). The genre and register strata correspond to the content and expression planes. In terms of Hjelmslev's distinction between connotative and denotative semiotics, language can be viewed as a denotative semiotic realizing context, and context as a connotative semiotic realized through language (Martin and Rose 2008: 16). Figure 3.9 shows the strata involved in language and context.

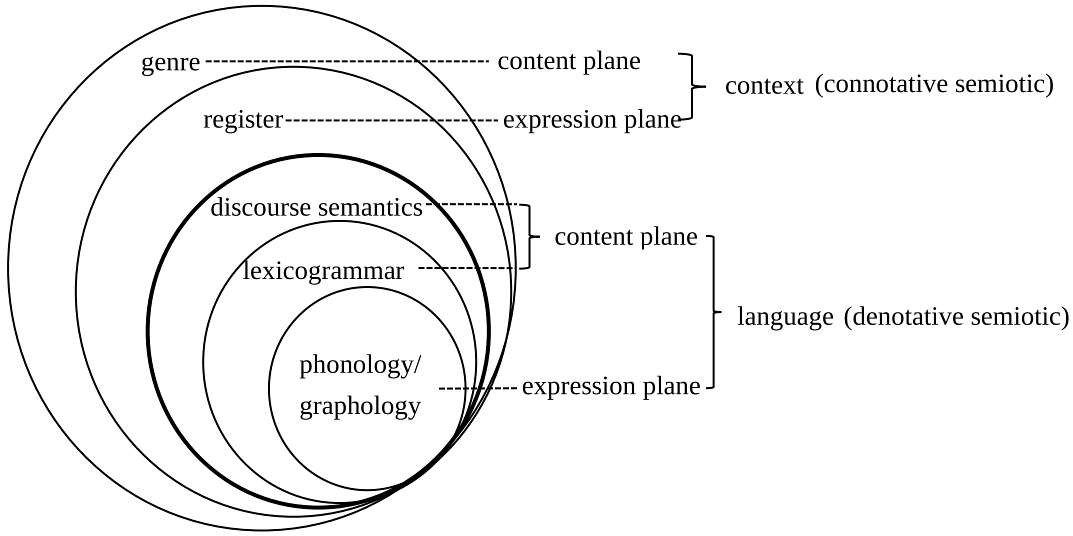


Figure 3.9 Stratal organization of language and context (adapted from Martin 1992)

3.1.1.1.4 Rank

Within strata, SFL develops another set of levels known as '**ranks**'. In contrast to strata which are levels of abstraction, ranks are constituency hierarchies – one rank comprises one or more units from the rank below (Halliday 1961). For example, clause, the highest rank at English's lexicogrammar stratum, comprises one or several groups at the rank below, i.e., the group rank. Ranks occur across the three strata of language. In SFL's model of English language, the semantics stratum includes three ranks: element, figure, and sequence (Halliday and Matthiessen 1999); the lexicogrammar stratum involves four ranks: morpheme, word, group/phrase, and clause (Halliday 1985); the phonology stratum includes four ranks: phoneme, syllable, feet, and tone group (Halliday 1963b).

Rank has several significant features. One is that they must be **obligatory**. Take the ranks at English's lexicogrammar stratum as an example. The clause, group, word, and morpheme ranks are obligatory in interpreting the grammar of English. A clause must be interpreted as comprising groups, which in turn comprise words, which in turn comprise morphemes. For instance, Example (3:5) involves a configuration of three functions: Actor (*the boy*), Process (*kicked*), and Goal (*the ball*). These functions are realized by units at the group rank - the Actor is realized by a nominal group (*the boy*), the Process is realized by a verbal group (*kicked*), and the Goal is realized by another nominal group (*the boy*). These groups have their own functional structure, which are in turn realized by words. For example, the nominal group *the boy* includes two functions: Deictic and Thing, which are realized by the two respective words: *the* and *boy*.

(3:5) *The boy kicked the ball*.

Another significant feature of rank is that it allows ‘**rankshift**’. Halliday (1961: 251) defines rankshift as “the transfer of a (formal realization of a) given unit to a lower rank”. That is, rankshift involves a unit of a rank realizing a function at the rank below. For example, in Example (3:6), the nominal group *playing football* is a rankshifted clause functioning as an Actor. The shift from the clause rank to the group rank is the most frequent rankshift in English’s lexicogrammar.

(3:6) *Playing football makes the boy happy*

The third critical feature of rank is that it is not only a part-whole relation but a **multivariate constituency** (Huddleston 1965). In Matthiessen and Halliday’s (2009: 68) term, units of a rank realize a configuration of functions. A rank consists of elements performing distinct functions with respect to the others, forming a multivariate structure at the rank above. For example, the nominal and verbal groups in Example (3:5) realize three distinct functions: Actor (*the boy*), Process (*kicked*), and Goal (*the ball*). These units at the group rank form the multivariate structure of ‘Actor ^ Process ^ Goal’ at the higher rank of clause.

In contrast to the multivariate-based rank scale, SFL also proposes a different scale based on univariate structures, usually known as “**nesting**” (Halliday 1981[1965]). The nesting scale arises from the iteration of a unit into a complex. For example, a clause can iterate to form a clause complex, such as Example (3:7).

(3:7) *After finishing homework, the boy starts playing football*.

The complex does not form a higher rank but remains at the clause rank. Complexing occurs at all the ranks in English’ lexicogrammar, as exemplified in Table 3.1.

Table 3.1 Complexing at the four ranks in English lexicogrammar

Ranks	Complexes (examples)
clause	*After finishing homework*, *the boy starts playing football*.
group/phrase	*the red flowers and the green grass*
word	*flowers and grass*
morpheme	*one- and two- dimensional*

In contrast to the multivariate-based rank, iterative structures are interpreted through **interdependency** associated with the logical metafunction (Halliday 1985). For example, the two clauses in Example (3:7) are in "hypotactic relation" (Halliday 1981[1965]). The distinctions between ranks and nesting and between multivariate structure and univariate structure are crucial for describing the chemical formalisms' grammars in Chapter 4. We will return to them periodically as they become relevant.

3.1.1.2 Field

The SFL dimension of field is a significant theoretical concept for capturing the content meaning realized by chemical formalisms, images, and language in this study. In SFL, field is one of the three register variables and is concerned with the **experiential meaning** of semiosis (Martin 1992). In common sense terms, field can be broadly understood as content meaning (Doran 2019: 337).

As introduced in Section 2.1.1.2, SFL distinguishes between two types of field: **common-sense field** and **uncommon-sense field**, arranged as the two poles of a continuum (Martin 1992: 544). The common-sense field tends to be everyday knowledge in domestic life, whereas the uncommon-sense field is typically technical knowledge involved in specialized activities. For example, the everyday phenomenon that water boils in a kettle is a common-sense field, while the chemical knowledge that the thermal motion of water molecules becomes intense after heating is an uncommon-sense field.

The above classification of fields by Martin captures the differences in the technicality of different fields. However, this does not allow for describing the different types of field meanings construed by semiotic resources. To offer a tool that enables grasping the different field meanings, Doran and Martin (2021) propose a model of field of language. In the model, field meanings can be viewed in broad terms **statically** as a set of items related to each other in taxonomies or **dynamically** as a set of events (known as activities) oriented to some global institutional purpose. In addition, each of these items and activities may have certain **properties** that may be graded and/or measured numerically.

From a static perspective, there are two main relations that can occur between items: classification and composition. **Classification** is the relation between items in terms of class and sub-class. For example, chemistry distinguishes between two types of ions: cation and anion. The items ‘cation’ and ‘anion’ are thus sub-classes in relation to the more general item ‘ions’. **Composition**, on the other hand, is the part-whole relations among items. For example, the explanation of atomic structure depends on multiple levels of composition - an atom is composed of one nucleus and at least one electron, and a nucleus is in turn composed of neutrons and protons.

From a dynamic perspective, field construes phenomena as **activities** that organize events and changes. Activities can be described in terms of two simultaneous dimensions – whether they are momented and what types of change they undergo. Momenting indicates that the activities are presented as a series of smaller events. For example, the change of water’s physical states can be described through **momented activities**, *ice melts into liquid water, which then evaporates to become water vapour*. In contrast, **unmomented activities** present the event as a single whole, such as *the ice sublimates*. In terms of the types of change these activities undergo, an activity can be categorized as cyclical or linear. **Cyclical activities** involve an event that can recur indefinitely, for example, *electrons orbit the nucleus*. **Linear activities** involve a linear unfolding event, for example, *a neutron strikes a nucleus*. The static and dynamic perspectives of field can be summarised as the network presented in Figure 3.10.

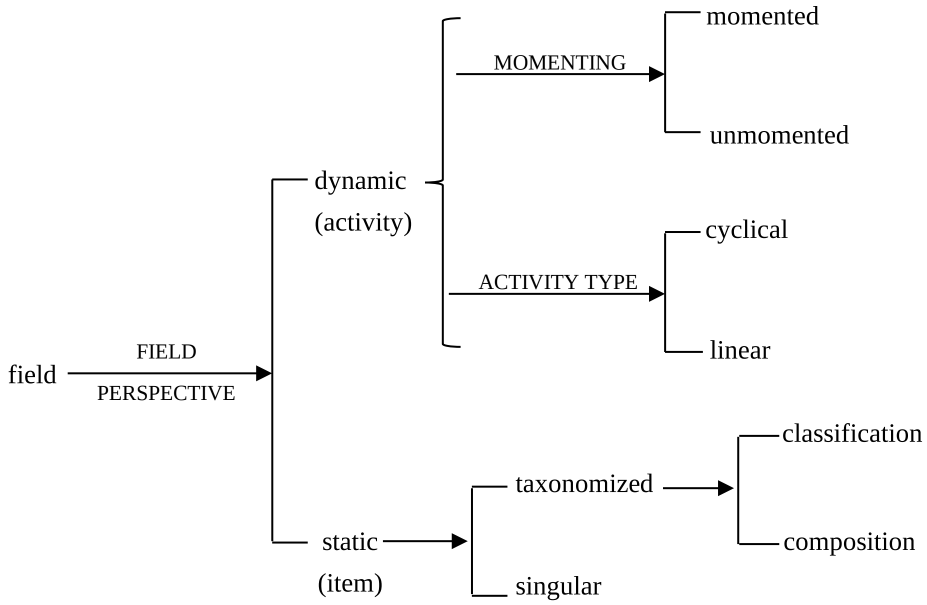


Figure 3.10 Dynamic and static perspectives on field (Doran and Martin 2021)

In addition to activities and items, field can be organized in terms of the **properties** they show. Properties can be understood as potentially gradable qualities or positions that enable rich descriptions of phenomena (Doran and Martin 2021: 116-117). They can characterize both items, for example, *ice is <u>hard</u>*, and activities, for example, *ice melts into liquid water <u>rapidly</u> under heating*. In terms of property types, properties may involve **qualitative** description, such as *electrons are <u>negatively</u> charged*, or offer **spatio-temporal** position, such as *electrons orbit <u>around the nucleus</u>*. In addition, these properties may be graded and potentially ordered into **arrays** in relation to other properties. For example, *electrons in the outer shell have the <u>highest</u> energy*. These properties in turn can also be measured or quantified, which is called **gauged** properties, for example, *an electron in the first energy level has -<u>13.6</u> eV of energy*. Figure 3.11 summarizes all the above properties.

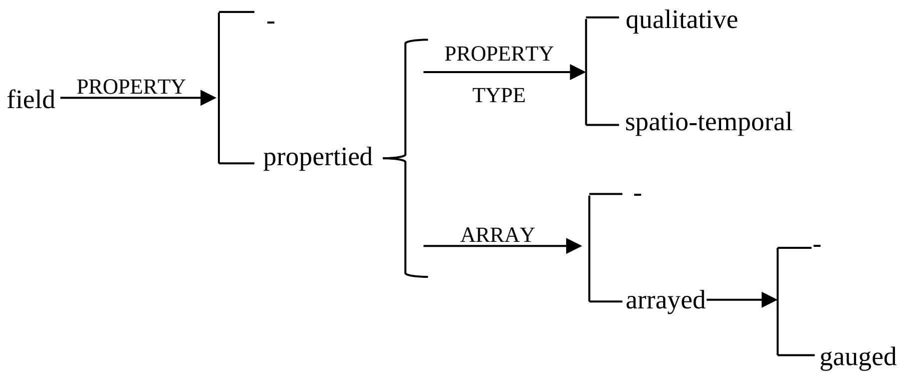


Figure 3.11 Network of PROPERTY (Doran and Martin 2021)

The FIELD PERSPECTIVE and PROPERTY systems introduced above constitute the basis of describing chemistry knowledge construed by the different semiotic resources in this study. Figure 3.12 summarizes the two components in the model of field we will use to capture the types of content meaning chemical formalisms, images, and language construe in Chapters 4, 5, and 6.

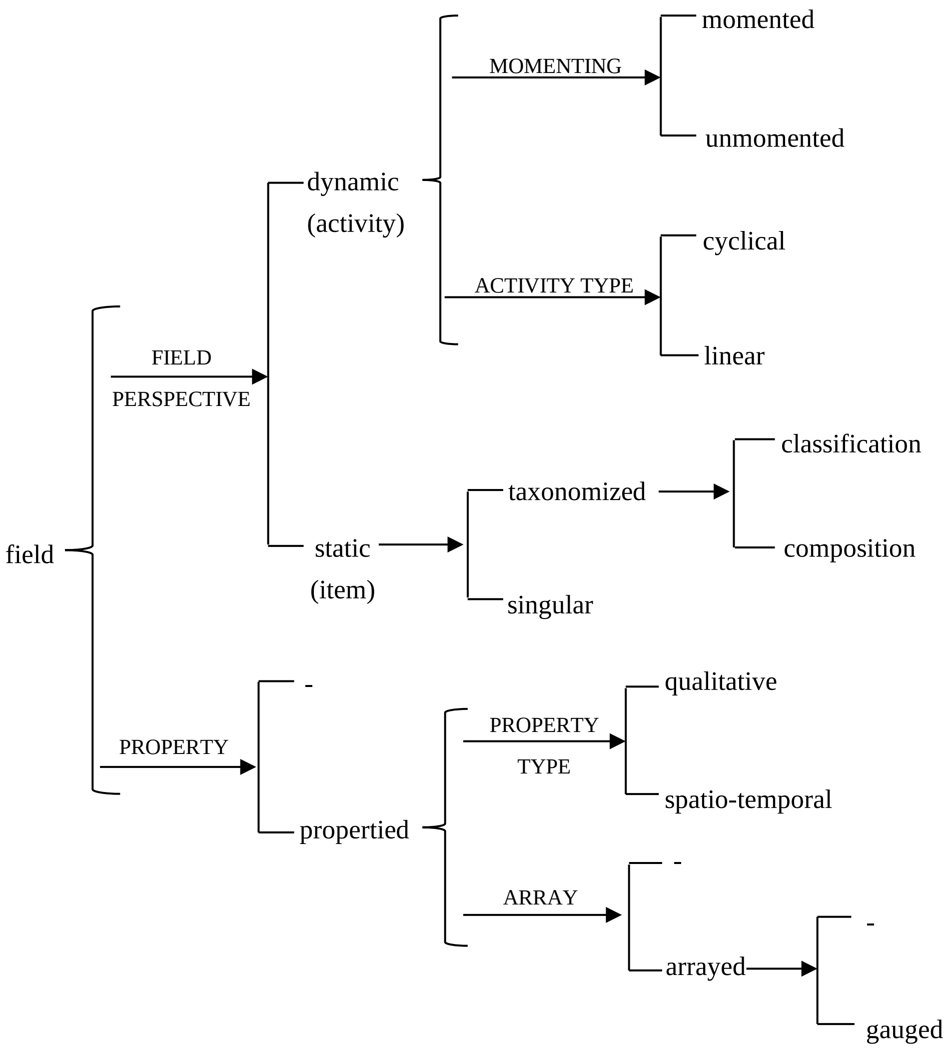


Figure 3.12 The field system used in this study (Doran and Martin 2021)

3.1.2 Legitimation Code Theory

Complementary to the SFL theoretical tools that enable analyzing the construal of meanings through the multisemiotic chemistry discourse, this study draws on theoretical tools from Legitimation Code Theory (LCT) to explore how the meanings develop across learning stages to build the hierarchical knowledge structure of chemistry.

LCT is a framework for researching knowledge practice. It comprises several dimensions that reveal organizing principles of knowledge. Among these dimensions, **Semantics** is the most pertinent to this study. It explores the organization of meaning within knowledge practices in terms of their context-dependence and complexity (Maton 2011, 2013, 2020). This dimension of LCT has been proven rather useful for analyzing a diverse range of knowledge practices, including academic writing (Brooke

2017, Clarence 2017a, Kirk 2017), musical performance (Richardson 2020; Walton 2020), and dance (Lambrinos 2020). The present study will extend this growing body of work to embrace the knowledge-building of chemistry through the multisemiotic chemistry discourse.

Semantics comprise two crucial concepts: semantic density and semantic gravity. **Semantic density** (SD) refers to the degree of complexity of meaning or practices (Maton 2014). Semantic density can be stronger or weaker along a continuum of strengths, where the stronger the semantic density (SD+), the more complex the meaning, and the weaker the semantic density (SD–), the simpler the meaning. Put another way, the more relations with other meanings enjoyed by a practice, the stronger its semantic density (Maton and Doran 2017a). For example, the word *salt* in everyday usage refers to small white crystals often used to add flavor to food – a relatively small number of relations among experiential meanings, such as its flavor, shape, and uses. In contrast, as a technical word in the field of chemistry, *salt* refers to a compound produced by the reaction of an acid with a base and involves relations with numerous chemical concepts, such as cations, anions, and ionic bonds, which themselves relate to a large number of other meanings. Thus, in chemistry, the term *salt* is situated within a relatively complex constellation of meanings that imbues the term with relatively strong semantic density. Moving from the everyday use of salt (weaker semantic density) to its technical use (stronger semantic density) is what Maton (2014) called **condensation**.

Semantic gravity (SG) is concerned with the degree to which meaning depends on its context (Maton 2014: 110). Semantic gravity can be stronger or weaker along a continuum of strengths, where the stronger the semantic gravity (SG+), the more context-dependent the meaning, and the weaker the semantic gravity (SG–), the less context-dependent the meanings. For example, in chemistry, a specific description of an instance of a chemical reaction (for example, *hydrogen gas reacts with oxygen gas to form water*) shows stronger semantic gravity than a generalized theoretical term describing a range of such chemical reactions (for example, *combination reactions*). This is because *hydrogen gas reacts with oxygen gas to form water* describes only a specific instance of *combination reactions*. The chemical term *combination reactions* applies to a much wider range of contexts, thereby embodying weaker semantic gravity. The movement from meaning with weaker semantic gravity (*combination reactions*) to stronger semantic gravity (the specific instance) is called **gravitation** (Maton 2014). On the other hand, moving from stronger semantic gravity (the specific instance) to weaker semantic gravity (combination reactions) is termed **levitation**.

Semantic density and semantic gravity are crucial for analyzing how the meanings construed by the multisemiotic chemistry discourse develop through schooling to build the hierarchical knowledge structure of chemistry in the present study. For chemistry discourse to build a hierarchical knowledge structure, it needs to generate increasingly technical meanings and relating theoretical concepts to empirical data. That is, it needs to build theoretical concepts from the lower order to higher ones and, at the same time, connect its theories to worldly phenomena. The Semantics dimension of LCT offers operable tools making explicit how such a hierarchical knowledge structure is built through the multisemiotic chemistry discourse. Semantic density helps reveal how the meanings construed by the chemistry discourse develop through schooling to build increasingly complex chemistry knowledge, and semantic gravity makes explicit the connection between the technical chemistry knowledge and empirical phenomena. Therefore, Semantics will be used to analyze how the different semiotic resources develop through schooling to build the hierarchical knowledge structure of chemistry in Chapters 4, 5, and 6.

3.1.3 Complementarity between the two theoretical tools

The collaborations between SFL and code theory originated by Basil Bernstein that has recently been developed into Legitimation Code Theory (Maton 2014) can be traced back to the 1960s. In recent years, the long-standing collaborations have intensified as SFL and LCT have become increasingly used together in joint analyses of shared data (for example, Doran 2017, 2018a, 2018c; Hood 2016, 2017; Martin and Maton 2013; Vidal 2014). The two frameworks are often used jointly in research because they are highly **complementary** to each other. SFL and LCT often operate side by side as analytic frameworks providing complementary analyses (Maton and Doran 2017b: 613). More specifically, in these studies, SFL tends to be used to analyze the meaning-making of semiotic phenomena, while LCT is enacted to analyze how the knowledge is developed. They are related together to more fully explore the shared problem-situation (ibid.: 613).

In the case of the present study, the complementarity between the two theoretical tools perfectly addresses the issues involved in the analysis of knowledge-building through the multisemiotic chemistry discourse. SFL explores the construal of meaning by the chemistry discourse, while LCT examines how the discourse develops through schooling to build the hierarchical knowledge structure of chemistry. Figure 3.13 illustrates how the two frameworks complement each other to achieve the research objectives of this study.

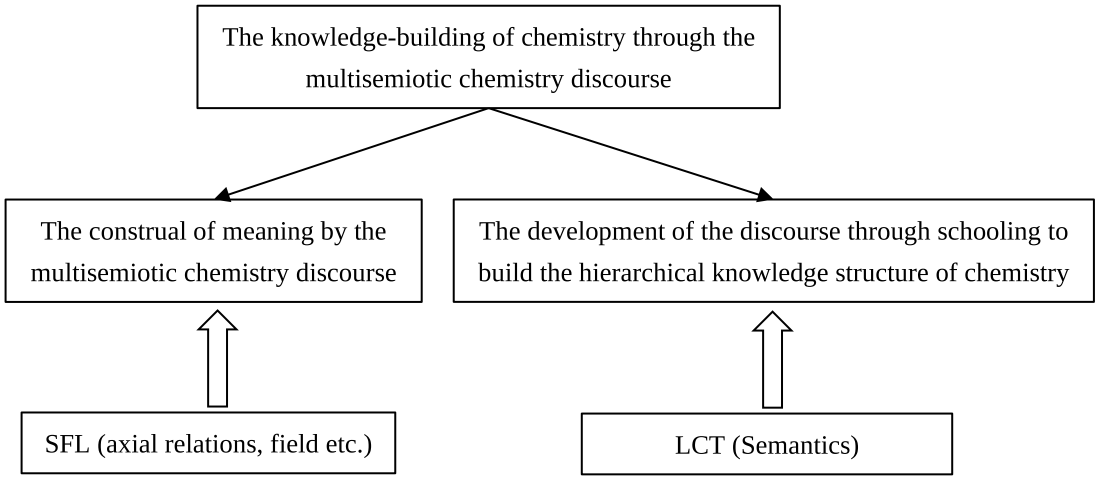


Figure 3.13 SFL and LCT complement each other to achieve the research objectives of this study

3.2 Methodology

Section 3.1 has introduced the theoretical tools that will be used to address the research questions. This section will describe the methodology of this study. As one of the major objectives of this study is to describe the grammar of chemical formalisms, it will first introduce the methodology underpinning the description (Section 3.2.1). Following this, it will explicate the method of analyzing how the multisemiotic chemistry discourse develops through schooling to build the hierarchical knowledge structure of chemistry (Section 3.2.2).

3.2.1 Semiotic description

Since one of this study’s objectives is to describe chemical formalisms’ grammatical systems, we need to consider the methodology underlying semiotic description. As noted in Section 2.2.1.1, the last few decades have seen an explosion of descriptions of semiotic systems other than language, for example, images (Kress and van Leeuwen 2006; O'Toole 1990; Painter et al. 2013), sound and music (van Leeuwen 1999, 2009), bodily action (Martinec 1998, 2000, 2001), three-dimensional space (Ravelli and McMurtrie 2016; Stenglin 2009), and animation (He 2020). These descriptions assume different SFL theoretical categories as their points of departure, some of which even conflict with each other. For example, some of the above works hold that the three metafunctions developed from English apply to all semiotic systems and take them as the departure point, whereas numerous other studies, for example, van Leeuwen (1999, 2009) on sound, Djonov and van Leeuwen (2011) on texture, Zhao (2010a) on web-based multimodal information texts, and Doran (2018c) on mathematics, show that metafunctions are not universal but specific to each semiotic system. Another conflict is that some studies assume a rank scale, such as O'Toole (1990) on images and O'Hallraon (2005) on mathematics, while others do not, such as Kress and van Leeuwen (2006) on images and Doran (2018c) on mathematics. These differences suggest a lack of consensus in what SFL categories to be assumed for describing non-linguistic semiotic systems, necessitating a discussion of the descriptive principles underpinning semiotic description.

In light of the absence of an agreed-upon methodology for semiotic description, Doran (2018c) usefully proposes three principles as the overarching goals to guide semiotic description. These principles are explicated in the following quote from Doran (2018c: 56)

> "First, the description must, in some way, bring out the specific functionality of the resource under study. This involves accounting for the possible variation within the resource and proposing varying degrees of generalization so as to push beyond a simple inventory of discrete possibilities. Second, the description must be able to be compared with descriptions of other resources (such as gesture, image, English, Pitjantjatjara and Tagalog), and in doing so show similarities and differences in organization. Third, the description must be based upon explicit methods of argumentation that allow it to be compared and judged in relation to competing descriptions of the same resource."

All of the three principles point to a key aspect of semiotic description - each semiotic resource must be described on its own terms, and the description of one semiotic resource cannot be unquestioningly transferred to another. Similar cautions have also been raised for even the descriptions of different languages (Caffarel et al. 2004). In this sense, the three metafunctions developed from the description of English cannot be taken as the point of departure for semiotic description as it unproblematically assumes that all semiotic resources share the same metafunctional organization as English, which inevitably waters down the intrinsic functionalities of these resources (see Section 2.2.1.1.1 for a detailed discussion of this). Moreover, it does not allow the descriptions of these resources to be compared and contrasted in a principled nature as they are already assumed to embody the same functionalities. Thus, instead of taking the three metafunctions as the point of departure, semiotic description should take the axial relations of system and structure, from which the metafunctions are derived, as the more fundamental descriptive principle (see Section 2.2.1.1.1 for a detailed introduction of how Halliday generates the three metafunctions based on his description of English and see also Martin 2013b for a demonstration of how the metafunctions of English are derived from the axial relations). In addition to metafunctions, Martin (2013b, 2015a) also demonstrates how ranks and strata in the model of English can be derived from the theoretical primitive of axis. The axial justifications for metafunctions, ranks, and strata are summarized in Table 3.2.

Table 3.2 Axial justifications for metafunction, rank, and strata (adapted from Doran 2018c: 48)

	Axial justification
metafunction	Each metafunction displays: ➢ relative paradigmatic independence to other metafunctions ➢ relative paradigmatic interdependence within metafunctions ➢ a distinct type of structure: • ideational: particulate structure • interpersonal: prosodic structure • textual: periodic structure
rank	➢ Each rank displays distinct paradigmatic options and syntagmatic structures. ➢ System-structure cycles are related through constituency.
stratification	➢ Each stratum displays distinct paradigmatic options and syntagmatic structures. ➢ System-structure cycles are related through abstraction.

Following this methodology, this study will take the axial relations between system and structure as the departure point for the description of chemical formalisms. This methodology starts with describing the grammatical structures of a semiotic system and then formalizes these structures as grammatical systems in the form of system networks. This system/structure cycle gives rise to metafunctions and ranks of that semiotic system. The following illustrates this with Halliday's description of

English.

First, it will illustrate how Halliday generates metafunctions based on the relative paradigmatic independence of English's clausal systems and the distinct types of syntagmatic structure realizing them (Halliday 1969, 1978, 1979). This can be demonstrated with the TRANSITIVITY and MOOD systems of English clauses, which are independent of each other and realized by distinct structures. Paradigmatically, choices in TRANSITIVITY are independent of choices in MOOD. For example, the following clauses show that the three major transitivity structures can occur with both indicative and imperative moods.

<table>
<tr><td>(3:8)</td><td>The boy</td><td colspan="2">kicked</td><td>the ball.</td></tr>
<tr><td>[material]</td><td>Actor</td><td colspan="2">Process</td><td>Goal</td></tr>
<tr><td>[indicative]</td><td>Subject</td><td>past
Finite</td><td>kill
Predicator</td><td>Complement</td></tr>
</table>

<table>
<tr><td>(3:9)</td><td>Kick</td><td>the ball!</td></tr>
<tr><td>[material]</td><td>Process</td><td>Goal</td></tr>
<tr><td>[imperative]</td><td>Predicator</td><td>Complement</td></tr>
</table>

<table>
<tr><td>(3:10)</td><td>The boy</td><td colspan="2">remembers</td><td>how to play football.</td></tr>
<tr><td>[mental]</td><td>Senser</td><td colspan="2">Process</td><td>Phenomenon</td></tr>
<tr><td>[indicative]</td><td>Subject</td><td>present
Finite</td><td>remember
Predicator</td><td>Complement</td></tr>
</table>

<table>
<tr><td>(3:11)</td><td>Remember</td><td>how to play football!</td></tr>
<tr><td>[mental]</td><td>Process</td><td>Phenomenon</td></tr>
<tr><td>[imperative]</td><td>Predicator</td><td>Complement</td></tr>
</table>

<table>
<tr><td>(3:12)</td><td>The boy</td><td colspan="2">is</td><td>Naughty.</td></tr>
<tr><td>[relational]</td><td>Carrier</td><td colspan="2">Process</td><td>Attribute</td></tr>
<tr><td>[indicative]</td><td>Subject</td><td>present
Finite</td><td>be
Predicator</td><td>Complement</td></tr>
</table>

(3:13)	*Be*	*Nice!*
[relational]	Process	Attribute
[imperative]	Predicator	Complement

SFL represents the two independent choices as two simultaneous and independent systems, as shown in Figure 3.14. The relative independence between the two systems suggests that they are two distinct functional components, termed ideational and interpersonal.

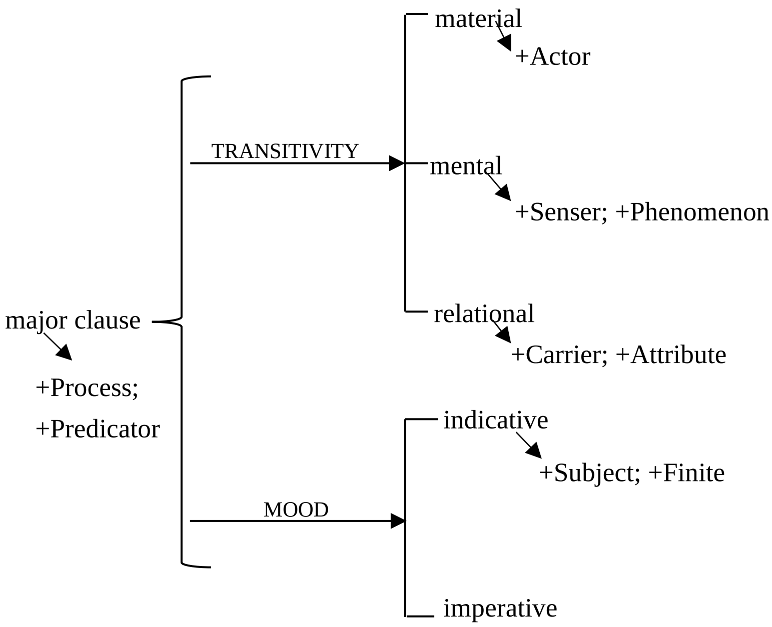


Figure 3.14 The simplified TRANSITIVITY and MOOD systems

In contrast to TRANSITIVITY being independent of MOOD, MODALITY is interdependent with MOOD. Choices in MODALITY are dependent on choices in MOOD but independent of choices in TRANSITIVITY. For instance, Examples (3:14-15) show that choices in MODALITY occur with only indicative mood, suggesting that MODALITY bundles with [indicative] (the asterisk * indicates that the example is impossible). Examples (3:16-18), on the other hand, show that choices in MODALITY can occur with all the three major TRANSITIVITY types.

(3:14) *The boy <u>may</u> play football after school* indicative & modality

(3:15) *<u>May</u> play football after school** imperative & modality (impossible)

(3:16) *The boy <u>may</u> play football after school* material & modality

(3:17) *The boy <u>may</u> remember how to play football* mental & modality

(3:18) *The boy may be right* relational & modality

MODALITY is thus interdependent with MOOD and independent of TRANSITIVITY, formalized as the network in Figure 3.15. Being interdependent with MOOD and independent of TRANSITIVITY suggests that MODALITY forms the same functional component as MOOD, a distinct component from TRANSITIVITY.

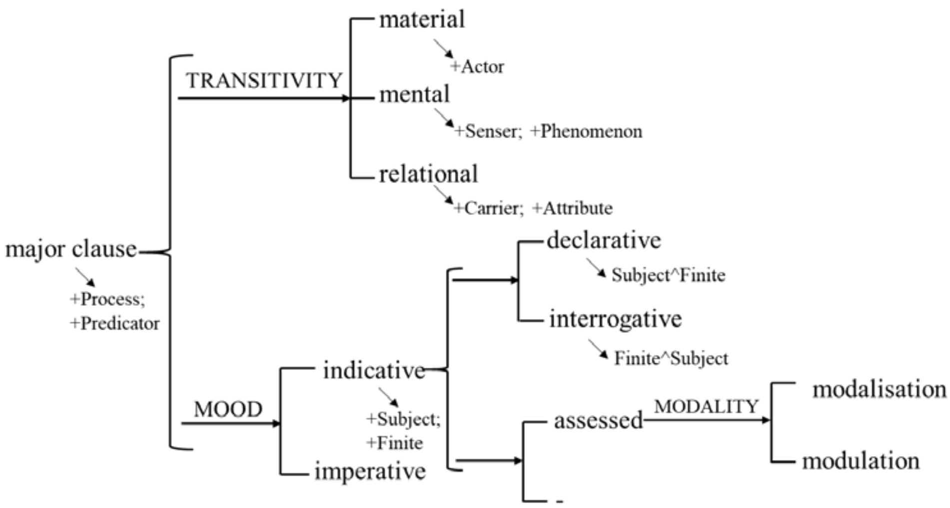


Figure 3.15 Simplified TRANSITIVITY, MOOD, and MODALITY

The above illustrates the generation of metafunctions from the paradigmatic independence of systems. Another evidence for suggesting metafunctions is the types of structure realizing these systems. As introduced in Section 3.1.1.1.2, different functional components tend to be realized by different types of structure. TRANSITIVITY is realized by multivariate particulate structures. MOOD and MODALITY are associated with prosodic structures (see Section 3.1.1.1.2 for a detailed introduction of these two types of structure). For example, the clause *if you don't get no publicity you don't get no people at the fight* involves the modality of negative polarity cutting across the clause. The meaning is distributed like prosody throughout a continuous stretch of discourse, working together to achieve a cumulative effect (Halliday 1979: 66). For MOOD, the agreement between the central interpersonal functions of Subject and Finite can be interpreted as a prosodic structure (Doran 2018c: 43). Syntagmatically, the Subject and Finite that determine moods must agree in terms of number and person. The following examples illustrate this agreement.

(3:19) ***I am*** *going to school*. First person, singular

(3:20) ***You are*** *going to school*. Second person, plural

(3:21) ***He/she is*** *going to school*. Third person, singular

(3:22) ***We/you/they are*** *going to school*. First/second/third person, plural

MOOD and MODALITY's association with prosodic structure suggest that they should be considered part of the same functional component, in contrast to TRANSITIVITY realized by particulate structure. Based on both the paradigmatic interdependence or independence and syntagmatic similarities or differences, the description reveals that MOOD and MODALITY form one functional component distinct from the one constituted by TRANSITIVITY. It is through this axial argumentation that metafunctions are derived.

The system/structure cycle also gives rise to rank. Like metafunctions, ranks derive from distinct bundles of systems, which, however, are interrelated through constituency. To illustrate this, the following will show how clause and group ranks are developed in English's lexicogrammar. For example, the clause *the boy kicked the ball* involves an Actor realized by the nominal group *the boy*, which itself involves the functional configuration of 'Deictic ^ Thing'. This nominal group may be expanded to include other functional elements but still performs Actor. For example, *the two boys* adds the functional element of Numerative; *the two naughty boys* inserts an Epithet; *the two naughty Chinese boys* adds a Classifier. These options form a system for nominal groups, as shown in Figure 3.16.

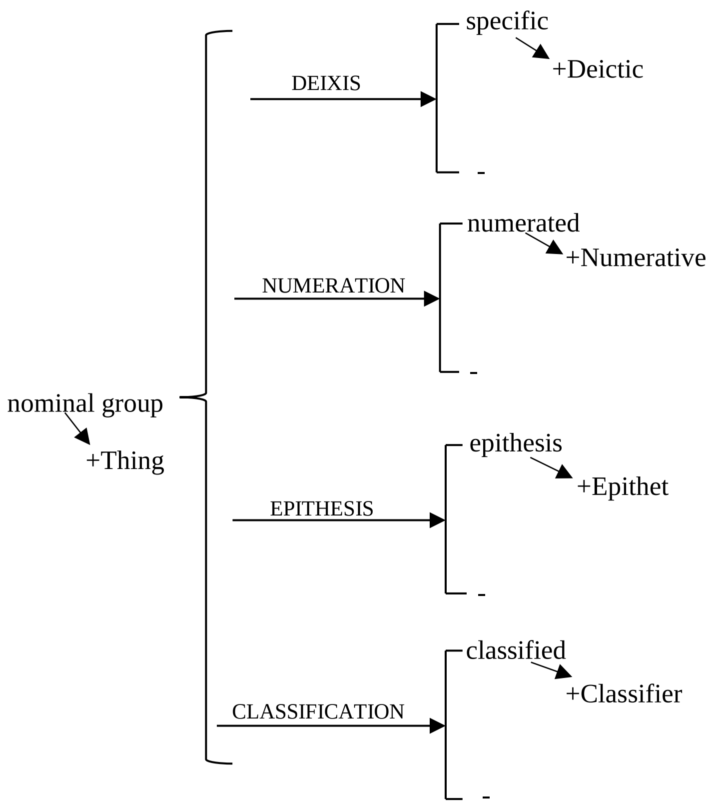


Figure 3.16 A simplified system for English nominal groups (adapted from Halliday and Matthiessen 2014)

The above nominal groups perform not only Actor in material clauses but also Senser in mental clauses and Carrier in relational clauses. However, they realize only parts of the functional elements involved in these clauses. Process, the key function in transitivity structures, is realized by verbal groups, which also have their own system (see the system network presented in Halliday and Matthiessen 2014: 410). With nominal and verbal groups each realizing parts of the functions in the clause system's transitivity structures, the group system should be viewed as independent of the clause system, interrelated through constituency. Figure 3.17 shows the constituency relation between the two grammatical units.

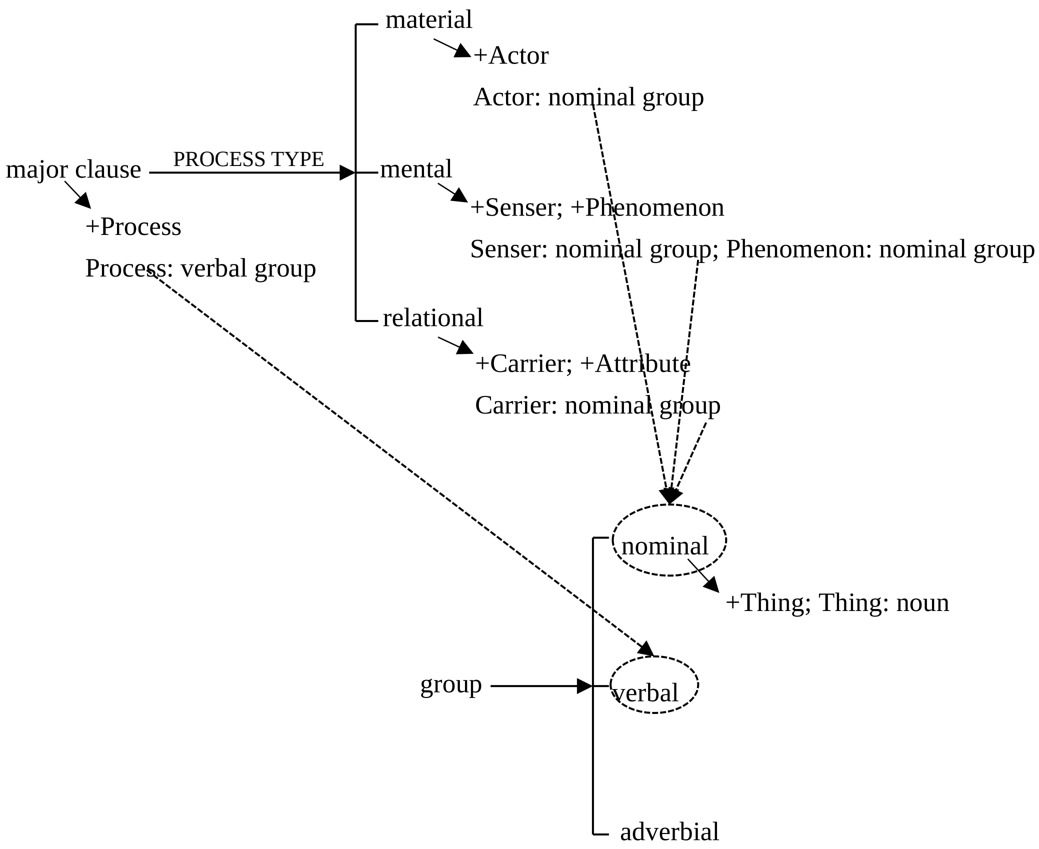


Figure 3.17 The simplified clause and group systems are arranged at different hierarchies and related through constituency

Figure 3.17 shows that the clause and group systems are independent of each other and related through constituency. The simplified PROCESS TYPE system is realized by three functional structures, of which each function is in turn realized by a group. The groups are in fact features from systems at the lower hierarchy. These hierarchies of grammatical units derived from independent systems interrelated through constituency are ranks.

The above demonstrates that taking the axial relations between system and structure as the departure point and starting with the actual uses of a semiotic resource allows for bringing out the resource's intrinsic functionalities and organizations. This '**bottom-up**' methodology enables Halliday to describe the systemic functional grammar of English (Halliday 1967a, 1967b, 1968, 1970). This study will follow this methodology to describe the grammar of chemical formalisms in Chapter 4.

3.2.2 Qualitative research

In addition to the above methodology unique to semiotic description, this study also adopts a qualitative research method to analyze how the multisemiotic chemistry discourse develops through schooling to build the hierarchical knowledge structure of chemistry. A qualitative research approach is selected because it affords deep analysis of data in particular contexts (Patton 2002). In the case of the present study, the method will be used to trace the development of the different semiotic resources in terms of their Semantics (primarily semantic density) and analyze how they develop to build the hierarchical knowledge structure of chemistry. To do this, it will first develop models that allow for 'seeing' the different strengths of Semantics embodied by the different semiotic resources and then enact them to trace their development through schoolings in secondary school chemistry textbooks. The changes in the semiotic resources' Semantics through schooling then make explicit how the knowledge of secondary school chemistry is built through the multisemiotic chemistry discourse.

3.3 Data

This section will introduce the data used in this study, including data collection (Section 3.3.1) and data analysis (Section 3.3.2).

3.3.1 Data collection

As stated in Chapter 1, this study investigates how the multisemiotic resources in secondary school chemistry textbooks build chemistry knowledge. The reason for focusing on chemistry textbooks is that they are the primary resource for chemistry instruction by teachers and the core reading material for students to learn chemistry (Chiappetta and Koballa 2002; Gkitzia et al. 2011; Upahi and Jimoh 2015). Specifically, the English chemistry textbooks used by secondary schools in New South Wales (NSW) Australia constitute the present study's data.[12] The Australian textbooks are selected for three reasons. Firstly, Australian science education has been one of the most advanced in the world. It has been reported by the International Association for Evaluation of Educational Achievement (2019) that Australian students' science achievement has stably ranked around top 10 in the world in the last two decades. Since it has been noted above that textbooks are the primary sources for teaching, the high achievement indicates that the textbooks selected can be viewed as representative of high-quality teaching material. Secondly, studies also show that the Australian science

[12] Once again, it must be noted that the term "English chemistry textbooks" does not mean the textbooks use only one type of semiotic sesource, i.e., language. The word "English" here merely denotes what type of language the textbooks use.

textbooks are high-quality pedagogical tools supporting teaching and learning in Australian secondary school science classes (McDonal 2016; Rennie et al. 2001). They are thus reliable data for the knowledge-building analysis in this study. The third reason is the availability of data. The present research was conducted in Sydney, Australia, where the chemistry textbooks used by secondary schools in NSW are more easily to be collected.

The curriculum of secondary school chemistry in NSW includes six years that are categorized into three stages: Stage 4 (Years 7 and 8), Stage 5 (Years 9 and 10), and Stage 6 (Years 11 and 12). Table 3.3 lists the textbooks used for each year.

Table 3.3 Data used in this study (see Appendix A for full details)

Stage	**Year**	**Textbook**
4	7	Oxford Insight Science 7 student book
	8	Oxford Insight Science 8 student book
5	9	Oxford Insight Science 9 student book
	10	Oxford Insight Science 10 student book
6	11	Pearson Chemistry 11 New South Wales Student Book
	12	Pearson Chemistry 12 New South Wales Student Book

Note that each of the Year 7-10 textbooks comprises four components: Physical World, Earth and Space, Living World, and Chemical world (Board of Studies NSW. 2012: 90). Chemistry is incorporated with other components in these textbooks but constitutes the Year 11 and 12 textbooks by itself. The corpus presented in Table 3.3 involves high usage of chemical formalisms and images, providing rich data for the descriptions of these resources in Chapters 4 and 5. In addition, the data ranging from Year 7 to 12 allow this study to trace the development of these resources in building the knowledge of secondary school chemistry. Therefore, the corpus provides rich data for both the descriptive work and the knowledge-building analysis to be conducted by this study.

3.3.2 Data analysis

As noted in Chapter 1, this study aims to explore two major issues: the construal of meaning by the multisemiotic chemistry discourse and how the discourse develops across schooling levels to build the hierarchical knowledge structure of chemistry. To address the first issue, the data will be analyzed descriptively. For example, the different chemical formalisms will be described in terms of their meaning-making patterns.

Specifically, as discussed in Section 3.2.1, the data will be described based on the axial relations between system and structure, which allows for generating chemical formalisms' metafunctional organization and grammatical hierarchies in Chapter 4. As for the second issue, the data will be processed analytically. As noted in Section 3.2.2, models of the semantic density of the different semiotic resources will be developed from the data and they will be enacted to analyze how the resources develop through schooling in terms of their semantic density. This analytical work would reveal the role each resource play in building the chemistry knowledge in Chapters 4, 5, and 6.

In a nutshell, the data will be processed both descriptively and analytically, which enables addressing the research questions to be explored in Chapters 4-6.

3.4 Summary

This chapter has outlined the research design for this study, including the theoretical framework, methodology, and data. It has introduced the theoretical concepts from both SFL and LCT and explained how these theoretical tools will complement each other to address the research questions. The methodology section has introduced the approach to be taken for the semiotic descriptions of chemical formalisms and the qualitative research method for analyzing how the multisemiotic chemistry discourse develops to build the hierarchical knowledge structure of chemistry. Finally, this chapter has introduced the data used in this study and how they are processed. Based on the theoretical tools, methodology, and data, the following substantive chapters will explore each of the research questions.

Chapter 4 Knowledge-building through Chemical Formalisms

Chemistry is characterized by its high use of formalisms (Gilbert and Treagust 2009; Treagust and Chittleborough 2001). For example, in secondary school chemistry textbooks, diverse chemical formalisms, including chemical formulas, chemical equations, and structural formulas, permeate different schooling levels. Chemical formalisms thus constitute an essential '**knowledge-builder**' that students need to master to succeed in learning chemistry. Therefore, understanding how the formalisms build chemistry knowledge is crucial for teachers to develop effective pedagogy that enables students to learn both the technical knowledge of chemistry and the chemical formalisms organizing this knowledge. As discussed in Chapter 1, examining the knowledge-building of chemistry from a semiotic perspective entails understanding both the construal of meaning by semiotic resources and the development of the resources across learning levels to build the hierarchical knowledge structure of chemistry. This chapter will thus first explore how chemical formalisms are organized to make meaning (Section 4.1) and then trace their development across schooling levels in secondary school chemistry textbooks (Section 4.2). Following this, Section 4.3 provides a summary of the findings found in this chapter.

4.1 Grammars of chemical formalisms

Chemistry is characterized by a wide range of semiotic resources (Gilbert and Treagust 2009). Among these resources, the use of formalisms relatively specific to chemistry provides one of the key ways of organizing chemistry knowledge. Indeed Johnstone (1991), in his seminar classification of chemistry knowledge, argues they constitute one of the three main levels of chemistry knowledge, known as the "symbolic level" (see Section 2.3.1 for a detailed introduction to this level of chemistry knowledge). For students learning chemistry, this means chemical formalisms make up an essential component of chemistry literacy. However, the uncommon-sense nature of chemical formalisms causes potential impediments for students accessing chemistry. To address such literacy challenges, explicit literacy pedagogy of the formalisms is crucial (Rose and Martin 2012; The New London Group 1996). However, it has long been recognized that to develop a literacy pedagogy for a particular type of discourse, it is necessary to first understand how that discourse works (Rose and Martin 2012). An approach that has put disciplinary literacy as the center of its pedagogic practice is the educational linguistics informed by SFL. Focusing on language, numerous SFL studies have examined how language works in different disciplines in terms of, for example, grammar, discourse semantics, register, and genre (Christie and Martin 2007; Halliday and Martin 1993; Hao 2020; Martin and Rose 2008; Rose and Martin 2012; Veel 1997) In the last two decades, this has been expanded into multimodal literacies including images (Kress and van Leeuwen 2006; O'Toole 1990; Unsworth 2001), mathematical symbolism (Doran 2018c; O'Halloran 2005), and animation (He 2020). However, at this stage, multimodal studies have developed only a nascent understanding of chemical formalisms (Liu 2011) that has not yet reached the depth of studies for other semiotic resources. Upon this background, this section will explore in detail how chemical formalisms work. That is, it will describe the grammars of chemical formalisms from the perspective of Systemic Functional Semiotics.

Chemical formalisms include three major types: chemical formulas, chemical equations, and structural formulas. Since chemical formulas are part of chemical equations, we will describe them together in Section 4.1.1. Following this, the grammar of structural formulas will be accounted for in Section 4.1.2. To understand the similarities and differences between these chemical formalisms in making meaning, Section 4.1.3 will develop a functional semiotic typology of the formalisms.

4.1.1 Grammar of chemical equations

Chemical equations are one of the central forms of representation for teaching and learning chemistry (Taber 2009: 84). They are ubiquitous in secondary school chemistry and form a crucial ‘builder’ of chemistry knowledge. Therefore, they are part of the high-stakes reading that students need to do to succeed in learning chemistry. However, it has been widely reported that students face significant difficulties in understanding the chemical equations (Jong and Taber 2014; Laugier and Dumon 2004; Taber 2009). To support students better understand chemical equations, it is essential to grasp the way the equations organize meaning. This section will thus focus on the meaning-making of chemical equations in secondary school chemistry. It will describe the chemical equations’ grammar in terms of the theoretical primitive of **axis** – the interaction between paradigmatic and syntagmatic relations (Martin 2013b) (see Section 3.1.1.1.1 for a detailed introduction of axial relations). It will first distinguish different types of chemical equations and specify the scope of this description (Section

4.1.1.1). Following this, it will then outline the grammatical units used to understand chemical equations (Section 4.1.1.2), before describing the grammatical systems themselves at each rank (Section 4.1.1.3). Based on the systems and structures, it will generate the metafunctions in the grammar of chemical equations (Section 4.1.1.4). Finally, it will briefly discuss the grammar of structural equations in terms of their similarities and differences with chemical equations (Section 4.1.1.5).

4.1.1.1 Scope of description

Chemical equations are characterized by their sheer diversity. Even for secondary school chemistry, our data suggest that it uses numerous types of chemical equations. However, for semiotic description, we need to focus on chemical equations that organize meaning in similar ways, i.e., chemical equations that belong to one semiotic system. Therefore, this section will specify the scope of the description. It will show that the description takes molecular and ionic equations as its objects of study due to their structural similarities and excludes structural equations and chemical equations that interact with mathematical equations. The following will explain the inclusion and exclusion of these equations.

In chemistry, a chemical equation is often defined as a type of “symbolic representation of a chemical reaction in terms of chemical formulas” (Ebbing and Gammon 2008: 73). Under this definition, there are two major types: molecular equations and ionic equations. A **molecular equation** is “a chemical equation in which chemical reactants and products are written as if they were molecular substances, even though they may actually exist in solution as ions’ (ibid.: 129).[13] For example, Equation (4:1) illustrates a molecular equation.

(4:1) $2H_2(g) + O_2(g) \rightarrow 2H_2O(l)$

This is a typical molecular equation involving three molecular formulas representing hydrogen gas molecules (H_2), oxygen gas molecules (O_2), and water molecules (H_2O).[14] The number ‘2’ attached to the formulas are called coefficients, showing the quantities of the molecules involved in the reaction. The signs ‘(g)’ and ‘(l)’ on the right sides of the formulas are known as state symbols, representing the physical states of the molecules. The two state symbols indicate two respective physical states: gas and liquid.

It should be noted that the term ‘molecular equation’ is also used as a cover term

[13] An ion is an atom or molecule with electric charges.
[14] Molecular formulas are formulas that typically represent molecules.

for those involving formulas representing what are known as ionic compounds.[15] Equation (4:2) gives an example of a molecular equation involving formulas representing an ionic compound.

(4:2) $Na(s) + 2H_2O(l) \rightarrow 2NaOH(aq) + H_2(l)$

In Equation (4:2), the formula NaOH represents an ionic compound, but it is written in the form of molecular formulas. In chemistry, chemical formulas used for representing ionic compounds are called “empirical formulas” (Brown et al. 2012: 68).[16] However, semiotically speaking, molecular formulas and empirical formulas do not show any formal distinctions – they are both composed of chemical symbols and subscript numbers.[17] ‘Molecular equation’ as an overarching term thus does not mean that the represented chemical species are necessarily molecules, but rather that the constituent formulas share the formal similarity with molecular formulas.[18] That is, they all consist of chemical symbols and subscript numbers. In contrast, **ionic equations** involve a type of formula that is formally distinct from molecular formulas. They are called ionic formulas in chemistry and are marked by a plus or minus sign that shows the polarity of electric charges, representing the dissolved ions in solution (Ebbing and Gammon 2008: 130). For example, in Equation (4:3), the dissolved sodium hydroxide is written as Na^+ and OH^-.

(4:3) $Na(s) + 2H_2O(l) \rightarrow 2Na^+(aq) + 2OH^-(aq) + H_2(l)$

Semiotically speaking, ionic equations are distinct from molecular equations in that the former involves ionic formulas while the latter does not. Overall, the two equations are almost the same except for a minor difference in the formulas used, suggesting that they can be accounted as one semiotic system for description.

Our data suggest another type of equation that is used primarily in organic chemistry. What is special about these equations is that they represent reactants and products through structural formulas (see discussion in Section 4.1.2).[19] For example, Equation (4:4) represents an organic reaction between methane and chlorine gases

---

15 An ionic compound is a compound composed of cations and anions. Cations are atoms or molecules with positive electric charges, whereas anions are the opposits, atoms or molecules with negative electric charges.

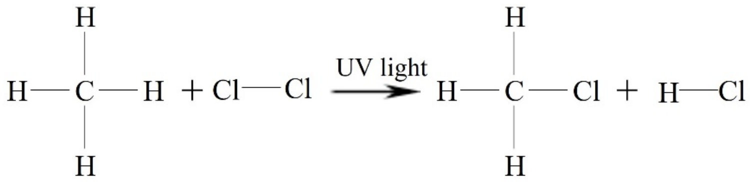
16 Note that empirical formulas represent ionic compounds only when the constituent chemical symbols involve metal symbols. For example, NaOH representing sodium hydroxide involve the metal symbol Na, which represents the sodium metal.

17 Chemical symbols are the symbols representing chemical atoms. For example, H is a chemical symbol representing hydrogen atoms. Subscript numbers, for example ‘2’ in the formula ‘H_2’, represent the quantities of the atoms they are attached to.

18 Chemical species are the substances involved in a chemical reaction. For example, the reaction represented by Equation (4:1) includes three chemical species: hydrogen gas, oxygen gas, and water.

19 Structural formulas is a type of formula in chemistry that represents molecules’ structure. They typically comprise chemical symbols connected by lines. For example, the structural formulas of hydrogen gas molecules is H—H.

under the presence of UV light.



(4:4)

All reactants and products are represented through structural formulas. The purpose of representing chemical species by structural formulas is to show which parts of the molecules undertake the reaction. Equation (4:4) implicitly shows that the chlorine atom (Cl) displacing the hydrogen atom (H) leads to the reaction. With respect to the naming of this type of equation, it appears that chemists do not have a formal name for them. Given the particular formulas constituting the equations, we refer to them as **structural equations**.

Overall, our data suggest three types of chemical equations, each designed for a particular communicative purpose. However, due to the significant structural differences in structural equations from molecular and ionic equations, it seems impossible to include all of them in one grammatical description. As discussed above, molecular equations and ionic equations are similar in grammatical organizations at all hierarchies, except for one minor distinction between the formulas involved, i.e., one shows signs representing electric charges and the other does not. It is thus feasible to include them in one description. Structural equations, however, differ from the other two equations in various aspects. Most strikingly, the organization of structural formulas is significantly different from the other two formulas as they employ two-dimensional space, or even three-dimensional space, to make meaning. In addition, structural equations do not show coefficients and state symbols as molecular and ionic equations do. For example, Equation (4:4) does not involve the coefficients and state symbols as Equation (4:3) does. In light of these distinctions, we exclude structural equations from the present description.

Finally, in order to understand the intrinsic functionality of chemical equations, the description in this chapter does not consider the interaction between chemical equations and other semiotic systems such as language and mathematics. The data under study shows one particular interaction in this regard – the interaction between chemical equations and mathematical equations as illustrated in (4:5) from the Year 11 textbook.

(4:5) $C(s) + O_2(g) \rightarrow CO_2(g) \qquad \Delta H = -393.5 kJmol^{-1}$

The equation on the left, $C(s) + O_2(g) \rightarrow CO_2(g)$, is a chemical equation representing a combustion reaction of carbon, while the equation on the right, $\Delta H = -393.5 kJmol^{-1}$, is a mathematical equation that denotes the energy released per mole carbon. Since mathematical equations have their own grammatical systems, they should be treated as a semiotic system in their own right (Doran 2018c; O'Halloran 2005). Therefore, the mathematical equation in this example should not be taken as part of the chemical equation. Instead, it is best to consider the issue at stake as intermodality (Painter et al. 2013) between chemical equations and mathematical symbolism. Not considering the interaction between chemical equations and other semiotic resources allows chemical equations to be described in their own right without running the risk of transferring categories from one semiotic resource to another, and so it provides access to see what chemical equations themselves do when interacting with other semiotic systems.

With the scope of description demarcated, the following sections will build up the architecture of chemical equations' grammar. To start with, we will first outline the grammatical units used to understand chemical equations.

4.1.1.2 Grammatical units for description

Similar to language, chemical equations organize meaning at different levels of grammatical units. In language, these levels are called ranks, a **multivariate constituency** relation between different hierarchies (Halliday 1961; Huddleston 1965) (for a detailed introduction of ranks see Section 3.1.1.1.4). Our data and description show that chemical equations include five ranks: equation, expression, term, formula, and symbol,[20] which will be illustrated through Equation (4:6).

(4:6) $CH_4(g) + 2O_2(g) \rightarrow CO_2(g) + 2H_2O(l)$

The highest rank is **equation**.[21] It refers to the entire chemical equation. As shown by Equation (4:6), chemical equations use the arrow sign '→' to relate two sides known as the reactants on the left and products on the right. This is the basic structure of a chemical equation at the highest rank. The second rank is **expression**.[22] Expression is a class term borrowed from mathematical symbolisms (O'Halloran 2005, Doran 2018), referring to the two sides of a chemical equation, for example, '$CH_4(g) + 2O_2(g)$' and

[20] Rather than a rank scale, Doran (2018c) proposes what he calls a '**nesting scale**' for mathematical symbolism, due to the overarching univariate organizations of the grammar. As chemical equations are largely **multivariate**, which will be shown in Section 4.1.1.3, we simply use the term 'rank' scale here.
[21] Liu (2011) terms this rank 'clause'. The term equation is preferred here as it is more oriented to chemical equations and manifests the object under description, rather than transferring categories over from language.
[22] It must be noted that the term 'expression' here is NOT Hjelmslev's (1969) concept of 'expression'.

‘$CO_2(g) + 2H_2O(l)$’ in Equation (4:6). It should be noted that the arrow sign is not assigned a class in the description because it involves only two variants: ‘→’ and ‘⇌’, which can be lexicalized through delicacy in the grammatical system at the equation rank (see Section 4.1.1.3.1 for a detailed discussion). The rank below expression is **term**, which refers to the components on either side of the plus sign, for example, ‘$2H_2O(l)$’. A term describes three things: the number, types, and physical states of chemical species. The term ‘$2H_2O(l)$’ refers to two units of water molecules in liquid state. Similar to the arrow sign not being assigned a class at the expression rank, the plus sign at the term rank is not assigned any classes as it does not involve any variations and can be lexicalized directly as ‘+’ within the grammatic system (see Section 4.1.1.3.3.1 for a detailed discussion). The rank below term is **formula**, referring to the chemical formulas representing chemical species. For example, CH_4, O_2, CO_2, and H_2O in Equation (6) represent four respective chemical substances: methane, oxygen gas, carbon dioxide, and water. Finally, the lowest rank is **symbol**.[23] It refers to the chemical symbols or the symbols and their attached subscript numbers, for example, ‘O’ or ‘H_2’ (see Section 4.1.1.3.5 for a detailed discussion of why taking the symbol and its attached subscript number as one unit). It is important to note that the formula and symbol ranks form the grammar of chemical formulas, which themselves belong to a **distinct semiotic system**. For the convenience of description, we will not separate the description of chemical formulas from chemical equations, but they will be taken as distinct semiotic systems when developing a functional semiotic typology for all the chemical formalisms involved in secondary school chemistry in Section 4.1.3. Table 4.1 illustrates all the grammatical units that will be used to describe the grammar of chemical equations.

Table 4.1 Grammatical units used to describe chemical equations

Grammatical unit	**Example**
equation	$CH_4(g) + 2O_2(g) \rightarrow CO_2(g) + 2H_2O(l)$
expression	$CO_2(g) + 2H_2O(l)$
term	$2H_2O(l)$
formula	H_2O
symbol	O or H_2

In summary, this section has introduced a preliminary set of ranks in the grammar

[23] It must be noted that the term ‘symbol’ here simply refers to chemical symbols and does NOT refer to Peirce’s term of ‘symbol’.

of chemical equations – equation, expression, term, formula, and symbol. As the description develops, some ranks may be revised – in particular, the rank of expression, which, however, forms a useful starting point for the discussion. The next section will begin to describe the grammatical systems of chemical equations.

4.1.1.3 Grammatical systems of chemical equations

This section will be devoted to building up systems and structures of chemical equations. The description will start from the equation rank through to the symbol rank. It will begin with the structures of equations and their systemic organization (Section 4.1.1.3.1), before examining expressions (Section 4.1.1.3.2). A key feature of this sub-section will be that expression as a unit is in fact not needed in a grammatical description. Following this, this section will consider the internal structures of terms and their complexing relations into expressions (Section 4.1.1.3.3). Finally, the description will then focus on the grammar of chemical formulas (Section 4.1.1.3.4) – one of the most complicated grammatical areas in chemical equations – and then the internal variation and complexing of symbols (Section 4.1.1.3.5).

4.1.1.3.1 Equation

As noted in Section 4.1.1.1, Chemical equations are a type of symbolic representation of chemical reactions, which typically include three parts: the left expression representing reactants, the arrow indicating reaction directions, and the right expression representing products. For example, as shown in Equation (4:7), the left expression includes two reactants: methane (CH_4) and oxygen gas (O_2), and the right expression consists of two products: carbon dioxide (CO_2) and water (H_2O). The arrow in the middle suggests that the reaction starts from the left and moves towards the right.

(4:7) $CH_4(g) + 2O_2(g) \rightarrow CO_2(g) + 2H_2O(l)$

Chemical equations represent the transformation of an initial state into a final state (Arnold 2011: 9), with the left expression showing the initial state and the right showing the final state after the reaction. The transformation of the two states is denoted by the reaction arrow, which implies a direction to the chemical change – from reactants (left expression) to products (right expression) (Taber 2009: 96). To account for this structure, we set up three distinct functions for the three components: Reactant, Relator, and Product, as shown in the table below.[24]

[24] Following Martin's (2013b) notations for systemic functional description, we use words with initial capital letters to represent **functions** and those with initial letters in lowercase to represent **classess**. For example, "Reactant" is a function while "expression" is a class.

$CH_4(g) + 2O_2(g)$	→	$CO_2(g) + 2H_2O(l)$
Reactant	Relator	Product

The left and right expressions perform distinct functions and so are labeled differently as Reactant and Product. This is because the two expressions cannot be swapped without changing the meaning of the equation. If changed to equation (4:8) (an impossible equation signified by the *), the two expressions of the equation change the meaning of the original equation by swapping the beginning and final elements.

(4:8) *$CO_2(g) + 2H_2O(l) \rightarrow CH_4(g) + 2O_2(g)$

The 'Reactant ^ Relator ^ Product' configuration is a **multivariate** structure, which is in contrast to the highest level for mathematics that Doran (2018c) described in terms of univariate structure. Doran's argument is based on the fact that mathematical equations can be expanded indefinitely by iterating the number of expressions, for example, Equations (4:9), (4:10), and (4:11).

(4:9) $x=1$

(4:10) $y=x+1=2$

(4:11) $y=x+1=2=a + b$

However, in chemical equations, the iteration is not possible at the highest rank (though see Sections 4.1.1.3.2 and 4.1.1.3.4 where this does occur at lower ranks). For example, Equations (4:12) and (4:13) are not possible. For this reason, chemical equations at the highest rank are considered to be a multivariate structure.

(4:12) *$CH_4(g) + 2O_2(g) \rightarrow CO_2(g) + 2H_2O(l) \rightarrow H_2CO_3(aq)$

(4:13) *$CH_4(g) + 2O_2(g) \rightarrow CO_2(g) + 2H_2O(l) \rightarrow H_2CO_3(aq) \rightarrow H_2SO_4\,(aq)$

Besides equations with uni-directional arrows (→) shown in Equation (4:7), chemical equations can also involve bi-directional arrows (⇌), for example, Equation (4:14).

(4:14) $N_2(g) +3H_2(g) \rightleftharpoons 2NH_3(l)$

Equations with uni-directional arrows are called **irreversible equations** and those with bi-directional arrows are called "**reversible equations**" (Brown et al. 2012: 613). As indicated by its name, reversible equations present reactions that can go forward and backward, which are called "forward reactions" and "reverse reactions" (ibid.: 587). For example, Equation (4:14) represents a reversible reaction between nitrogen gas, hydrogen gas, and ammonia liquid. It indicates that there is a sequential order between the forward reaction (nitrogen gas first reacts with hydrogen gas to form ammonia liquid) and then the reverse reaction (ammonia liquid decomposes into nitrogen gas and hydrogen gas). That is, the forward reaction always begins first and then the reverse reaction. The irreversible and reversible equations form a systemic contrast in terms of their reversibility, as shown in Table 4.2.

Table 4.2 Irreversible and reversible equations

Equation type	Arrow	Example
irreversible	$\rightarrow$	$2H_2(g) + O_2(g) \rightarrow 2H_2O(l)$
reversible	$\rightleftharpoons$	$N_2(g) + 3H_2(g) \rightleftharpoons 2NH_3(l)$

Reversible equations are similar to irreversible equations in that there is a definite ordering where the left is the starting point. That is, the left expression always performs Reactant and the right expression always performs Product, which is evidenced by the fact that swapping the two expressions in, for example, Equation (4:14) changes the meaning of the original equation, as shown by Equation (4:15).

(4:15) $2NH_3(l) \rightleftharpoons N_2(g) + 3H_2(g)$

Equation (4:15) means that ammonia liquid first decomposes into nitrogen gas and hydrogen gas, which then react to form ammonia liquid. This is a different chemical reaction from the one represented by Equation (4:14), suggesting that the ordering of the expressions matters for both reversible and irreversible equations. We can thus use the 'Reactant ^ Relator ^ Product' configuration developed above to account for reversible equations, as shown in the following table.

$N_2(g) + 3H_2(g)$	$\rightleftharpoons$	$2NH_3(l)$
Reactant	Relator	Product

However, it is important to note that this does not account for the fact that the process is reversible or more directly, that both the left and right expressions perform the same function when viewed at in terms of reversibility. That is, the left expression '$N_2(g) + 3H_2(g)$' construes reactants in the forward reaction and products in the reverse reaction, and the right expression '$2NH_3(l)$' represents the product in the forward reaction and the reactant in the reverse reaction. To account for this, we can alternatively view both expressions performing Actant – the same function, as shown in the table below.

$N_2(g) + 3H_2(g)$	$\rightleftharpoons$	$2NH_3(l)$
Actant	Relator	Actant

The above analysis suggests that both structures are needed to fully account for the grammatical organization of reversible equations. This results in a two-layered conflated structure, in which Reactant and Product are conflated with Actant, as shown in the table below.

$N_2(g) + 3H_2(g)$	⇌	$2NH_3(l)$
Reactant	Relator	Product
Actant	Relator	Actant

The distinction between the two types of chemical equations can be formalized in a system known as REVERSIBILITY, shown in Figure 4.1. Reading from left to right, this system network indicates that an equation includes a Reactant, a Relator, and a Product (indicated by '+') ordered from left to right (shown by the '^': Reactant ^ Relator ^ Product). It also indicates there are two types of equations – an irreversible equation, where the Relator is lexicalized by the uni-directional arrow (→), and a reverse equation, where the Relator is lexicalized by the bi-directional arrow (⇌). In addition, the Reactants and Products in reversible equations are conflated with the function Actant, so as to offer a complementary perspective that accounts for these equations.

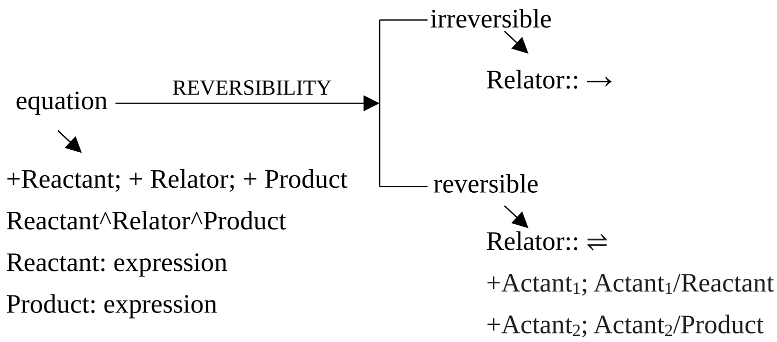


Figure 4.1 The system of REVERSIBILITY

In addition to the meaning potential of construing reversibility of chemical reactions, chemical equations can also make meaning about the conditions under which chemical reactions happen. For example, calcium carbonate decomposes under heating, and photosynthesis happens under the presence of light, as shown by Equations (4:16) and (4:17).

(4:16) $CaCO_3(s) \xrightarrow{\Delta} CaO(s) + CO_2(g)$

(4:17) $6CO_2(g) + 6H_2O(l) \xrightarrow{\text{sunlight}} C_6H_{12}O_6(g) + 6O_2(g)$

As these examples show, to indicate the conditions, chemical equations often place particular signs or words representing the conditions above the arrow. For instance, "Δ" in Equation (4:16) represents 'heating', and 'sunlight' in Equation (4:17) refers to the condition of 'under the presence of sunlight'. In contrast, some reactions happen without any particular requirements for environmental conditions. In these instances, there is no marking on the equations. For example, Equation (4:18) represents the reaction between sodium metal and liquid water, requiring no particular environmental conditions.

(4:18) $Na(s) + 2H_2O(l) \rightarrow 2NaOH(aq) + H_2(g)$

Whether or not there are conditions thus forms a systemic contrast between the two equations, as displayed more clearly in Table 4.3. Note that this contrast is not whether there are necessary conditions for any particular reaction in the field of chemistry, but rather whether these conditions are marked on the equations. In some instances, conditions may be necessary, but are not, for whatever reason, explicitly marked.

Table 4.3 Conditioned and non-conditioned chemical equations

equation types	arrows	examples
conditioned	attached with condition signs	$CaCO_3(s) \xrightarrow{\Delta} CaO(s) + CO_2(g)$
unconditioned	no such attachment	$Na(s) + 2H_2O(l) \rightarrow 2NaOH(aq) + H_2(l)$

To account for this, we can build upon the model put forward by Liu (2011: 152-153) who proposed the function Condition to cover these markings.[25] As Conditions are always attached to the arrow, the simplest means of accounting for this is by taking the Condition as conflating with the Relator. This is illustrated in the table below.

$CaCO_3(s)$	$\xrightarrow{\Delta}$	$CaO(s) + CO_2(g)$
Actant	Relator/Condition	Actant

Conditions can occur on both irreversible equations and reversible equations, illustrated by Equations (4:19) and (4:20), indicating that chemical equations can make the meaning about reversibility and conditions of chemical reactions simultaneously.

(19) $2H_2(g) + O_2(g) \rightarrow 2H_2O(l)$

(20) $H_2(g) + I_2(g) \overset{\Delta}{\rightleftharpoons} 2HI$

These meaning potentials are systematized in Figure 4.2. The curly bracket ({) indicates that chemical equations can be both reversible or irreversible and conditioned or unconditioned. If they are conditioned, a Condition function is conflated with Relator.

[25] Liu considered Conditions as a category of Circumstance developed from English. Since we are not transferring categories from language description but describing chemical equations on their own, the category of Circumstance is not needed.

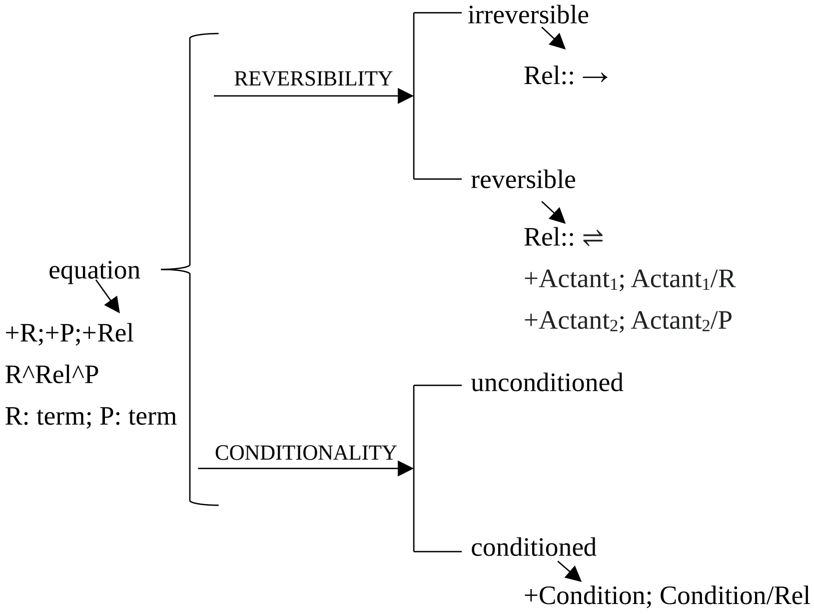


Figure 4.2 The full system at the equation rank

(To make the network concise, we use several abbreviations: R" for Reactant, "Rel" for Relator, and "P" for Product.)

Chemical equations are a relatively designed system with strictly defined possibilities for variation at the equation rank. As the description shows, the variations are explicitly afforded by the directionality of the arrows, '→' or '⇌', and whether there are conditions or not. It is important to note that the directionality of the arrows determines the structural roles of the left and right expressions, which thus does not allow for textual variations. For example, swapping the two expressions in Equation (4:21) produces a chemically wrong Equation (4:22).

(4:21) $Na(s) + 2H_2O(l) \rightarrow 2NaOH(aq) + H_2(g)$

(4:22) $*2NaOH(aq) + H_2(g) \rightarrow Na(s) + 2H_2O(l)$

Although swapping the two expressions of a reversible equation can produce a chemically correct equation, for example, Equations (4:23) and (4:24), they represent different chemical reactions.

(4:23) $N_2(g) + 3H_2(g) \rightleftharpoons 2NH_3(l)$

(4:24) $2NH_3(l) \rightleftharpoons N_2(g) + 3H_2(g)$

Equation (4:23) represents the reversible reaction that nitrogen gas first reacts with hydrogen gas to form ammonia liquid, which then decomposes into nitrogen gas and hydrogen gas. In contrast, Equation (4:24) represents a distinct reversible reaction − ammonia liquid first decomposes into nitrogen gas and hydrogen gas, which then combine to form ammonia liquid. According to Halliday and Mattheissen's (2014) description of English clauses, textual variation should have no bearing on a clause's experiential meaning. For example, the clauses *I will go fishing tomorrow* and *tomorrow I will go fishing* differ only in the point of departure, with the former being *I* and the latter being *tomorrow*. This distinction does not influence the clauses' experiential meaning. The fact that swapping the two expressions in chemical equations produces either chemically wrong equations or equations representing different chemical reactions indicates that chemical equations at the equation rank do not allow for textual variation.

Importantly for our discussion to this point, the two systems – REVERSIBILITY and CONDITIONALITY – account for all the variation at the highest rank of the grammar – the equation rank. As we will see below, despite this relative simplicity, at lower ranks there is a significant expansion in the possibilities of meaning. So it is to this we now turn.

4.1.1.3.2 Expression as term complex

As introduced above, expressions describe the elements that occur as Reactant and Product. The simplest expression consists of only one term, for example, '$2H_2O(l)$' in the right expression of Equation (4:25). It can also include two terms, for example, '$2H_2(g)$' and '$O_2(g)$' in the left expression of Equation (4:25).

(4:25) $2H_2(g) + O_2(g) \rightarrow 2H_2O(l)$

An expression may also consist of three terms, for example, '$2CaO(s)$', '$2SO_2(g)$', and '$O_2(g)$' in the right expression of Equation (4:26).

(4:26) $2CaSO_4(s) \xrightarrow{\Delta} 2CaO(s) + 2SO_2(g) + O_2(g)$

It can also involve as many as four terms, for example, '$2KCl(aq)$', '$2MnCl_2(aq)$', '$5Cl_2(g)$', and '$8H_2O(l)$' in the right expression of Equation (4:27). Theoretically, there can be no limit to the number of terms that can occur within an expression, meaning that terms in expressions are potentially indefinitely iterative.

(4:27) $2KMnO_4(aq) + 16HCl(aq) \rightarrow 2KCl(aq) + 2MnCl_2(aq) + 5Cl_2(g) + 8H_2O(l)$

In contrast to Liu (2011) viewing the plus sign ('+') as signaling reactions and taking expressions as realizing 'reactive processes' defined as "the grammatical resources to interpret material phenomena as interaction between invisible chemical entities involving the motion of valency electrons" (ibid.: 154), we take the plus sign as simply functioning to add in terms to form term complexes. There are three reasons against positing reactive processes for expressions. Firstly, if the left expression in, for example, Equation (4:28) is viewed as a reactive process, the right expression in the equation should also be taken as realizing a reactive process since they share the same structure, i.e., both 'CH_4(g) + $2O_2$(g)' and 'CO_2(g) + $2H_2O$(l)' should be taken as reactive processes. However, there is no so-called 'interaction between invisible chemical entities involving the motion of valency electrons' in the products represented by the right expression 'CO_2(g) + $2H_2O$(l)' because carbon dioxide (CO_2) and oxygen gas (H_2O) are products in the reaction.

(4:28) CH_4(g) + $2O_2$(g) → CO_2(g) + $2H_2O$(l)

Secondly, it is not clear how the concept of reactive processes can adequately explain expressions involving more than two terms, for example, the left expression in Equation (4:29). It includes three terms and two plus signs. Should it be analyzed as a complex of two reactive processes, i.e., a complex of '$2NH_3$(g)+ H_2O(l)' and 'H_2O(l) + $2CO_2$(g)'? What about '$2NH_3$(g)' and '$2CO_2$(g)'? They also involve the so-called 'motion of valency electrons' according to the definition of reactive processes.

(4:29) $2NH_3$(g) + H_2O(l) + $2CO_2$(g) → $2NH_4HCO_3$(aq)

The third is a terminology issue. Liu (2011: 156-157) noted that the terms on both sides of a plus sign are of equal status and play the same functional role in an expression. This is shown by the fact that swapping the two terms in the left expression of Equation (4:28), i.e., from 'CH_4(g) + $2O_2$(g)' to '$2O_2$ + CH_4(g)', does not change the meaning of the expression. That is, they represent the same reactants involved in the reaction. An expression can involve as many as four terms realizing the same function, as shown in the right expression of Equation (4:27). However, in SFL, the term 'process' is indicative of multivariate structures, in which the same variable can occur only once (Halliday 1981[1965]). For example, the material process involves a multivariate configuration of 'Actor ^ Process ^ Goal'. Each of the functional elements occurs only once. Calling the structure involving repetitions of the same functional elements 'process' seemly contradicts the term's definition.

All the above problems brought about by positing 'reactive processes' suggest that it is best to view the plus sign as the device for adding in terms that perform the same function. This is similar to the complexing of English words into a group, where the words can be indefinitely iterative and fulfill the same function, for example, *cats, dogs, and horses*. This structure is accounted for as univariate by Halliday (1981[1965]). In a similar vein, the possibility for indefinite repetition of terms performing the same function indicates that the grouping of terms into expression should be modeled as a univariate structure. Therefore, an expression can be viewed as a term or a term complex, suggesting that expression as a grammatical unit is not needed. The complexing of terms will be described in more detail in the next section.

4.1.1.3.3 Term

As discussed above, an expression can be either a term or a term complex. This section will first probe the variation within expressions in terms of the complexing of terms (Section 4.1.1.3.3.1) and then examine the internal organization of terms (Section 4.1.1.3.3.2). To start with, we will first focus on the complexing of terms into expressions.

4.1.1.3.3.1 The complexing of terms

As noted above, a term can occur on its own or iterate to form a term complex. We can show this variation through the tables below using the examples mentioned in Section 4.1.1.3.2.

$2H_2O(l)$
1

$2H_2(g)$	+	$O_2(g)$
1		2

$2CaO(s)$	+	$2SO_2(g)$	+	$O_2(g)$
1		2		3

$2KCl(aq)$	+	$2MnCl_2(aq)$	+	$5Cl_2(g)$	+	$8H_2O(l)$
1		2		3		4

Following Halliday (1981[1965]), a series of numbers are used to refer to the iteration of the same functional elements in paratactic univariate structures. As discussed above, akin to the ‘and’ in English nominal groups, the plus sign is the device that creates equal status among the elements and allows them to iterate. Its functional label is not shown in the above tables, but it is labeled Operator in the system network to be presented later.

In addition to the capability of iterating terms performing the same function, the ordering of the terms has no bearing on the expressions’ content meaning. For example, ‘$O_2(g) + 2H_2(g)$’ share the same meaning as ‘$2H_2(g) + O_2(g)$’. This is because expressions serve for listing reactants or products involved in chemical reactions, and the variation in the terms’ sequences does not produce ideational distinctions. All the evidence suggests that the complexing of terms into expressions is purely univariate. This univariate structure allows expressions to involve as many terms as possible.

Chemically, this univariate organization is crucial for construing reactions that may involve a large number of reactants or products. We can formalize this grammatical organization as the system of TERM TYPE shown in Figure 4.3.

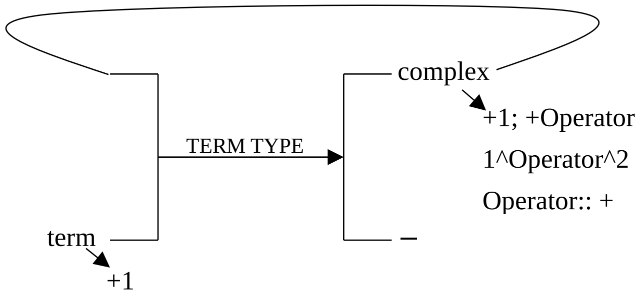


Figure 4.3 The system of TERM TYPE

The dash [–] indicates that the system is optional, i.e., iterate or not.[26] If [–] is chosen in the first round, the choice ends, producing a single term (labeled ‘1’). If [complex] is chosen, an addition of Operator (lexicalized as ‘+’) and a term (labeled ‘2’) are added to form a term complex. The wire from [complex] back into the system of TERM TYPE indicates a recursive loop, with the possibility of indefinite iteration. Choosing [complex] means that the system is entered again to determine whether a third term is to be added in the second round. If [–] is chosen, the recursion stops. If [complex] is chosen again, the recursion goes on and produces a longer term complex.

The TERM TYPE system well accounts for the complexing relation between terms. However, it is not yet clear if the ordering of terms in term complexes makes meaning. That is, we need to examine whether term complexes involve **information organization**. As discussed in Section 4.1.1.3.2, swapping terms in an expression produces a chemically correct expression, which represents the same reactants or products involved in chemical reactions as the original expression does. For example, the left expressions in Equations (4:30) and (4:31) are grammatical and represent the same reactants involved in the reaction.

(4:30) $Mg(s) + 2HCl(aq) \rightarrow MgCl_2\ (aq) + H_2(g)$

(4:31) $2HCl(aq) + Mg(s) \rightarrow MgCl_2\ (aq) + H_2(g)$

In these decontextualized examples, it seems that the ordering of terms is arbitrary and does not make meaning. However, in some texts where these equations occur, the first term in an expression appears playing the function of emphasizing and orienting to the chemical species introduced in the texts. Compare Texts 4.1 and 4.2:

[26] Note that we use the notation ‘[ ]’ to indicate system features, i.e., choices in a system.

Text 4.1 Chan et al. (2018: 357)

When **metals** react with dilute acids, there is a transfer of electrons from metal atoms to hydrogen ions. For example, when **magnesium** reacts with dilute hydrochloric acid, each **magnesium atom** loses two electrons, and two hydrogen ions from the acid gain two electrons to form hydrogen gas. The chloride ions from the hydrochloric acid are not involved in the reaction.

$$Mg(s) \rightarrow Mg^{2+}(aq) + 2e^{-}$$

$$2H^{+}(aq) + 2e^{-} \rightarrow H_2(g)$$

The ionic equation for this reaction is:

$$\mathbf{Mg(s)} + 2H^{+}(aq) \rightarrow Mg^{2+}(aq) + H_2(g)$$

The full K(overall) equation is:

$$\mathbf{Mg(s)} + 2HCl(aq) \rightarrow MgCl_2\ (aq) + H_2(g)$$

(bold fonts are my emphasis)

Text 4.2 (Chan et al. 2019: 157)

When **acids** are added to reactive metals, bubbles of hydrogen gas are released and a salt is formed. Because there is no transfer of protons, it is not a Bronsted-Lowry acid-base reaction. Instead, it is a redox reaction. Copper, silver and gold do not react with acids because their respective ions are stronger oxidizing agents than hydrogen ions so a spontaneous reaction does not occur.

The general equation for the reaction is:

acid + reactive metal → salt + hydrogen

For example, the reaction between **hydrochloric acid** and magnesium metal can be represented by the chemical equation below.

$$\mathbf{2HCl(aq)} + Mg(s) \rightarrow MgCl_2\ (aq) + H_2(g)$$

(bold fonts are my emphasis)

The above texts show that Equations (4:30) and (4:31) are used in Texts 4.1 and 4.2 respectively. Text 4.1 occurs in a chapter from the Year 11 textbook in our data introducing properties of metals. As shown by the words in bold fonts (my emphasis), metal and magnesium are emphasized throughout the text. This emphasis on magnesium metal is reflected in placing the term 'Mg(s)' first in the expression 'Mg(s)

+ 2HCl(aq)’. In contrast, Text 4.2 occurs in a chapter in the Year 12 textbook in our data introducing properties of acids. Throughout the text, ‘acids’ are the core topic, as shown by the words in bold fonts (my emphasis). The emphasis on acids is shown by placing first the term ‘2HCl(aq)’ in the expression ‘2HCl(aq) + Mg(s)’. This suggests that the first term in an expression has the function of orienting to its co-text and emphasizing the core chemical species introduced in the text. We will follow Halliday’s (1985) description of English clauses to refer to the function realized by the first term in an expression as Theme. The Theme of an expression is used to express the **thematic prominence** of the text it orients to. The remainder of the expression functions as Rheme, which has minimal effect on the textual pattern, aside from not being the Theme. The Theme-Rheme structure is illustrated in the table below.

Mg(s)	+	2HCl(aq)
Theme	Rheme	

In contrast to English clauses, the Theme-Rheme structure does not occur in all expressions. For example, the right expression of Equation (4:32) is thematized in Text 4.3 but not so in Text 4.4.

(4:32) $H_2SO_4(l) + H_2O(l) \rightarrow H_3O^+(aq) + HSO_4^-(aq)$

Text 4.3 (Chan et al. 2019: 164)

Diprotic acids, such as sulfuric acid (H_2SO_4) and carbonic acid (H_2CO_3), can donate two protons. A diprotic acid, for example sulfuric acid, dissociates in two stages.

Stage 1:

$$H_2SO_4(l) + H_2O(l) \rightarrow \mathbf{HSO_4^-(aq)} + H_3O^+(aq)$$

Sulfuric acid is described as a strong acid in water because it readily donates a proton, so this stage occurs almost to completion. Virtually no H_2SO_4 molecules are found in an aqueous solution.

Stage 2:

The **HSO_4^-** ion formed can also act as an acid. In a 1.0molL^{-1} solution, only a small proportion of those ions reacts further to produce H3O$^+$ ions and SO4^{2-} ions.

$$\mathbf{HSO_4^-(aq)} + H_2O(l) \rightleftharpoons SO_4^{2-}(aq) + H_3O^+(aq)$$

HSO_4^- is described as a **weak acid** because it is only partially dissociated. A double (reversible) arrow indicates that an incomplete reaction occurs. (you will learn more about strong acids later in this chapter.) Therefore a solution of sulfuric acid contains hydrogen ions, hydrogen sulfate ions and sulfate ions.

(bond fonts are my emphasis)

Text 4.4 (Chan et al. 2019: 151)

As you saw previously, hydrogen chloride gas (HCl) dissociates completely when it is bubbled through water, virtually no HCl molecules remain in the solution (Figure 7.1.2a). Similarly, pure HNO_3 and H_2SO_4 are covalent molecular compounds that also dissociate completely in water:

$$HCl(g) + H_2O(l) \rightarrow H_3O^+(aq) + Cl^-(aq)$$

$$H_2SO_4(l) + H_2O(l) \rightarrow H_3O^+(aq) + \mathbf{HSO_4^-}(aq)$$

$$HNO_3(l) + H_2O(l) \rightarrow H_3O^+(aq) + NO_3^-(aq)$$

The single reaction arrow (→) in each equation above indicates that the dissociation reaction is complete.

Acids that readily donate a proton are called strong acids. Strong acids donate protons easily. Therefore solutions of strong acids contain ions, with virtually no unreacted acid molecules remaining. Hydrochloric acid, sulfuric acid and nitric acid are the most common strong acids.

(bond fonts are my emphasis)

Text 4.3 develops around the core chemical species HSO_4^-, which is thematized in the expression '$HSO_4^-(aq) + H_3O^+(aq)$'. However, in Text 4.4 that focuses on introducing strong acids, the weak acid HSO_4^- is not thematized, as shown in the expression '$H_3O^+(aq) + HSO_4^-(aq)$'. This suggests that thematization is optional for expressions, which can be systemized as the network shown in Figure 4.4. The system says that an expression constituted by a term complex can be thematized or not. If thematized, it involves a Theme-Rheme structure.

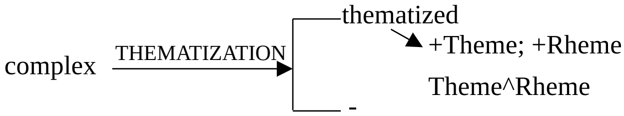


Figure 4.4 The system of THEMATIZATION

This section has discussed the complexing of terms and the information organization of term complexes. We have not yet analyzed the internal organization of terms, to which the next section will turn.

4.1.1.3.3.2 The internal organization of terms

Before probing the internal organization of terms, it is necessary to clarify the chemical equations selected from our data for the description of terms. In our data, some incomplete equations occur in early school-year textbooks when they first introduce chemical equations. For example, in the Year 8 textbook, Equation (4:33) is used to represent the reaction between hydrogen gas and oxygen gas.

(4:33) $H_2 + O_2 \rightarrow H_2O$

This incomplete equation elides coefficients and state symbols. In early schooling levels, these components are omitted to avoid learning difficulty caused by chemical equations' complexity. With the knowledge being built cumulatively, the equation is fully developed in year 10 to its complete form, i.e., Equation (4:34).

(4:34) $2H_2(g) + O_2(g) \rightarrow 2H_2O(l)$.

As pointed out by Taber (2009: 76), students are expected to progress through increasingly complex chemical equations in learning. Thus, it is understandable that the equation is simplified at the early schooling level so that students will not be overwhelmed by the manifold components within chemical equations. However, for semiotic description, an issue then arises - which equation to take as the description object. The simplified one or the complete one? Our choice is the latter. We can analogize the two equations to the language systems of children and adults. To describe the grammatical systems of English, adults' usage of English is taken as the object for description because they include the full meaning potential of the language (Halliday 1985, Martin 1992). In a similar vein, to describe the grammar of chemical equations, we must take the fully developed equations as the description object as they encode all the meaning potential of the writing system. Therefore, the incomplete chemical equations are not considered in the present description, but they are a rather crucial part of the knowledge-building analysis in Section 4.2.

Clarifying the types of equations chosen as the object of our study is crucial for the description of terms because the incomplete and complete chemical equations differ primarily in terms – the incomplete equations do not include coefficients and state symbols but the complete equations do. Since the coefficients and state symbols are two crucial components of terms, we need to focus on the complete terms that embody the full meaning potential.

A complete term is composed of three parts: coefficients, chemical formulas, and state symbols (Ebbing and Gammon 2008: 73). For example, in '$2H_2O(l)$', '2' is the coefficient that specifies the number of the molecule's units, 'H_2O' is the chemical formula that indicates the chemical species (water in this case), and '(l)' is the state symbol that indicates the physical state. '$2H_2O(l)$' thus shows that there are two units of water molecules in liquid state. The three components are always in a definite sequence – the coefficients and state symbols are on the left and right of the chemical formulas. The strict ordering does not allow rearranging the components. Placing the coefficients between the chemical formulas and state symbols, for example, '*$H_2O2(l)$', or placing the state symbols between the coefficients and chemical formulas, for example, '*$2(l)H_2O$', is chemically wrong. In addition, these components cannot iterate within a term. For instance, Examples (4:35) and (4:36) are chemically wrong.

(4:35) *$22H_2O(l)$

(4:36) *$2H_2O(l)(l)$

Example (4:35) is incorrect because there can be only one coefficient (the number '2' attached to 'H_2O') in a term. Example (4:36) is not acceptable because a term allows only a single state symbol. With the components being non-iterative and arranged in a definite sequence, terms should be best modeled as a multivariate structure. This is augmented by the evidence that the elements are in a particular semantic relation to each other. Regarding the chemical meaning of a term, the chemical formula plays a central role since both the coefficient and state symbol describe the represented chemical species' properties. As introduced above, in the example '$2H_2O(l)$', '2' specifies water molecules' quantities, and '(l)' describes water's physical state. This echoes Halliday's description of the experiential structure of English nominal groups, in which Thing is the semantic core (Halliday 1985: 167). The resonance between the internal structure of terms and English nominal groups suggests that the elements in a term are in a multivariate structural relation to each other. To account for this structure, we devise the functional configuration of 'Quantity ^ Species ^ State' for the example '$2H_2O(l)$', as shown more clearly in the table below.[27]

[27] Note that Liu (2011: 160) takes the coefficient and chemical formula as a rank-shifted clause. Drawing on O'Halloran's (2005) description of mathematical symbolisms, Liu takes, for example, '$2H_2$', as a rank-shifted operative process of the multiplication '$2\times H_2$'. In mathematics, it is true that a multiplication operator can be elided in a mathematical expression, for example, from '$2 \times a$' to '$2a$'. However, this is not the case in chemical terms because '$2\times H_2$' as a term is unacceptable in chemical equations. There is no evidence showing that '$2H_2$' is developed from "$2\times H_2$" in chemistry. Although the coefficient specifies the number of units of the chemical species, it does not necessarily mean that it involves a structure of multiplication as mathematical expressions do. This structure is not seen in the writing system of chemical terms (see Klein 2001 for an introduction to the historical development of the writing system). In contrast to Liu's approach, we do not assume the description of mathematical expressions but describe chemical terms in terms of their own grammatical organization. This is because assuming the description of other semiotic systems would inevitably homogenize the description of chemical terms.

2	H_2O	(l)
Quantity	Species	State

The example discussed above includes all the three elements. However, our data suggest that not all terms show coefficients, for example, 'H_2O(l)'. The absence of the coefficient means that the unit of the chemical species is one (Chan et al 2018: 207). 'H_2O(l)' thus means one water molecule in liquid state. For semiotic description, however, the absence of coefficients suggests a lack of the function Quantity in the structure. This is because there is no reason to posit a function if the form itself does not exist. The functional structure of 'H_2O(l)' can thus be illustrated as the following table.

H_2O	(l)
Species	State

The distinction between the two types of terms can be formalized in a system known as QUANTITY, as shown in Figure 4.5. This system shows that the minimal configuration of a term is 'Species ^ State'. It then shows that it can optionally have a Quantity (indicated by '+Quantity') if there are multiple units in the term (indicated by the feature [multiple]). Quantity is placed first in the functional structure (indicated by '#^Quantity') and realized by natural numbers, such as 2 and 3.[28]

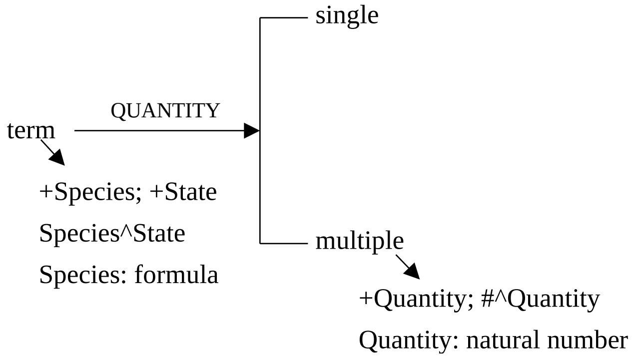


Figure 4.5 The system of QUANTITY

Terms can also be distinguished in terms of the co-occurrence of chemical formulas and state symbols. As introduced in Section 4.1.1.1, formulas within terms can be broadly categorized as molecular formulas and ionic formulas. Molecular formulas show compositions of chemical species through chemical symbols (and subscript numbers), for example, H_2O, whereas ionic formulas show compositions and

[28] Following Martin (2013b), we use the notation '#^Function' to indicate that the function is placed first in the functional structure.

electric charges of chemical species through chemical symbols, charge signs ('+' or '–') (and subscript and superscript numbers), for example, 'CO_3^{2-}'. Although they both perform the function of Species within terms, they differ in the state symbols that they can go with. In general, molecular formulas can occur with all the four types of state symbols, '(g)', '(l)', '(s)', and '(aq)', as shown in Equation (4:37).

(4:37) $Na(s) + 2H_2O(l) \rightarrow 2NaOH(aq) + H_2(g)$

In contrast, the possibility of ionic formulas occurring with state symbols is relatively restricted, as shown in Equations (4:38) and (4:39).

(4:38) $Na(s) + 2H_2O(l) \rightarrow 2Na^+(aq) + 2OH^-(aq) + H_2(l)$

(4:39) $NaCl(s) \rightarrow Na^+(g) + Cl^-(g)$

As these equations indicate, only '(aq)' and '(g)' can occur with ionic formulas. This is because, generally, ions exist as free ions only in aqueous and gas states. The distinction of the co-occurrences of chemical formulas and state symbols between the two types of terms can be systemized as the STATE system shown in Figure 4.6. The system shows that there are two types of terms: [molecular] and [ionic]. The function of Species is realized by molecular formulas in [molecular] terms and by ionic formulas in [ionic] terms. This links the formula system at the rank below term in Section 4.1.1.3.4. The system then shows that the [molecular] terms have four further subtypes: [gas], [liquid], [solid], and [aqueous], in which the function of State is lexicalized as '(g)', '(l)', '(s)', and '(aq)'. [ionic] terms, on the other hand, show only two variations: [gas] and [aqueous], where the function of State is lexicalized as '(g)' and '(aq)'.

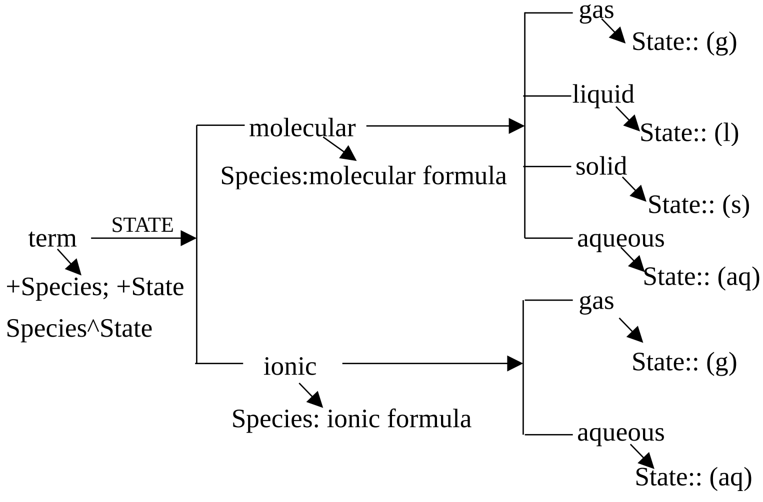


Figure 4.6 The system of STATE

The systems of QUANTITY and STATE together capture the grammatical organization of elements within terms. As the choices in each system are independent of each other, we can describe them systemically as being simultaneous. Putting together the system that accounts for the complexing relation between terms and the systems grasping the internal organization within terms produces the full system at the term rank in Figure 4.7.

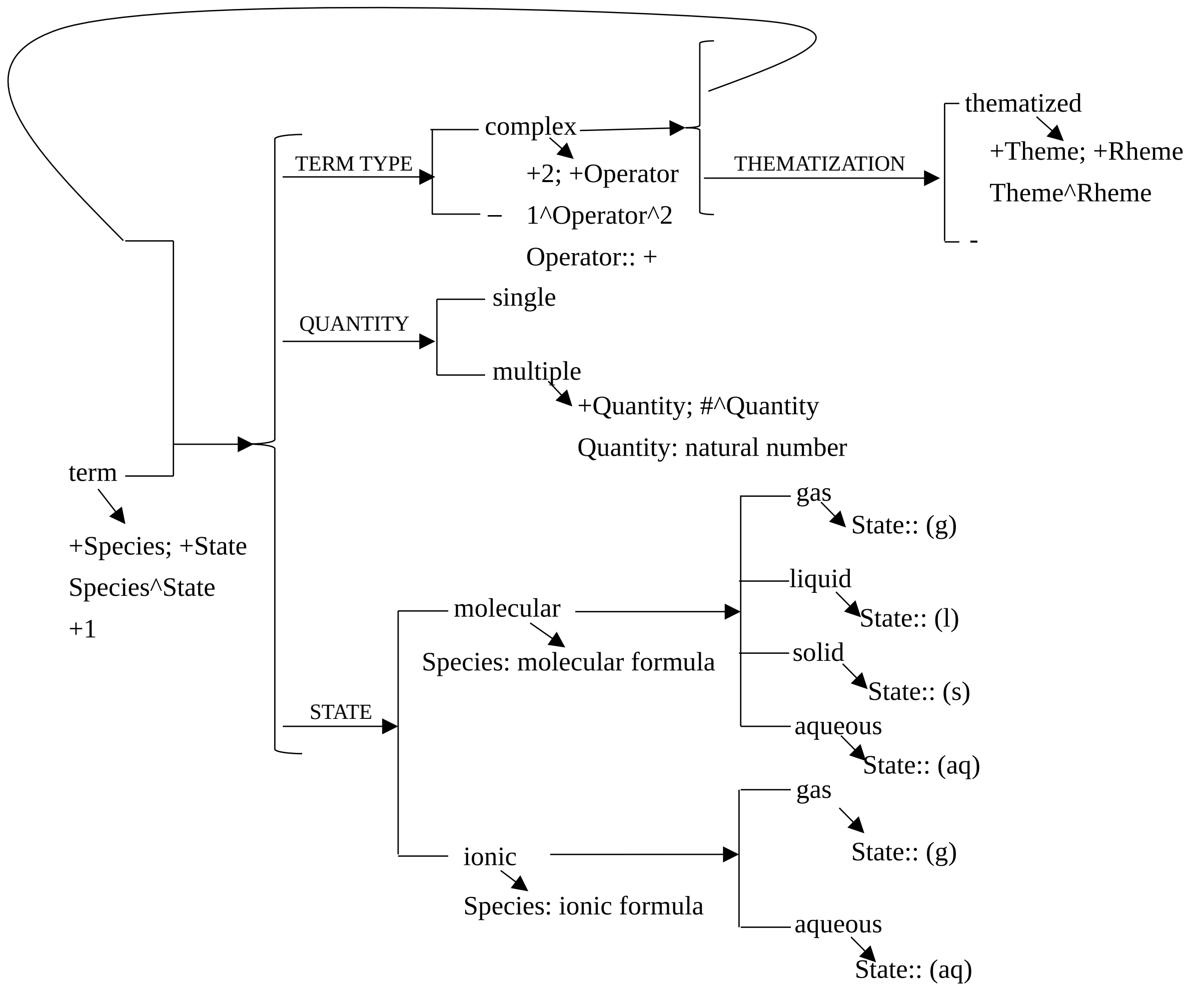


Figure 4.7 The full system at the term rank

4.1.1.3.4 Formula

Having completed the description of chemical equations at the equation and term ranks, we now move to the formula rank. To have a clearer view of where we are in the description, we bring in again the grammatical units involved in chemical equations shown in Table 4.4 (the current formula rank is highlighted in bold font). Before our description, we need to emphasize that the formula and symbol ranks are hierarchies in both chemical equations and chemical formulas, which are two distinct semiotic systems. Therefore, the description at the formula and symbol ranks constitutes both chemical equations and chemical formulas' grammar.

Table 4.4 Grammatical units involved in chemical equations

Grammatical unit	**Example**
equation	$CH_4(g) + 2O_2(g) \rightarrow CO_2(g) + 2H_2O(l)$
expression	$CO_2(g) + 2H_2O(l)$
term	$2H_2O(l)$
formula	H_2O
symbol	O or H_2

In chemistry, chemical equations typically involve three formulas: molecular formulas, empirical formulas, and ionic formulas. Molecular formulas describe the compositions of a compound through chemical symbols and subscript numbers, representing the constituent atoms and their numbers. For example, the molecular formula CO_2 shows that carbon dioxide molecules are composed of one carbon atom and two oxygen atoms. Empirical formulas, on the other hand, present only the relative number of atoms of each type in a compound. For example, Na_2O is an empirical formula, showing that sodium atom and oxygen atom exist in sodium oxide in the ratio of 2 to 1. As noted in Section 4.1.1.1, in molecular equations, empirical formulas represent ionic compounds, while molecular formulas represent single substances or molecular compounds.[29] For example, in Equation (4:40), NaOH is an empirical formula representing the ionic compound of sodium hydroxide,[30] and Na, H_2, and H_2O are molecular formulas representing two single substances – sodium metal and hydrogen gas – and one molecular compound – water molecules.

(4:40) $Na(s) + 2H_2O(l) \rightarrow 2NaOH(aq) + H_2(g)$

Semiotically speaking, molecular formulas and empirical formulas do not show any formal distinctions. They are both composed of chemical symbols and subscript numbers. The only feature that offers an implicit distinction is that molecular formulas comprise only non-metal symbols when representing compounds. For example, H and O in H_2O are non-meal symbols.[31] Empirical formulas, however, can involve both metal symbols and non-metal symbols. For example, Na is a metal symbol and H and O are non-metal symbols in the empirical formula NaOH.

[29] Single substances are chemical matter comprising only one type of chemical elements. For example, 'H_2' represents hydrogen gas, which is a single substance made of hydrogen elements only. Molecular compounds are chemical matter made of at least two types of non-metal elements. For example, 'H_2O' represents the molecular compound water, which consists of two chemical elements: hydrogen (H) and oxygen (O).

[30] Ionic compounds are chemical matter made of cations and anions. For example, 'NaCl' represents the ionic compound of sodium clorine, comprising the sodium cation and chlorine anion.

[31] Non-metal symbols are chemical symbols representing non-metal atoms. For example, H is a non-metal symbol representing hydrogen atoms.

Ionic formulas, in contrast, represent electrically charged atoms or molecules. The formulas are typically composed of chemical symbols, subscript numbers, and superscript numbers, and charge signs (‘+’ or ‘–’). For instance, the ionic formula ‘CO_3^{2-}’ shows that a carbonate anion is composed of one carbon atom and three oxygen atoms that, as a whole, bear two negative electric charges. Ionic formulas are markedly distinct from molecular formulas and empirical formulas in that they involve superscripts representing electric charges. Given this, we will describe ionic formulas separately from the other two formulas. Since molecular formulas and empirical formulas share formal similarities, we will describe them together in Section 4.1.1.3.4.1 and then ionic formulas in Section 4.1.1.3.4.2. Section 4.1.1.3.4.3 brings together these descriptions.

4.1.1.3.4.1 Molecular formulas and empirical formulas

As noted above, molecular formulas and empirical formulas are formally similar when representing chemical compounds. For example, in Equation (4:41), the molecular formula H_2O represents water molecules constituted by two hydrogen atoms and one oxygen atom, whereas the empirical formula NaOH represents an ionic compound called sodium hydroxide, composed of a sodium cation and a hydroxy anion.

(4:41) $2Na(s) + H_2O(l) \rightarrow NaOH(aq) + H_2(g)$

The two formulas are basically the same as they are constituted by chemical symbols and subscript numbers. However, the two formulas sharing similar forms but representing different types of compounds provoke concerns over whether they embody the same functional structure. Here we argue that they embody two distinct structures, which can be tested and revealed through a treatment of the formula – transforming the molecular equation to its ionic version, as shown in Equation (4:42).

(4:42) $Na(s) + H_2O(l) \rightarrow Na^+(aq) + OH^-(aq) + H_2(g)$

The empirical formula NaOH is rewritten as Na^+ and OH^- when the equation is transformed into its ionic form, whereas the molecular formula H_2O remains the same. This is because ionic compounds are in fact composed of cations and anions, but the charge signs are elided when represented in the form of empirical formulas. That is, to more explicitly denote the cation and anion an ionic compound comprises, we could use ‘Na^+OH^-’. However, such representation is not accepted in chemistry. Instead, it elides the charge signs and simply uses NaOH, turning the formula into an unmarked form. When the equation is transformed into its ionic form, the markedness is made explicit. In contrast, water molecules represented by H_2O are composed of atoms and do not involve such distinction between cation and anion. The formula thus remains the same after the transformation. This test suggests that the empirical formula involves a different configuration from molecular formulas, with the former embodying two distinct functional elements while the latter showing iteration of a single element. Such a test that makes explicit the markedness of a category is called "reactance" by Whorf (1945: 2) and the empirical formulas are known as a "**cryptotype**".[32] These cryptic formulas involve two distinct functional elements that we term Cation and Anion, as illustrated more clearly in the table below.

Na_2	O
Cation	Anion

The functional elements are distinct from each other and can occur only once, suggesting that the configuration is a multivariate structure. In terms of the realization of these functions, our data suggest that the Cation is typically realized by a single metal symbol (for example, Na in NaCl), while the Anion can be realized by either a single non-metal symbol or a complex of non-metal symbols, which will be discussed in detail at the symbol rank in Section4.1.1.3.5.

Molecular formulas representing molecular single substances and molecular compounds, in contrast, embody a univariate structure, as shown in the tables below.[33]

H_2
1

H_2	O
1	2

H	Cl	O
1	2	3

Following Halliday (1981[1965]), a series of numbers are used to refer to the iteration of the same functional element of Atom.[34] The functional element Atom can occur only once to represent a single substance, for example, H_2 (hydrogen gas), or iterate multiple times to represent molecular compounds, for example, H_2O (water) and HClO (hypochlorous acid). Note that chemical symbols and subscript numbers as a whole are given the function label 'Atom', for example, H_2. Here it is necessary to clarify why treating the two components as one unit. That is, why not the functional structure shown in the table below.

[32] According to Whorf (1945), a cryptotype is a covert category, in which the grammatical feature is unmarked. For example, the transitivity of English verbs is a cryptotype, which can be tested by a passive construction. 'Go', for example, is an intransitive verb because it cannot occur in a pasive construction like '*he is went to school'.

[33] Molecular single substances are chemical matter made of one type of non-metal elements. For example, H_2 represents hydrogen gas molecules comprising only the non-metal elements of hydrogen.

[34] Note that the iterative structure works for molecular formulas comprising non-metal chemical symbols only.

H	$_2$	O
$Atom_1$	Quantity	$Atom_2$

There are two arguments against this structural pattern. One is that this structure allows the chemical symbol and the subscript number to change independently since they perform distinct functions at the same level. It means that the chemical symbol can vary by itself, while the subscript number remains the same, which may lead to chemically wrong formulas like *C_2O (H is turned into C while the subscript number '2' remains unchanged). This indicates that the chemical symbol and subscript number should be viewed as one unit. The other argument is that confusion would arise from this pattern over which chemical symbols' quantity the subscript number describes. Since the function of Atom and Quantity are distinct functions at the same level as posited in the structure, it is not clear if the subscript number '2' describes the quantity of H or O. To avoid such confusion, it is best to view the symbol and subscript number as an intact whole at the formula rank.

A significant feature of the univariate structure is that only non-metal symbols can iterate. For example, H, Cl, and O in the formula of HClO (hypochlorous acid) are non-metal symbols. If metal symbols occur, there is only one possibility – it occurs by itself and represents a metal single substance.[35] For example, Na represents sodium metal. The metal symbol cannot iterate, and if added a non-metal symbol, the formula becomes an empirical formula representing an ionic compound. For example, Na_2O involves a metal symbol Na and a non-metal symbol O, forming a multivariate structure as analyzed above, i.e., 'Cation ^ Anion'. Therefore, a molecular formula includes the following options: (1) if a metal symbol occurs, it cannot iterate, and the formula is constituted by a single metal symbol; (2) if the formula comprises only non-metal symbols, it can iterate multiple times to form complexes, as shown in the tables below. Following Halliday (1981[1965]), a series of numbers are used to refer to the iteration of the same functional elements in paratactic univariate structures.

H_2
1

H	Cl
1	2

H	Cl	O
1	2	3

[35] Metal single substances are chemical matter made of only one type of metal elements. For example, Na represents sodium metal comprising only the metal elements of sodium.

We can formalize the above grammatical organization of molecular formulas as the MOLECULAR TYPE system shown in Figure 4.8. Reading from the left to the right, the system says that the functional element of the simplest molecular formulas is an Atom, indicated by '+Atom'. It then says that a molecular formula can be either [metallic], a molecular formula constituted by a single metal symbol, in which the function Atom is realized by a metal symbol (indicated by 'Atom: metal symbol'), or [non-metallic], a molecular formula composed of non-metal symbols, in which the function Atom is realized by non-metal symbols (denoted by 'Atom: non-metal symbol'). The system of NON-METALLIC TYPE captures the iteration of [non-metallic] molecular formulas. A [non-metallic] can occur by itself, indicated by [-], or iterate to form a complex, indicated by '+2' and '1^2'.[36] The wire from [complex] back into the system of NON-METALLIC TYPE indicates a recursive loop, with the possibility of indefinite iteration.

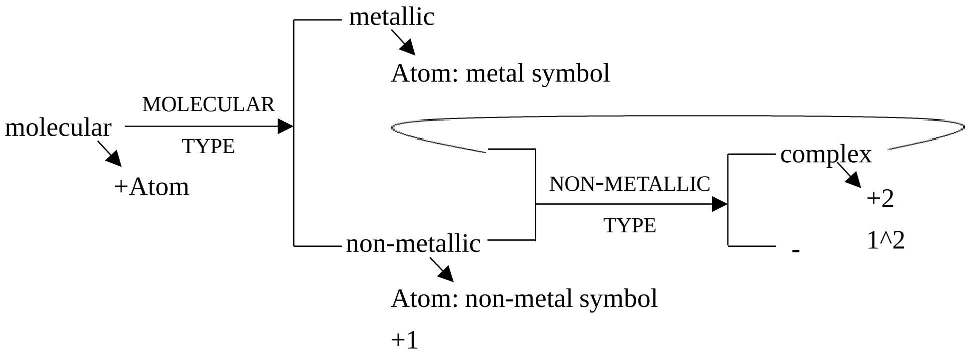


Figure 4.8 The system of MOLECULAR TYPE

Note that in the realization statements in Figure 4.8, the function Atom is realized by metal symbols in [metallic] and by non-metal symbols in [non-metallic], which leads to a distinction between chemical symbols at the symbol rank (the rank below formula). As will be shown in Section 4.1.1.3.5, chemical symbols include two subtypes: metal and non-metal symbols. The two realization statements thus link the two systems at the two different ranks, as shown in Figure 4.9.

[36] The numbers '1' and '2' are the simplified notations for the same functional element of Atom. '1^2' shows that the formula is a complex of two Atoms.

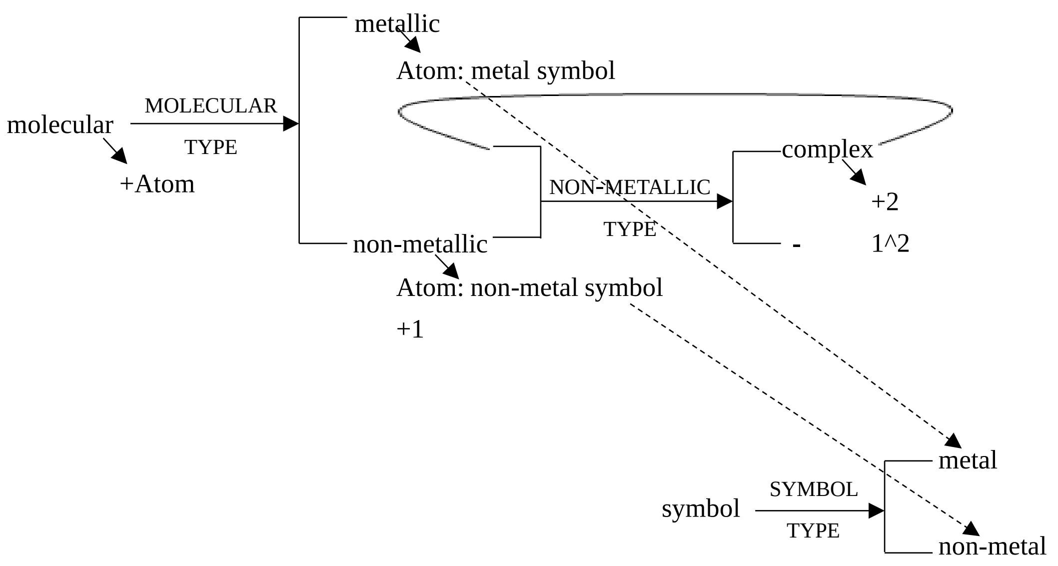


Figure 4.9 The link between the MOLECULAR TYPE system at the formula rank and the SYMBOL TYPE system at the symbol rank

Thus far, we have described the grammatical organization of molecular formulas and empirical formulas. To have an overall view of the two formulas' systemic organization, we bring together the descriptions in Figure 4.10. The two formulas are categorized as [non-ionic], as opposed to ionic formulas, which will be described in the next section.

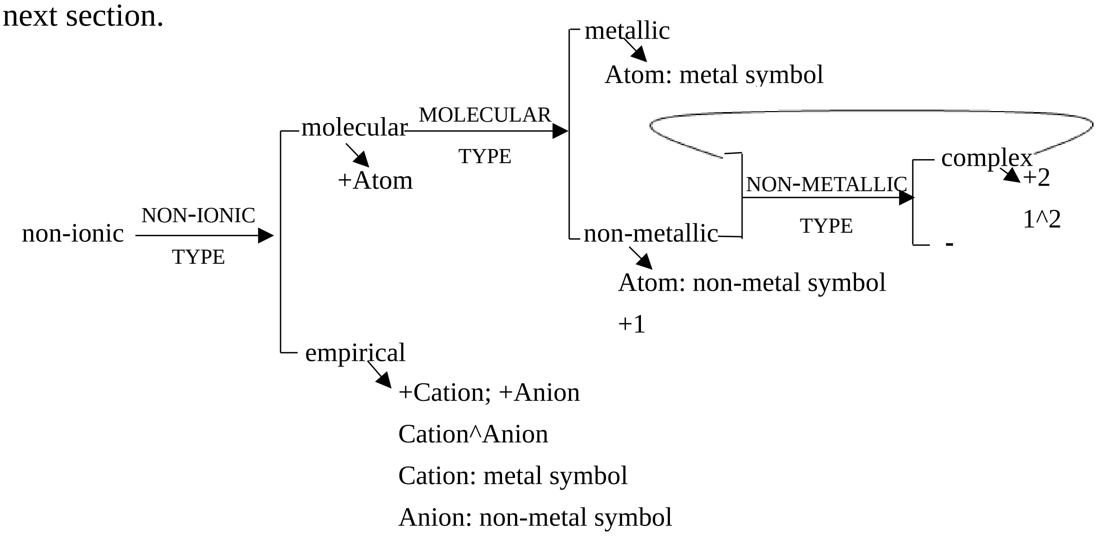


Figure 4.10 The full system of NON-IONIC TYPE at the formula rank

4.1.1.3.4.2 Ionic formulas

Ionic formulas comprise two major components: superscripts including signs representing polarities of electric charges and numbers (elided when the number is '1') representing the charge's quantities, and chemical symbols and subscript numbers (elided when the number is '1') representing particles bearing the electric charges. For example, in CO_3^{2-} (the ionic formula for carbonate anion), the superscript '2' represents the quantities of electric charges, '$^-$' denotes the charges' negative polarity, and 'CO_3' shows the particle that bearing the charges. To account for the functional structure of ionic formulas, we devise a configuration of 'Particle^Quantity^Charge', as illustrated in the table below.

CO_3	^{2}	$^-$
Particle	Quantity	Charge

The functional elements are distinct from each other and cannot iterate, suggesting that the configuration is a multivariate structure. In contrast to ionic formulas that involve numbers in the superscripts, some others elide the number when they represent particles bearing only a single electric charge. For example, OH^- represent hydroxyl ions bearing only one negative electric charge. The functional structure of this type of ionic formula is illustrated in the table below.

OH	$^-$
Particle	Charge

The distinction between the two structures can be formalized as the system of CHARGE QUANTITY shown in Figure 4.11. Reading from left to right, the system says that an ionic formula comprises two basic functional elements: Particle and Charge. For ionic formulas that embody single electric charges, the functions are configured as 'Particle ^ Charge'. In contrast, for ionic formulas construe ions bearing multiple electric charges, the function Quantity is inserted and realized by natural numbers. The three elements are configured as 'Particle ^ Quantity ^ Charge'.

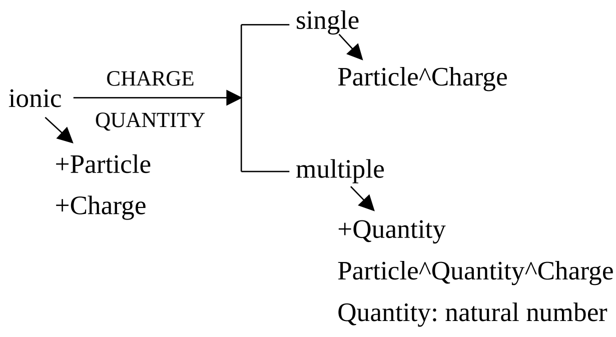


Figure 4.11 The system of CHARGE QUANTITY

In addition to the above grammatical variation, ionic formulas are also distinct in terms of the cooccurrence of chemical symbols and signs representing electric charges' polarities. In ionic formulas, metal symbols tend to occur with signs representing only positive charges ('+'), while non-metal symbols typically occur with signs representing negative charges (‘⁻’). For example, in Na^+ and Cl^-, Na is a metal symbol for sodium metal and Cl is a non-metal symbol for chlorine. This suggests that the choice of signs representing charge polarities preselects the choice of chemical symbols. The above grammatical distinction between the two types of ionic formulas can be formalized as the POLARITY system shown in Figure 4.12. It shows that when an ionic formula represents ions bearing positive charges, the function Charge is lexicalized as ‘+’, and the function Particle is realized by metal symbols. In contrast, when an ionic formula represents ions bearing negative charges, Charge is lexicalized as ‘⁻’ and Particle is realized by non-metal symbols.

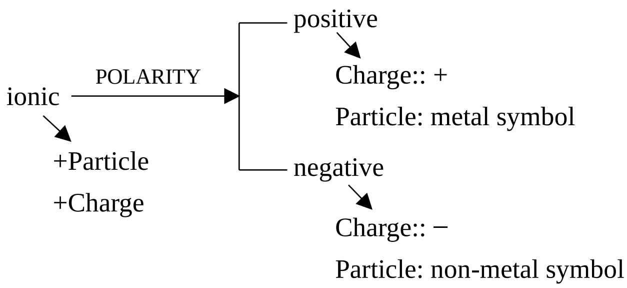


Figure 4.12 The POLARITY system

The above two systems account for all grammatical variations in ionic formulas. As the two systems describe two different aspects of ionic formulas’ organization, they form two simultaneous subsystems of ionic formulas, as shown in Figure 4.13.

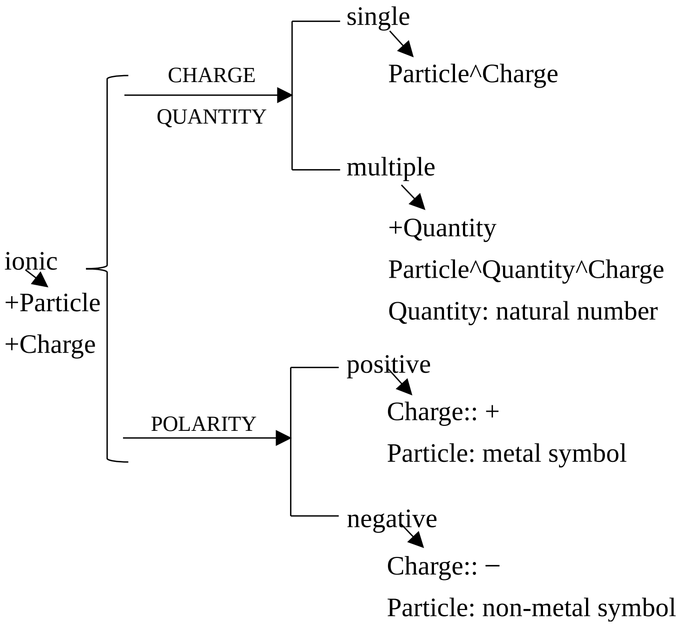


Figure 4.13 The full system of ionic formulas

4.1.1.3.4.3 The full system of chemical formulas

Sections 4.1.1.3.4.1 and 4.1.1.3.4.2 have described the grammatical systems of non-ionic formulas (molecular formulas and empirical formulas) and ionic formulas. Bringing together these systems produces the full grammatic system of chemical formulas presented in Figure 4.14.

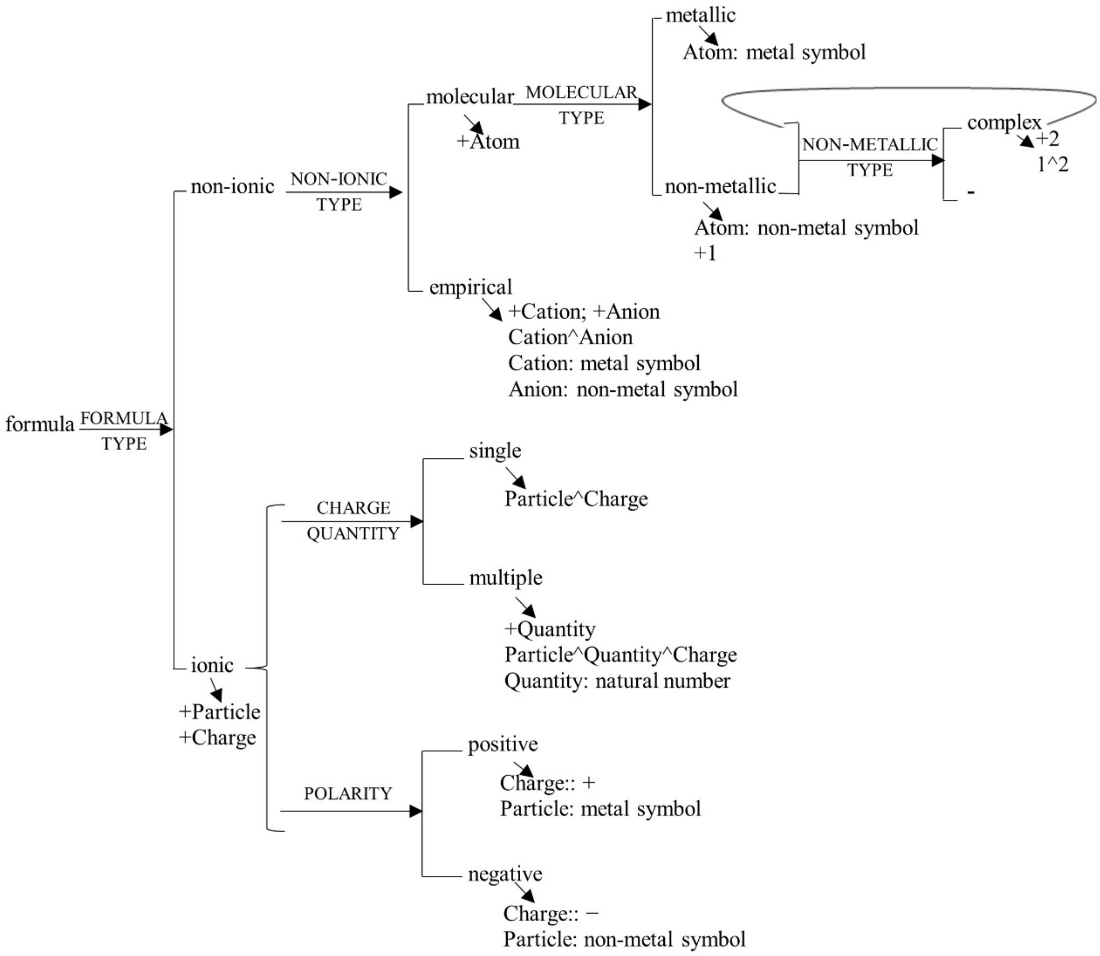


Figure 4.14 The full system at the formula rank

As shown in the grammatical system of chemical formulas, the cryptogrammatical distinction between metal and non-metal symbols and their properties of iteration are crucial for formulas because they determine most of the chemical formulas' grammatical organization: (1) in molecular formulas, metal symbols occur alone by itself and cannot iterate, while non-metal symbols can iterate to form complexes; (2) in empirical formulas representing ionic compounds, the function of Cation is typically realized by metal symbols and Anion is realized by non-metal symbols, and (3) in ionic formulas, metal symbols typically occur with the sign representing positive electric charges only, while non-metal symbols tend to occur with the sign representing negative electric charges. These grammatical distinctions between metal and non-metal symbols reflect, to some extent, the difference in chemical properties between metal and non-metal chemical elements. This is manifested more explicitly in the periodic table, where metal symbols are grouped on the left and non-metal symbols are grouped on the right, each of which shows distinct chemical properties. A key distinction in the chemical properties is that metal atoms tend to lose electrons, while non-metal atoms tend to gain electrons, meaning that the former typically becomes cations and the latter becomes anions when forming an ionic compound. This explains the second and third grammatical distinctions listed above. The grammatical distinction between metal and non-metal symbols thus grammaticalizes certain chemical properties of chemical substances.

4.1.1.3.4.4 The information organization of chemical formulas

Chemical symbols are written in a definite order in chemical formulas. For example, H is always sequenced before O in the chemical formula of water H_2O. This is because the ordering of chemical symbols follows the pattern shown in Figure 4.15.

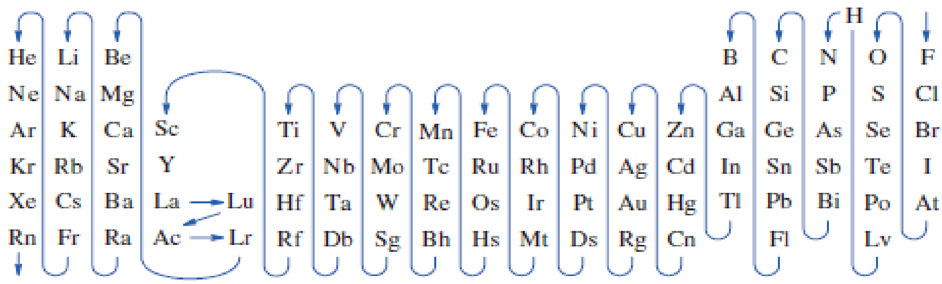


Figure 4.15 Element sequence (IUPAC 2005: 266)

Figure 4.15 presents the relative electronegativity of chemical elements represented by the symbols. Electronegativity is the ability of an atom in a compound to attract electrons to itself (Brown et al. 2012: 299). Chemical symbols representing elements with the least electronegativity are placed first, followed by those with stronger electronegativity. Typically, if symbol A is sequenced after symbol B in Figure 4.15, A comes before B in a chemical formula. For example, H being sequenced after O determines that H comes before O in the formula of water molecules H_2O. For another example, C comes before O in the formula of carbon dioxide CO_2 since C is ordered after O in Figure 4.15. Both obey the element sequence rule. This means that variations in symbol sequence are generally not allowed in chemical formulas. For instance, *OH_2 is chemically wrong and never seen written in this way. In other words, it would be viewed as ideationally unacceptable.

However, we do find one particular type of formula that partially violates the element sequence rule. They represent a chemical kind called inorganic acids. For example, HNO_3 and H_3PO_4 represent nitric acid and phosphoric acid. Although H is sequenced before N and P in the element sequence, these formulas place H first, violating the ordering principle. The motivation for putting H first lies in that substances that dissociate in water and produce hydrogen anions (H^+) are identified as acids (Chan et al. 2018: 147). The presence of H in an inorganic acid is responsible for the identity of the compound as an acid. H is thus placed first to emphasize its status. However, this does not necessarily mean that textual variation in formulas is possible because turning the above formulas into those that obey the element sequence rule is unacceptable in chemistry. For example, *NHO_3, transformed from HNO_3, is chemically wrong, though obeying the sequence rule. This indicates that the chemical formulas have strict rules for ordering symbols and do not allow textual variations.

4.1.1.3.5 Symbol

Section 4.1.1.3.4 shows that many variations at the formula rank are accounted for at the rank below, which is called symbol. This section will thus focus on the symbol rank. Symbol is the lowest grammatical unit. As discussed in Section 4.1.1.3.4.1, the unit symbol refers to either a chemical symbol by itself or a chemical symbol attached with a subscript number, for example, O or H_2 in H_2O. A symbol does not involve any internal structure if occurring by itself. For those occurring with subscript numbers, they involve a special structure. As noted in Section 4.1.1.3.4.1, a chemical symbol and its subscript number are viewed as a unit performing the function Atom at the formula rank. For example, H_2 in H_2O plays the role of Atom. The key issue to be addressed here is how to model the two elements, "H" and "$_2$", that together perform one function. To grasp this particular structure, we may gain some insight from a similar but not quite the same structure in English clauses. Consider the following material clause:

<table>
<tr><td colspan="2">Frank but not Jack</td><td>went into the room.</td></tr>
<tr><td colspan="2">Actor</td><td></td></tr>
<tr><td>α</td><td>β</td><td></td></tr>
</table>

As shown in the above table, the two elements, *Frank* and *but not Jack*, together function as Actor in the clause, sharing some similarity to the structure that "H" and "$_2$" together function as Atom in the formula H_2O. Denoted by the labels α and β, the two elements realizing Actor form a hypotactic structure, in which β is dependent on α (Halliday 1981[1965]). That is, "but not Jack" is dependent on "Frank". This is comparable to the relation between 'H' and '$_2$', in which the subscript number "$_2$" is dependent on the symbol H. The number cannot occur if the symbol is absent, but the symbol can occur without the number, for example, H in HCl. However, there is one significant difference between the two structures. The dependent element in English can iterate to form a more extended complex, as illustrated in the table below.

Frank but not Jack, but not Mike			*went into the room.*
Actor			
α	β	γ	

The dependent element in H_2, however, cannot iterate. For example, $^*H_{234}$ is impossible. This suggests that the two elements form a multivariate structure. Considering these particular features of the structure, we call this type of structure **subjacency**: a relation of non-iterative nuclear dependency (personal communication with James Martin and Yaegan Doran). Instead of using α and β, we use A to represent the nuclear functional element and B the non-iterative dependent element. Table 4.5 illustrates this structure.

Table 4.5 The subjacency structure of H_2

	H	2
function	Atom	
class	symbol	
function	A	B

The distinction between symbols that occur by themselves and those that occur with subscript numbers can be formalized as the QUANTITY system shown in Figure 4.16. The system says that if there is only one symbol, it does not involve internal structure, indicated by [single] without a realization statement. In contrast, if there are multiple symbols, it embodies a subjacency structure, indicated by [multiple] and A_B in the realization statement.

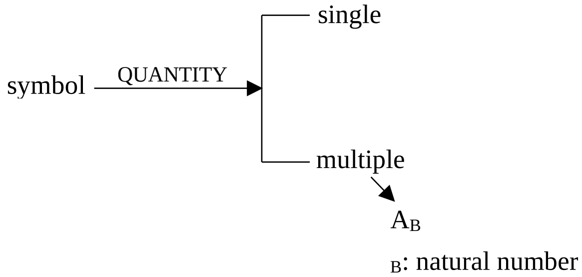


Figure 4.16 The QUANTITY system

In addition to the QUANTITY system, the system of SYMBOL TYPE described in Section 4.1.1.3.4.1 also belongs to the symbol rank. It distinguishes between metal and non-metal symbols, preselected by numerous realization statements at the formula rank. Apart from this preselection, the formula rank also preselects simple and complex non-metal symbols. For example, in empirical formulas discussed in Section 4.1.1.3.4.1, a non-metal symbol can occur by itself or iterate to form a complex to perform the function of Anion, for instance, $Na_2\underline{O}$, $Na\underline{OH}$, and $Na\underline{HCO_3}$, illustrated by the tables below.

Na_2	O
Cation	Anion
	1

Na	O	H
Cation	Anion	
	1	2

Na	H	C	O_3
Cation	Anion		
	1	2	3

These examples indicate that non-metal symbols can iterate to form complexes, performing the function of Anion at the formula rank. This grammatical organization can be formalized as the system of NON-METAL TYPE presented in Figure 4.17. It shows that a non-metal symbol can be [simple] if occurring by itself or [complex] if iterating.

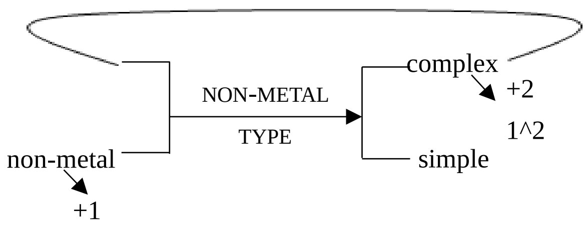


Figure 4.17 The system of NON-METAL TYPE at the symbol rank

Bringing together the systems of QUANTITY, SYMBOL TYPE, and NON-METAL TYPE produces the full system at the symbol rank shown in Figure 4.18. This system captures all the meaning potential at the symbol rank.

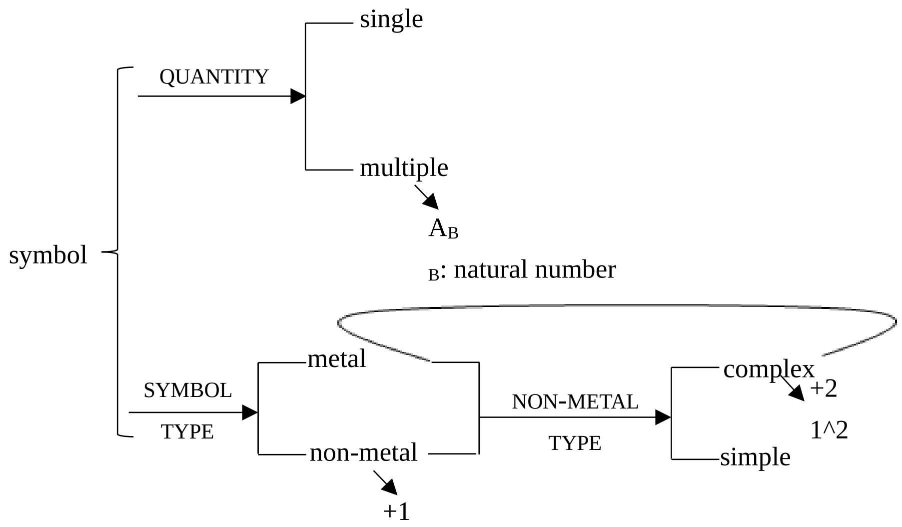


Figure 4.18 The full system at the symbol rank

4.1.1.4 Metafunctions in the grammar of chemical equations

Based on the axial relations of system and structure, we have described the grammar of chemical equations on their own terms. The axial principle will continue to form the basis of determining metafunctions in the grammar of chemical equations. As introduced in Section 3.1.1.1.2, the evidence for suggesting metafunctions lies in two aspects: relative paradigmatic independence or interdependence, and structural similarity or dissimilarity (Halliday 1969, 1970, 1979, 2002[1970]; Martin 2013b, 2015a). In other words, examining metafunctionality entails probing both the systemic and structural organization of chemical equations. If a part of the grammar shows relatively independent variation and a distinct type of structure, this gives evidence for a metafunctional component. Based on this, we will see that chemical equations are dominated by the ideational metafunction. In particular, the experiential component pervades every rank of the grammar. Besides, the grammar also organizes a logical component, though not as pervasive as the experiential component. In addition to the ideational metafunction, there is a rather limited textual component at the term rank. The following will explicate these observations in turn.

4.1.1.4.1 The experiential component

As shown in the description of chemical equations in Section 4.1.1.3, the systems at all ranks are predominantly realized by multivariate structure. At the equation rank, the REVERSIBILITY system is realized primarily by a multivariate configuration of 'Reactant ^ Relator ^ Product', and the CONDITIONALITY system by a similar structure but with an extra functional element of Condition that conflates with Relator. The system at the term rank accounts for both the internal structure of terms and their complexing into expressions. The QUANTITY system is realized by a multivariate configuration of '(Quantity^) Species ^ State', while the TERM TYPE system by a univariate structure. At the formula rank, the systems are also dominated by multivariate structures. They can be realized by a multivariate configuration of 'Cation ^ Anion' or 'Particle (^ Quantity) ^ Charge'. At the lowest rank of symbol, the system of QUANTITY is realized by a special multivariate structure of subjacency of 'A_B'. The above indicates that multivariate structure permeates systems at all ranks. The structural similarity of each of these systems, being multivariate, suggests that they could be part of a similar functional component. This is also evidenced by the fact that in the grammar of chemical equations, systems realized by multivariate structures are independent of the recursive systems that are based on univariate structure. For example, at the term rank, the QUANTITY system is independent of the TERM TYPE system. The choice of the number of terms within an expression has no bearing on the choice of quantities of chemical species. At the formula rank, the system of CHARGE QUANTITY is independent of the system of MOLECULAR TYPE. The choice of the number of symbols in a formula is independent of the choice of charge quantity. This means that the systems realized by multivariate structure are almost entirely independent of those organized univariately, suggesting that they are structurally similar but paradigmatically independent of other systems, which meets the criteria for being grouped into a distinct functional component, i.e., the experiential metafunction. The grammatical systems that constitute the experiential component at each rank are summarized as Table 4.6.

Table 4.6 The experiential metafunction in the grammar of chemical equations

Metafunction / **Rank**	**Experiential**
equation	REVERSIBILITY CONDITIONALITY
term	QUANTITY STATE
formula	NON-IONIC TYPE MOLECULAR TYPE CHARGE QUANTITY POLARITY
symbol	QUANTITY SYMBOL TYPE

The experiential metafunction permeating every rank in the grammar of chemical equations suggests that it is key to the construal of chemistry knowledge. It enables chemical equations and chemical formulas to express a range of meanings concerning chemical reactions and chemical species involved in the reactions. For example, at the formula rank, the systems of REVERSIBILITY and CONDITIONALITY allow chemical equations to construe the reversibility of chemical reactions and react conditions. At the term rank, the systems of QUANTITY and STATE construe the number of units of chemical species involved in chemical reactions and their physical states. Systems at the formula and symbol ranks are also the grammar of chemical formulas, construing the knowledge of chemical matter. The systems of NON-IONIC TYPE and MOLECULAR TYPE construe the different types of chemical matter, and the systems of CHARGE QUANTITY and POLARITY express the knowledge of electric charges ions bear. At the symbol rank, the QUANTITY system allows chemical formulas to construe quantities of atoms in chemical matter, and the system of SYMBOL TYPE distinguishes between metal and non-metal chemical elements. All these experiential meanings are crucial components of the knowledge of chemical reactions and chemical matter.

4.1.1.4.2 The logical component

Along with the systems realized through multivariate structures, several independent systems are realized by univariate structures. At the term rank, the TERM TYPE system organizes terms into term complex through a univariate structure. As noted above, it is independent of the QUANTITY system at the term rank as the choice in one system is independent of the choice in the other. At the formula rank, there is also one recursive system that complexes symbol to formula – the NON-METALLIC TYPE system. This system is also independent of other systems organized multivariately. For example, the choice from the EMPIRICAL TYPE system has no bearing on the choice from the NON-METALLIC TYPE system. At the lowest rank of symbol, the recursive system of NON-METAL TYPE is also independent of the multivariately organized system of QUANTITY. These recursive systems are paradigmatically independent at their ranks and share the structural similarity of univariate structure, meaning that they are part of a similar functional component. Being recursive and organized through a univariate structure, these systems can be categorized as the logical component, as found in English (Halliday 1985). We thus can classify these systems as part of the logical metafunction, as shown in Table 4.7. The logical component is absent at the highest rank of equation because it does not involve an iterative system organized around a univariate structure.

Table 4.7 The logical metafunction in the grammar of chemical equations

Rank \ Metafunction	Logical
term	TERM TYPE
formula	NON-METALLIC TYPE
symbol	NON-METAL TYPE

Similar to the experiential component, the logical component is also key to the construal of chemistry knowledge. It enables chemical equations and chemical formulas to construe other significant aspects of the knowledge of chemical reactions and chemical matter. At the term rank, the iterative system of TERM TYPE allows chemical equations to represent chemical reactions that include as many reactants and products as possible. On the other hand, the iterative systems at the formula and symbol ranks enable chemical formulas to represent chemical matter involving as many types of chemical element as possible. Overall, the logical component serves to construe as complex chemical reactions and chemical mater as possible.

4.1.1.4.3 The textual component

As shown in Section 4.1.1.3.3.2, only term complexes allow for textual variation, which is accounted for by the THEMATIZATION system at the term rank. It is an independent system and shares similarities to periodic structure (Martin 1996). The system thus forms a component by itself, which we refer to as the textual component as found in English (Halliday 1985). The system and the textual component it realizes are summarized in Table 4.8.

Table 4.8 The logical metafunction in the grammar of chemical equations

Rank \ Metafunction	Textual
term	THEMATIZATION

Our finding that the textual component exits only at the term rank contrasts with Liu’s (2011) finding that chemical symbols and formulas also realize this component. Liu (2011: 170-171) observed that spatial positions of numbers within chemical symbols or formulas can realize different experiential meanings. For example, the subscript and superscript numbers in ‘$_{8}O$’and ‘^{16}O’ refer to the proton number and nucleon number of an oxygen atom. Liu took these textual strategies of spatial notations as resources realizing the textual metafunction. However, it should be noted that grammatical resources using textual strategies do not necessarily mean that they realize the textual metafunction. Textual metafunction mainly describes the discursive flow and cohesion of discourse (Halliday and Matthiessen 2014: 31). The spatial positions of numbers do not appear to be textual as they serve to construe different experiential meanings (the subscript number ‘8’ in ‘$_{8}O$’ and superscript ‘16’ in ‘^{16}O’ represent the number of protons and nucleons, respectively). Such a way of determining whether the grammatical resources realize the textual meaning seemly conflicts with the fundamental principle in SFL that textual variation should be independent of variations in experiential meaning (Halliday 1985). For example, the following two clauses show textual variation but no variations in experiential meaning: *I will go fishing tomorrow* and *tomorrow I will go fishing*. The point of departure changes, i.e., *I* in the former and *tomorrow* in the latter, but this does not change the experiential meaning. The change of a number’s spatial position in a symbol, however, changes the experiential meaning, as shown in the above two examples of ‘$_{8}O$’and ‘^{16}O’. This suggests that it is better not to take spatial positions of numbers as realizing the textual metafunction. Instead, they should be considered part of the experiential component. In our description, the variations in subscript numbers within symbols and superscript numbers within ionic formulas are accounted for by the QUANTITY system at the symbol rank and the system of CHARGE QUANTITY at the formula rank, respectively, which are parts of the experiential metafunction, as shown in Table 4.7.

4.1.1.4.4 The function-rank matrix for the grammar of chemical equations

With the discussion of metafunctions in chemical equations, we have completed the description of the grammar. We can now bring together the metafunctions and ranks in a single function-rank matrix shown in the Table 4.9.

Table 4.9 The function-rank matrix for the grammar of chemical equations

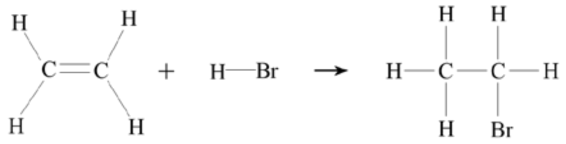
Rank \ Metafunction	Experiential	Logical	Textual
equation	REVERSIBILITY CONDITIONALITY		
term	QUANTITY STATE	TERM TYPE	THEMATIZATION
formula	NON-IONIC TYPE MOLECULAR TYPE CHARGE QUANTITY POLARITY	NON-METALLIC TYPE	
symbol	QUANTITY SYMBOL TYPE	NON-METAL TYPE	

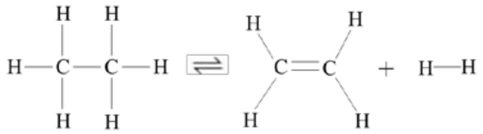
Table 4.9 shows that the experiential and logical metafuctions are predominant in the grammar of chemical equations. The experiential component permeates every rank, and the logical component occurs at every rank below the equation rank. The two components have a clear division of labor in construing chemistry knowledge. The multivariate structure realizing the experiential component enables chemical equations to construe a range of experiential meanings in relation to chemical change and chemical species. The univariate structure realizing the logical component, on the other hand, allows chemical equations to represent as complex chemical reactions and chemical matter as possible. The division of labor of the two components is closely related to the disciplinary affordance of chemistry in construing chemical reactions and chemical matter, which will be discussed in more detail in Section 4.1.3.3.

4.1.1.5 A brief note of the grammar of structural equations

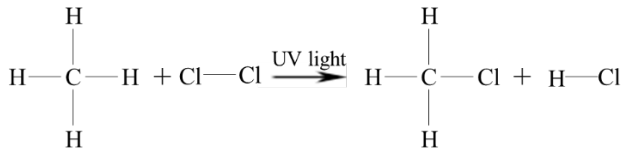
As discussed in section 4.1.1.1, structural equations are not included in the present grammatical description due to their significant structural difference from molecular and ionic equations. With the description of the two equations completed, we can now briefly examine structural equations' grammar by comparing their similarities to and differences from the other two equations.

At the equation rank, the grammatical organization of structural equations is the same as the other equations. Firstly, they also involve a systemic contrast between irreversible and reversible equations, for example, Equations (4:43) and (4:44). The two equations' functional structures are the same as those of molecular and ionic equations, i.e., a multivariate structure of 'Reactant ^ Relator ^ Product' and a two-layered structure in which Actant conflates with Reactant and Product.

(4:43) 

(4:44) 

Secondly, structural equations can also be distinguished in terms of whether a condition sign is attached to the arrow, for example, equations (4:44) and (4:45).

(4:45) 

The above shows that structural equations embody the same grammatical organization as the other two equations at the equation rank. However, structural formulas begin to diverge from the two equations at the ranks below equation. Firstly, term as a rank does not work for structural equations because they do not involve coefficients and state symbols. For structural equations, formula is the direct rank below equation. Secondly, at the formula rank, the organization of structural formulas is vastly distinct from molecular and ionic formulas. The primary distinction is that structural formulas employ two-dimensional space, or even three-dimensional, to organize meaning. At the symbol rank, chemical symbols in structural formulas do not show the subjacency structure in molecular formulas and ionic formulas. That is, it is impossible to have a subscript number attached to a symbol in structural formulas. For example, 'H_2—H' is impossible.

In summary, there are remarkable similarities in the grammatical organization between structural equations and the other two equations at the highest rank and vast differences at the formula and symbol ranks. The distinctions lie in that structural formulas whose organization is significantly different from that of molecular and ionic formulas. Therefore, it is necessary to understand how structural formulas are organized to make meaning. This is then an appropriate place to turn to the grammar of structural formulas in Section 4.1.2.

4.1.2 Grammar of structural formulas

Structural formulas are a type of chemical representation that shows how atoms are arranged and bonded to one another in molecules (Ebbing and Gammon 2008: 57). For example, the structural formula of water molecules presented in Figure 4.19 shows that the hydrogen atoms (H) are bonded to the oxygen atom (O). The lines represent single covalent bonds connecting the atoms.[37] The formula presents the structure of water molecules.

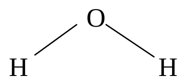


Figure 4.19 The structural formula of water molecules

A question remains contentious among chemists due to structural formulas' special organization as introduced above – are structural formulas symbolisms or images? Some argue that they are graphs as they retain resemblances to actual molecular structures (Weininger 1998), whereas others believe they are symbolic as they involve chemical symbols (Hoffmann and Laszlo 1991). The blurriness of the identification of structural formulas as a type of representation is a key question for multimodality as it attempts to understand the breadth of semiotic resources used in society. This blurriness suggests a need to explore how structural formulas make meaning. In contrast to chemists' non-semiotic view of structural formulas, our description of structural formulas in this section will show how structural formulas work as a semiotic system in their own right and how they organize their grammatical systems in ways that both similar to and distinct from images and symbolisms.

This section continues to base the description of structural formulas on the theoretical primitive of axis – the interaction between paradigmatic and syntagmatic relations (Martin 2013b) (see Section 3.1.1.1.1 for a detailed introduction of axial relations). It will first distinguish different types of structural formulas and specify the scope of this description (Section 4.1.2.1). It will then outline the grammatical units used to understand structural formulas (Section 4.1.2.2), before describing the grammatical systems themselves at each hierarchy (Section 4.1.2.3). Finally, based on the systems and structures, it will generate the metafunctions in the grammar of structural formulas (Section 4.1.2.4).

[37] Covalent bonds are pairs of shared electrons through which atoms are bonded together (Chan *et al*. 2018: 166).

4.1.2.1 Scope of description

Structural formulas are made up of three components: chemical symbols representing atoms in a molecule (for example H for hydrogen atoms), lines representing connections between atoms, and spatial organization of the symbols representing the spatial arrangement of atoms in a molecule (Goodwin 2008: 119). The structural formula of water molecules shown in Figure 4.19, for example, presents that water molecules comprise two types of atom − hydrogen and oxygen atoms − connected by two single covalent bonds (indicated by the two lines) that arrange the atoms in a diagonal position in relation to each other.[38]

Structural formulas such as the one presented in Figure 4.20 using chemical symbols and lines to represent molecular structures are called ‘Kekule formulas’ (Bruice 2017: 17), also informally referred to as ‘line-bond structural formulas’. These are the most frequently used structural formulas in secondary school chemistry (Chan et al. 2019). However, in addition to Kekule formulas, chemistry uses a range of other structural formulas designed to represent structures of organic molecules only. These formulas are developed based on Kekule formulas and, in some sense, can be seen as abbreviated forms of the Kekule formulas. One of these formulas is known as ‘condensed structural formulas’. Based on Kekule formulas, they omit some or all lines between chemical symbols (Bruice 2017: 17). For instance, both Example (4:46) and Figure 4.20 represent propane molecules.

(4:46) $CH_3CH_2CH_3$

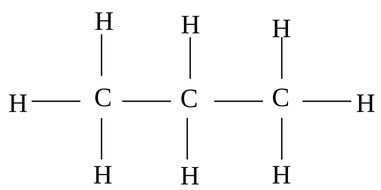


Figure 4.20 The structural formula of propane molecules

The condensed structural formula shown in Example (4:46) presents the same types of chemical symbol as Figure 4.20 and sequences them in one dimension from left to right. It does not show spatial arrangements of the symbols nor the lines representing covalent bonds between atoms. In contrast, the Kekule formula presented in Figure 4.20 displays all lines representing covalent bonds and the spatial arrangement of chemical symbols in two-dimensional space.

Another type of structural formula is called ‘skeletal structural formula’, which

38 A single covalent bond is a single pair of shared electrons between two atoms (Brown *et al.*: 2012: 189).

elides all carbon and hydrogen symbols and keeps only the lines, as shown in Figure 4.21. Each vertex in the formula denotes the elided groups of carbon and hydrogen symbols.[39]

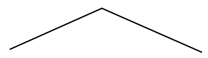


Figure 4.21 The skeletal structural formula of propane molecules

The two variations - condensed structural formulas and skeletal structural formulas - are developed primarily to save space and make it more convenient to represent structures of organic molecules (Brown et al. 2012). For the grammatical description in this section, however, we need to focus on Kekule formulas as they are most widely used, and the others are simplifications of the Kekule formulas. In the description, we will use the term ‘structural formulas’ for Kekule formulas throughout since our data shows that this is how they are referred to in secondary school chemistry.

With the scope of description demarcated, the following will build up the architecture of structural formulas’ grammar. To start with, we will first outline the grammatical units used to understand structural formulas.

4.1.2.2 Grammatical units for description

Structural formulas include two ranks: the overall formula and the symbols, at which structural formulas organize their meaning.[40] The highest rank is **formula**. It refers to the entire structure formula, which comprises chemical symbols and connectors (the lines between chemical symbols). For instance, the structural formula of hydrogen chloride molecules shown in Figure 4.22 consists of two chemical symbols, H and Cl, representing hydrogen and chlorine atoms, and one connector denoting a single covalent bond. This formula specifies hydrogen chloride molecules’ atomic composition and shows how the atoms are bonded together to form the molecules.

H—Cl

Figure 4.22 The structural formula of hydrogen chloride molecules

The rank below formula is **symbol**. It refers to the chemical symbols in structural formulas, for example, H and Cl in Figure 4.22. The connectors are not assigned a class

[39] The word ‘vertex’ here means the end points of each line.

[40] Rather than a rank scale, Doran (2018c) proposes what he calls a ‘nesting scale’ for mathematical symbolism, due to the overarching univariate organizations of the grammar. To suggest a nesting scale requires at least two obligatory hierarchies of grammatical units organized around univariate structures (Doran 2018c). As structural formulas involve only one level of univariate organization, which will be shown in Section 4.1.2.3, we simply use the term ‘rank’ scale here.

for the description below the formula rank because they involve only three variants: '—', '═', and '≡', which can be lexicalized in the systems at the formula rank (see Section 4.1.2.3.1 for a detailed discussion). Symbols, in contrast, involve relatively more complex variations and cannot be lexicalized at the formula rank, necessitating an independent rank for describing their systemic organization.

In comparison to chemical equations, the levels of grammatical units where structural formulas organize meaning are much simpler, with only two ranks - formula and symbol. The following will describe structural formulas' grammatical systems at these two ranks.

4.1.2.3 Grammatical systems of structural formulas

This section will be devoted to building up systems and structures of structural formulas. It will begin with exploring the structures of structural formulas and their systemic organization at the formula rank (Section 4.1.2.3.1) and then examine the grammatical variations at the symbol rank (Section 4.1.2.3.2).

4.1.2.3.1 Formula

As noted in Section 4.1.2.2, a minimal structural formula is composed of two chemical symbols and one connector. For example, the structural formula of hydrogen chloride molecules presented in Figure 4.23 is composed of a hydrogen symbol (H), a chlorine symbol (Cl), and a single connector (—).

H—Cl

Figure 4.23 The structural formula of hydrogen chloride molecules

The symbols represent hydrogen (H) and chlorine (Cl) atoms, and the connector denotes a single covalent bond. In structural formulas, each connector must necessarily have a symbol on both sides. In the formula shown in Figure 4.23, the H and Cl symbols on the two sides of the connector perform the same function and enjoy equal status, which is shown by the fact that swapping the two symbols does not change the represented molecular structure. That is, 'Cl—H' shows the same structure of hydrogen chloride molecules as 'H—Cl'.[41] To account for this structural organization, we term the function realized by chemical symbols **Atom**. Since the connector represents a pair of shared electrons through which atoms are bonded together (Brown et al 2012: 189), we term the function realized by the connector **Bond**. The table below shows the

[41] Note that this is unusual as chemistry does not put 'Cl' before 'H'. Here we use it to show that swapping the order of the symbols does not influence its ideational meaning.

functional structure of hydrochloric acid molecules.

H	—	Cl
Atom	Bond	Atom

Shown above is a minimal structural formula. It can be expanded through iterating connectors and symbols. For example, the structural formula of hydrogen cyanide molecules shown in Figure 4.24 involves three symbols: ‘H’, ‘C’, and ‘N’, and two connectors: ‘—’ and ‘≡’.[42]

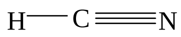


Figure 4.24 The structural formula of hydrochloric acid molecules

Akin to the structural formula of hydrochloric acid, the chemical symbols in Figure 4.24 perform the same function of Atom and there is no functional centrality in the structural formula. The following table shows the functional structure of the structural formula presented in Figure 4.24.

H	—	C	≡	N
Atom	Bond	Atom	Bond	Atom

The above analysis indicates that the expanded structural formula can be viewed as a complex of the minimal structural formula. That is, the configuration of ‘Atom ^ Bond ^ Atom’ realized by the minimal structural formula is the basic unit, upon which expanded structural formulas are developed. A special feature of this complexing relation is that it allows structural formulas to expand indefinitely on a two-dimensional space. For example, the structural formula of polyethylene molecules presented in Figure 4.25 shows that the expansion can go both vertical and horizontal.

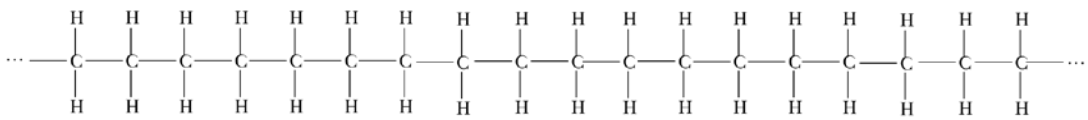


Figure 4.25 A structural formula of polyethylene molecules (Chan et al. 2019: 378)

The possibility of indefinite repetition of chemical symbols and connectors without any functional central entity suggests that connecting symbols into structural formulas is best modeled as a univariate structure. This means that each symbol has the same function as all the others, though what may differ is the bonding relations between symbols. For example, in Figure 4.24, there is a single connector between the hydrogen (H) and carbon (C) symbols but a triple connector between the carbon (C) and nitrogen

[42] This connector represents a triple covalent bond formed by three pairs of shared electrons.

(N) symbols in the structural formula of hydrogen cyanide molecules.

The above grammatical distinction between minimal structural formulas and expanded ones can be formalized as the system network shown in Figure 4.26. The system indicates that a simple structural formula includes two Atoms and one Bond, configured as 'Atom$_1$ ^ Bond ^ Atom$_2$'. It then shows that a simple structural formula can be expanded through complexing, indicated by the realization statement under [complex] – '+Bond' and '+Atom$_3$'. The wire from [complex] back to the FORMULA TYPE system indicates that the complexing can be indefinite.

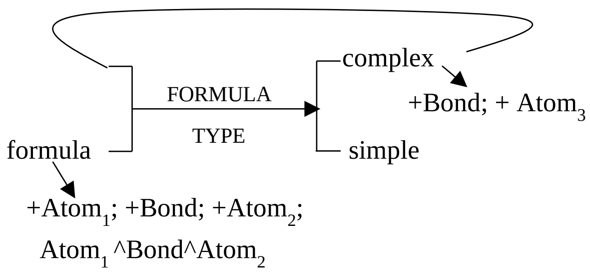


Figure 4.26 The partial FORMULA TYPE system

To this point, we have described only structural formulas that are arranged along one horizontal dimension. However, our data show that structural formulas can also be arranged along multiple spatial dimensions, as illustrated in Figure 4.27.

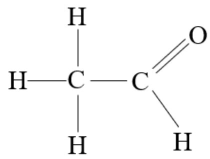


Figure 4.27 The structural formula of propanoic acid

In Figure 4.27, the chemical symbols are arranged in horizontal, vertical, and diagonal directions. To account for the spatial arrangement of chemical symbols in structural formulas, we first need to understand why the symbols are spatially arranged in relation to each other. Since structural formulas represent, to some extent, molecular structures (Brown et al. 2012: 54), understanding the meaning of the relative spatial locations of symbols is, in some sense, to see why atoms are spatially arranged in molecules. To grasp this, it is necessary to introduce the chemical theory known as 'valence-shell electron-pair repulsion theory'. According to this theory, the valence-shell electron pairs surrounding an atom form regions called 'electron domains', which repel each other to adopt an arrangement that minimizes the repulsion (Brown et al.

2012: 335).[43] For example, in the structural formula of carbon dioxide molecules shown in Figure 4.28, the two oxygen symbols are arranged horizontally on both sides of the carbon symbol because this arrangement minimizes the repulsion between the two electron domains between carbon and oxygen atoms.

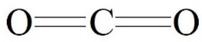


Figure 4.28 The structural formula of carbon dioxide molecules

For another example, the symbols in the structural formula of boron trifluoride shown in Figure 4.29 are arranged to form a trigonal planar shape so that the repulsion among the three electron domains can be minimized.

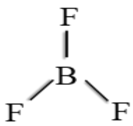


Figure 4.29 The structural formula of boron trifluoride molecules

In this sense, the spatial arrangement of symbols in relation to each other in structural formulas is meaningful and reflects the repulsion between electron domains surrounding an atom, which determines molecules' geometry. That is, relative locations of symbols in structural formulas themselves construe particular meaning and need to be described.

Structural formulas can be arranged spatially in three types of dimensions. For example, the symbols in the structural formula of carbon dioxide molecules shown in Figure 4.28 are arranged in one horizontal dimension, construing the technical meaning that the electron domains between the carbon and oxygen atoms are linearly arranged. The symbols in the structural formula of boron trifluoride molecules shown in Figure 4.29, by contrast, are arranged in two dimensions, construing planarly distributed electron domains within the molecules. The symbols in the structural formula of ammonia molecules shown in Figure 4.30 are arranged in three dimensions, construing a stereo distribution of electron domains between nitrogen and hydrogen atoms. The three dimensions are realized by a set of conventionalized types of lines: the solid line (—) represents the relative locations of symbols in the plane of page; the dashed wedge ( ) represents that the hydrogen symbol is arranged into the plane of page in relation to the nitrogen symbol; the solid wedge ( ), in contrast, denotes that the hydrogen symbol is positioned out of the page in relation the nitrogen symbol.

[43] Valence-shell electron pairs are pairs of electrons on the outer shell of an atom.

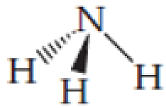


Figure 4.30 A structural formula of ammonia molecules

These examples show that each arrangement of symbols in structural formulas corresponds to a particular construal of the distribution of electron domains within a molecule. In this sense, the spatial arrangement of symbols is a special grammatical resource for construing the technical meanings. Symbols arranged in one, two, and three dimensions construe three respective distributions of electron domains within molecules: linear, planar, and stereo. Grammatically, the linear includes chemical symbols arranged in one dimension on the horizontal direction, involving Bond lexicalized as the solid horizontal line ‘—’. The planar comprises chemical symbols arranged in two dimensions, involving Bond lexicalized by the solid horizontal line ‘—’, the solid vertical line ‘｜’, and the solid diagonal line ‘／’. The stereo, by contrast, includes chemical symbols arranged in three dimensions, involving Bond lexicalized as out of the page ‘ ’ and into the page ‘ ’. To formalize these grammatical variations, we set up the system of SPATIAL ARRANGEMENT shown in Figure 4.31.

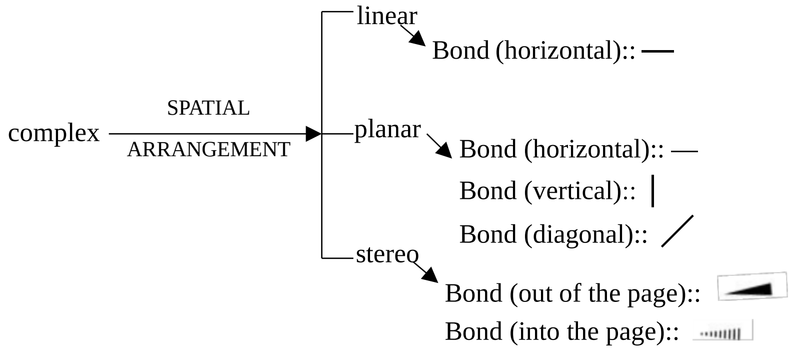


Figure 4.31 The system of SPATIAL ARRANGEMENT

As shown in Figure 4.31, the system includes three disjunctive choices: [linear], [planar], and [stereo], referring to structural formulas composed of chemical symbols arranged in one, two, and three dimensions. As noted above, they construe three different distributions of electron domains within molecules. The realization statements indicate that the dimensions of the connectors are the grammatical resources organizing the meaning. Note that this is different from the realization statement of [formula] shown in Figure 4.26, which includes the insertion of functional elements and their configuration. This is because spatial arrangement as a kind of grammatical resource does not involve any insertion of functions or their configurations. Instead, they are realized by the different types of connector displaying different dimensionalities, a type of visual grammar that organizes meaning concerning the spatial distribution of electron domains.

Bringing together the systems that describe the complexing of symbols and connectors and the spatial arrangement of symbols produces the full system of FORMULA TYPE shown in Figure 4.32. Since the entry condition of the system of SPATIAL ARRANGEMENT is [complex], this system is bundled under the feature [complex]. The iterative system of FORMULA TYPE organizes iteration of symbols, and the system of SPATIAL ARRANGEMENT specifies the arrangement of the symbols in space, construing molecules' compositions and spatial distribution of electron domains within the molecules, respectively.

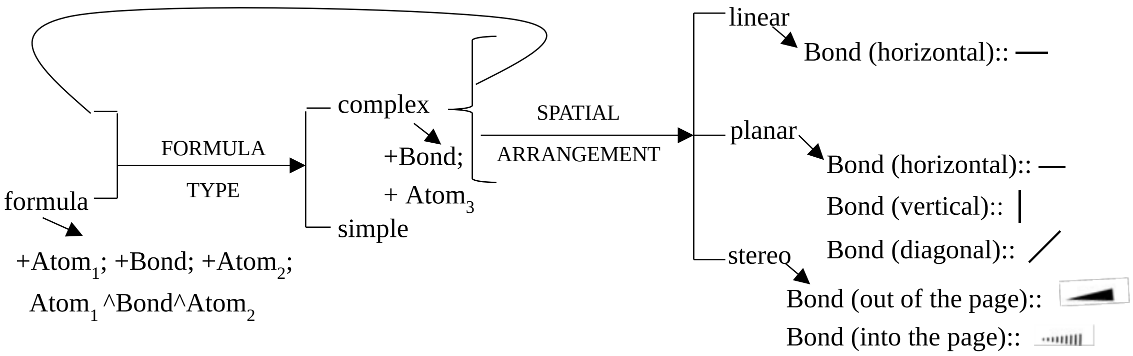


Figure 4.32 The full system of FORMULA TYPE

In addition to molecular composition and spatial arrangement of atoms, structural formulas specify another crucial descriptive content – connectivity between atoms (Goodwin: 2008: 119). In chemistry, different atoms can be bonded together through different covalent bonds according to the number of valence-shell lone electrons that each atom has (Brown et al. 2012: 297). For example, hydrogen atoms have only one valence-shell lone electron and can form only a single covalent bond with other atoms. Nitrogen atoms, in contrast, have three valence-shell lone electrons and can be bonded with three types of covalent bonds: single, double, and triple. Grammatically, these meanings are realized by the agreement between symbols and connectors. That is, the choice of connectors prescribes the choice of symbols (and vice versa). For example, the connector '≡' in the structural formula 'H—C≡N'occurs with only symbols representing atoms that bear three or more valence-shell lone electrons, i.e., 'C' and 'N' in the formula. In this sense, there is no 'free choice' of any connectors with any symbols. There is in fact a **cryptogrammar** (Whorf 1945) organizing the configuration of connectors and symbols. ‘H’ and ‘N’, for example, are not markedly distinct, but their difference is shown by the fact that the symbol ‘H’ can occur with only ‘—’ but ‘N’ with all the three types of connectors: ‘—’, ‘═’, and ‘≡’. To account for these cryptogrammatical grammatical variations, we need to build up a system that specifies the co-occurrence of symbols and connectors.

Developing this system entails first understanding what symbols and connectors can occur in structural formulas. Since structural formulas are devised for representing the arrangement of atoms in molecules (Brown et al. 2012: 68), only a small number of symbols can occur in structural formulas. These symbols represent primarily non-metallic chemical elements, including ‘H’ (hydrogen), ‘B’ (boron), ‘C’ (carbon), ‘N’ (nitrogen), ‘O’ (oxygen), ‘F’ (fluoride), ‘Si’ (silicon), ‘P’ (phosphorus), ‘S’ (sulfur), ‘Cl’ (chlorine), ‘Br’ (bromine), and ‘I’ (iodine). The number of connectors occurring in structural formulas is even smaller - only three types of connectors: ‘—’, ‘═’, and ‘≡’ (single, double, and triple connectors). According to valence-shell lone electrons that the atoms (represented by the symbols) have, different symbols can occur with different connectors, as shown in Table 4.10.

Table 4.10 Configuration of symbols and connectors

symbols	**connectors**
H, F, Cl, Br, and I	—
O and S	—
	═
B, N, P, C, and Si	—
	═
	≡

To account for the above configurations, we term the configurations that symbols occur with single connectors [single], that symbols occur with double connectors [double], and that symbols occur with triple connectors as [triple]. The configurations of symbols with connectors in structural formulas can be formalized as the system of COVALENCY shown in Figure 4.33.

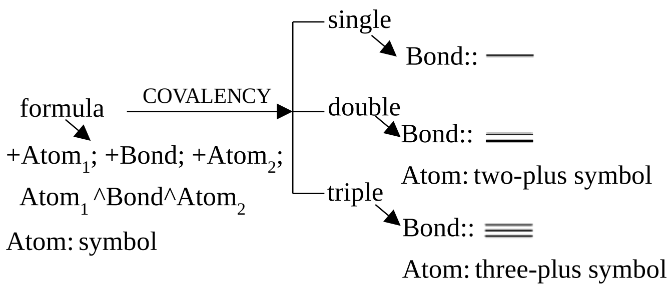


Figure 4.33 The system of covalency

Figure 4.33 shows that the system consists of three choices: [single], [double], and [triple]. Choosing [single] leads to the function Bond being lexicalized as '—', with Atom potentially being realized by any symbols. Choosing [double] means that Bond is lexicalized as '═' and Atom is realized a certain subtype of symbol that we will call 'two-plus' symbols, as shown in Table 4.11. Here two-plus denotes the number of valence-shell lone electrons that particular atoms have. Two-plus symbols represent atoms with two or more valence-shell lone electrons and thus can occur with both single and double connectors. Similarly, in the realization statement of [triple], the function Bond is lexicalized as '≡' and the Atom is realized by 'three-plus' symbols that denote atoms with three or more (usually four at most) valence-shell lone electrons, which can occur with all the three types of connector. The agreement between symbols and connectors forms a prosodic structure but **ideationalized** for construing the bonding relations between atoms, which will be discussed in more detail in Section 4.1.2.4 and Section 4.1.3.

The realization of the function of Atom by a particular set of symbols in the system of COVALENCY leads to a system at the rank of symbol (the rank below formula). This system distinguishes between different types of symbols in terms of the 'valences' an atom embodies. Here valences refer to the number of covalent bonds an atom can form with others. Table 4.11 lists the different types of symbol and their instances.

Table 4.11 Types of symbols and their instances

<table>
<tr><th colspan="3">types</th><th>symbols</th></tr>
<tr><td colspan="3">monovalent</td><td>H, F, Cl, Br, I</td></tr>
<tr><td rowspan="3">two-plus</td><td colspan="2">bivalent</td><td>O, S</td></tr>
<tr><td rowspan="2">three-plus</td><td>trivalent</td><td>B, N, P</td></tr>
<tr><td>quadrivalent</td><td>C, Si</td></tr>
</table>

As shown in the table, chemical symbols can be categorized into four types in terms of their valences – monovalent, bivalent, trivalent, and quadrivalent. 'Two-plus' includes bivalent, trivalent, and quadrivalent symbols, and ‘three-plus’ covers trivalent and quadrivalent symbols. The different types of symbols can be systematized as the network shown in Figure 4.34.

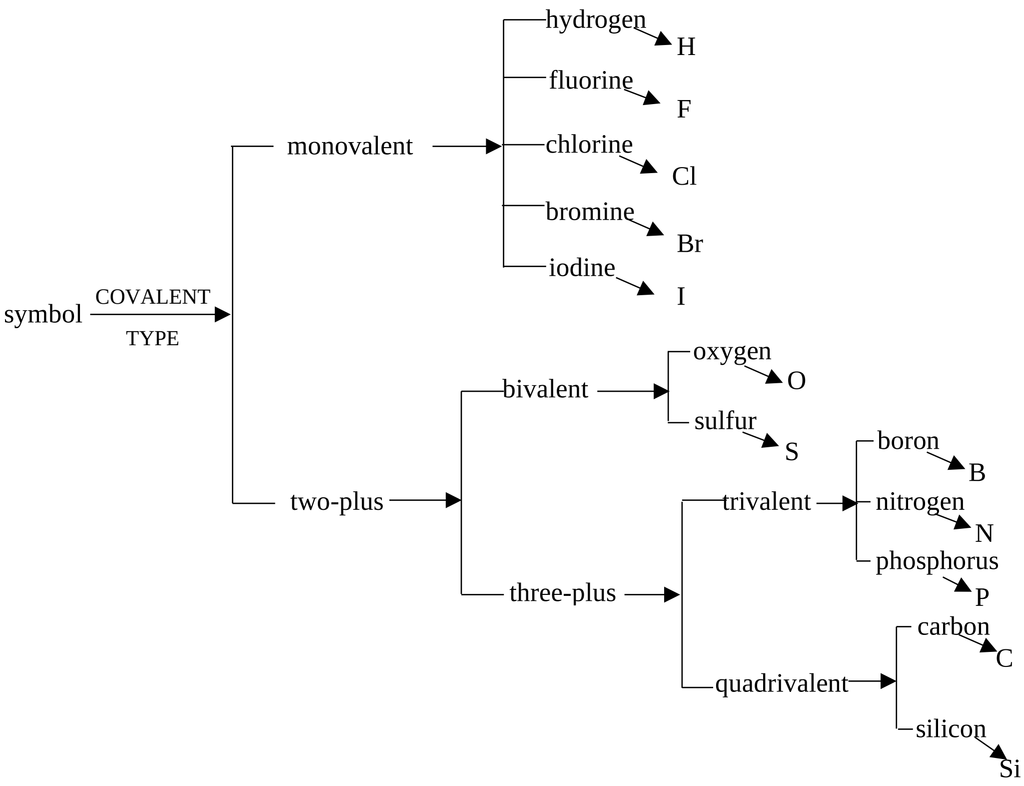


Figure 4.34 The system of COVALENT TYPE at the rank of symbol

The two system features [two-plus] and [three-plus] at the symbol rank are interconnected with the realization of the function Atom in the COVALENCY system at the formula rank, which is shown more clearly in Figure 4.35.

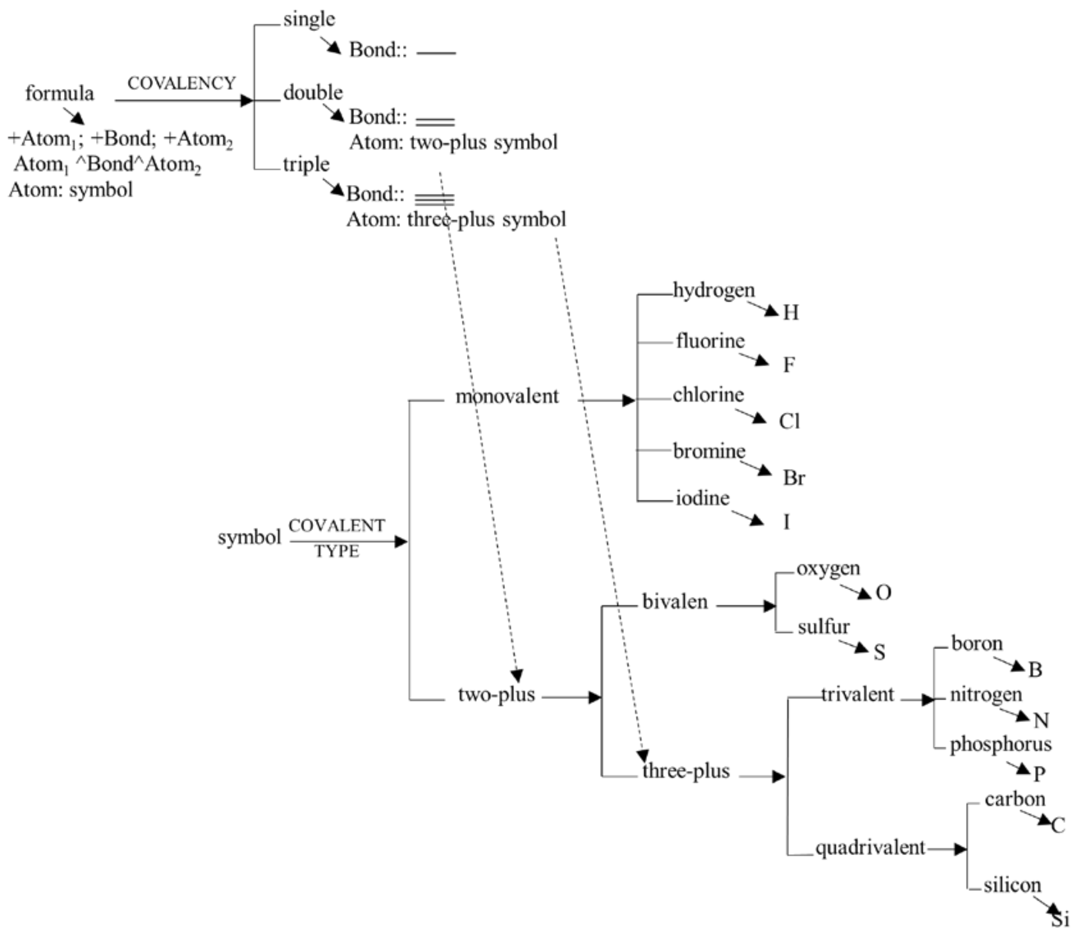


Figure 4.35 The COVALENCY system at the formula rank and the COVALENT TYPE system at the symbol rank

Note that the COVALENCY system describes the configuration of one connector with two symbols, i.e., the minimal form of structural formulas as noted above. It accounts for the connectivity between two atoms in diatomic molecules. To account for all structural formulas representing molecules with multiple atoms, this system needs to work together with the FORMULA TYPE system shown in Figure 4.26 that describes the complexing of symbols and connectors. For example, in the structural formula of hydrogen chloride 'H—Cl', one connector is configured with two monovalent symbols. We take this configuration as a simple bonding relation. [complex] structural formulas, where there are multiple connectors and more than two symbols, can be viewed as complexes of such bonding relations. For instance, the structural formula of hydrogen cyanide 'H—C≡N' is a complex of the two bonding relations: 'H—C' and 'C≡N'. This means that together with the iterative system in FORMULA TYPE, the COVALENCY system accounts for connectivity between atoms in all structural formulas.

The cryptogrammar organizing the configuration of connectors and symbols determines that chemical symbols are sequenced in certain orders in structural formulas. That is, structural formulas do not allow variations in the sequence of chemical symbols. A prominent feature of the ordering is that chemical symbols representing atoms with higher valency tend to occur in the center, and those representing atoms with lower valency are more likely to occur in the marginal positions. For example, in the structural formula of methane shown in Figure 4.36, the chemical symbol ‘C’ representing carbon atoms embodying four valencies is positioned in the center, while the symbols ‘H’ representing hydrogen atoms with only one valency are arranged at each pole.

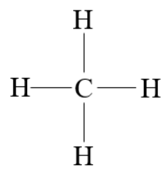


Figure 4.36 The structural formula of methane

In addition to the connectivity between atoms in molecules, structural formulas also specify molecular polarity in contexts where it needs to be made explicit.[44] The molecular polarity is indicated by attaching signs representing partial electric charges to chemical symbols. For example, the structural formula of hydrogen fluoride shown in Figure 4.37 includes hydrogen (H) and fluorine (F) symbols attached with two respective signs – ‘δ+’ and ‘δ-’, representing partial positive charge and partial negative charge, respectively. The formula shows that the hydrogen atom embodies partial positive charge, while the fluorine atom bears partial negative charge. They denote that the fluorine atom has a greater electronegativity than the hydrogen atom, indicating that hydrogen fluoride molecules are polar.[45]

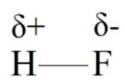


Figure 4.37 The structural formula of hydrogen fluoride molecules showing the polarity

Molecules can also be non-polar, shown by a set of signs representing different partial charges attached to a group of chemical symbols. For example, the structural formula of carbon dioxide shown in Figure 4.38 comprises two oxygen symbols attached with signs representing partial negative charge and one carbon symbol attached with a sign representing partial positive charge. This arrangement of the signs

[44] Molecular polarity refers to differences between electronegativities of atoms involved in covalent bonds in molecules (Chan *et al*. 2018: 177).

[45] ‘Electronegativity’ is a measure of the tendency of an atom to attract a shared pair of electrons (Chan *et al*. 2018: 176).

denoting partial charges is symmetrical, indicating that the whole molecule is non-polar. Note that the arrangement of the set of signs in the formula forms a prosodic structure ideationalized for construing molecular polarity, which will be discussed in more detail in Section 4.1.2.4 and Section 4.1.3.

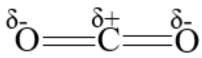


Figure 4.38 The structural formula of carbon dioxide molecules showing the polarity

Grammatically, the meaning of molecular polarity is realized by symbols attached with signs denoting partial charge, suggesting a preselection of symbols at the formula rank. As will be shown in Section 4.1.2.3.2, these symbols are called ‘charged symbols’. Symbols can also occur without the partial charge signs because the descriptive content of molecular polarity is optional for structural formulas. That is, structural formulas show molecular polarity only in contexts where it needs to be denoted. We can formalize the above grammatical organization through an optional system presented in Figure 4.39. It shows that when molecular polarity is displayed, the function Atom is realized by charged symbols. The feature [-] indicates the system is optional.

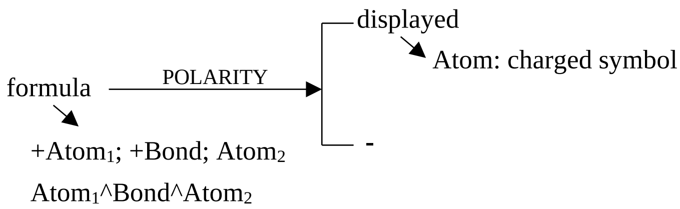


Figure 4.39 The system of POLARITY

The systems of FORMULA TYPE, SPATIAL ARRANGEMENT, COVALENCY, and POLARITY describe all grammatical variations of structural formulas suggested by our data. Bringing them together produces the full system at the formula rank, as shown in Figure 4.40. Note that our data suggest that triple covalent bonds occur in only linear structural formulas. This is indicated by the ‘I/T’ notation in the system. That is, if [triple] is chosen, then [linear] must be chosen.

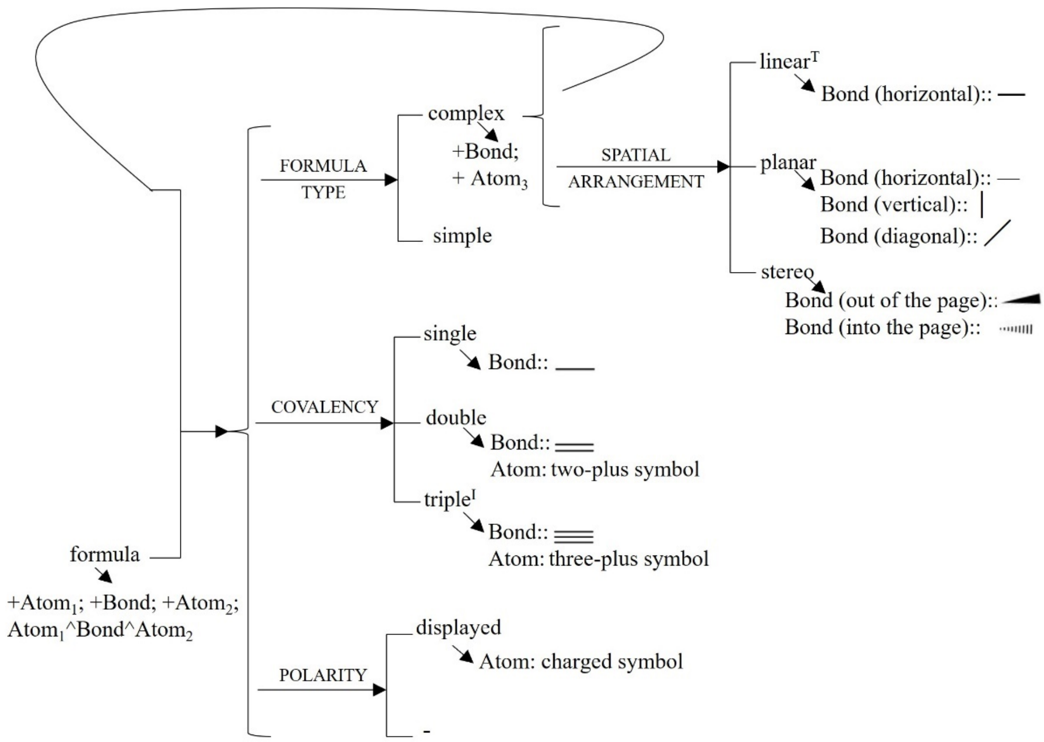


Figure 4.40 The full system at the formula rank

The FORMULA TYPE system accounts for the complexing of symbols and connectors, the SPATIAL ARRANGEMENT system describes the spatial arrangement of atoms within molecules, and the COVALENCY system accounts for the connectivity between atoms in molecules. The optional system of POLARITY, on the other hand, describes molecular polarity in contexts where it needs to be denoted. Together they describe the whole meaning potential of structural formulas at the formula rank.

4.1.2.3.2 Symbol

The rank of symbol is the grammatical unit below formula and the lowest rank in the grammar of structural formulas. Since structural formulas represent molecular structure, symbols refer to only the chemical symbols representing non-metallic elements as discussed in Section 4.1.2.3.1. Usually, they occur by themselves in structural formulas and do not involve internal structure. However, as discussed in Section 4.1.2.3.1, in contexts where molecular polarity is represented, structural formulas construe this meaning by assigning two signs representing partial charges to chemical symbols. For example, the signs ‘δ+’ and ‘δ-’ attached to the symbols H and F in Figure 4.41 symbolize partial positive and negative electronic charges, respectively.

$$\overset{\delta+}{H}—\overset{\delta-}{F}$$

Figure 4.41 The structural formula of hydrogen fluoride molecules showing partial charges each atom bears

For symbols occurring with signs representing partial charges, they involve the same structure as the symbols attached with subscript numbers in chemical formulas discussed in Section 4.1.1.3.5 – the subjacency structure. The symbols and the signs representing partial charges together perform the function of Atom at the formula rank. For example, ‘H’ and ‘δ+’ in Figure 4.41, as a whole, perform the function of Atom. This structure is similar but not quite the same as the structure in English clauses shown in the following table.

Frank but not Jack		*went into the room.*
Actor		
α	β	

As shown in the table, the two elements, *Frank* and *but not Jack*, together function as Actor in the clause, sharing some similarity to the structure that ‘H’ and ‘δ+’ together function as Atom at the formula rank. Denoted by the labels α and β, the two elements realizing Actor form a hypotactic structure, in which β is dependent on α (Halliday 1981[1965]). That is, *but not Jack* is dependent on *Frank*. This is comparable to the relation between ‘H’ and ‘δ+’, in which the superscript “δ+” is dependent on the symbol H. The superscript cannot occur if the symbol is absent, but the symbol can occur without the number, for example, H in ‘H—F’. However, there is one significant difference between the two structures. The dependent element in English can iterate to form a more extended complex, as illustrated in the table below.

Frank but not Jack, but not Mike			*went into the room.*
Actor			
α	β	γ	

The dependent element in ‘$^{\delta+}$H’, in contrast, cannot iterate. For example, ‘$^{\delta+\delta+}$H’ is impossible. This suggests that the two elements form a multivariate structure, although together they perform one function at the formula rank. Here we call this type of structure **subjacency**: a relation of non-iterative nuclear dependency. We use A to represent the nuclear element and B for the non-iterative dependent element. Table 4.12

illustrates this structure.

Table 4.12The subjacency structure of ‘$^{\delta+}$H’

	$\delta+$	H
function	Atom	
class	symbol	
function	B	A

Those with charges can either be positive charges lexicalized as ‘δ+’ or the negative charges lexicalized as ‘δ-’. The above grammatical variations can be formalized as the system shown in Figure 4.42. It indicates that if a symbol represents atoms that do not embody partial electric charges, it does not involve any internal structure; if it does, it exhibits a subjacency structure, indicated by the realization statement ‘^{B}A’. If there is a charge, this may be positive, in which case it is marked by ‘δ+’, or negative, in which case it is marked by ‘δ-’.

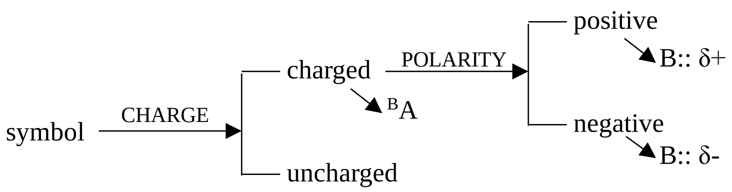


Figure 4.42 The full system of CHARGE

In addition to the CHARGE system, the symbol rank includes a system describing the different types of non-metallic symbols in terms of their valences, which connects to the COVALENCY system at the formula rank. This system has already been described and presented in Figure 4.34. Bringing them together produces the full system at the symbol rank, as shown in Figure 4.43. Here the curly bracket ‘{‘ indicates that the systems are simultaneous – charge can occur for any types of symbol.

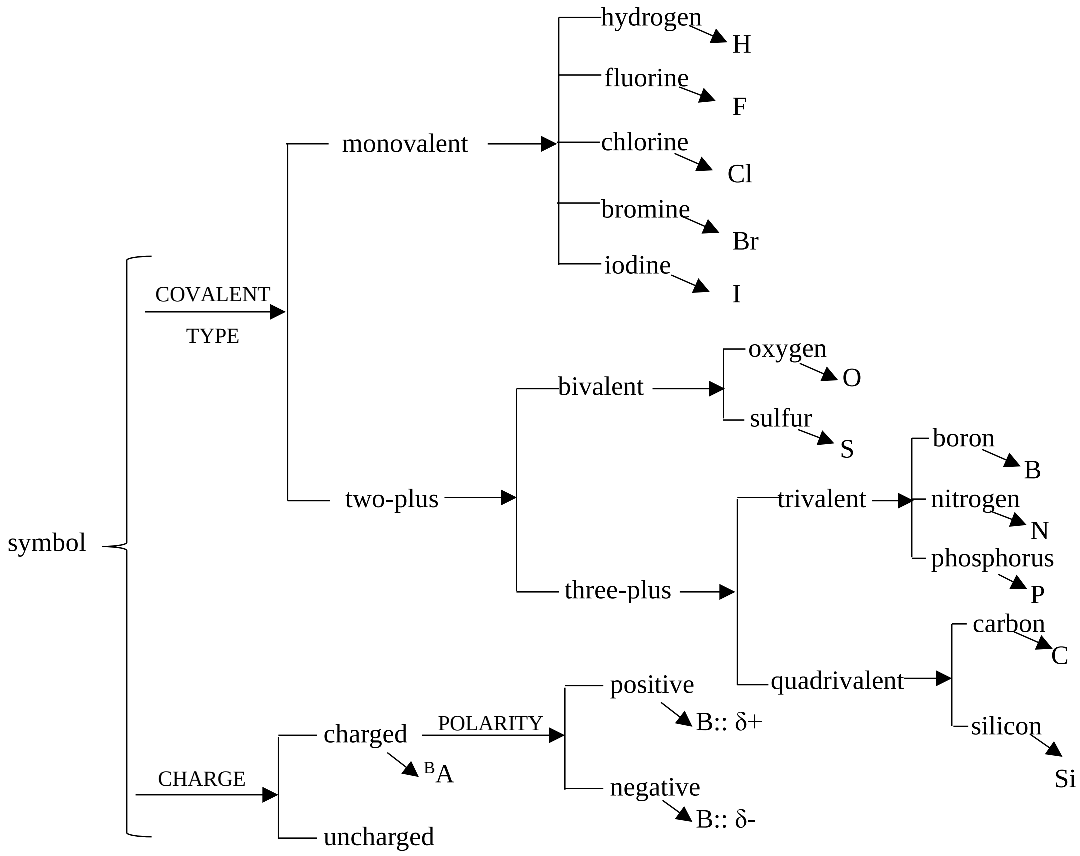


Figure 4.43 The full system at the symbol rank

4.1.2.4 Metafunctions in the grammar of structural formulas

Based on the axial relations of system and structure, we have completed the description of the grammar of structural formulas on their own terms. This axial principle will continue to be drawn as the basis for suggesting metafunctions in the grammar of structural formulas. As introduced in Sections 3.1.1.1.2 and 3.2.1, the evidence for suggesting metafunctions lies in two aspects: relative paradigmatic independence or interdependence, and structural similarity or dissimilarity (Halliday 1969, 1970, 1978, 1979, 2002[1970]; Martin 2013b, 2015a). In other words, examining metafunctionality entails probing both structural formulas' systemic variations and structural realizations. If a part of the grammar shows relatively independent variation and a distinct type of structure, that part should be considered a metafunctional component. Based on this, we will see that structural formulas are dominated by ideational metafunction. In particular, the experiential component pervades both the formula and symbol ranks. In addition, the grammar at the formula rank organizes a logical component. The following will discuss these observations in turn.

As shown in Section 4.1.2.3.1, the system of FORMULA TYPE is realized by a univariate structure that organizes the iteration of symbols and connectors in [complex] structural formulas. It is an independent system realized by a distinct structure from other systems, suggesting that it should be considered a distinct metafunctional component. Being recursive and realized by a univariate structure, the system can be categorized as the logical component found in English. Since the system of SPATIAL ARRANGEMENT depends on the iterative choice being chosen in the FORMULA TYPE system and only specifies the particular lexical realizations within the univariate structure in [complex], it should be taken as the logical component as well. In contrast to this iterative system organized by univariate structure, the CHARGE system at the symbol rank is non-iterative and realized by multivariate structures, indicating that it should be grouped into a distinct functional component, the experiential metafunction.

The COVALENCY system at the formula rank is independent and organized by a special structure – prosodic structure. In English clauses, prosodic structure organizes interpersonal meanings that cut across units (Halliday 1979). For example, the clause complex *if you don't get no publicity, you don't get no people at the fight* (cf. Martin 1996: 43) involves negation as a prosody, diffusing across the two clauses. In structural formulas, the agreement between connectors and symbols also involves meaning cutting across multiple units. For example, in the structural formula of ethyne molecules shown in Figure 4.44, there is a prosody across the formula that organizes the types of covalent bonds between atoms, i.e., single - triple - single.

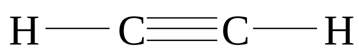


Figure 4.44 The structural formula of ethyne molecules

However, this prosodic structure is not associated with interpersonal meaning as English does. Instead, they are **ideationalized** because, looking from above, it construes field meaning of bonding relations between atoms in molecules and does not enact interpersonal relations through realizing typical speech functions, appraisal, involvement, etc. This is similar to the prosodic structure in chemical equations, where the quantities of chemical species realized by coefficients cut across equations. For example, in the equation '$2KMnO_4(aq) + 16HCl(aq) \rightarrow 2KCl(aq) + 2MnCl_2(aq) + 5Cl_2(g) + 8H_2O(l)$', the different coefficients, 2, 16, 2, 2, 5 and 8, diffuse across the equation, forming a prosody that construes different quantities of chemical species.

Similarly, the POLARITY system at the formula rank is also realized by a prosodic structure. As shown in Section 4.1.2.3.1, a set of signs representing partial charges attached to chemical symbols constitute a prosody cutting across the formula, for example, the structural formula of carbon dioxide shown in Figure 4.45. ‘δ-’ denotes partial negative charge the oxygen atoms bear and ‘δ+’ indicates that the carbon atom embodies partial positive charge. The negative and positive charges diffuse across the formula, constituting a prosody that construes the polarity of carbon dioxide molecules.

$$\overset{\delta-}{O}=\overset{\delta+}{C}=\overset{\delta-}{O}$$

Figure 4.45 The structural formula of carbon dioxide molecules showing the polarity

The above indicates that both the prosodic structures in structural formulas and chemical equations are not associated with interpersonal meaning, but instead, they are ideationalized. The backgrounding of interpersonal meaning has enabled structural formulas to employ typically interpersonal prosodic structures to realize experiential meaning. Based on this, we group the systems of COVALENCY and POLARITY as the experiential component. Since the COVALENT TYPE system at the symbol rank is interrelated with the COVALENCY system, i.e., the two systems are paradigmatically dependent, they should be grouped into the same functional component. That is, the COVALENT TYPE system should also be categorized as the experiential component.

With the discussion of metafunctions in structural formulas, we have completed the description of the grammar. We can now bring together the metafunctions and ranks in a single function-rank matrix shown in Table 4.13.

Table 4.13 The function-rank matrix for the grammar of structural formulas

Metafunction / **Rank**	**Experiential**	**Logical**
formula	COVALENCY POLARITY	FORMULA TYPE SPATIAL ARRANGEMENT
symbol	COVALENT TYPE CHARGE	

Table 4.13 shows that structural formulas are an independent semiotic system with their own meaning potential and ways of organizing meaning. Although the grammar involves spatial layout (accounted for by the system of SPATIAL ARRANGEMENT) associated with mode and aligned with the two-dimensional layout of images, it includes numerous other grammatical systems construing field meanings concerning molecules. The label ‘image’ is thus too general and common-sense for structural formulas that their meaning-making cannot be appropriately articulated. This then

answers the question raised at the beginning of Section 4.1.2 – structural formulas are a semiotic system by themselves, and image as a term for this semiotic system is too vague and general.

As shown in Table 4.13, the meaning potential includes two major components – experiential and logical. These two components exhibit a clear division of labor for construing chemistry knowledge. The logical component realized by univariate structures enables structural formulas to represent molecular structures that can be indefinitely large and show as many bonding relations between atoms as possible. The experiential components realized by multivariate structures, on the other hand, allow structural formulas to construe the spatial arrangement of atoms in molecules and partial electric charges the atoms bear. In addition, the ideationalized prosodic structures that realize the systems of COVALENCY and POLARITY enable structural formulas to show different types of covalent bonds between atoms and molecular polarity. The division of labor of the two functional components is closely related to the descriptive content required by the chemistry discipline, i.e., molecular compositions, spatial arrangement of atoms, the connectivity between the atoms, and molecular polarity (Goodwin 2008: 119).

Thus far, we have described the systemic functional grammar of chemical formulas, chemical equations, and structural formulas and grasped how the different chemical formalisms organize their meaning. However, it is not yet clear that why chemistry needs such diverse formalisms to organize its knowledge. To see this, we need to understand the similarities and differences between the formalisms in their functionalities for construing the particular technical knowledge that chemistry needs. To achieve this, we will establish a **functional semiotic typology** for the chemical formalisms to compare them based on both their internal grammatical organization and meanings they construe, which is the issue the next section will focus on.

4.1.3 A functional semiotic typology of chemical formalisms

Chemical formalisms are characterized by their sheer diversity. As we have seen in Sections 4.1.1 and 4.1.2, they include chemical formulas, chemical equations, and structural formulas. To understand why secondary school chemistry employs such diverse formalisms to construe its knowledge, we need to grasp their similarities and differences in making meaning. A fruitful avenue to this is to develop a “functional semiotic typology” (Doran 2019) for chemical formalisms. A functional semiotic typology compares semiotic systems in terms of their similarities and differences in making meaning. A key principle for the comparison is that the description of the meaning-making should be based on the semiotic systems' internal grammatical organizations, rather than assuming the description of English (Doran 2019: 355). This is because it is impossible to develop a functional typology for semiotic systems if we already assume they share similar functionalities and grammatical organizations as English. Our description of chemical formulas, chemical equations, and structural formulas in Sections 4.1.1 and 4.1.2 constitutes a solid basis for establishing a semiotic typology for the formalisms since it does not assume the description of English and the secondary categories such as metafunction and rank, but instead, it bases the description on the more fundamental theoretical primitive of axis. This section will thus take a step toward developing a functional typology for chemical formalisms based on the descriptions completed in Sections 4.1.1 and 4.1.2.

To achieve this goal, Section 4.1.3.1 analyzes the predominant types of structure (Halliday 1981[1965]) the formalisms show in their grammar and the meaning they realize. Section 4.1.3.2 establishes a functional semiotic typology for the chemical formalisms based on the analysis in Section 4.1.3.1. Finally, Section 4.1.3.3 discusses the semiotic typology in terms of chemistry's disciplinary affordance.

4.1.3.1 Grammatical organization and meaning realized

To develop a functional typology for a set of semiotic systems, we need to consider the similarities and differences in their grammatical organization and meaning realized (Doran 2019). As shown in Sections 4.1.1 and 4.1.2, the three chemical formalisms are similar in that they are dominated by the ideational metafunction. However, the specific ideational meanings construed by each formalism vary, and the grammatical structures realizing these meanings may be similar or distinct from one to another. To grasp the similarities and differences in their grammatical structures and the ideational meanings realized, we examine the predominant types of structure (Halliday 1981[1965]) that organize chemical formalisms and the types of field-specific meanings realized (Doran and Martin 2021). The following will go through the three chemical formalisms in terms of the two aspects.

4.1.3.1.1 Chemical formulas

According to the grammatical description in Section 4.1.1.3.4, chemical formulas, including molecular formulas and ionic formulas, are dominated by univariate structures. As shown in the grammatical description, iterative systems prevail through different ranks. They allow chemical formulas to add in multiple chemical symbols to represent chemical compounds as complex as possible. A chemical formula can have only one chemical symbol to represent a single substance, for example, H_2 (hydrogen gas), two chemical symbols to represent a compound, for example, H_2O (water), three symbols to represent a more complex compound, for example, HClO (hypochlorous acid) or in principle, indefinitely more, so long as the represented chemical species exist.[46] For example, $C_{10}H_{16}N_2O_3S$, the chemical formula of amidephrine (an organic compound), involves five chemical symbols. The iterative structure enables chemical formulas to represent a wide range of chemical species that include indefinitely many elements.

In terms of field, chemical formulas and the chemical symbols representing different chemical elements form a whole-part relation, i.e., a compositional taxonomy. For example, H_2 is composed of only two hydrogen atoms, H_2O comprises two hydrogen atoms and one oxygen atom, and HClO consists of one hydrogen atom, one chlorine atom, and one oxygen atom. The iterative structure of the grammar means that chemical formulas can indefinitely expand the compositional taxonomy it realizes. This expansion occurs in terms of **breadth** – formulas and the element symbols that compose them form a compositional taxonomy (the whole chemical species and the atoms), but there can be an indefinite number of atoms, progressively expanding the taxonomy's breadth.

In addition to the predominant iterative structures, chemical formulas include a small component of elements that are non-iterative but realize important meaning for chemistry. Ionic formulas involve a multivariate structure of 'Ion ^ Number ^ Charge' at the formula rank (see Section 4.1.1.3.4.2), and molecular formulas include a subjacency structure of 'A_B' at the symbol rank (see Section 4.1.1.3.5). The functional elements in these structures are non-iterative. In terms of field, the element 'Number' in the multivariate structure and 'B' in the subjacency structure construe gauged properties as they describe the number of electric charges and atoms. The functional element 'Charge', on the other hand, realizes a qualitative property of charge polarities. Table 4.14 summarizes the relation between the grammatical organization and field-specific meanings realized in chemical formulas.

46 It should be noted that in chemical formulas, normally, only non-metallic element symbols can iterate. H, O, and Cl all represent non-metallic elements. This is important for certain aspects of the grammar, but it does not affect the argument here.

Table 4.14 Grammatical organization and field-specific meanings realized in chemical formulas

Grammatical organization	**Field-specific meanings realized**
univariate structure at the formula rank (iteration of chemical symbols)	the breadth of composition taxonomy
multivariate structure at the formula rank	qualitative properties of charge polarities and gauged properties of charge quantities
subjacency structure at the symbol rank	gauged properties of atom quantities

4.1.3.1.2 Structural formulas

Like chemical formulas, structural formulas are also primarily organized around univariate structures. As shown in Section 4.1.2.3.1, chemical symbols and connectors can iterate indefinitely to form an extremely broad structural formula. The iterative structure, however, is distinct from that of chemical formulas in that chemical symbols can iterate in different directions in two-dimensional space. This spatial organization is manifested by the iteration of the symbols in three primary directions: horizontal, vertical, and diagonal (see Section 4.1.2.3.1).

In terms of field, like chemical formulas, the symbols and structural formulas form a part-whole relation – a compositional taxonomy. This means that similarly, the univariate structure in structural formulas also expands the breadth of compositional taxonomy. However, in contrast to chemical formulas that iterate symbols only in one dimension, structural formulas allow the compositional taxonomy to be expanded on a two-dimensional space through iteration of symbols in different directions. This gives a spatial organization of the compositional taxonomy that is termed a **spatial property** within the model of field put forward by Doran and Martin (2021). This property enables structural formulas to show not just what atoms occur in a molecule, but also their arrangement. In other words, the spatial arrangement of symbols, to some extent, reflects how chemistry conceives of the molecule's actual structure, giving structural formulas the mixed feature of being "half-symbolic and half-iconic" (Hoffmann and Laszlo 1991: 10). The spatial property afforded by structural formulas also plays a significant role in making explicit molecule's chemical properties. For example, the different spatial arrangements of '—O—H' (circled red) in Figure 4.46 and Figure 4.47 form two different molecules, butan1-ol and butan-2-ol, with different physical and chemical properties (Chan et al. 2019: 286). However, this meaning potential is not afforded by chemical formulas because the two molecules share the same formula

$C_4H_{10}O$. That is, although from the perspective of field, Figure 4.46 and Figure 4.47 both show the same compositional taxonomy, and are thus represented by the same chemical formula, they have different spatial arrangements and are thus distinguished by structural formulas.

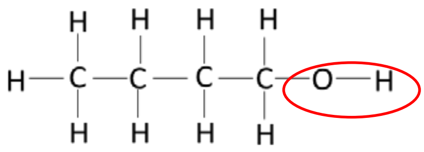


Figure 4.46 The structural formula of butan-1-ol

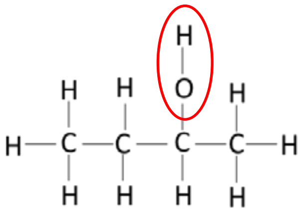


Figure 4.47 The structural formula of butan-2-ol

In addition to the iterative structure, structural formulas include a small component of elements that are non-iterative at the symbol rank. As shown in Section 4.1.2.3.2, structural formulas involve a subjacency structure of '^{B}A', in which B describes partial charge polarities and A denotes types of atom. None of these functional elements are iterative. In terms of field, this grammatical structure realizes a qualitative property of partial charge polarities that atoms bear.

Besides the dimension of univariate and multivariate structures, the meaning-making of structural formulas can also be captured by another dimension of describing structures – particulate, prosodic, and periodic (Halliday 1979). As discussed in Section 4.1.2.3.1, structural formulas involve two prosodic structures ideationalized for construing the knowledge of chemical matter. One is the agreement between symbols and connectors, construing a set of bonding relations cutting across a molecule. The other is the arrangement of a set of partial charge signs, representing partial electric charges diffusing across a molecule. In terms of field, these prosodic structures construe qualitative properties of molecules' connectivity and polarity.

The grammatical organization and field-specific meanings realized discussed in this section can be summarized as Table 4.15.

Table 4.15 Grammatical organization and field-specific meanings realized in structural formulas

Grammatical organization	**Field-specific meanings realized**
univariate structure (iteration of symbols along different dimensions)	the breadth of composition taxonomy
	spatial properties
prosodic structures	qualitative properties of molecular connectivity and polarity
subjacency structure at the symbol rank	qualitative properties of partial charge polarities

4.1.3.1.3 Chemical equations

Whereas chemical formulas and structural formulas are primarily organized around univariate structures, chemical equations tend to be dominated by multivariate structures. To discuss these structures, we focus on the equation and term ranks because the formula rank includes the chemical formulas' grammatical structures discussed in Section 4.1.3.1.1.

At the equation rank, chemical equations show two types of multivariate structures: a structure of 'Reactant ^ Relator ^ Product' for irreversible equations and a two-layered conflated structure of 'Reactant ^ Relator ^ Product' and 'Actant ^ Relator ^ Product' for reversible equations (see Section 4.1.1.3.1). The first structure describes irreversible reactions, involving only forward reactions, while the second construes reversible reactions, including both forward and reverse reactions that can go back and forth until they reach "an equilibrium state" (Taber 2009: 96). Put differently, the irreversible reactions are "non-cyclical activities" (Doran and Martin 2021), whereas the reversible reactions involve two activities sequenced one after another, which, as a result, are cyclical. In terms of field, they are modeled as two respective activities in Doran and Martin's (2021) field model - linear unmomented activities and cyclical momented activities.

The equation rank also involves a prosodic structure ideationalized for construing the number of units of chemical species involved in chemical reactions. For example, in the equation '$2KMnO_4(aq) + 16HCl(aq) \rightarrow 2KCl(aq) + 2MnCl_2(aq) + 5Cl_2(g) + 8H_2O(l)$', the different coefficients, 2, 16, 2, 2, 5 and 8, diffuse across the equation, forming a prosody that construes different quantities of chemical species involved in that chemical reaction. In terms of field, the ideationalized prosodic structure realizes a gauged property of the number of units of chemical species involved in a reaction.

At the term rank, terms realize a multivariate configuration of 'Quantity ^ Species ^ State' (see Section 4.1.1.3.3.2). In contrast to the multivariate structure of the broader equations that realize activities, the multivariate structure of terms construes two types of properties. The function Quantity specifies the number of units of a chemical species, construing what Doran and Martin (2021) refer to as a gauged property. This property allows chemical equations to realize one significant purpose – establishing quantitative relations of reactions (Simon 1926: 1306). More specifically, it specifies the molar ratios in which the chemical species react and are formed. For example, in the chemical equation $CH_4(g) + 2O_2(g) \rightarrow CO_2(g) + 2H_2O(l)$, methane gas molecules and oxygen gas molecules react in a ratio of one to two, and carbon dioxide molecules and water molecules are formed in the same ratio. In addition to gauged properties, terms realize another type of properties – qualitative properties, realized by the function of State. '(l)' in '$2H_2O(l)$', for example, represents that the water is in the liquid state. Whereas the gauged property enables chemical equations to describe the quantitative relations of a reaction at the microscopic level (the molecule level), the qualitative property allows for construing chemical reactions at the macroscopic level (the material world that can be sensed), i.e., it describes the physical states of chemical species observed in reactions. Viewed in terms of Johnstone's (1991) chemical triplet, the two field properties thus bridge two levels of chemistry knowledge: the macroscopic and microscopic, all the while construing these in terms of Johnstone's third level of symbolic knowledge. (See Section 2.3 for a detailed introduction of the three levels of chemical knowledge - microscopic, macroscopic, and symbolic)

Note that in addition to the multivariate structure within terms, the term rank also involves a univariate structure that groups terms into term complex (see Section 4.1.1.3.3.1). In terms of field, each side of a chemical equation (the left or right of the reaction arrow) and the terms within it form a whole-part relation, i.e., a compositional taxonomy. For example, the left side of the equation $C(s) + O_2(g) \rightarrow CO_2(g)$ shows that the reactants for the chemical reaction comprise two chemical species, i.e., carbon and oxygen gas. The iterative structure means that term complexes can indefinitely expand the breadth of the compositional taxonomy and allow chemical equations to construe chemical reactions that involve as many reactants or products as possible, enabling them to represent highly complex chemical reactions.

Taking into account the univariate structure at the formula rank, chemical equations construe three levels of composition taxonomies, two of which can expand indefinitely. The first level of part-whole relation is between reactants and products (represented by the left and right sides of a chemical equation), and the whole substances involved in the chemical reaction. Since chemical equations are non-iterative, this level of compositional taxonomies cannot expand. The second level of compositional taxonomies has been discussed in the paragraph above. The univariate structure of terms gives reactants and products the potential of including an infinite number of chemical species, meaning that the breadth of the compositional taxonomies can be expanded indefinitely. The third level has been discussed in Section 4.1.3.1.1. Chemical species (represented by chemical formulas) can comprise an indefinite number of atoms, construing compositional taxonomies that can be expanded indefinitely. The three levels of compositional taxonomies are illustrated in Figure 4.48.

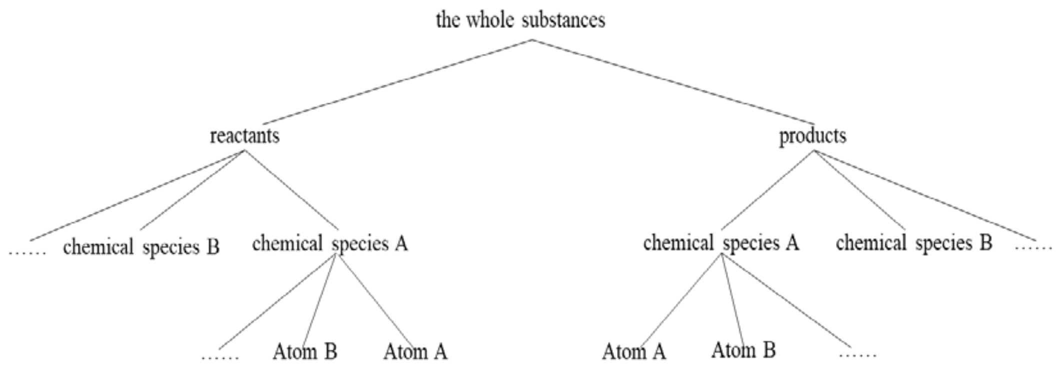


Figure 4.48 The three levels of compositional taxonomies by chemical equations

Take the equation $CH_4(g) + 2O_2(g) \rightarrow CO_2(g) + 2H_2O(l)$ for an example. At the first level, the whole substances consist of reactants (CH_4 and O_2) and products (CO_2 and H_2O), which themselves, in turn, comprise chemical species at the second level. For example, the products consist of CO_2 and H_2O. At the third level, the chemical species themselves are composed of atoms, for example, H_2O consisting of two hydrogen atoms and one oxygen atom. Overall, chemical equations show deeper compositional taxonomies than the other two formalisms.

Thus far, we have analyzed the types of structure and the field-specific meanings realized at the equation and term ranks. It should be noted that since chemical formulas are a part of chemical equations, the structures and meanings realized in chemical formulas are themselves involved in chemical equations. Table 4.16 presents the relations between the grammatical organization and the field-specific meanings realized in chemical equations.

Table 4.16 Grammatical organization and field-specified meanings realized in chemical equations

Grammatical organization	Field-specific meanings realized
multivariate structure at the equation rank	linear unmomented activities cyclical momented activities
prosodic structure at the equation rank	gauged properties of units of chemical species involved in chemical reactions
multivariate structure at the term rank	gauged properties of quantities of substances qualitative properties of physical states
univariate structure at the term rank	the breadth of composition taxonomy of chemical species within reactants or products
univariate structure at the formula rank	the breadth of composition taxonomy of elements within chemical species
multivariate structure at the formula rank	qualitative properties of charge polarities and gauged properties of charge quantities
subjacency structure at the symbol rank	gauged properties of atom quantities

4.1.3.2 A functional semiotic typology

With the analysis of the grammatical organization and field-specific meanings completed, we are now in a position to have a typological view of the three chemical formalisms based on their similarities and differences in their types of structure and field meanings realized.

In terms of types of structure, chemical formulas and structural formulas are similar in that both are dominated by univariate structures, each capable of iterating chemical symbols multiple times. However, in contrast to chemical formulas, structural formulas can iterate in a two-dimensional space, which affords a reading of spatial layout. This leads to a distinction in the field-specific meanings they realize. Whereas chemical formulas realize only a compositional taxonomy, the two-dimensional organization of the structural formulas realizes both a compositional taxonomy and spatial properties. The two formulas are also similar in that they both involve subjacency structures. However, they employ these structures to realize different field-specific meanings: gauged properties of atom quantities for chemical formulas and qualitative properties of partial charge polarities for structural formulas. In contrast to chemical formulas and structural formulas dominated by univariate structures, chemical equations are organized largely around multivariate structures. At the equation rank, they construe activities in terms of different types of reaction, while at the term rank, these structures put forward various properties, including the quantity relations of reactions and physical states of chemical species. However, as noted above, chemical equations also include two univariate components – the first is the complexing of terms, which realizes the first level of the breadth of compositional taxonomy, while the second is that afforded by the chemical formulas within equations, which in turn realize the second level of the breadth of compositional taxonomy. In contrast to chemical formulas and structural formulas that involve only one level of compositional taxonomy, chemical equations include deeper compositional taxonomies – three levels. In addition to the discussion of similarities and differences in meaning-making in terms of univariate and multivariate structures, chemical formalisms can also be compared in terms of whether they involve **ideationalized prosodic structures**. Sections 4.1.3.1.2 and 4.1.3.1.3 have shown that structural formulas and chemical equations employ prosodic structures to construe qualitative properties of molecular connectivity and polarity and gauged properties of the number of units of chemical species involved in chemical reactions respectively, in contrast to chemical formulas that do not involve such prosodic structures.

Overall, Table 4.17 summarizes the similarities and differences of the three chemical formalisms in the grammatical organization and field-specific meanings realized.

Table 4.17 Similarities and differences of the three chemical formalisms in their grammatical organization and field meanings realized

		Types of structure		
		univariate	**multivariate**	**prosodic**
Field meanings	**composition**	➢ chemical formulas • composition breadth (elements within chemical species) ➢ structural formulas • composition breadth (elements within chemical species) ➢ chemical equations • the term rank: composition breadth (chemical species within reactants or products) • the formula rank: composition breadth (elements within chemical species)		
	activity		➢ chemical equations • activities – irreversible and reversible reactions	
	property	➢ structural formulas • spatial properties of elements within chemical species	➢ chemical equations • the term rank: (1) gauged (the quantities of chemical species), (2) qualitative (physical states of chemical species) ➢ chemical formulas • the formula rank: (1) gauged (the quantities of electric charges), (2) qualitative (electric charge polarities) • the symbol rank: gauged (the quantities of atoms) ➢ structural formulas • the symbol rank: qualitative (partial charge polarities)	➢ Structural formulas • qualitative properties of molecular connectivity and polarity ➢ chemical equations • gauged properties of the number of units of chemical species involved in chemical reactions.

In terms of field-specific meanings realized, Table 4.17 suggests that in some sense, structural formulas and chemical equations are ‘building’ upon chemical formulas – chemical formulas construe the basic compositional taxonomy, structural formulas add in spatial properties, and chemical equations add in extra levels of compositional taxonomies, two properties, and activities. From chemical formulas through structural formulas to chemical equations, the field-specific meanings realized aggregate. In this sense, the table makes clear the importance of composition for chemistry knowledge, i.e., each formalism focuses heavily on building the breadth of its compositional taxonomies. More specifically, each formalism presents the organization of elements within chemical specifies. What differs is how these elements in chemical species are related to other meanings – spatial organization for structural formulas, extra levels of composition, activities, and properties for chemical equations, and nothing more for chemical formulas.

4.1.3.3 Disciplinary affordance of chemical formalisms

As has been pointed out by Doran (2019), academic formalisms are developed to organize particular meanings that are crucial for the development of knowledge in specific disciplines. Doran’s study shows that univariate structures are prevalent across academic formalisms, but the field-specific meanings realized vary depending on what is needed in the disciplines. For example, mathematical symbolism used across the sciences and social sciences, tree diagrams and system networks in linguistics, and nuclear equations used in physics are largely organized around univariate structures, but the field-specific meanings they realize differ greatly with their disciplines. Mathematical symbolisms construe interdependencies between properties, tree diagrams in linguistics realize an indefinite depth and breadth of compositional taxonomy, system networks in linguistics organize indefinite depth and breadth of classificational taxonomies and nuclear equations construe compositional taxonomies and indefinitely long activities. In comparison, chemical formalisms involve a number of univariate structures as well as several multivariate structures, construing activities, properties, and compositional taxonomies. With three types of field-specific meanings, chemical formalisms enjoy more flexibility to shift between perspectives as needed for construing the technical knowledge in chemistry. This section will discuss how the semiotic typology of chemical formalisms meets chemistry’s disciplinary affordance.

According to Airey et al. (2014), disciplinary affordance is defined as the agreed meaning-making functions that a semiotic resource fulfills for a particular disciplinary community. In the context of this thesis, the disciplinary affordance of chemical formalisms is their meaning potential for construing secondary school chemistry knowledge. The disciplinary affordance is largely shaped by the nature of that discipline. Chemistry as a discipline can be broadly understood as “the science of the composition and structure of materials and of the changes that materials undergo” (Ebbing and Gammon 2008: 2). Interpreted in terms of field, this suggests chemistry knowledge centers around the static organization of items and dynamic unfolding of activities which then drive the disciplinary affordance of chemical formalisms and determines in part how they work. The three chemical formalisms discussed above meet the disciplinary needs of representing the technical knowledge of chemical substances’ composition and internal structure, and chemical reactions between the substances. All three formalisms construe compositional taxonomies that can expand their breadth through iterative structures. This is in line with chemistry’s need to describe the atomic composition of chemical compounds that can involve multiple different chemical elements. This knowledge allows chemists not only to differentiate compounds sharing the same chemical formulas (Goodwin 2008: 119) but also to account for the compounds’ chemical behavior (Weininger 1998: 19). Chemical equations, on the other hand, construe another significant part of chemistry knowledge – the changes that matter undergoes during a reaction. The multivariate structure allows the equations to represent the two respective chemical reactions: irreversible reactions and reversible reactions, which, in terms of field, are linear unmomented activities and cyclical momented activities.

Overall, the formalisms explored in this thesis are complementary to each other, with each presenting chemical substances’ composition while potentially adding other meanings. The typological analysis of the formalisms in terms of their grammatical organization and field-specific meaning thus provides a useful understanding of how similar or different the components of chemistry knowledge could be, in ways that can be related to similar analyses for other semiotic resources including language. This is important for informing the development of literacy pedagogy that can facilitate teaching and learning chemistry across language, images, and this range of formalisms.

4.2 The development of chemical formalisms across schooling levels

Section 4.1 has offered a comprehensive description of chemical formalisms, considering the construal of meaning by the formalisms. We are now in a position to understand their role in building the hierarchical knowledge structure of chemistry. To do this, however, we need to examine how their meaning develops across learning stages. The description in Section 4.1 accounts for their possibilities of use, i.e., their meaning-potential, but does not yet show their actualities of use. To understand the knowledge-building of chemistry through chemical formalisms, we need to push a step further by investigating how the formalisms are used in schooling and how they develop and change across schooling levels. This section will thus focus on the development of chemical formalisms through schooling in NSW secondary school chemistry textbooks. Since chemical matter and chemical reactions are two major components of chemistry knowledge (Ebbing and Gammon 2008: 2), this section will explore how the formulas, including chemical formulas and structural formulas, develop across schooling levels in the textbooks to build the knowledge of chemical matter (Section 4.2.1) and how chemical equations develop in the textbooks to build the knowledge of chemical reactions (Section 4.2.2).

4.2.1 The development of formulas across schooling levels to build the knowledge of chemical matter

As noted in Section 4.1.2.3, various types of formulas are employed to construe the knowledge of chemical matter. These formulas embody different degrees of complexity and are used in different schooling levels. To understand their role in building the knowledge of chemical matter, this section will first examine their varying levels of complexity in terms of LCT's semantic density (Section 4.2.1.1) and then trace their development across schooling levels in secondary school chemistry textbooks to explore how the formulas build the hierarchical knowledge structure of chemistry (Section 4.2.1.2).

4.2.1.1 Semantic density of formulas in secondary school chemistry

As shown in Section 4.1.3, the meanings construed by formulas have been well accounted for in terms of field. Given these meanings are all technical, the formulas' semantic density can be measured in terms of how many field-specific meanings they include. The following will thus step through the field meanings realized by the various formulas used in secondary school chemistry and establish a model for their semantic density that can be enacted to trace the formulas' development in Section 4.2.1.2.

As discussed in Section 4.1.3.1.1, the most basic formulas are molecular formulas, which, in terms of field, construe compositional taxonomies of chemical matter. For example, the molecular formula O_3 shows that an ozone molecule comprises three oxygen atoms. Based on molecular formulas, ionic formulas add in two field properties: a qualitative property (polarities of electric charges) and a gauged property (the number of electric charges). For example, the ionic formula O^{2-} shows that an oxygen anion bears two negative electric charges (shown by '$^{2-}$'), in addition to the composition information that the anion is composed of one oxygen atom. Thus, ionic formulas embody stronger semantic density than molecular formulas.

Structural formulas push the semantic density of formulas to an even higher level as they bring in the meanings of the spatial arrangements of atoms and covalent bonds. As Section 4.1.2.3.1 has shown that structural formulas involve several subtypes, structural formulas themselves exhibit different levels of semantic density. Formulas that show atoms and covalent bonds in one or two dimensions are usually referred to as 'conventionalized structural formulas' (Chan et al. 2019). For example, Figure 4.49 is a conventionalized formula of ozone molecules showing the spatial arrangements of oxygen atoms and covalent bonds connecting the atoms. In terms of field, the spatial arrangements of atoms construe a spatial property (Doran and Martin 2021). The covalent bonds are new items added and constitute a classificational taxonomy, i.e., the single bond (the single line) and the double bonds (the double lines) are sub-classes of covalent bonds. Compared with ionic formulas, the conventionalized structural formulas add more field meanings based on molecular formulas and hence embody stronger semantic density.

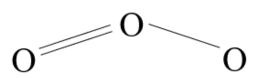


Figure 4.49 The structural formula of ozone molecules

As shown in Section 4.1.2.3.1, conventionalized structural formulas can denote molecular polarities by attaching signs representing partial electric charges to chemical symbols. For example, Figure 4.50 shows that the two hydrogen atoms bear partial positive electric charges (H represents hydrogen atoms and 'δ+' represents partial positive charges), and the oxygen atom embodies a partial negative charge (O denotes oxygen atoms and 'δ-' signifies partial negative charges). The asymmetrical distribution of the partial electric charges indicates that methanal molecules are polar.[47] In terms of field, conventionalized structural formulas showing molecular polarities include all the field-specific meanings construed by the formula shown in Figure 4.49 and add a qualitative property of partial electric charges. Therefore, conventionalized structural formulas showing molecular polarities exhibit stronger semantic density than those that do not.

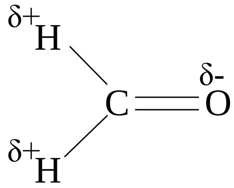


Figure 4.50 The structural formula of methanal molecules showing molecular polarities

There is still one further variation of structural formulas – stereo formulas. As shown in Section 4.1.2.3.1, these formulas show arrangements of atoms in three-dimensional space. For example, Figure 4.51 is the stereo formula of methane molecules, presenting the molecules' three-dimensional arrangements of the atoms.

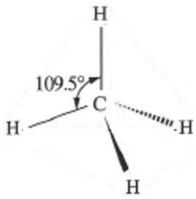


Figure 4.51 The stereo formula of methane molecules

The solid wedge ( ) linking C with H indicates an arrangement out of the page, while the dashed wedge ( ) linking C with H indicates an arrangement into the page. The solid line (−), on the other hand, represents an arrangement in the plane of page. Such formulas contrast with conventionalized structural formulas that depict only the relative positions of atoms to each other in a two-dimensional plane. For instance, the conventionalized structural formula of methane molecules in Figure 4.52 does not show the actual relative locations of each atom.

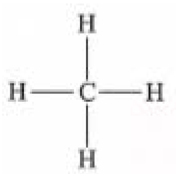


Figure 4.52 The conventionalized structural formula of methane molecules

[47] In asymmetrical molecules, the partial electric charges borne by the atoms do not cancel one another, resulting in an overall polarity (Chan et al. 2018: 178).

In terms of field, both formulas construe spatial properties. However, they realize two different subtypes. In Kress and van Leeuwen's (2006) term, the former tends to be 'topographical' while the latter is more of 'topological'. Topographical structures can accurately represent the physical spatial relations between elements, while topological ones present only 'logical' relations between them (Kress and van Leeuwen 2006: 98–99). Thus, we specify two subtypes of spatial property - **topographical** spatial property and **topological** spatial property. This enriches the original field model proposed by Doran and Martin regarding the description of spatial organizations of items.

In addition to the topographical spatial property, stereo formulas construe another property: a qualitative property of angle measurement. Figure 4.51, for example, construes a qualitative property of angle measurement between the carbon-hydrogen covalent bonds, realized by the double-ended arrow and the numeral labeling '109.5°'. It means that the angle between the carbon-hydrogen atoms is 109.5°. Thus, compared with conventionalized structural formulas with partial charge signs that add only one field-specific meaning based on conventionalized structural formulas showing atoms and covalent bonds in one or two dimensions, stereo formulas bring in two additional field meanings - a topographical spatial property and a qualitative property of angel measurement. This indicates that stereo formulas embody the strongest semantic density among the various formulas.

Based on the above analysis, we can sequence the semantic density of the formulas used in secondary school chemistry along a strength of continuum, as shown in Table 4.18.

Table 4.18 semantic density of formulas used in secondary school chemistry

Semantic density (SD)	**Formulas**
+++++	stereo formulas
↕	conventionalized structural formulas showing molecular polarities
	conventionalized structural formulas
	ionic formulas
+	molecular formulas

For ease of enacting the above model for the semantic density of formulas to compare the various formulas, we assign different scales of semantic density to them: molecular formulas (SD+), ionic formulas (SD++), conventionalized structural formulas (SD+++), conventionalized structural formulas showing molecular polarities (SD++++), and stereo formulas (SD+++++) (SD is the abbreviation of semantic density). '+' denotes the strength of semantic density. The more '+', the stronger the semantic density. These scales constitute a tool for comparing and tracing the development of formulas in terms of their semantic density in secondary school chemistry textbooks, which allows for revealing how the formulas develop across schooling levels to build the field of chemical matter.

4.2.1.2 The development of formulas in secondary school chemistry textbooks

To understand how the knowledge of chemical matter is built through formulas in chemistry, this section will draw on the discussion of semantic density in Section 4.2.1.1 to trace the development of formulas across schooling levels in chemistry textbooks used by secondary schools in NSW, Australia.

As introduced in Chapter 3, the curriculum of NSW secondary school chemistry includes six years categorized as three stages: Stage 4 (Years 7 and 8), Stage 5 (Years 9 and 10), and Stage 6 (Years 11 and 12). In Stage 4, chemical formulas do not occur until Year 8, when molecular formulas are introduced to describe molecules' composition. For example, when teaching molecular compositions of carbon dioxide molecules, the Year 8 textbook uses a clause complex in which the molecular formula CO_2 is incorporated: *carbon dioxide has the molecular formula* CO_2*, which means it has 1 carbon atom and 2 oxygen atoms* (Zhang et al. 2014b: 184). The attributive clause explains the meaning of the molecular formula, i.e., the composition of carbon dioxide molecules. In terms of the scales of formulas' semantic density discussed in Section 4.2.1.1, the molecular formulas in Year 8 construe compositional taxonomies and embody the weakest semantic density (SD+). The formulas used in Stage 4 and their field-specific meanings and semantic density are shown in Table 4.19.

Table 4.19 The formulas used in Stage 4 and their field meanings and semantic density

Schooling levels		**Formulas**	**Field-specific meanings**	**Semantic density**
Stage 4	Year 7	-	-	none
	Year 8	Molecular formulas	Compositional taxonomies (composition of chemical matter)	From none to SD+

In Stage 5, the Year 9 textbook uses molecular formulas throughout the year to represent molecular compounds.[48] For example, H_2O is used to represent water molecules in the following text: *water is an example of a molecule compound. Its formula is* H_2O (Zhang et al. 2014b: 75). In addition to molecular formulas, the Year 9 textbook introduces a new type of formula, ionic formulas, to construe the knowledge of ions. For example, ionic formulas in the following text represent a cation and an anion constituting an ionic compound: *Calcium carbonate contains calcium ions,* Ca^{2+}*, and carbonate ions,* CO_3^{2-}*. These ions must be present in the ratio 1:1 so the total positive charge equals the total negative charge* (Zhang et al. 2014b: 70). The ionic formulas Ca^{2+} and CO_3^{2-} represent calcium cations and carbonate anions that bear two positive and negative charges. As discussed in Section 4.2.1.1, these ionic formulas construe composition taxonomies, a gauged property of the number of electric charges, and a qualitative property of charge polarities in terms of field, exhibiting stronger semantic density (SD++) than the molecular formulas used in Year 8. Therefore, from Year 8 to 9, there is a growth of semantic density in the formulas used to construe the field of chemical matter. As the Year 10 textbook uses molecular and ionic formulas throughout and does not introduce new formulas, the semantic density of formulas in Year 10 remains at the same level as Year 9. Table 4.20 summarizes the formulas used in Stage 5 and their field meanings and semantic density. It shows that the semantic density of formulas ranges from SD+ to SD++ in both Years 9 and 10.

Table 4.20 The formulas used in Stage 5 and their field meanings and semantic density

Schooling levels		**Formulas**	**Field-specific meanings**	**Semantic density**
Stage 5	Year 9	Molecular formulas Ionic formulas	Compositional taxonomies (compositions of chemical matter) A gauged property (the number of electric charges) A qualitative property (polarities of electric charges)	From SD+ to SD++
	Year 10	Molecular formulas Ionic formulas	Compositional taxonomies (compositions of chemical matter) A gauged property (the number of electric charges) A qualitative property (polarities of electric charges)	From SD+ to SD++

In Stage 6, the Year 11 textbook includes all formulas used in Year 10. For instance,

[48] Molecular compounds are chemical compounds that take the form of discrete molecules, for example, water (H_2O) and carbon dioxide (CO_2).

it uses molecular formulas to represent molecular compounds, as shown in *the element oxygen, for example, forms* O_2 *molecules consisting of two oxygen atoms held together by a covalent bond*, and ionic formulas to denote ions, as illustrated in *losing one electron makes it positively charged, and it is given the symbol* Na^+. This suggests that the field meanings of chemical matter construed in Year 10 continue to constitute a crucial part of the knowledge of chemical matter in Year 11. In addition to this, Year 11 introduces a new type of formula – structural formulas – to represent molecular structure. This concerns the various covalent bonds and the spatial arrangement of atoms. For example, when teaching ozone molecules' structure, the Year 11 textbook employs the conventionalized structural formula presented in Figure 4.53 to show how the oxygen atoms are bonded together.

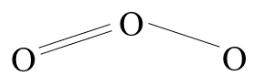


Figure 4.53 The structural formula of ozone molecules

It depicts that an ozone molecule comprises a central oxygen atom bonded to two other oxygen atoms through a single covalent bond (represented by the single line) and a double covalent bond (the double line). The two oxygen atoms are arranged in diagonal directions to the central oxygen atoms. As discussed in Section 4.2.1.1, such conventionalized structural formulas construe compositional taxonomies, classificational taxonomies, and a typological spatial property, embodying stronger semantic density (SD+++) than ionic formulas.

In addition to the bonding relations between atoms and their spatial arrangements, the latter part of the Year 11 textbook also introduces the knowledge of molecular polarities. As discussed in Section 4.1.2.3.1, this is realized by attaching partial charge signs 'δ+' and 'δ-' to chemical symbols in conventionalized structural formulas. For example, the textbook illustrates polar molecules through the conventionalized structural formula of methanal molecules presented in Figure 4.54.

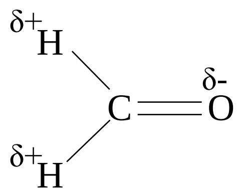


Figure 4.54 The structural formula of methanal molecules showing molecular polarities

The partial charge signs denote that the hydrogen atoms (H) bear partial positive charges (δ+) while the oxygen atom (O) bears a partial negative electric charge (δ-). The asymmetrical distribution of the partial electric charges indicates that methanal molecules are polar. As discussed in Section 4.2.1.1, in terms of field, this type of formula adds a qualitative property of partial electric charges in addition to the spatial property and classificational taxonomies construed in the conventionalized structural formula shown in Figure 4.53, indicating that they embody stronger semantic density (SD++++).

In Year 12 Stage 6, the textbook uses all the formulas occurring in previous years and introduces a new type of structural formula. These formulas are introduced to teach organic chemistry, a key component of Year 12 chemistry. Significantly, since organic molecules' structure largely determines their chemical properties (Weininger 1998: 19), structural formulas at this stage need to present more information about the structural organization of atoms in organic molecules. In particular, they need to show their three-dimensional structure and the angles between each covalent bond – what we called in Section 4.1.2.3.1 stereo formulas. This is illustrated in Figure 4.55 that presents a stereo formula of methane, showing that two hydrogen atoms are positioned out of and into the plane of the page, denoted by the solid wedge ( ) and dashed wedge ( ), respectively. The solid lines (－) mean that the two hydrogen atoms are positioned in the plane of the page.

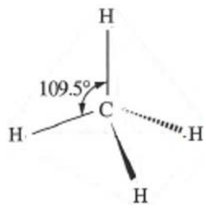


Figure 4.55 The stereo formula of methane molecules

As discussed in the previous section, these spatial arrangements of atoms construe a topographical spatial property in terms of field, showing more information about atoms' location in a molecule than as presented in the conventionalized structural formulas that present only relative logical locations of atoms in one- or two-dimensional space. Besides this topological spatial property, the stereo formula also construes a qualitative property of angle measurement between the carbon-hydrogen covalent bonds, realized by the numeral labeling '109.5°'. As discussed in Section 4.2.1.1, this type of formula embodies the strongest semantic density (SD+++++) among the formulas used in secondary school chemistry.

Based on the above analysis, we can summarize the formulas used in Stage 6 and their field meanings and semantic density as Table 4.21. It shows that the semantic density of formulas ranges from SD+ to SD++++ in Year 11 and from SD+ to SD+++++ in Year 12.

Table 4.21 The formulas used in Stage 6 and their field meanings and semantic density

Schooling levels		**Formulas**	**Field-specific meanings**	**Semantic density**
Stage 6	Year 11	Molecular formulas Ionic formulas Structural formulas	Compositional taxonomies (compositions of chemical matter) A gauged property (the number of electric charges) A qualitative property (polarities of electric charges) A typological spatial property (two-dimensional arrangements of atoms in molecules) A qualitative property (partial electric charges atoms bear)	From SD+ to SD++++
	Year 12	Molecular formulas Ionic formulas Structural formulas Stereo formulas	Compositional taxonomies (compositions of chemical matter) A gauged property (the number of electric charges) A qualitative property (polarities of electric charges) A typological spatial property (two-dimensional arrangements of atoms in molecules) A qualitative property (partial electric charges atoms bear) A topological spatial property (three-dimensional arrangements of atoms) A qualitative property (angle measurements between covalent bonds)	From SD+ to SD+++++

The overall development of formulas in terms of field-specific meanings and semantic density across schooling levels can be summarized as Table 4.22. The words in bold font mean that the field meanings are newly developed compared with the year before. For example, Year 9 includes two new types of properties compared with Year 8.

Table 4.22 The development of formulas in terms of field-specific meanings realized and their semantic density

<table>
<tr><th colspan="2">Schooling levels</th><th>Formulas</th><th>Field-specific meanings</th><th>Semantic density</th></tr>
<tr><td rowspan="2">Stage 4</td><td>Year 7</td><td>-</td><td>-</td><td></td></tr>
<tr><td>Year 8</td><td>Molecular formulas</td><td>Compositional taxonomies
(compositions of chemical substances)</td><td>+</td></tr>
<tr><td rowspan="2">Stage 5</td><td>Year 9</td><td>Molecular formulas
Ionic formulas</td><td>Compositional taxonomies
(compositions of chemical substances)
A gauged property
(the number of electric charges)
A qualitative property
(polarities of electric charges)</td><td>From
SD+ to
SD++</td></tr>
<tr><td>Year 10</td><td>Molecular formulas
Ionic formulas</td><td>Compositional taxonomies
(compositions of chemical substances)
A gauged property
(the number of electric charges)
A qualitative property
(polarities of electric charges)</td><td>From
SD+ to
SD++</td></tr>
<tr><td rowspan="2">Stage 6</td><td>Year 11</td><td>Molecular formulas
Ionic formulas
Structural formulas</td><td>Compositional taxonomies
(compositions of chemical matter)
A gauged property
(the number of electric charges)
A qualitative property
(polarities of electric charges)
A typological spatial property
(two-dimensional arrangements of atoms in molecules)
A qualitative property
(partial electric charges atoms bear)</td><td>From
SD+ to
SD++++</td></tr>
<tr><td>Year 12</td><td>Molecular formulas
Ionic formulas
Structural formulas
Stereo formulas</td><td>Compositional taxonomies
(compositions of chemical substances)
A gauged property
(the number of electric charges)
A qualitative property
(polarities of electric charges)
A typological spatial property
(spatial arrangements of atoms)
A qualitative property
(partial electric charges atoms bear)
A topological spatial property
(three-dimensional arrangements of atoms)
A qualitative property
(angle measurements between covalent bonds)</td><td>From
SD+ to
SD+++++</td></tr>
</table>

The table shows an increase in field-specific meanings realized by the formulas from Year 8 to 9, from Year 10 to 11, and from Year 11 to 12, indicating a growth in semantic density through schooling. The range of the semantic density of the formulas in each year level can be presented as Figure 4.56.

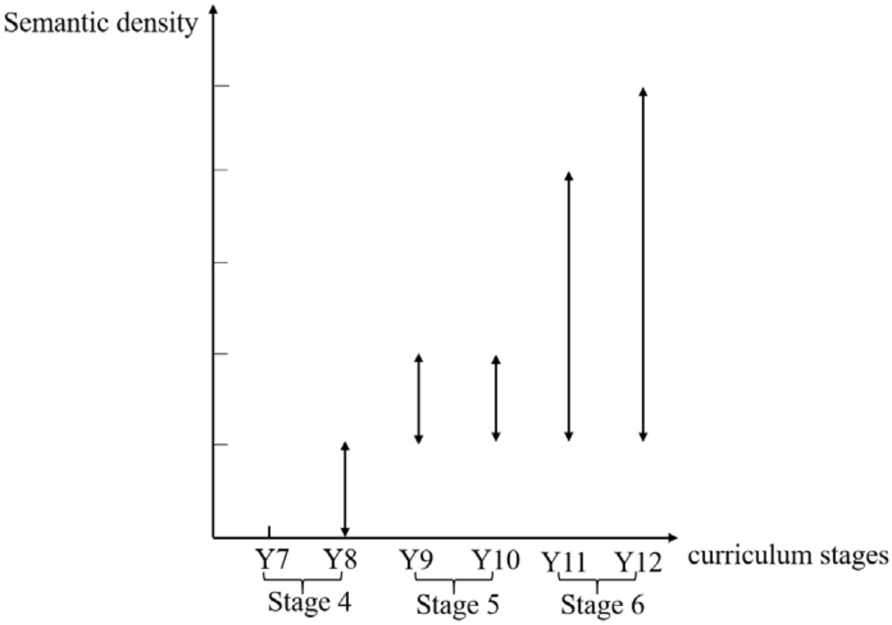


Figure 4.56 The range of semantic density of formulas across curriculum stages in secondary school chemistry textbooks ('Y' is short for 'Year')

As shown in Figure 4.56, the molecular formulas used in Year 8 constitute the basis where Stages 5 and 6 expand the field of chemical matter. Overall, through the curriculum stages, the field construed by formulas in secondary school chemistry textbooks reaches increasingly stronger semantic density, indicating that as learning progresses, the knowledge of chemical matter presented through formulas becomes increasingly complex. As well as this, except for Year 7, which does not use formulas, each year uses formulas occurring in the previous years, allowing them to maintain connections to relatively weak semantic density. In this sense, it is not simply that there is a strengthening of semantic density through schooling, but rather there is a **growing range** of semantic density.[49] This means that the field regarding chemical matter becomes increasingly technical each year, and the higher the schooling level, the wider range of the technicality. The increasingly technical field of chemical matter presented by formulas at each year level suggests that the meanings are organized into the hierarchical knowledge structure of chemistry.

[49] Note that within Stage 5, Year 9 and 10 do not show any distinctions in the range of semantic density that formulas exhibit. This is because Year 10 does not introduce new knowledge of chemical matter and does not introduce new formulas.

In this section, we have mapped out the development of formulas across curriculum stages in terms of their field meanings and semantic density, which reveals how the meanings develop to build the knowledge of chemical matter. As mentioned earlier, in addition to the knowledge of chemical matter, chemical reactions are also a critical part of chemistry knowledge. Having completed tracing the development of formulas, we will now turn to chemical equations to examine how they develop to build the knowledge of chemical reactions.

4.2.2 The development of chemical equations across schooling levels to build the knowledge of chemical reactions

Chemical reactions are another significant component of chemistry knowledge. As noted in Section 4.1.1.3.1, diverse chemical equations are used to construe the knowledge of chemical reactions. These equations exhibit different degrees of complexity and are used in different schooling levels. To grasp their role in building the knowledge of chemical reactions, this section will first examine their varying degrees of complexity in terms of LCT's semantic density (Section 4.2.2.1) and then trace their development across schooling levels in secondary school chemistry textbooks to explore how chemical equations build the hierarchical knowledge structure of chemistry (Section 4.2.2.2).

4.2.2.1 Semantic density of chemical equations

As noted in Section 4.1.3.1.3, the meanings construed by chemical equations have been well accounted for in terms of field. Given these meanings are all technical, semantic density of chemical equations can be measured in terms of how many field-specific meanings they include. The following will thus step through the field meanings realized by the various equations used in secondary school chemistry and establish a model for their semantic density that can be enacted to trace the equations' development in Section 4.2.2.2.

As noted in Section 4.1.1.3.3.2, the simplest chemical equations are incomplete molecular equations without coefficients and state symbols. For example, '$H_2 + O_2 \rightarrow H_2O$' represents the reaction that hydrogen gas and oxygen gas react to yield water. In terms of field, it realizes an activity. As molecular formulas, for example 'H_2O', are components of chemical equations, they also construe compositional taxonomies as discussed in Section 4.2.1.1. Chemical equations with relatively stronger semantic density than the above equations are incomplete molecular equations with coefficients. For example, '$2H_2 + O_2 \rightarrow 2H_2O$' means that two units of hydrogen gas react with one unit of oxygen gas to form two units of water. This equation includes all field meanings realized by the above incomplete molecular equations without coefficients and state symbols and adds a property - a gauged property of the quantities of chemical species (for example, the coefficient '2' attached to H_2O shows that there are two units of water).

Pushing the semantic density to a higher level, complete molecular equations (with both coefficients and state symbols) add a qualitative property of physical states. For instance, in the complete molecular equation '$2H_2(g) + O_2(g) \rightarrow 2H_2O(l)$', the symbols '(g)' and '(l)' denote the physical states of gas and liquid, describing the qualities of the chemical species. Ionic equations, on the other hand, push the sematic density to a further higher level as they bring in two properties - a gauged property of the quantities of electric charges and a qualitative property of polarities of the charges. Take the ionic equation '$Mg(s) + 2H^+(aq) \rightarrow Mg^{2+}(aq) + H_2$' as an example. This equation represents the reaction that magnesium metal reacts with acids to yield hydrogen gas, in which one of the reactants and one of the products exist as ions, i.e., hydrogen cations (H^+) and magnesium cations (Mg^{2+}). The superscript number '2' represents the quantities of the electric charges the ions bear and '+' denotes the positive polarity of the charges. These two added properties imbue ionic equations with stronger semantic density than the complete molecular equations.

Chemical equations with stronger semantic density than ionic equations are reversible molecular equations that bring in a new activity – reverse reactions, which significantly raise the level of semantic density. For example, the reversible molecular equation '$N_2(g) + 3H_2(g) \rightleftharpoons 2NH_3(l)$' construes the reaction that nitrogen gas reacts with hydrogen gas to form ammonia liquid (the forward reaction), which then decomposes into nitrogen and hydrogen gases (the reverse reaction). Compared with the two minor properties brought in by ionic equations based on complete chemical equations as discussed above, the activity (reverse reactions) introduced in reversible equations imbues them with stronger semantic density because it also involves the field meaning of items constituting the activity, i.e., the reactant ammonia liquid (NH_3) and the products nitrogen gas (N_2) and hydrogen gas (H_2) in the reverse reaction.

Chemical equations embodying even stronger semantic density than the reversible molecular equations are reversible ionic equations, which include all field meanings realized by the reversible molecular equations and add two properties – a gauged property of the quantities of electric charges and a qualitative property of polarities of the charges. For example, the reversible ionic equation '$NH_3(aq) + H_2O(l) \rightleftharpoons NH_4^+(aq) + OH^-(aq)$' denotes the reaction that aqueous ammonia reacts with liquid water to form ammonium cations (NH_4^+) and hydroxyl anions (OH^-), which in turn react with each other to form aqueous ammonia and liquid water. In the ionic formulas, the elided number '1' construes the gauged property of the number of electric charges, and '+' represents the polarity of the charges. These two properties invest reversible ionic equations with stronger semantic density than reversible molecular equations.

Based on the above analysis, we can sequence the various semantic density of chemical equations in terms of field meanings they realize as Table 4.23.

Table 4.23 Semantic density of chemical equations used in secondary school chemistry

Semantic density (SD)	**Chemical equations**
++++++	reversible ionic equations
↑	reversible molecular equations
	ionic equations
	complete molecular equations
↓	incomplete molecular equations with coefficients
+	incomplete molecular equations without coefficients and state symbols

For ease of enacting the above model for the semantic density of chemical equations to compare the various equations, we assign different scales of semantic density to them: incomplete molecular equations without coefficients and state symbols (SD+), incomplete molecular equations with coefficients (SD++), complete molecular equations (SD+++), ionic equations (SD++++), reversible molecular equations (SD+++++), and reversible ionic equations (SD++++++). '+' denotes the strength of semantic density. The more '+', the stronger the semantic density. These scales constitute a tool for comparing and tracing the development of chemical equations in terms of their semantic density in secondary school chemistry textbooks, which allows for revealing how the equations develop across schooling levels to build the field of chemical reactions.

4.2.2.2 The development of chemical equations in secondary school chemistry textbooks

As pointed out by Taber (2009: 76), to succeed in learning the knowledge of chemical reactions, students are expected to progress through increasingly complex types of chemical equations that reflect subtle variations in form and meaning. Indeed, across the curriculum stages, different chemical equations occur in different schooling levels. As discussed in Section 4.2.2.1, these diverse chemical equations construe different knowledge of chemical reactions in a way that varies their complexity of meaning. To understand how chemical equations develop across schooling levels to build the knowledge of chemical reactions, this section will employ the tool developed in Section 4.2.2.1 to trace their development in terms of the field meanings they realize and their semantic density.

In Stage 4, chemical equations do not occur until Year 8. In this year, incomplete molecular equations without coefficients and state symbols are used to replace word equations, as shown in the following example:

> For example, the chemical equation for the reaction between iron and sulfur can be rewritten as:
>
> Word equation: iron + sulfur $\rightarrow$ iron sulfide
>
> Symbol equation: $Fe + S \rightarrow FeS$
>
> (Zhang et al. 2014a: 211)

The chemical equations at this stage are an alternative way to represent chemical reactions, which specify only the chemical species involved in the reactions, for example, iron (Fe), sulfur (S), and iron sulfide (FeS). As discussed in Section 4.2.1.1, in terms of field, these incomplete molecular equations construe compositional taxonomies and an activity, embodying the weakest semantic density (SD+). The chemical equations used in Stage 4 and their field meanings and semantic density can be summarized as Table 4.24.

Table 4.24 Chemical equations used in Stage 4 and their field-specific meanings and semantic density

Schooling levels		**Chemical equations**	**Field-specific meanings**	**Semantic density**
Stage 4	Year 7	-	-	From none to SD+
	Year 8	Incomplete molecular equations without coefficients and state symbols	Compositional taxonomies (compositions of chemical matter) An activity (a forward reaction)	

In Stage 5, the Year 9 textbook does not use any chemical equations as it focuses on knowledge of atomic structures and radioactivity. The Year 10 textbook, on the other hand, introduces new components of chemical equations based on the equations used in Year 8 and expands the complexity of the field construed by chemical equations. It first uses incomplete molecular equations with coefficients to introduce quantitative relations of reactants and products in chemical reactions, as illustrated in the following text:

> Methane gas (CH_4) is the main gaseous compound present in natural gas, which is used in the home for cooking and heating. When it burns, it combines with oxygen (O_2) in the air to form carbon dioxide (CO_2) and water (H_2O), which can be represented as:
>
> $CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O$
>
> The equation shows that one methane molecule combines with two oxygen molecules to produce one carbon dioxide molecule and two water molecules.
>
> (Zhang et al. 2015: 114)

As discussed in Section 4.2.1.1, in terms of field, this equation realizes all field meanings construed by the incomplete molecular equations without coefficients and state symbols (compositional taxonomies and an activity) and adds a gauged property of the number of chemical species involved in reactions, imbuing it with stronger semantic density (SD++). Building upon the incomplete molecular equations with coefficients, the latter part of the Year 10 textbook introduces the complete molecular equations that include both coefficients and state symbols to add another component of the knowledge of chemical reactions – physical states of chemical species involved in the reactions. For example, the equation in the following text specifies that iron and ferric hydroxide are in solid state [‘(s)’], water is in liquid state [‘(l)’] and oxygen molecules are in gas state [‘(g)’].

> Corrosion is generally associated with a metal reacting with oxygen and water. The corrosion you will be most familiar with is the corrosion of iron or steel known specifically as rusting.
>
> The word equation for this reaction is:
>
> iron + water + oxygen → rust
>
> The chemical equation is:
>
> $4Fe(s) + 6H_2O(l) + 3O_2(g) \rightarrow 4Fe(OH)_3(s)$
>
> (Zhang et al. 2015: 125)

As discussed in Section 4.2.2.1, the added qualitative property of physical states of chemical species invests the complete molecular equations stronger semantic density (SD+++) than the incomplete molecular equations. Table 4.25 summarizes chemical equations used in Stage 5 and the field-specific meanings construed and their semantic density.

Table 4.25 Chemical equations used in Stage 5 and the field-specific meanings realized and their semantic density

Schooling levels		**Chemical equations**	**Field-specific meanings**	**Semantic density**
Stage 5	Year 9	-	-	From SD ++ to SD+++
	Year 10	Incomplete molecular equations with coefficients	Compositional taxonomies (compositions of chemical matter) An activity (a forward reaction) A gauged property (quantities of chemical species)	
		Complete molecular equations	Compositional taxonomies (compositions of chemical matter) An activity (a forward reaction) A gauged property (quantities of chemical species) A qualitative property (physical states of chemical species)	

In Stage 6, the Year 11 textbook continues to use the complete molecular equations occurring in Year 10 and introduces a new type of chemical equations - ionic equations - to represent ionic reactions. For example, the following text uses an ionic equation to represent the disposition of silver cations (Ag^+) by copper (Cu).

> When a coil of copper metal is placed in a solution of silver nitrate, the copper becomes coated with a black deposit, as shown in the flask on the right in Figure 11.1.10. After a while, the solution turns a pale blue colour. When the deposit is tested, it can be shown to be silver. The blue colour of the solution is due to the presence of copper (II) ions (Cu^{2+}). The copper metal has displaced the Ag^+ ions from the solution of silver nitrate. The ionic equation for this reaction is:
>
> $$Cu(s) + 2Ag^+(aq) \rightarrow Cu^{2+}(aq) + 2Ag(s)$$
>
> (Chan et al. 2018: 357)

As discussed in Section 4.2.2.1, ionic formulas, for example Cu^{2+}, add two field properties - a gauged property of the number of electric charges and a qualitative property of polarities of the charges - to ionic equations, imbuing them with stronger semantic density (SD++++) than the complete molecular equations.

In Stage 6, the Year 12 textbook uses all the chemical equations occurring in Year 11 and introduces a new type of chemical equation to construe the knowledge of reversible reactions, i.e., reversible equations. Year 12 includes two subtypes of reversible equations. One is reversible molecular equations, which represent reversible reactions involving only molecular substances. For example, the following text uses a reversible molecular equation to represent the reversible reaction between dinitrogen tetroxide (N_2O_4) and nitrogen dioxide (NO_2).

> The decomposition of dinitrogen tetroxide (N_2O_4) to nitrogen dioxide (NO_2) is an example of a reversible reaction that reaches a dynamic equilibrium. The progression of this reaction from pure N_2O_4 to the equilibrium mixture containing both N_2O_4 and NO_2 can be monitored through the changing colour of the gases in the reaction vessel. N_2O_4 is colourless and NO_2 is dark brown.
>
> The reaction occurs according to the equation:
>
> $$N_2O_4 (g) \rightleftharpoons 2NO_2 (g)$$
>
> (Chan et al. 2019: 61)

In terms of field, the reversible molecular equation construes two activities: a forward reaction, i.e., dinitrogen tetroxide molecules (N_2O_4) decompose into nitrogen dioxide molecules (NO_2), and a reversible reaction, i.e., nitrogen dioxide molecules (NO_2) combine to form dinitrogen tetroxide molecules (N_2O_4). As discussed in Section 4.2.2.1, this reversible molecular equation embodies stronger semantic density (SD+++++) than ionic equations used in Year 11.

The Year 12 textbook also employs another type of reversible equation - reversible ionic equations, which are used to construe reversible reactions involving ions. For example, the following text uses the reversible ionic equation to represent the dissociation of ammonia.

Ammonia is a covalent molecular compound that dissociates in water by accepting a proton. This dissociation can be represented by the equation:

$$NH_3(aq) + H_2O(l) \rightleftharpoons NH_4^+(aq) + OH^-(aq)$$

(Chan et al. 2019: 165)

In terms of field, the equation involves two activities: the forward reaction that ammonia (NH_3) reacts with water (H_2O) to form ammonium cations (NH_4^+) and hydroxyl anions (OH^-) and the reverse reaction that the two ions react to yield ammonia and water. As discussed in Section 4.2.1.1, the ionic formulas invest this equation with stronger semantic density (SD++++++) than the reversible molecular equations. The chemical equations used in Stage 6 and the field-specific meanings realized and their semantic density are summarized as Table 4.26.

Table 4.26 Chemical equations used in Stage 6 and the field-specific meanings realized and their semantic density

<table>
<tr><th colspan="2">Schooling levels</th><th>Chemical equations</th><th>Field-specific meanings</th><th>Semantic density</th></tr>
<tr><td rowspan="6">Stage 6</td><td rowspan="2">Year 11</td><td>Complete molecular equations</td><td>Compositional taxonomies (compositions of chemical matter)
An activity (a forward reaction)
A gauged property (quantities of chemical species)
A qualitative property (physical states of chemical species)</td><td rowspan="2">From SD+++ to SD+++++</td></tr>
<tr><td>Ionic equations</td><td>Compositional taxonomies (compositions of chemical matter)
An activity (a forward reaction)
A gauged property (quantities of chemical species)
A qualitative property (physical states of chemical species)
A gauged property (the number of electric charges)
A qualitative property (polarities of electric charges)</td></tr>
<tr><td rowspan="4">Year 12</td><td>Complete molecular equations</td><td>Compositional taxonomies (compositions of chemical matter)
An activity (a forward reaction)
A gauged property (quantities of chemical species)
A qualitative property (physical states of chemical species)</td><td rowspan="4">From SD +++ to SD++++++</td></tr>
<tr><td>Ionic equations</td><td>Compositional taxonomies (compositions of chemical matter)
An activity (a forward reaction)
A gauged property (quantities of chemical species)
A qualitative property (physical states of chemical species)
A gauged property (the number of electric charges)
A qualitative property (polarities of electric charges)</td></tr>
<tr><td>Reversible molecular equations</td><td>Compositional taxonomies (compositions of chemical matter)
An activity (a forward reaction)
A gauged property (quantities of chemical species)
A qualitative property (physical states of chemical species)
An activity (reverse reaction)</td></tr>
<tr><td>Reversible ionic equations</td><td>Compositional taxonomies (compositions of chemical matter)
An activity (a forward reaction)
A gauged property (quantities of chemical species)
A qualitative property (physical states of chemical species)
An activity (a reverse reaction)
A gauged property (the number of electric charges)
A qualitative property (polarities of electric charges)</td></tr>
</table>

Based on the above analysis, we can bring the three stages in one table, which enables an overview of the development of chemical equations across schooling levels in terms field meanings realized and their semantic density. As shown in Table 4.27, the words in bold fonts mean that the field meanings are newly developed compared with the year before.

Table 4.27 The development of chemical equations across the three curriculum stages in terms of their field-specific meanings and semantic density

Schooling levels		**Chemical equations**	**Field-specific meanings**	**Semantic density**
Stage 4	Year 7	-	-	None
	Year 8	Incomplete molecular equations without coefficients and state symbols	Compositional taxonomies (compositions of chemical matter) An activity (a forward reaction)	From none to SD+
Stage 5	Year 9	-	-	None
	Year 10	Incomplete molecular equations with coefficients	Compositional taxonomies (compositions of chemical matter) An activity (a forward reaction) **A gauged property (quantities of chemical species)**	From SD ++ to SD+++
		Complete molecular equations	Compositional taxonomies (compositions of chemical matter) An activity (a forward reaction) A gauged property (quantities of chemical species) **A qualitative property (physical states of chemical species)**	
Stage 6	Year 11	Complete molecular equations	Compositional taxonomies (compositions of chemical matter) An activity (a forward reaction) A gauged property (quantities of chemical species) A qualitative property (physical states of chemical species)	From SD+++ to SD++++
		Ionic equations	Compositional taxonomies (compositions of chemical matter) An activity (a forward reaction) A gauged property (quantities of chemical species) A qualitative property (physical states of chemical species) **A gauged property (the number of electric charges)** **A qualitative property (polarities of electric charges)**	
	Year 12	Complete molecular equations	Compositional taxonomies (compositions of chemical matter) An activity (a forward reaction) A gauged property (quantities of chemical species) A qualitative property (physical states of chemical species)	From SD +++ to SD++++++
		Ionic equations	Compositional taxonomies (compositions of chemical matter) An activity (a forward reaction) A gauged property (quantities of chemical species) A qualitative property (physical states of chemical species) A gauged property (the number of electric charges) A qualitative property (polarities of electric charges)	
		Reversible molecular equations	Compositional taxonomies (compositions of chemical matter) An activity (a forward reaction) A gauged property (quantities of chemical species) A qualitative property (physical states of chemical species) **An activity (reverse reaction)**	
		Reversible ionic equations	Compositional taxonomies (compositions of chemical matter) An activity (a forward reaction) A gauged property (quantities of chemical species) A qualitative property (physical states of chemical species) An activity (a reverse reaction) **A gauged property (the number of electric charges)** **A qualitative property (polarities of electric charges)**	

The above table shows that from Stage 4 to 5, the chemical equations develop from incomplete molecular equations without coefficients and state symbols to the complete ones, adding two field properties and exhibiting a consecutive growth in semantic density. Years 11 and 12 in Stage 6 build upon the complete molecular equations and introduces new chemical equations, expanding the range of field meanings concerning chemical reactions. Year 11 introduces ionic equations and adds two field properties, while Year 12 brings in reversible equations and adds one activity and two field properties. The field construed by chemical equations becomes increasingly complex and the semantic density strengthens as learning progresses, which is shown more clearly in Figure 4.57.

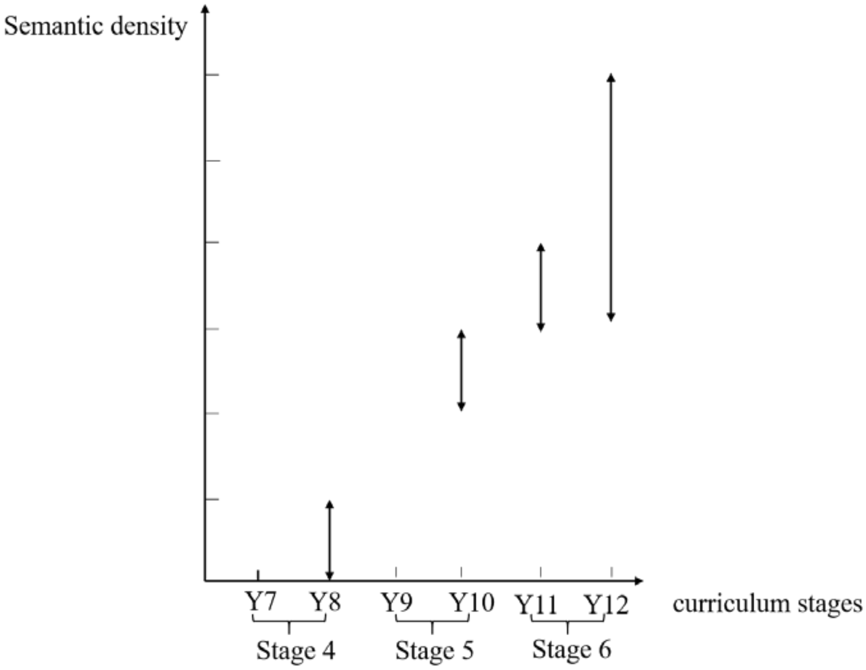


Figure 4.57 The development of chemical equations' semantic density across the curriculum stages

As shown in Figure 4.57, there is a consecutive progression in semantic density from Stage 4 to 5 and a widening range of semantic density from Year 11 to 12 in Stage 6. This suggests a consecutive growth in the complexity of the field of chemical reactions from Stage 4 to 5 and a widening range of the field's complexity in Stage 6. From Year 8 to 10, the field construed by chemical equations develops from compositional taxonomies and an activity (incomplete molecular equations) to compositional taxonomies, an activity, a gauged property, and a qualitative property (complete molecular equations). In Stage 6, the field construed by the complete molecular equations constitutes the basis on which Year 11 and 12 expand the range of the field meanings, with Year 12 bringing in significantly more meanings. Overall, with the growth of field meanings throughout the schooling levels, chemical equations progress to their increasingly complex forms and build the increasingly complex knowledge of chemical reactions. Such a development of chemical equations across schooling levels builds the hierarchical knowledge structure of chemistry.

4.3 Summary

This chapter has examined the first research question of this study – how do chemical formalisms in secondary school chemistry textbooks build the knowledge of chemistry. This question includes two sub-questions: (1) how are chemical formalisms organized to make meaning? (2) how do the formalisms develop across schooling levels to build the hierarchical knowledge structure of chemistry in the textbooks? To address the first sub-question, this chapter has probed the grammars of chemical equations, chemical formulas, and structural formulas in Sections 4.1.1 and 4.1.2. It has established grammatical systems for these formalisms and generated their metafunctions based on the axial principle – the interaction between the paradigmatic and syntagmatic relations (system and structure). It has also developed a functional semiotic typology for the different chemical formalisms in terms of their similarities and differences in making meaning.

The description of the chemical formalisms shows that the grammar of chemical equations includes four ranks: equation, term, formula, and symbol, at which the grammatical systems organize primarily the experiential and logical metafunctions. The grammar also realizes a textual component but only at the term rank. Since chemical formulas are a part of chemical equations, the grammatical systems at the formula and symbol ranks are also chemical formulas' grammar. The description of structural formulas' grammar shows that it involves two ranks: formula and symbol. Similarly, the grammatical systems of structural formulas also organize primarily the experiential and logical metafunctions. The two descriptions suggest that chemical formalisms are designed mainly for construing the technical knowledge of chemistry. In terms of grammatical analysis, this study finds two special structures. One is the subjacency structure, a relation of non-iterative nuclear dependency. This structure tends to occur in both chemical formulas and structural formulas. The other is the ideationalized prosodic structures, typically occurring in structural formulas and chemical equations to construe ideational meaning of bonding relations between atoms, molecular polarities and quantitative relations of chemical species involved in chemical reactions.

The functional semiotic typology analysis of the chemical formalisms is based on the comparison of the similarities and differences in their grammatical organizations, i.e., univariate structures or multivariate structures, and the field-specific meanings they realize. The finding shows that chemical formulas and structural formulas are similar in that both are dominated by univariate structures, each capable of iterating chemical symbols multiple times. However, in contrast to chemical formulas, structural formulas can iterate in a two-dimensional space, which affords a reading of spatial layout. This leads to a distinction in the field-specific meanings they realize. Whereas chemical formulas realize only compositional taxonomies, the two-dimensional organization of the structural formulas realizes both a compositional taxonomy and their spatial properties. The two formulas are also similar in that they both involve subjacency structures. However, they employ these structures to realize different field-specific meanings: gauged properties of the number of atoms for chemical formulas and qualitative properties of partial charge polarities for structural formulas. In contrast to chemical formulas and structural formulas dominated by univariate structures, chemical equations are organized largely around multivariate structures. In terms of field, these multivariate structures construe activities of different types of reactions at the equation rank and various properties and quantitative relations of reactions at the term rank. In addition to the discussion of similarities and differences in meaning-making in terms of univariate and multivariate structures, chemical formalisms can also be compared in terms of whether they involve ideationalized prosodic structures. The finding shows that both structural formulas and chemical equations employ prosodic structures to construe qualitative properties, in contrast to chemical formulas that do not involve such prosodic structures. The functional semiotic typology analysis of the chemical formalisms indicates that they complement each other to construe the chemistry knowledge of chemical substances and chemical reactions.

To address the second sub-question – how do chemical formalisms develop across schooling levels to build the hierarchical knowledge structure of chemistry in secondary school chemistry textbooks, Section 4.2 has traced the development of the formalisms' meanings accounted for in Section 4.1 across the three curriculum stages in the textbooks. In specific, it has analyzed how formulas and chemical equations develop across the stages to build the two crucial components of chemistry knowledge -

chemical matter and chemical reactions – in Sections 4.2.1 and 4.2.2 .

Section 4.2.1 has first established a model for the semantic density of formulas used in secondary school chemistry in terms of the field-specific meanings they realize and then used the model to trace the formulas' development in terms of their semantic density. The finding shows that the semantic density of the various formulas used in the textbooks can be sequenced as five scales, with molecular formulas embodying the weakest strength and stereo formulas exhibiting the strongest strength. This continuum of scales constitutes a tool for tracing the development of the formulas' semantic density across curriculum stages in the NSW secondary school chemistry textbooks. Enacting the tool reveals that formulas with different degrees of semantic density are used in different curriculum stages. They exhibit a widening range of semantic density across the three curriculum stages, with the bottom of the range being the most basic molecular formulas construing the compositions of chemical matter and the top reaching increasingly strong semantic density that construes more and more complex fields of chemical matter. This indicates that to build the knowledge of chemical matter, the formulas used in each stage maintain connections to the simpler technical field meanings and establish increasingly technical fields of chemical matter. From the simpler to the more complex field, the meanings presented by formulas are organized into the hierarchical knowledge structure of chemistry.

Section 4.2.2 has examined how chemical equations develop across schooling levels to build the knowledge of chemical reactions. Similar to Section 4.2.1, it has first established a model for the semantic density of the various chemical equations used in secondary school chemistry in terms of field and enacted the model to analyze how the equations develop across schooling levels to build the knowledge of chemical reactions. The finding shows that the semantic density of the chemical equations can be sequenced as six scales along a continuum, with incomplete molecular equations embodying the weakest strength and reversible ionic equations exhibiting the strongest strength. This continuum of scales constitutes a tool for tracing the development of the semantic density of chemical equations across curriculum stages in the NSW secondary school chemistry textbooks. Enacting the tool to analyze the development of the equations reveals that there is a consecutive increase in chemical equations' semantic density from Stage 4 to 5, expanding the field of chemical reactions that involves only compositional taxonomies (compositions of chemical species) and an activity (a forward reaction) to the field that includes two extra properties concerning chemical reactions (quantities of chemical species involved in the reactions and their physical states). Building upon the complete molecular equations used in Stage 5, Years 11 and 12 in Stage 6 exhibit a widening range of semantic density, with Year 11 introducing ionic equations and Year 12 bringing in reversible equations, construing increasingly complex fields of chemical reactions. The growth of chemical equations' semantic density across the school levels builds the hierarchical knowledge structure of chemistry.

Chapter 5 Knowledge-building through Images

Images are ubiquitous in chemistry. Photographs, diagrams, graphs, and charts are widely used to represent chemistry knowledge and form a crucial component of the texts through which students learn that knowledge. Therefore, together with chemical formalisms, images constitute one of the key ‘knowledge builders’ of chemistry. To better inform the development of pedagogy that enables students to learn the technical knowledge of chemistry by reading the images, exploring the knowledge-building of chemistry through images is crucial. As discussed in Chapter 1, examining the knowledge-building of chemistry from a semiotic perspective entails first investigating the construal of meaning by a semiotic resource and then exploring how it develops through schooling to build the hierarchical knowledge structure of chemistry. This chapter will thus approach the knowledge-building of chemistry through images by first examing the meanings construed by the images in terms of field in SFL (Doran and Martin 2021) and the grammatical resources realizing these meanings (Section 5.1), and then probe how the images develop to build the hierarchical knowledge structure of chemistry (Section 5.2).

5.1 Construing meaning: a model of images in chemistry from the perspective of field

As a key type of visual semiotic resource, images are used in various disciplines. They afford a particular set of meaning potentials specific to each discipline. In other words, images have “disciplinary affordance” (Airey and Linder 2017). For example, images afford distinct meaning potentials between fine art and science, with the former typically presenting aesthetics or beauty while the latter tending to show logic relations and data. To capture the disciplinary affordance of images in chemistry, we explore the meanings construed by images in chemistry in terms of field (Doran and Martin 2021) and the grammatical resources realizing these meanings. It will show that chemical images realize three broad types of field-specific meanings: item, activity, and property, each of which involves more delicate field meanings realized by a particular set of grammatical resources. This section will gradually build up the model by first examining the static item (Section 5.1.1), then the dynamic activity (Section 5.1.2), and finally, the property describing both item and activity (Section 5.1.3). Section 5.1.4 brings these categories into one complete network that links with the grammatical resources realizing the meanings.

5.1.1 Item

As noted in Section 4.2, chemical matter is a significant component of chemistry knowledge. In terms of field, the most pertinent field-specific meaning is item. To construe the knowledge of chemical matter, chemical images organize different types of meaning in relation to item. Firstly, they can construe "**singular item**" (Doran and Martin 2021). Chemical images realizing this specific-specific meaning tend to be diagrams showing a single particle in the microscopic world. For example, Figure 5.1, a diagrammatic representation of a helium atom, displays a single entity. Grammatically, the field-specific meaning of singular item is typically realized by images presenting an intact entity that does not involve internal structures.

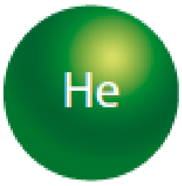


Figure 5.1 A diagrammatic representation of a helium atom

In addition to singular item, chemical images display a strong capacity for construing taxonomic relations between multiple items. One of these relations is **compositional taxonomies**, which view relations between items in terms of part-whole relations. Chemical images construing this taxonomy often illustrate compositions of chemical matter. For example, Figure 5.2, a diagrammatic representation of an oxygen gas molecule, shows that the molecule is composed of two oxygen atoms.

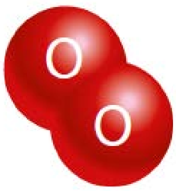


Figure 5.2 A diagrammatic representation of an oxygen gas molecule

Grammatically, this composition taxonomy is realized by an analytical structure including a Carrier (the whole molecule, denoted in yellow in Figure 5.3) and two Possessive Attributes (the compositional atoms, shown in green) (Kress and van Leeuwen 2006: 87). This structure realizes one level of composition taxonomy, with the molecule being the whole and the oxygen atoms being the parts.

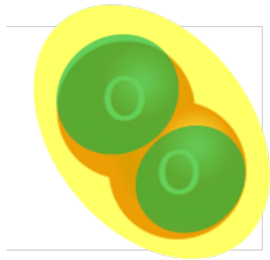


Figure 5.3 Composition analysis of Figure 6.2 (Carrier is shown in yellow and Possessive Attributes in green)

Chemical images can also construe multiple levels of compositional taxonomies. This is particularly important when representing chemical matter involving compositions with one embedded in another. Figure 5.4, for example, shows a lithium atom's atomic structure, which can be interpreted as two levels of composition - the lithium atom is composed of three electrons (the blue balls) and one nucleus (the whole of the red and green balls); the nucleus is in turn composed of three protons (the red balls) and four neutrons (the green balls).

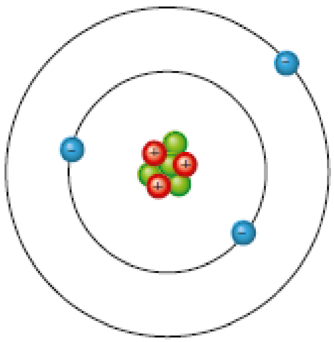


Figure 5.4 A diagrammatic representation of the atomic structure of lithium atoms

Grammatically, these two levels of compositional taxonomies are realized by one analytical structure embedded with another. The first-level analytical structure is constituted by one Carrier (the entire atom, shown in yellow in Figure 5.5) and two Possessive Attributes (the electrons and nucleus, denoted in light blue). At the second level, one of the Possessive Attributes itself, the nucleus, involves an analytical structure constituted by one Carrier (the entire nucleus, shown in light blue) and two Possessive Attributes (protons and neutrons, shown in black).

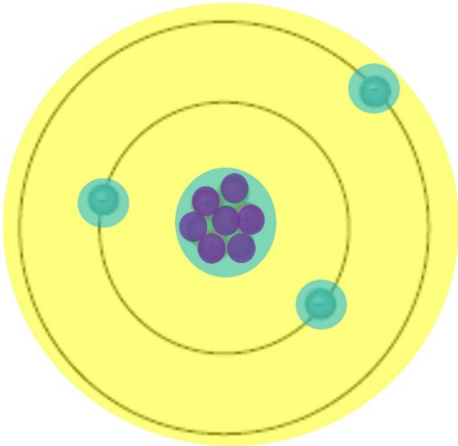


Figure 5.5 Composition analysis of Figure 5.4 (Carrier is shown in yellow and Possessive Attributes in light blue at the first level; Carrier is shown in light blue and Possessive Attributes in black at the second level)

In term's of Kress and van Leeuwen's (2006) classifications of analytical structures, the first level is a conjoined analytical structure (the parts are separated), while the second level is a compounded analytical structure (the parts are welded together, while at the same time retaining their distinct entities). Kress and van Leeuwen identified these different analytical structures but did not explicitly propose that one analytical structure can be recursively embedded within another. Here we argue that **recursive embeddings** of different analytical structures exist, and they are crucial for chemistry to construe the **depth** of compositional relations within chemical matter.

These embedded structures show not only chemical matter's different levels of composition but also the components' relative locations. Figure 5.4, for instance, shows that electrons circle the nucleus, and protons and neutrons are connected within the nucleus at the center. These spatial relations are modeled as "spatial properties" in terms of field (Doran and Martin 2021), which will be discussed in more detail in Section 5.1.3. A key point to note here is that when we attempt to present compositions of things using images, we must necessarily show their physical relations. This is because images have their **epistemological commitment** (Kress 2003). As a visual mode, they have to show the relative locations between the components. Language, in contrast, does not involve such unavoidable affordance. It can describe compositions without presenting their spatial relations, for example, *a lithium atom is composed of three electrons and one nucleus, which in turn is composed of three protons and four neutrons*.

In addition to compositional taxonomies, chemical images also construe **classificational taxonomies** between items. These images tend to present classifications of chemical matter. For instance, Figure 5.6 illustrates three types of matter status at the microscopic level. Particles constituting a substance are close to each other in solid status, less so in liquid status, and distanced from each other in gas status. These three classifications form the basic status of a substance.

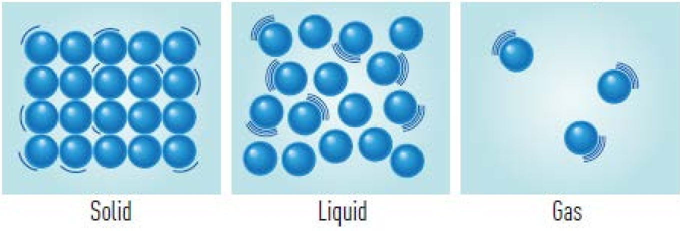


Figure 5.6 Three types of matter status

Grammatically, the classification relation is realized by a "covert taxonomy" structure (Kress and van Leeuwen 2006: 79), in which the superordinate (matter status) is not shown, and the three subordinates (solid, liquid, and gas) are distributed along the horizontal axis at equal distance from each other. This covert taxonomy structure realizes one level of classification, with the covert 'matter status' as the superordinate and the three statuses as the subordinates.

In addition to the one-level classification, chemical images are rather powerful in construing multiple levels of classificational taxonomies. Figure 5.7, for instance, construes two levels of classification. The first level involves one superordinate ('pure substances') and two subordinates ('elements' and 'compounds'). Each subordinate itself is the superordinate at the second level. 'Elements' have three subordinates: 'monatomic', 'diatomic', and 'lattice', and 'compounds' also have three subordinates: 'molecular', 'polymer', and 'lattice'. Grammatically, the two levels of classifications are realized by a "multi-leveled overt taxonomy structure" (Kress and van Leeuwen 2006: 86), which connects superordinates and subordinates through a tree structure, explicitly showing the hierarchical relations between the different types of substances.

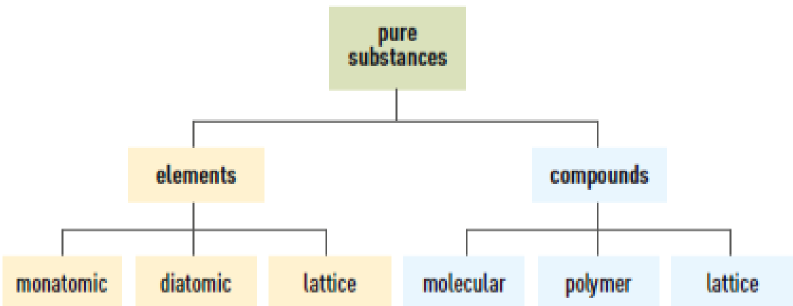


Figure 5.7 Classifications of pure substances

The two levels of classification can also be realized by a **multi-leveled covert taxonomy structure**. A typical example is the periodic table of chemical elements, as shown in Figure 5.8.

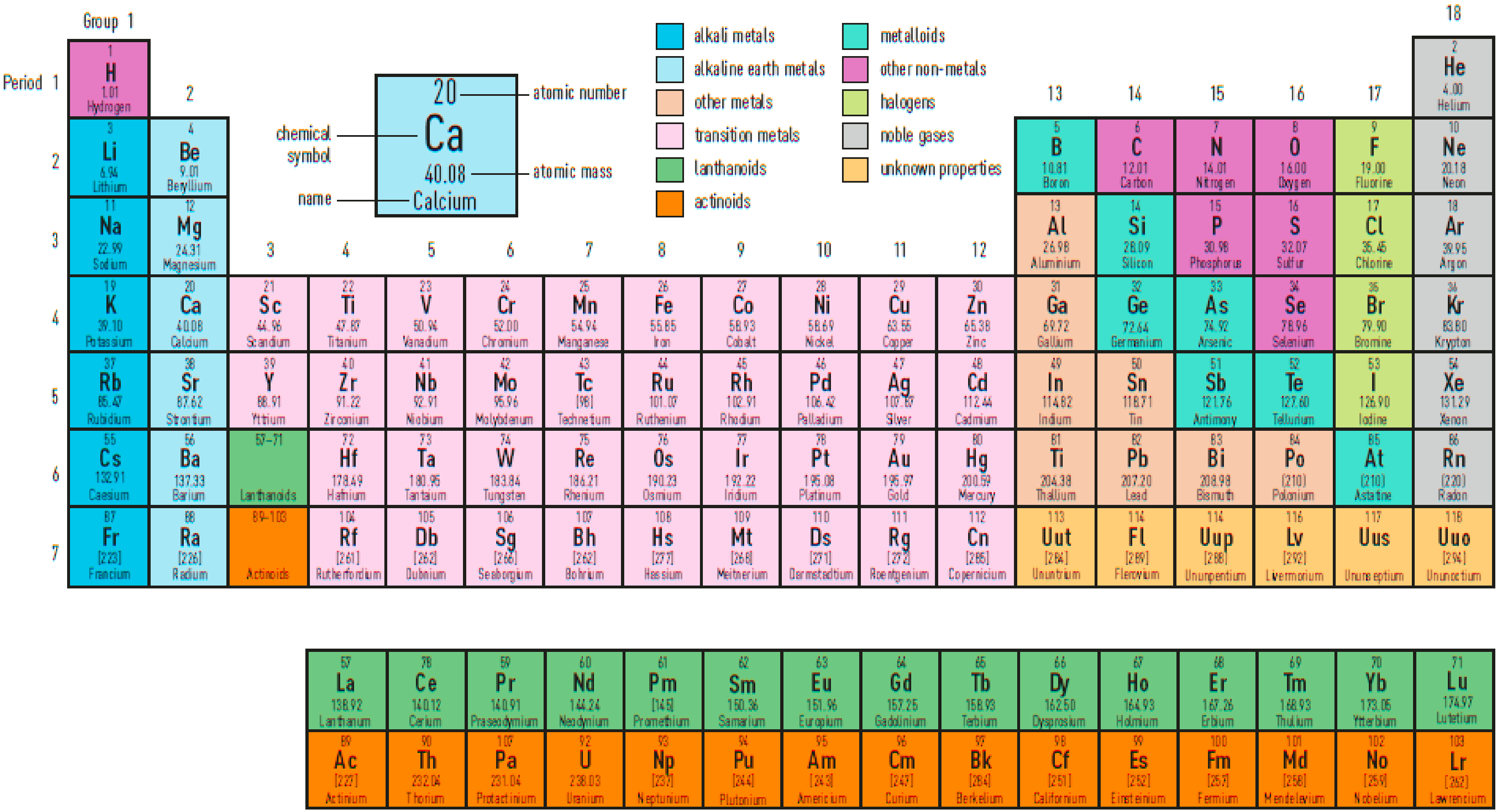


Figure 5.8 The periodic table of chemical elements

A periodic table is a tabular display of chemical elements arranged by atomic numbers and recurring chemical properties (Brown et al. 2012: 50). Figure 5.8 marks chemical elements belonging to one group as the same color. The column in the middle-upper part lists the type of chemical element each color represents. In terms of field, this periodic table realizes two levels of classificational taxonomies. At the first level, the superordinate is 'chemical elements', which is covert and not shown in the table. This superordinate has four subordinates: 'metals', 'metalloids', 'non-metals', and 'unknown properties'.[50] Among the four subtypes, 'metalloids' and 'unknown properties' are overt, shown in and , while 'metals' and 'non-metals' are covert and not explicitly shown in the table. At the second level, 'metals' and 'non-metals' become the superordinates. 'Metals' includes six subordinates: 'alkali metals' (shown in), 'alkaline earth metals' (shown in), 'lanthanoids' (shown in), 'actinoids' (shown in), 'transition metals' (shown in), and 'other metals' (shown in). 'Non-metals' includes three subtypes: 'halogens' (shown in), 'noble gases' (shown in), and 'other non-metals' (shown in).

Grammatically, the two levels of classification are realized by two covert taxonomy structures organized along a cline of delicacy. In the bottom covert taxonomy structure, chemical elements categorized as one group are in the same color and arranged together primarily on the vertical axis (some are also arranged on both vertical and horizontal axes, for example, 'transition metals'). These groups form the subordinates of 'metals' and 'non-metals', which are not made explicit in the table. In the top covert taxonomy structure, the two covert types, 'metals' and 'non-metals', are separated by the two overt types, 'metalloids' and 'unknown properties', in the table. The boxes representing each of the four types are arranged together as a cluster, becoming subordinates of the highest covert superordinate - 'chemical elements'. The two covert taxonomy structures are embedded within the periodic table, presenting increasing specific classifications of chemical elements.

In Kress and van Leeuwen's (2006: 79–87) description of classificational structures, they distinguish between single-leveled and multi-leveled only within overt taxonomy structures. The above analysis of the periodic table suggests that covert taxonomy structures can also involve a distinction between single-leveled and multi-leveled. Based on this, we propose the system of **CLASSIFICATIONAL STRUCTURE** for images, as shown in Figure 5.9. The network comprises two simultaneous subsystems:

[50] A metalloid is a type of chemical element which has properties in between those of metals and non-metals (Brown et al. 2012: 51).

TYPE and DELICACY. TYPE consists of two features: covert taxonomy and overt taxonomy, and DELICACY includes two topological features: single-leveled and multi-leveled. Together they describe images' different types of classificational structures.

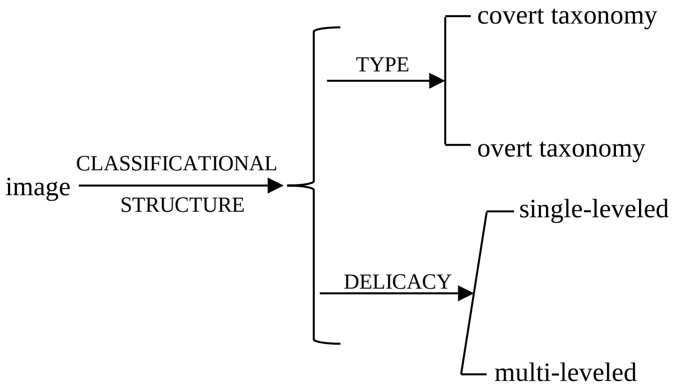


Figure 5.9 The system of CLASSIFICATIONAL STRUCTURE for images

Thus far, we have discussed the two taxonomic relations between items chemical images construe. These taxonomies occur independently in the images. However, this does not mean that chemical images cannot construe both taxonomies at the same time. In some chemical images, composition and classification relations can be realized simultaneously in a single snapshot. For example, Figure 5.10, a diagrammatic representation of oxygen atoms' atomic structure, construes two-level compositions and a one-level classification. We will not analyze the compositional taxonomies here since they are similar to Figure 5.4, which has already been analyzed. The one-level classification is between 'charged particles' and the two subtypes: 'protons' (positively charged) and 'electrons' (negatively charged). This classification is realized by a covert taxonomy structure, in which the superordinate 'charged particles' is covert, and the two subordinates 'protons' and 'electrons' are arranged on the vertical axis on the top-left part of the diagram.

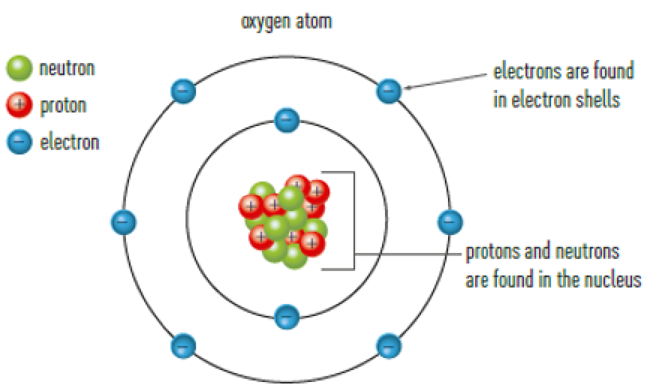


Figure 5.10 A diagrammatic representation of oxygen atoms' atomic structure

In this section, we have discussed the types of field-specific meanings chemical images construe in terms of item and the grammatical structures realizing these meanings. The analysis shows that chemical images construe either singular item or taxonomic relations between multiple items. Notably, they are rather powerful in construing multiple levels of composition and classification taxonomies. In terms of composition, chemical images can establish multiple levels of part-whole relations between components that constitute a chemical matter. These multiple levels of composition are realized by a spatial analytical structure recursively embedded within another, allowing for construing as complex compositions of chemical matter as possible. In terms of classification, chemical images can set up a cline of type-subtype relation between different types of chemical matter. These classification relations can be realized by either multi-leveled overt taxonomy structures or multi-leveled covert taxonomy structures. The composition and classification taxonomies can occur simultaneously in a single chemical image. The network shown in Figure 5.11 summarizes the field-specific meanings construed by chemical images in terms of item and the grammatical structures realizing these meanings.

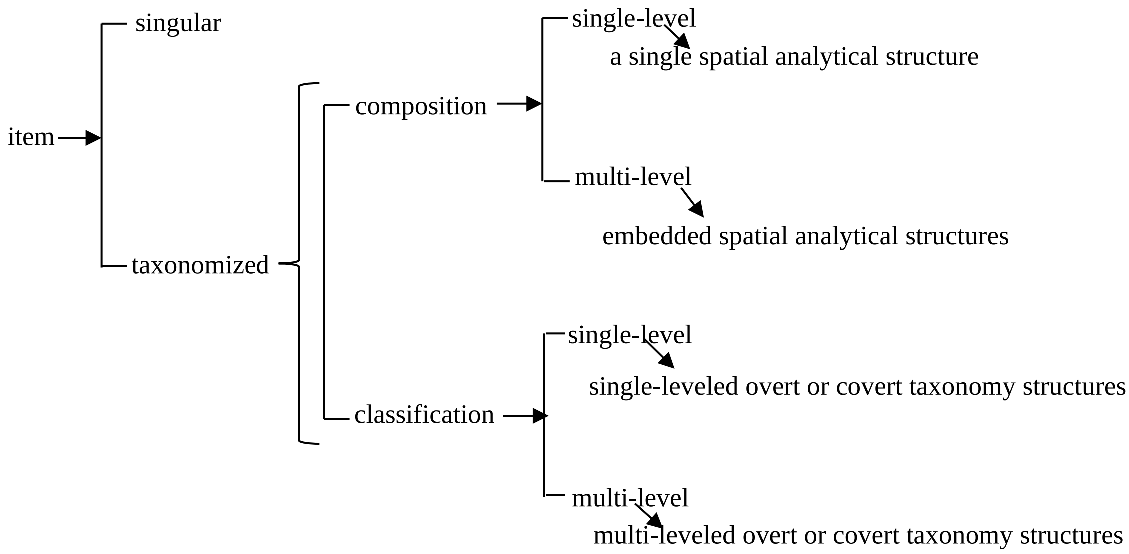


Figure 5.11 The network of field-specific meanings in terms of item realized by chemical mages (means that 'composition' and 'classification' can be selected either disjunctively or conjunctively)

5.1.2 Activity

In addition to the above static perspective of field, chemical images also realize activities. This dynamic perspective of field constitutes a crucial part of the knowledge of chemical reactions. This section will thus focus on what types of activities images in chemistry construe and the grammatical resources realizing these activities.

Among the activities construed by chemical images, the simplest is a single activity, termed “unmomented activity” in Doran and Martin’s (2021) model of field. Chemical images construing unmomented activities tend to illustrate a single chemical reaction. Figure 5.12, for example, presents a generalized synthesis reaction: reactants ‘A’ and ‘B’ combine to form product ‘AB’.

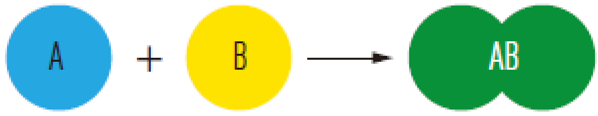


Figure 5.12 A generalized synthesis reaction

Grammatically, the unmomented activity is realized by a non-agentive narrative process, which does not construe actions but processes of change. It does not involve Actor and Goal because the vector (shown in red in Figure 5.13) does not signify ‘do to’ or ‘aim at’. Instead, it realizes a ‘transformation’ from the Participant on the left (shown in yellow) to the one on the right (shown in blue). In Kress and van Leeuwen’s (2006: 68–69) description of non-agentive narrative processes, they discuss only multiple activities that relay one another to form a cycle, which they term conversion processes. The Participants involved in these processes are called Relay. However, it is not yet clear how a single non-agentive narrative process expressing transformation could be modeled. To analyze this image, it is useful to make an addition to their grammar. This addition distinguishes between the Participants before and after the transformation. We term the Participants before and after the transformation **Source** and **Product** and refer to this non-agentive narrative process as a **transformation process**.

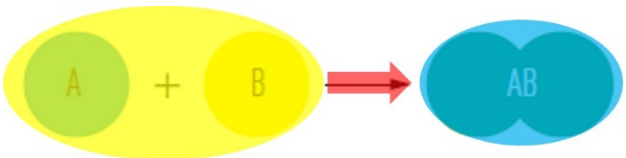


Figure 5.13 Analysis of the non-agentive process (Source is shown in yellow, Vector in red, and Product in blue)

Transformation processes are one of the key types of processes realized by chemical images in secondary school textbooks. This is because they construe the meaning of change, which is the essence of chemical reactions (Brown et al. 2012: 30). In contrast to the above chemical image involving a single transformation process, chemical images often include multiple transformation processes to show how reactants become products through a series of chemical reactions. In terms of field, such images construe "momented activities", a series of activities related through implication or expectancy (Doran and Martin 2021:114). For example, Figure 5.14 shows the two steps by which 'C(s) + O_2(g)' changes to 'CO_2(g)'. The two steps form momented activities - "C(s) + O_2(g)' first changes to 'C(s) + $\frac{1}{2}O_2$(g)', which in turn changes to 'CO_2(g)'. The two activities are related through expectancy, a relation of expectation that one event will follow another (Martin 1992: 324). That is, the activity that 'C(s) + $\frac{1}{2}O_2$(g)' changes to 'CO_2(g)' follows the activity that 'C(s) + O_2(g)' changes to 'C(s) + $\frac{1}{2}O_2$(g)'.

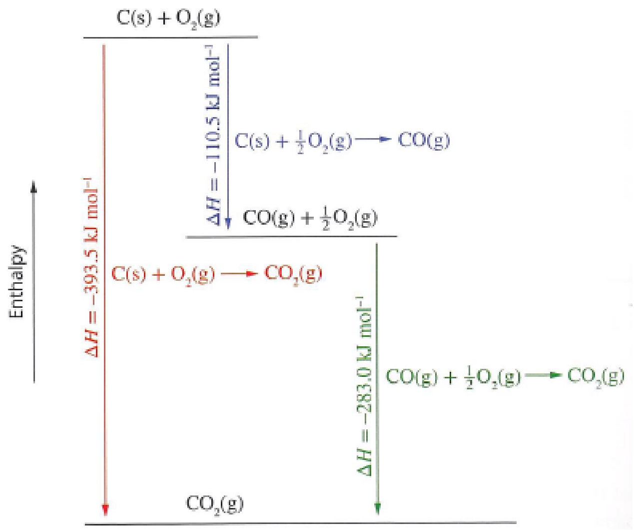


Figure 5.14 An energy level diagram for the formation of carbon dioxide from carbon and oxygen via carbon monoxide

Grammatically, the momented activities are realized by a **complex of transformation processes**. In the first transformation process, 'C(s) + O_2(g)' is Source (shown in yellow in Figure 5.15) and 'C(s) + $\frac{1}{2}O_2$(g)' is Product (shown in blue), connected through the Vector (shown in red) between them. In the second one, 'C(s) + $\frac{1}{2}O_2$(g)' is Source (shown in yellow) and 'CO_2(g)' is Product (shown in blue). It is interesting to note that 'C(s) + $\frac{1}{2}O_2$(g)' plays different functional roles in the two processes - Product in the first and Source in the second, meaning that the two functions are conflated. We follow Kress and van Leeuwen (2006: 68) to call this conflated function Relay. Therefore, the overall transformation process involves three Participants: Source, Relay, and Product.

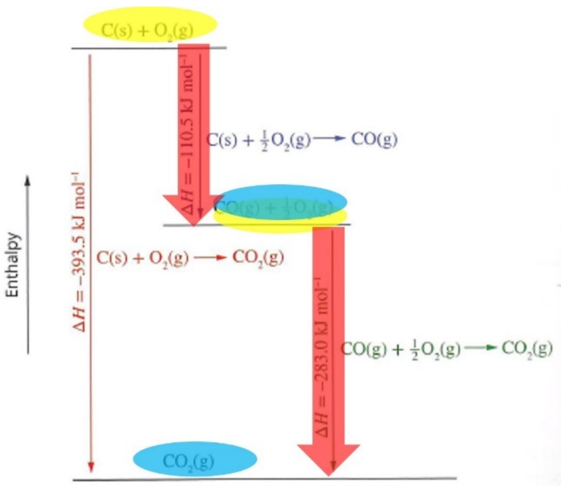


Figure 5.15 Analysis of Figure 5.14 (Source is shown in yellow, Vector in red, and Product in blue)

Construing momented activities is a significant affordance of chemical images. They can present as many momented activities as possible to illustrate extremely complex processes of producing certain chemical matter. Figure 5.16, for example, shows two reaction pathways for the production of ibuprofen, which is a kind of medicine. It illustrates how the initial reactant '2-maethylpropylbenzene' changes to the final product 'ibuprofen' through a series of reactions in two different pathways.

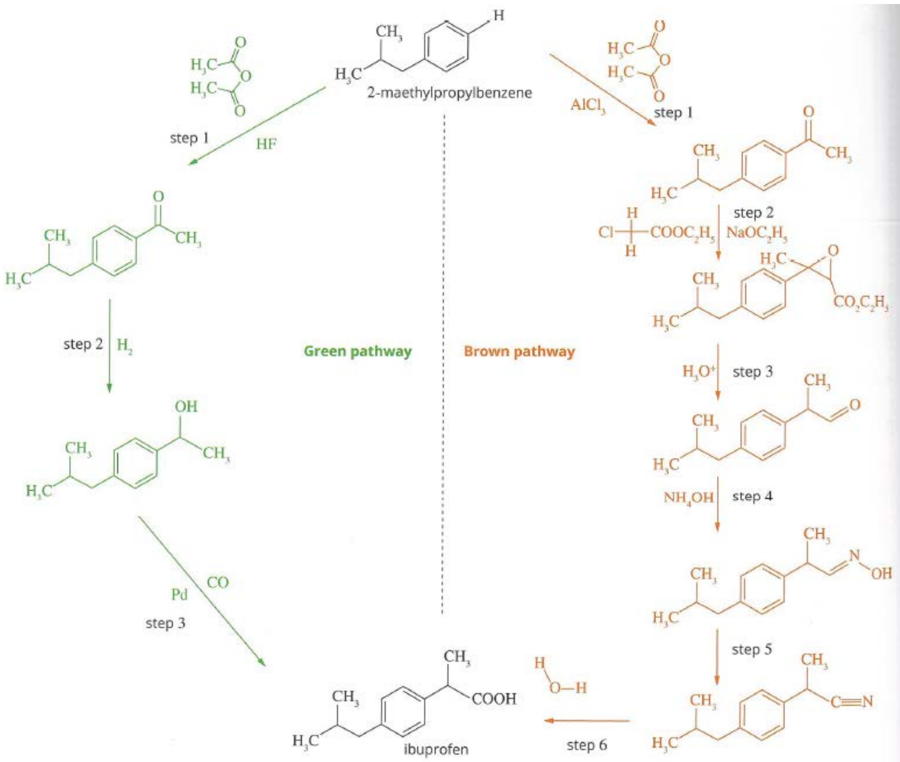


Figure 5.16 Two alternative reaction pathways for the production of ibuprofen

In terms of field, the image construes different numbers of momented activities between the green and brown pathways. The green pathway involves three activities momented from one to another, while the brown pathway involves six, indicating that the left is much more efficient than the right in producing the target product. Grammatically, the momented activities in the two pathways are realized by different numbers of transformation processes that complex with each other linearly. As shown in Figure 5.17, the two pathways share the same Source (shown in yellow) and Product (shown in blue). However, they involve different numbers of Relay (shown in purple) - two on the left and four on the right. The more Relays, the more complex the overall transformation process.

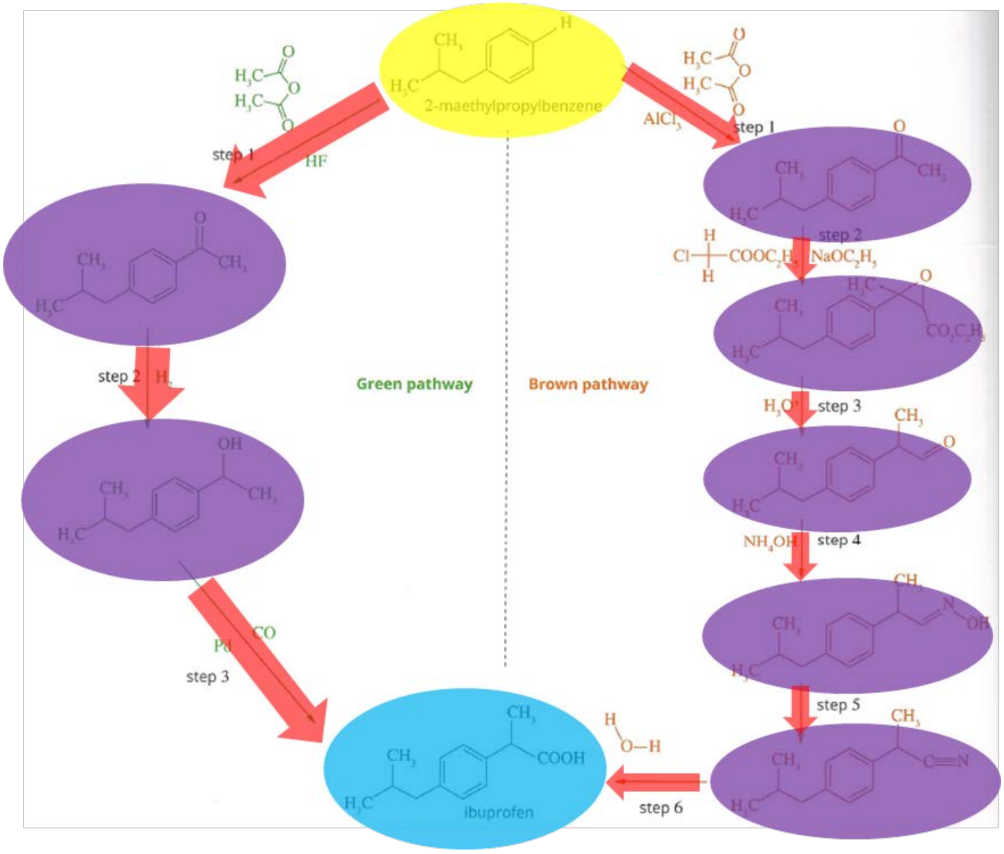


Figure 5.17 Analysis of Figure 6.16 (Source is shown in yellow, Relay in purple, and Product in blue)

Thus far, all the momented activities construed by chemical images analyzed above are linear. Chemical images also realize momented activities that form a cycle. These activities constitute a cyclical loop and can happen over and over again. Figure 5.18, for example, shows the cyclical transitions between the three states of water.

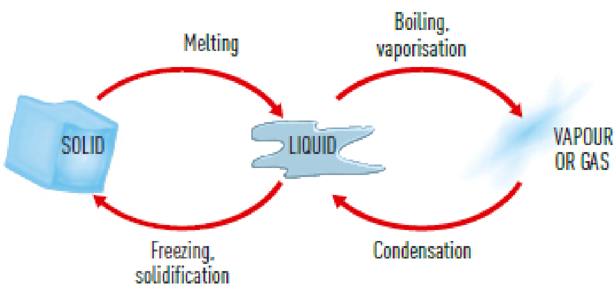


Figure 5.18 The transitions between water's states

It construes four momented activities that form three cyclical loops: (1) 'solid' to 'liquid' and 'liquid' to 'solid', (2) 'liquid' to 'vapour or gas' and 'vapour or gas' to 'liquid', and (3) 'solid' to 'liquid' to 'vapour or gas' and 'vapour or gas' to 'liquid' to 'solid'. Grammatically, the three cyclical loops are realized by three conversion processes in which all Participants are Relay (Kress and van Leeuwen 2006: 69). As shown in Figure 5.19, each Relay (shown in purple) plays the role of Source and Product at the same time. Put differently, it performs Source in one transformation process and Product in another. For instance, 'solid' is the Source in the process that 'solid' changes to 'liquid' and the Product in the process that 'liquid' changes to 'solid'. The series of Vectors (shown in red in Figure 5.19) connect these Relays to form cyclical loops.

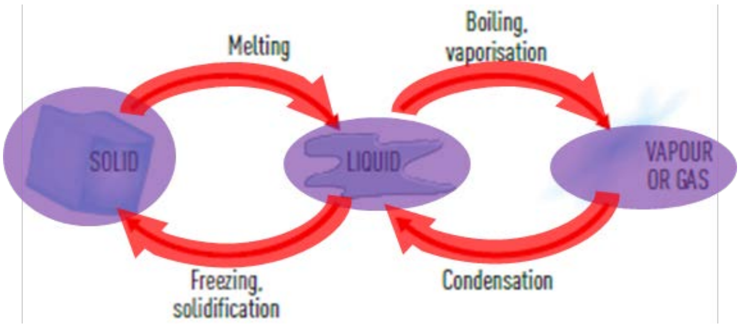


Figure 5.19 Analysis of Figure 6.18 (Relay is shown in purple and Vector in red)

All the above-discussed momented activities, linear or cyclical, are related through **expectancy**, which is the typical relation that connects consecutive chemical changes. Chemical images may also involve another type of relation – **implication** (Doran and Martin 2021). Implication sequences involve cause and effect relations between activities and 'explain how things are, or come to be the way they are' (Wignell et al. 1993: 174–175). Momented activities related through implication in chemical images tend to explain how a chemical reaction happens. For example, Figure 5.20 shows how sodium (Na) and chlorine (Cl) react to form sodium cation (Na^+) and chloride anion (Cl^-).

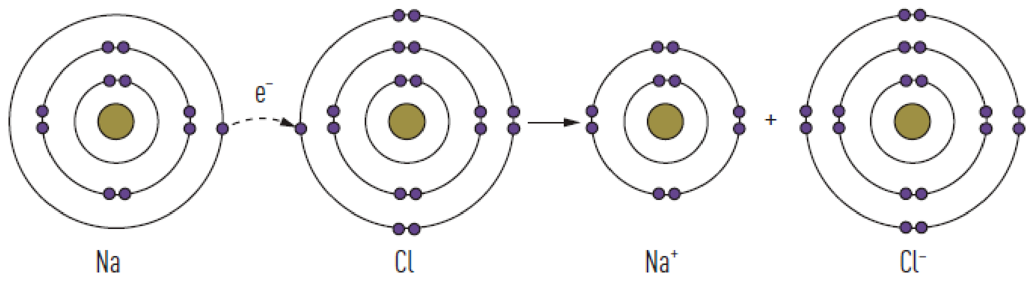


Figure 5.20 Sodium atoms react with chlorine atoms to form sodium cations and chloride anions

The image presents two activities. One is that the electron at the outer shell of the sodium atom moves to the chlorine atom's outer shell. The other is that the reactants (Na and Cl) on the left of the middle black arrow change to the products (Na^+ and Cl^-) on the right. It is the movement of the electron from the sodium atom to the chlorine atom that turns the two atoms into anions. That is, the first activity leads to the second activity, which is a typical implication relation. Grammatically, the two activities are realized by two different processes. The first one is by a transactional process, constituted by an Actor (the electron, shown in green in Figure 5.21), a Vector (the dashed arrow, shown in red), and a Goal (the outer electron shell in the chlorine atom, shown in pink). The second one is realized by a transformation process, comprising a Source (the whole of the left including both the sodium and chlorine atoms, shown in yellow in Figure 5.21), a Vector (the black arrow in the middle, shown in red), and a Product (the whole of the right including the sodium cation and chloride anion, shown in blue).

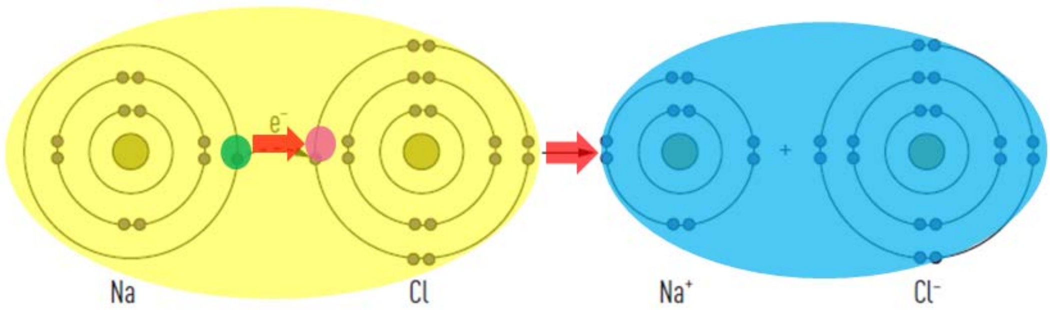


Figure 5.21 Analysis of Figure 5.20 (Actor is shown in green, Vector in red, Goal in pink, Source in yellow, and Product in blue)

In this section, we have analyzed the meaning potential of chemical images in terms of the activities they realize. The analysis shows that chemical images can construe either unmomented or momented activities, with the former conveying the knowledge of a single chemical change while the latter typically showing how initial reactants reach final products. Grammatically, the unmomented activities tend to be realized by a single transformation process. Momented activities, on the other hand, involve further complications. They can be further described in terms of their cyclicality and the relation connecting them. In terms of cyclicality, they can be either cyclical or non-cyclical. Grammatically, the cyclical momented activities are realized by conversion processes, and the non-cyclical ones by complexes of transformation processes. In terms of the relation linking the activities, they can be related by either implication or expectancy. The expectancy sequences tend to be realized by complexes of transformation processes or conversion processes, and the implication sequences are typically realized by complexes of transaction processes and transformation processes. The system network presented in Figure 5.22 summarizes the activities construed by chemical images in secondary school chemistry textbooks and the grammatical structures that realize these activities.

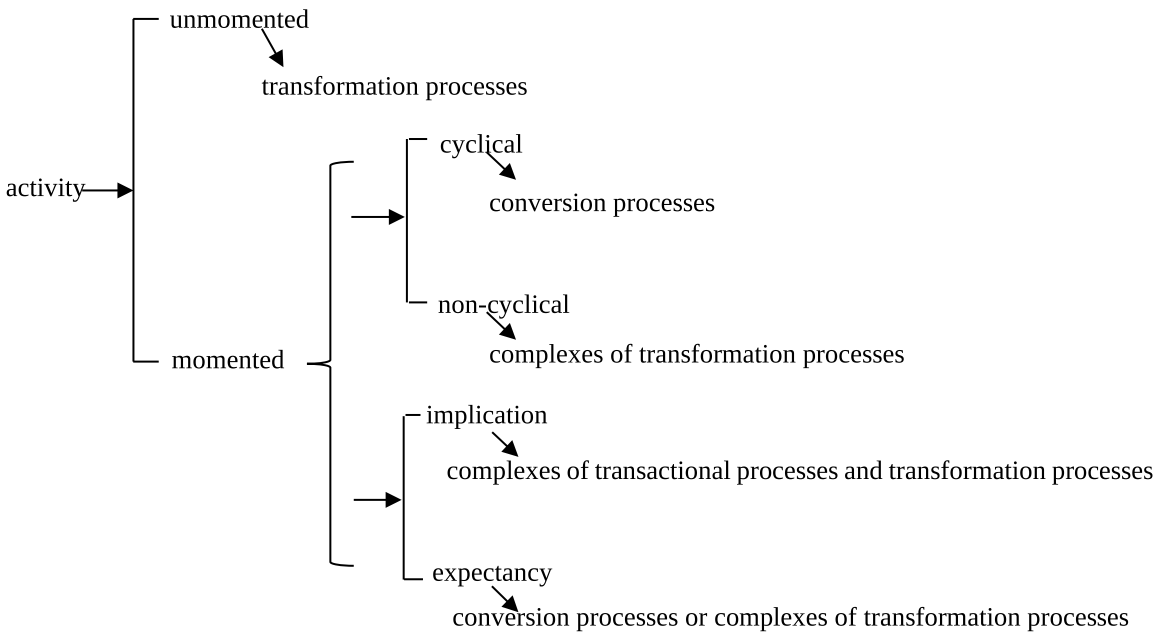


Figure 5.22 Activities construed by chemical images in secondary school chemistry textbooks and the grammatical structures realizing the activities

5.1.3 Property

Chemical images also realize the field-specific meaning of property, which characterizes both items and activities. This section will examine the properties afforded by chemical images in secondary school chemistry textbooks.

Doran and Martin (2021: 118-121) distinguish in language between two types of properties: qualitative and spatial-temporal. Qualitative properties describe the qualities of items or activities, while spatial-temporal properties tend to describe their spatial or temporal positions. In their description of these two properties, they are disjunctive. Chemical images, however, can construe both properties within a single snapshot. For example, Figure 5.23, a diagrammatic representation of lithium atoms' atomic structure, realizes the qualitative properties of the charge polarities that electrons (the blue balls) and protons (the red balls) embody and the spatial property of the relative positions of the electrons and the nucleus (the whole of the red and green balls) within the atom.

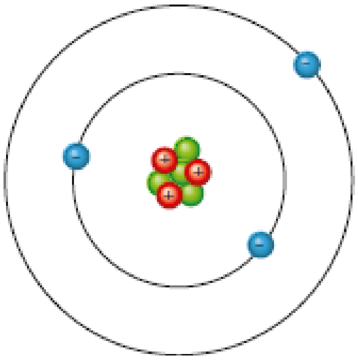


Figure 5.23 A diagrammatic representation of a lithium atom' atomic structure

It should be noted that the spatial property construed by images is distinct from that construed by language because images, in Lemke's (1998) term, afford topological meaning while language construes typological meaning. This is caused by the different semiotic modes' epistemological commitment (Kress 2003). As a visual mode, images must show the relative positions of the components. For example, Figure 5.23 depicts the relative locations of the electrons and nucleus. Language, in contrast, does not involve such epistemological commitment. It does not depict relations of relative nearness or connectedness or continuous change (Lemke 1998: 87). For example, the clause complex *the nucleus is at the center and electrons orbits around it* does not present how close the nucleus is to the electrons. In order to mark this distinction, we term the spatial property realized by images as **topological** spatial property, in contrast to the **typological** spatial property construed by language.

Grammatically, the qualitative properties of charge polarities are realized by symbolic labelings: the labels '+' and '-' represent positive and negative electronic polarities. The spatial property, on the other hand, is realized by the spatial arrangement of the entities within the plane of the page: the blue balls center around the red and green balls.

Qualitative and temporal properties can also co-occur in an image. For example, Figure 5.24 presents a radioactive decay curve for an isotope that has a half-life of 25 years. The vertical axe shows the isotope's mass, and the horizontal axe denotes time, construing a qualitative property and a temporal property, respectively. Like the topological spatial property discussed above, the temporal property construed by the figure is also **topological** because it employs the visual mode that allows for depicting infinite gradations of time on the horizontal axe, in contrast to the categorical gradations by the linguistic mode, such as *from 0 to 175 years*. Grammatically, both the qualitative property of isotope mass and the temporal property of radioactive decay time are realized by linguistic labelings on the axis.

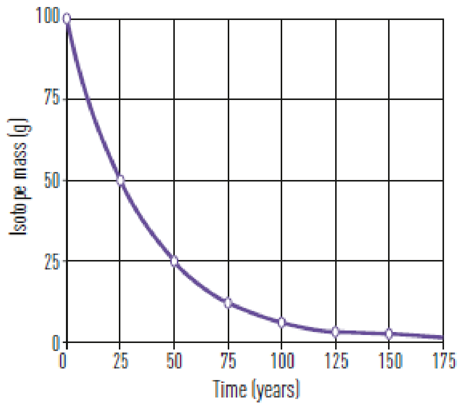


Figure 5.24 A radioactive decay curve for an isotope that has a half-life of 25 years

Based on the above analysis, the types of property construed by images in chemistry and the grammatical resources realizing these properties can be summarized as Figure 5.25. In the system network, ‘ ’ means that ‘qualitative’ and ‘spatial-temporal’ can be selected either disjunctively or conjunctively. That is, an image can construe either qualitative or spatial-temporal properties or both of them. ‘spatial-temporal’ includes two sub-choices: ‘topological spatial’ and ‘topological temporal’. The realization statements pointed to by the arrows show the grammatical resources realizing the properties.

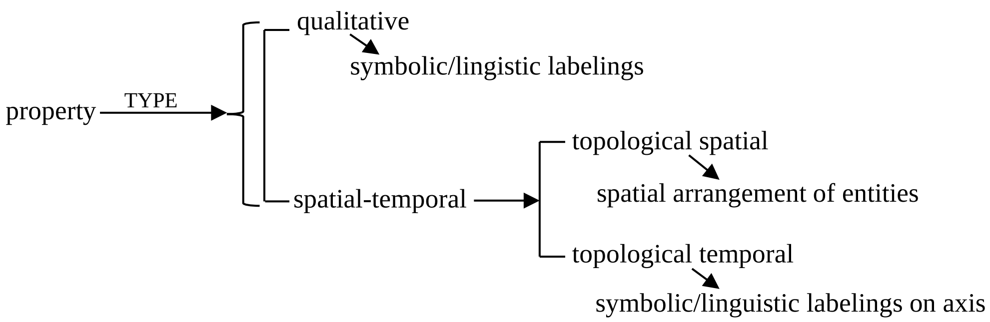


Figure 5.25 Types of property construed by images in chemistry (means that ‘qualitative’ and ‘spatial-temporal’ can be selected either disjunctively or conjunctively)

In addition to the classifications of property, another significant feature of the field-specific meaning of property is that it can be gradable (Doran and Martin 2021: 117). In Doran and Martin's model of field, they call the gradable properties "arrays", for example, *electrons in the outer shell have the* *__highest__* *energy*. They can also be measured or quantified, which are called "gauged" properties, for example, *an electron in the first energy level has -*__*13.6*__ *eV of energy*. These properties are also afforded by chemical images and can be presented simultaneously in a single snapshot. For instance, Figure 5.26 shows the relative pH of various substances.

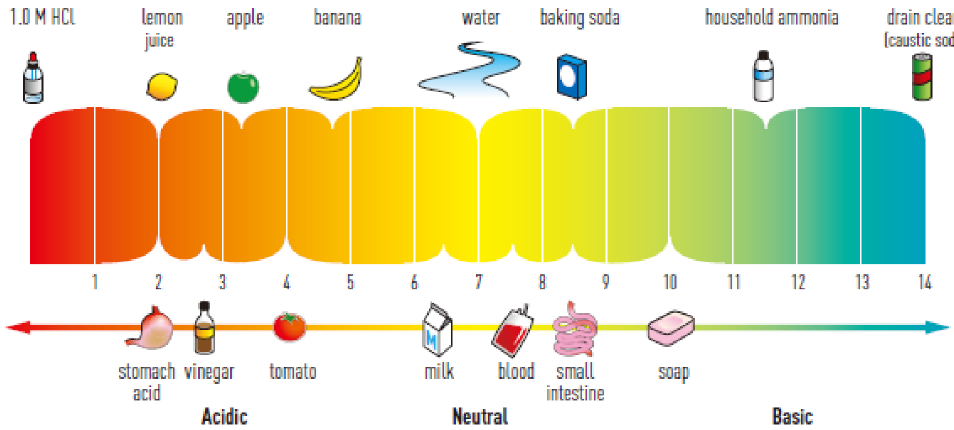


Figure 5.26 The pH scale and the relative pH of various substances

The ordinal numbers from 1 to 14 represent different pH values. The higher the value, the more basic the substances. In terms of field, these quantified scales construe a gauged property. In addition to the numbers, the gradations of pH values are also indicated by the gradual change of the colors on the bidirectional arrow. When the color changes toward red, the pH becomes lower, and the substances gain acidity. In contrast, when the color changes toward green, the pH becomes higher, and the substances get increasingly basic. In terms of field, the gradations of colors realize arrayed but non-gauged properties. It should be noted that this arrayed property is distinct from the one realized by language because the arrow is a visual mode that affords topological meaning while language does not (Lemke 1998). This means that the arrow can show indefinite gradations in pH through the gradual change of the colors and their visual connectedness, whereas language construes categorical gradations, for example, the numbers "1, 2, 3". To distinguish between the two arrayed properties, we term the one realized by the visual mode (the arrow) a **topological** arrayed property and the one realized by the linguistic mode (the ordinal numbers) a **typological** arrayed property. Therefore, Figure 5.26 construes both arrayed properties.

In addition to the gauged and non-gauged arrayed properties noted in Doran and Martin's description, we notice a further distinction within non-gauged properties in

Figure 5.26. The three categories ‘acidic’, ‘neutral’, and ‘basic’ sequenced from the left to the right below the bidirectional arrow also form an array. However, they are distinct from the color gradations since they are categorical. We thus term them **categorical** and the color gradations **non-categorical**. In terms of modes of meaning, the categorical non-gauged property tends to be typological meaning realized by the linguistic mode, whereas the non-categorical one is typically topological meaning realized by the visual mode.

Grammatically, the gauged arrayed properties are realized by linguistic labelings, the non-categorical arrayed properties by color gradations, and the categorical arrayed properties by linguistic labelings. The gradations of properties discussed above can be summarized as the system of ARRAYS presented in Figure 5.27.

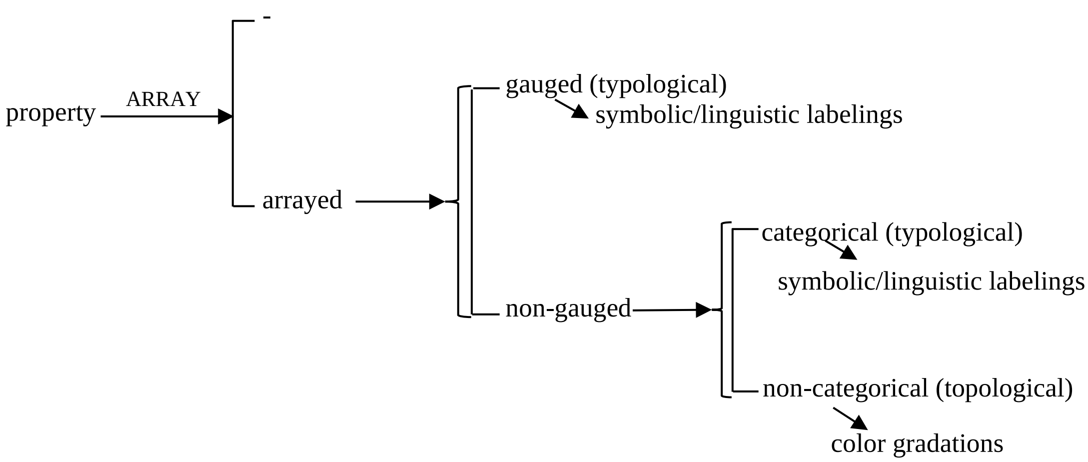


Figure 5.27 The system of ARRAYS (means that the choices can be selected either disjunctively or conjunctively)

As shown in the analysis of Figures 5.23, 5.24, and 5.26, both qualitative and spatial-temporal properties involve arrays, suggesting that the two systems of TYPE and ARRAY are conjunctive sub-systems of PROPERTY shown in Figure 5.28.

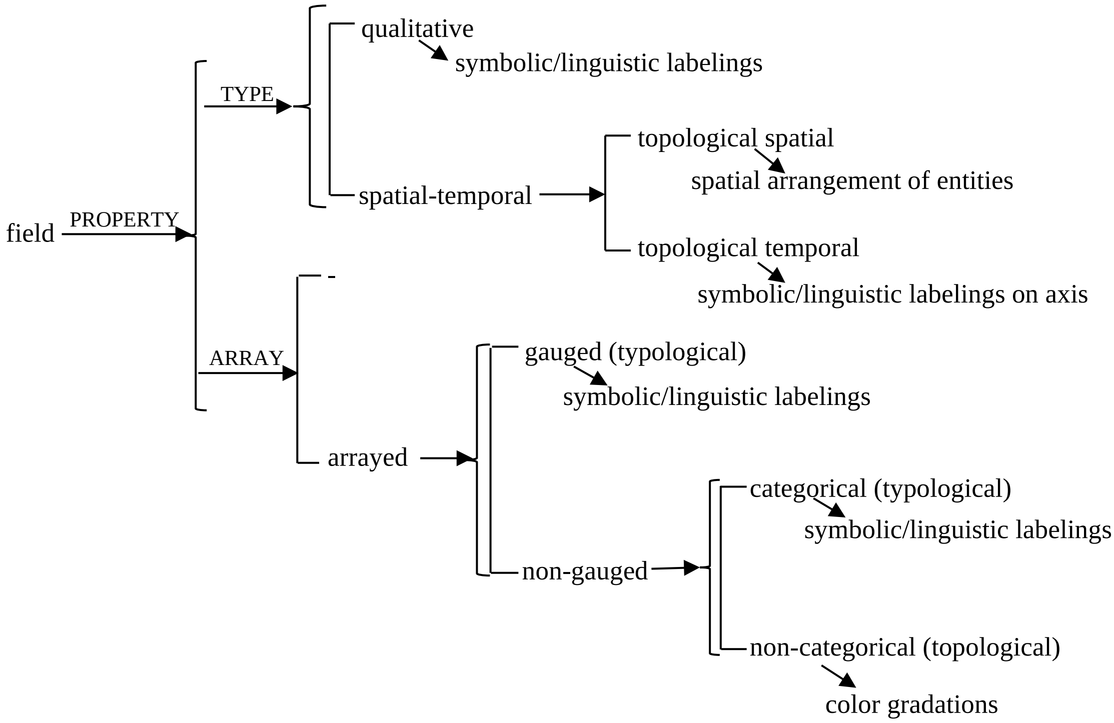


Figure 5.28 The full system of PROPERTY

Figure 5.28 shows that images in chemistry construe all the properties described in Doran and Martin 2021's model and include a further distinction within the non-gauged arrayed property: categorical and non-categorical. The former tends to be typological, while the latter is typically topological. What is also distinct from Doran and Martin's model is that the spatial and temporal properties construed by images can be topological, in contrast to the typological ones construed by language. Grammatically, the properties are not realized by particular structures by which items and activities are realized but by non-structural resources. The primary resource is symbolic labelings, including numbers, symbols (e.g. '+' and '-'), mathematical equations, etc. Another significant source for realizing properties, especially non-gauged arrayed properties, is color gradations. In addition, the spatial arrangement of entities in images is also a key grammatical resource realizing the topological spatial property.

5.1.4 The full model of chemical images in terms of field

Sections 5.1.1, 5.1.2, and 5.1.3 have described the field meanings realized by images in chemistry in terms of field perspective – static (item) and dynamic (activity) – and property. In fact, property can describe both item and activity. The analysis of Figure 5.26 above has already shown that the properties construed by chemical images can describe chemical matter (items). In addition to this, images in chemistry can also construe different properties of chemical reactions (activities). Our data suggest that these properties are primarily gauged qualitative and temporal properties. Figure 5.29, for example, realizes gauged properties of energy released during the reactions. The mathematical equations (highlighted in yellow) attached to the arrows indicate the amount of energy released by the reactions. For example, the mathematic equation '$\Delta H = -393.5 \mathrm{kJ/mol}$' indicates that the reaction that carbon reacts with oxygen gas to yield carbon dioxide releases 393.5kJ energy per mole. In terms of field, the mathematical equation realizes a gauged qualitative property.

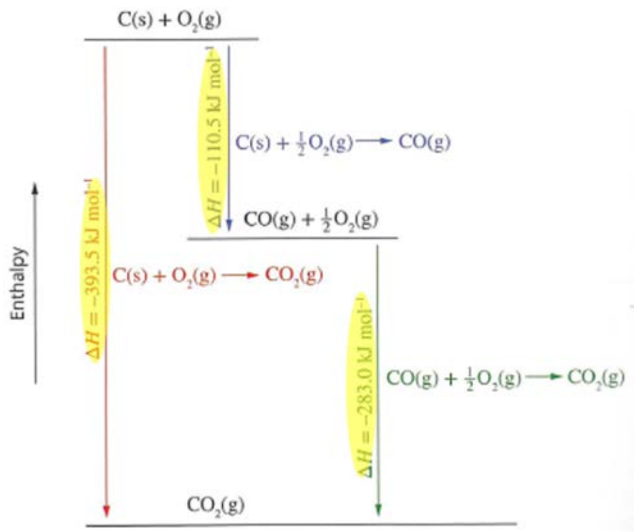


Figure 5.29 An energy level diagram for the formation of carbon dioxide from carbon and oxygen via carbon monoxide

Chemical images can also construe temporal properties of activities. Images realizing these properties tend to be line graphs that record reaction processes. For example, Figure 5.30 shows the concentration versus time plot for the decomposition of dinitrogen tetroxide (N_2O_4). The purple and blue lines record the changes of NO_2 and N_2O_4's concentration over the reaction time. The graph shows that it takes 6 seconds for the reaction to reach equilibrium.[51] In terms of field, this construes a gauged temporal property realized by numeral labelings on the horizontal axis.

[51] 'Equilibrium' is a state where the chemical species' concentration does not change anymore.

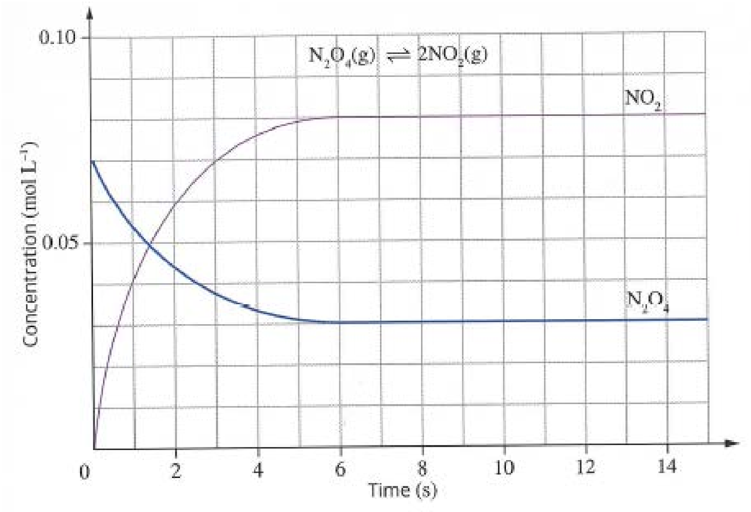


Figure 5.30 The decomposition of dinitrogen tetroxide

The above analysis suggests that PERSPECTIVE (the system constituted by item and activity) and PROPERTY are two conjunctive systems in the system network of field of images in chemistry, as shown in Figure 5.31. Distinct from the field model of language proposed by Doran and Martin (2021), in which item and activity are two disjunctive choices, the system of field of images in chemistry allows choosing the two features either disjunctively or conjunctively, as denoted by the notation of ‘{[’. That is, images in chemistry can construe either items or activities, or both in a single snapshot. The system network not only presents the types of field meanings but also specifies the grammatical resources realizing the meanings. It can be viewed as a model of images in chemistry from the perspective of field that links with the grammatical organization of images. This model maps out the disciplinary affordance of images in chemistry and scaffolds the analysis of the development of images across schooling levels in Section 5.2.

Figure 5.31 The full system of field of images in chemistry and the grammatical resources realizing the field meanings ( means that the choices can be selected either disjunctively or conjunctively)

5.2 Building the field of chemistry: the development of images across schooling levels

Section 5.1 has described the meanings afforded by images in chemistry in terms of field and how the images make those meanings. We are now in a position to examine how the images in secondary school chemistry textbooks develop through schooling to build the hierarchical knowledge structure of chemistry. This section will continue to approach this issue from the perspective of field in SFL because in terms of ideational meaning, building the knowledge of chemistry can be viewed as building the field of chemistry.

Images in secondary school chemistry textbooks show differing complexity at different schooling levels. They can be simple and everyday in the lower year levels and technical and complex at the higher year levels. To understand how the images develop across schooling levels, we need first to establish a model that allows for comparing the different complexity of meaning construed by the images. Therefore, this section will first develop a model that enables ‘seeing’ the different levels of complexity of ideational meaning construed by the images used in secondary school chemistry textbooks and then trace their development across schooling levels to explore how they build the hierarchical knowledge structure of chemistry. To achieve this objective, we will use the LCT concept of semantic density (Maton 2014) and integrate it with the field analysis of images to model their complexity of meaning (Section 5.2.1). The model will then be enacted to analyze the development of images’ semantic density across schooling levels in secondary school chemistry textbooks to investigate how the images build the hierarchical knowledge structure of chemistry (Section 5.2.2).

5.2.1 Semantic density of images in secondary school chemistry textbooks

This section is devoted to developing a model for the semantic density of images used in secondary school chemistry textbooks that allows for comparing the different levels of complexity of meaning. Specifically, it approaches this from the perspective of field and considers two essential factors that determine the semantic density of images – internal organization (Section 5.2.1.1) and external relations (Section 5.2.1.2). Internal organization describes the number of field-specific meanings realized in an image, and external relations describe the complexity of concepts represented in the image. Sections 5.2.1.1 and 5.2.1.2 will discuss the two respective factors, and Section 5.2.1.3 brings them together into one full model.

5.2.1.1 Internal organization

Internal organization is so termed for it concerns the number of field-specific meanings realized by the grammatical structures of images. To model this factor, we need to use the model of images in terms of field developed in Section 5.1. As shown in the model, the images in secondary school chemistry textbooks mainly construe four types of field-specific meanings: composition, classification, activity, and property. As suggested by our data, varying numbers of these field-specific meanings may be realized in an image. The simplest construes only one type of field-specific meaning. For example, Figure 5.32, a diagrammatic model of oxygen gas molecules, construes a composition taxonomy between the oxygen gas molecule (the whole) and the component oxygen atoms (the parts). According to the model of chemical images proposed in Section 5.1, this field-specific meaning is realized by a spatial analytical structure.

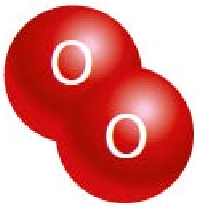


Figure 5.32 A diagrammatic representation of oxygen gas molecules

In contrast to Figure 5.32 involving only one type of field-specific meaning, some chemical images can present multiple field categories in a single snapshot. This brings multiple structures in a single image and greatly condenses the images' meaning. To illustrate this, we will focus on Figure 5.33, a diagram explaining the working mechanism of voltaic cells.

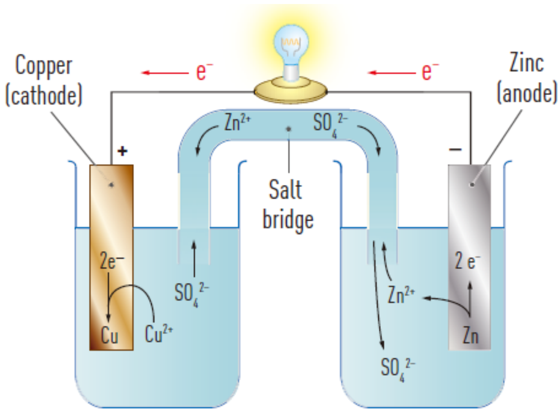


Figure 5.33 The working mechanisms of voltaic cells

In terms of field, this diagram construes all of the four types of field-specific meanings. Firstly, it construes a compositional taxonomy between the overall experimental set-up and its multiple components - the electrolytic tanks (the left and right beakers), the electrodes (the yellow and grey sticks), the salt bridge (the 'U' shaped tube connecting the two beakers), and the light bulb. This compositional taxonomy is realized by a spatial analytical structure. Secondly, the diagram construes two classificational taxonomies - the cathode (the yellow stick labeled 'copper') and the anode (the grey stick labeled 'zinc') are subclasses of electrodes, and the cations ('Cu^{2+}'and 'Zn^{2+}') and anions ('SO_4^{2-}') are subclasses of ions. Both classifications are realized by covert taxonomy structures discussed in Section 5.1.1. Thirdly, the diagram realizes two qualitative properties - the cathode and anode are positively and negatively charged (shown by the labels '+' and '-'), and the cations and anions bear positive and negative charges (shown by superscripts '$^{2+}$' and '$^{2-}$'). These properties are realized by symbolic labelings as discussed in Section 5.1.1. Fourthly, the diagram construes multiple activities. Using the model of images proposed in Section 5.1.4, we can see that the chemical formulas connected through arrows at the two electrodes construe two different activities. At the cathode, the arrow leading 'Cu^{2+}' to 'Cu' realizes an unmomented activity, denoting the reduction reaction that a copper cation is transformed into a copper atom. Grammatically, this activity is realized by a transformation process as we have discussed in Section 5.1.2. At the anode, the arrows leading 'Zn' to '$2e^-$' and 'Zn^{2+}', which is in turn connected with another arrow, realize momented activities. The first arrow connecting 'Zn' and 'Zn^{2+}' represents the oxidation reaction that a zinc atom is oxidized and becomes a zinc cation. The second arrow connecting 'Zn^{2+}' shows the movement of zinc cations. Grammatically, the first activity is realized by a transformation process and the second by a non-transaction process. In addition to the electrodes where activities occur, other parts of the apparatus also involve activities. In the left electrolytic tank (the left beaker), the chemical formula 'SO_4^{2-}' connected with an arrow construes an unmomented activity realized by a non-transactional process. This activity shows the movement of the sulfate cations (SO_4^{2-}). On the electric wire (the solid line) that connects the two electrodes, the symbol 'e^-' connected with arrows also construe unmomented activities of movement realized by non-transactional processes. The above analysis shows that chemical images can be rather powerful in bringing multiple structures in a single image to construe multiple field-specific meanings as an eyeful.

The above suggests that chemical images can be as simple as Figure 5.32 construing only one type of field-specific meaning and as complex as Figure 5.33 realizing all the four types of field-specific meanings. Chemical images construing all four field-specific meanings show stronger semantic density than those involving a single field-specific meaning. Most images will fall between the two poles. We can thus account for the different degrees of semantic density of images in terms of the number of field-specific meanings realized as the network shown in Figure 5.34.

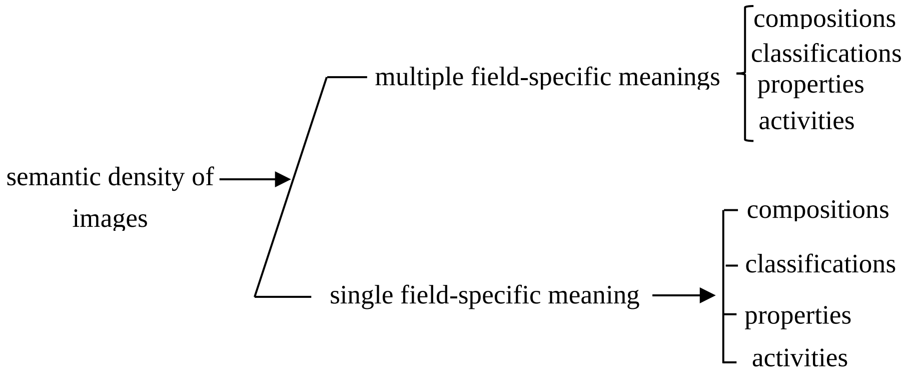


Figure 5.34 A partial framework for the semantic density of images

The slant square bracket ( ) denotes a topological relation between the two disjunctive features. That is, ‘single field-specific meaning’ and ‘multiple field-specific meanings’ are the two poles of a continuum. The square bracket ( ) on of right side of ‘single field-specific meaning’ represents a disjunctive choice from the four features, i.e., only one type of field-specific meaning can be chosen. The brace ( ) on the right side of ‘multiple field-specific meanings’, in contrast, means a conjunctive choice from the four options, i.e., all the four field-specific meanings are chosen. The network indicates that the more field-specific meanings construed in an image, the stronger its semantic density.

However, it should be noted that the network presented in Figure 5.34 constitutes only a partial framework for the semantic density of images. It considers the density of images only in terms of the number of field categories but does not take into account the complexity of the field categories themselves. That is, it does not consider the complexity of the concepts represented in the images. This would result in identifying an everyday image and a technical image that share the same number and types of field-specific meanings as embodying the same degree of semantic density. Take the two images in Figure 5.35 as an example.

Figure 5.35 Diagrams of a tree (left) and a water molecule (right)

In terms of field, they both construe a one-level composition taxonomy - a tree is composed of a trunk, branches, and leaves, and a water molecule is composed of one oxygen atom (the red ball labeled ‘O’), two hydrogen atoms (the grey balls labeled ‘H’), and two covalent bonds (the grey sticks). By applying the partial framework presented in Figure 5.34, we would arrive at the conclusion that the two diagrams embody the same degree of semantic density because they involve the same number and depth of composition taxonomies. However, this is an obvious contradiction to our common sense that the complexity of the structure of a water molecule is much stronger than that of the structure of a tree. The chemical concepts of atoms and covalent bonds themselves are rather complex meanings since they locate in a specialized field of chemistry where they connect with numerous other chemical concepts. For example, covalent bonds are pairs of shared lone electrons from atoms’ outer electronic shells (Chan et al. 2018: 164). Understanding the concept ‘covalent bonds’ requires knowing the series of related technical concepts - shared lone electrons, outer electron shells, and atoms. Although these concepts are not explicitly represented in the image, they are essential for understanding the knowledge of covalent bonds depicted in the diagram. We call concepts explicitly represented in images **explicit meanings** and those related concepts **implicit meanings**. Both meanings contribute to the complexity of the concepts represented in images. This is similar to technical words that involve much more meanings than a simple definition affords (Martin 1989; Wignell et al. 1993). They systematically relate to other technical terms through implicit taxonomies (Wignell et al. 1989: 378). These implicit taxonomies can be viewed as implicit meanings that imbue the technical words with relatively strong complexity. In a similar vein, the numerous implicit meanings related to the technical concepts represented in an image infuse it with relatively strong complexity.

The above analysis suggests that considering only the number of field categories is inadequate for assessing images’ semantic density. To complement this, we need to consider the complexity of the concepts represented in images. That is, an image’s overall field complexity is determined by both the complexity of concepts represented in an image and the number of field categories presented in the image. The former describes the technicality of the field categories and the latter describes the number of these categories. Both factors contribute to the semantic density of images because if we consider only the number of field categories we cannot distinguish between the semantic density of the two images in Figure 5.35, and if we consider only the complexity of concepts represented in images, it is difficult to distinguish between the semantic density of Figure 5.33 and the right image in Figure 5.35. Therefore, to complete the model for the semantic density of images, we still need to account for the factor of complexity of concepts represented in images. In contrast to ‘internal organization’, this factor is termed ‘external relations’ as they concern the semantic relations to other concepts.

5.2.1.2 External relations

External relations concern the complexity of a concept represented in an image and describe the degree of the concept’s technicality. This factor is crucial for assessing the semantic density of images because it determines whether the concepts construed by an image locates in a complex constellation of meanings. To model this, we draw on Maton and Doran’s (2017c) criterion for measuring the complexity of English words – **relationality** of meanings. Relationality describes the complexity of the constellation of which the concept under consideration is a constituent (ibid.: 57). The general principle is that the more relations a concept relates to other concepts, the stronger its complexity. Maton and Doran illustrate the notion of relationality through Figure 5.36. ‘Gwiffly’ is a newly created meaningless word represented by a solitary node (number 1). If we describe two types of Gwiffly as ‘A’ and ‘B’ (number 2), we have set up relations between ‘Gwiffly’ and two further terms. If we then describe their characteristics, such as ‘Gwiffly A is red and Gwiffly B is blue’, we are augmenting those terms by relating them to more meanings (number 3). This starts to establish a ‘constellation’ of meanings in relation to the concept ‘Gwiffly’. The more relations established with other meanings, the stronger the concept’s complexity.

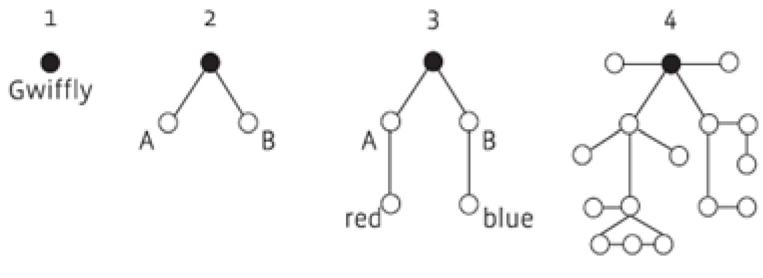


Figure 5.36 Constellating a Gwiffly (Maton and Doran 2017: 49)

We believe that relationalities as a criterion for measuring the complexity of English words also work for images. The more complex the constellation of meanings, the more relations the concept presented in an image enjoys with other concepts and hence the more complex its meaning. To measure the complexity of the constellation of meanings where concepts presented in an image locates entails identifying what meanings they relate to. However, Maton and Doran have not offered a guideline that can help more objectively identify what meanings a concept would relate to. In response to this, we propose that the field categories mentioned above (activities, composition, classification, and properties) can be useful dimensions for identifying the meanings a concept relates to. This can be illustrated through Figure 5.37, an energy level diagram for the formation of carbon dioxide from carbon and oxygen through carbon monoxide.

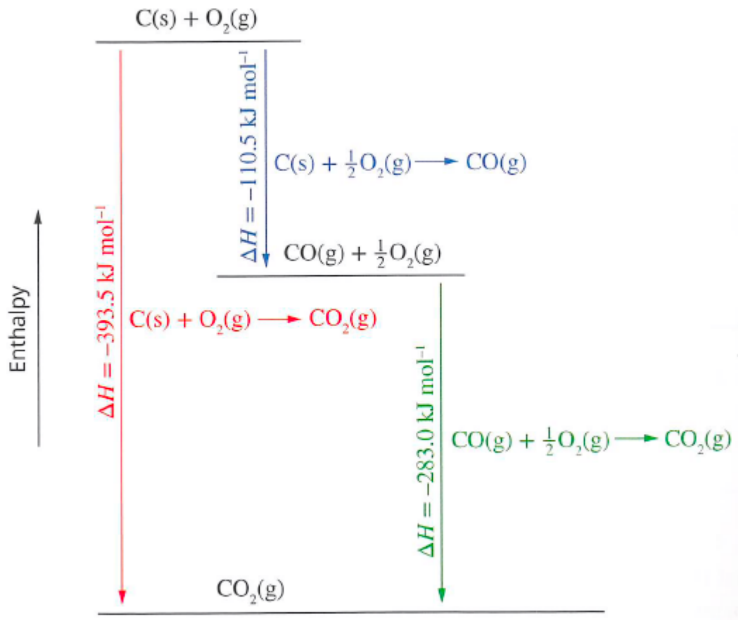


Figure 5.37 An energy level diagram for the formation of carbon dioxide from carbon and oxygen

As shown in the diagram, there are different pathways from the reactants '$C(s) + O_2(g)$' to the product '$CO_2(g)$'. The pathway from '$C(s) + O_2(g)$' to '$CO(s) + \frac{1}{2}O_2(g)$' (denoted by the blue arrow) is an incomplete combustion reaction of carbon while the pathway from '$C(s)+O_2(g)$' to '$CO_2(g)$' (denoted by the red arrow) indicates a complete combustion reaction of carbon. The mathematic equations attached to the arrows show the amount of energy released during the reactions. This diagram involves multiple components that relate to other technical concepts. We will demonstrate how the four field categories can help identify these related technical meanings.

Firstly, the two equations '$C(s) + \frac{1}{2}O_2(g) \rightarrow CO$' and '$C(s) + O_2(g) \rightarrow CO_2(g)$' relate to two chemical concepts. In terms of activities, these reactions involve transitions of electrons from carbon to oxygen atoms, relating to the chemical concept of 'oxidation reaction'. In terms of classification, as introduced above, they are two subtypes of combustion - incomplete and complete, which relates to the chemical concept of 'combustion reactions'. Secondly, the components within the equations, i.e., '$C(s)$', '$CO(g)$', and '$CO_2(g)$', relate to multiple technical meanings. In terms of compositions, the molecules, for example, carbon monoxide molecules (CO) and carbon dioxide molecules (CO_2), are composed of carbon and oxygen atoms, relating to the chemical concept 'atom'. In terms of classification, 'C' represents a monatomic molecule, while 'CO' and 'CO_2' denote polyatomic molecules, relating to the chemical concepts of 'single substance' and 'compounds'. In terms of properties, the state symbols '(s)' and '(g)' on the right side of the chemical formulas refer to two states - solid and gas, relating to physical properties of chemical substances. Thirdly, the mathematic equations, for example, '$\Delta H = -393.5 kJmol$', represent energy released during the reactions, which denote that the combustion reactions are exothermic. In terms of classification, they relate to the chemical concept of thermal reactions (exothermic reactions are a subtype of thermal reactions).

All the above related chemical concepts themselves are highly technical and relate to other meanings in chemistry. For example, in terms of classification, the concept of single substance includes two subtypes - metallic and non-metallic (for example, 'carbon' [C] and 'oxygen gas' [O_2] are both non-metallic single substances); the concept of chemical compounds involve two subclasses - covalent compounds and ionic compounds (for example, carbon monoxide molecules [CO] and carbon dioxide molecules [CO_2] are covalent compounds); thermal reactions include two subtypes: exothermic and endothermic (the reactions represented in the diagram are exothermic). The above analysis suggests that the field categories as an angle through which we identify related meanings are effective.

The relationality analysis of Figure 5.37 shows that the image relates to numerous chemical concepts. We can draw a broad constellation of meanings where this image locates, as shown in Figure 5.38. The map indicates that the image sits in a rather complex constellation of meanings and enjoys relatively strong relationalities to other meanings. This suggests that there are much more meanings than explicitly presented in the image and the technicality of the field construed by the image is rather strong. Reading it is particularly challenging for novel students because it entails understanding all the technical meanings related. To highlight the potential for greater strength of semantic density afforded by images such as Figure 5.37, we refer to them as **power images** (in correspondence to 'power words' proposed by Martin 2013a). Learning power images in chemistry requires a good mastery of the related concepts and hence poses significant challenges for students. Here we propose this concept to, hopefully, raise teachers' awareness of these images' significance in their teaching.

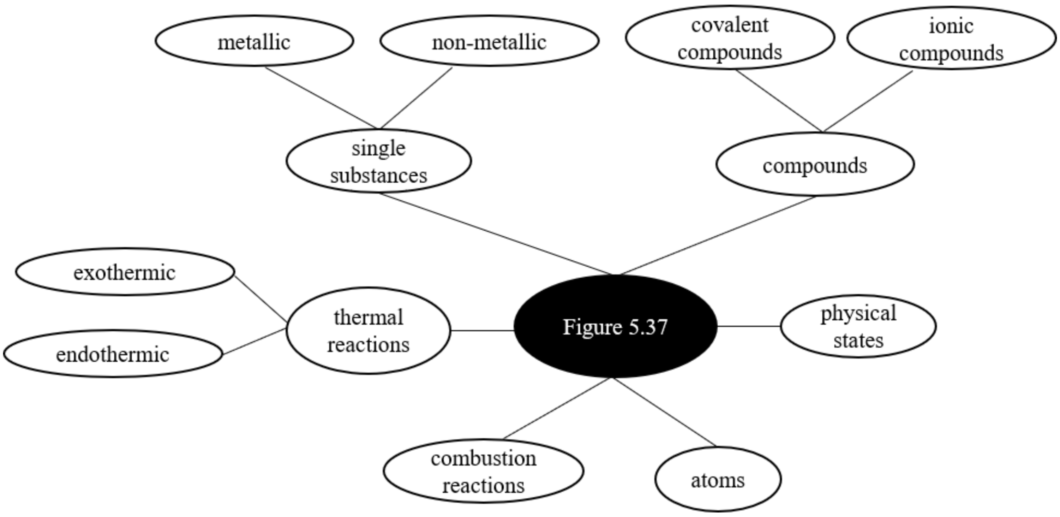


Figure 5.38 The broad constellation where Figure 5.37 locates

With the above method of identifying relationalities to other meanings, we can compare the relationalities of the two images in Figure 5.39, both of which involve only an entity.

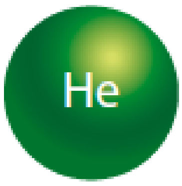


Figure 5.39 A diagram of a helium atom (left) and a photo of an apple (right)

The left image in Figure 5.39 is a diagrammatic representation of a helium atom. It relates to multiple technical meanings. In terms of composition, a helium atom is composed of two electrons and one nucleus, which is in turn composed of protons and neutrons. Understanding the image requires knowing all these technical concepts. In terms of properties, the image resonates out to two technical qualities. One is that helium atoms are rather stable and inactive, relating to the chemical concept of ‘noble gas’. The other is a gauged property – the relative atomic mass of a helium atom is 4.002602, which pushes the concept’s complexity further by relating to the technical concept of ‘relative atomic mass’. In terms of classification, the concept of helium atom relates to the chemical concept of ‘isotope’. Helium elements include numerous isotopes, such as ‘$^{3}_{2}He$’ ‘$^{4}_{2}He$’ and ‘$^{5}_{2}He$’.[52] The helium atom presented in the image could be any of these isotopes. All the above related concepts are rather technical

[52] The subscript ‘2’ represents the number of protons and the superscripts 3, 4, and 5 represent the sums of protons and neutrons.

themselves and can relate to other meanings. For example, noble gases are monatomic and have exceptionally low chemical reactivity. Relative atomic mass is obtained from comparing with the atomic mass constant (m_u), which is defined as being 1/12 of a carbon-12 atom's mass. Isotopes are variants of a particular chemical element that differ in neutron number and consequently in nucleon number.

In contrast to the diagram, the photo of an apple in Figure 5.39 relates to much fewer meanings. We can examine its relationality in terms of classification, composition, and properties. In terms of classification, apples are a subtype of fruit; in terms of composition, an apple is composed of peel, pulp, and pit; in terms of properties, apples are usually round and sweet. All these related meanings are not technical and do not relate to other meanings. Based on the above analysis, we can draw two broad constellation maps of the two images' semantic structure, as shown in Figure 5.40 and Figure 5.41.

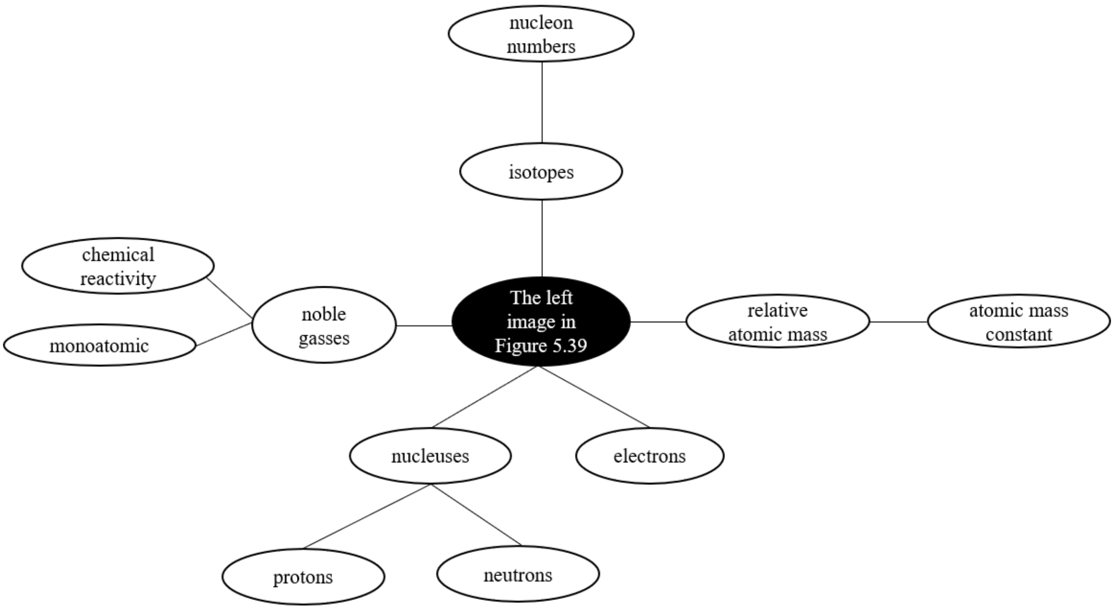


Figure 5.40 The broad constellation map for the diagram of a helium atom

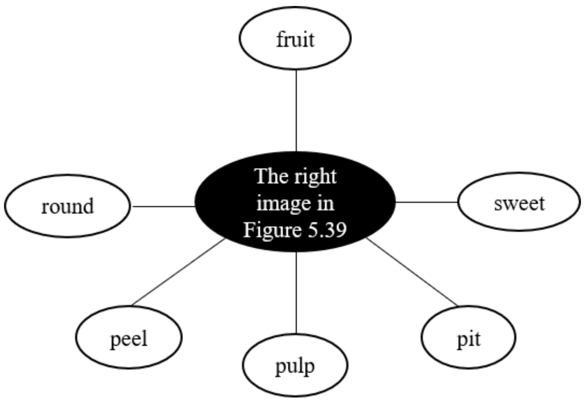


Figure 5.41 The constellation map for the photo of an apple

As shown in the two constellation maps, the diagram of a helium atom and the photo of an apple embody two distinct relationalities to other meanings – the diagram has much stronger relationalities than the photo, suggesting that the concept presented in the diagram is much more complex than that in the photo. This means the stronger the relationalities of the concepts presented in an image, the more complex the image's meaning. In terms of external relations, we can thus model the semantic density of images as a cline from relatively weak to relatively strong, as shown in Figure 5.42.

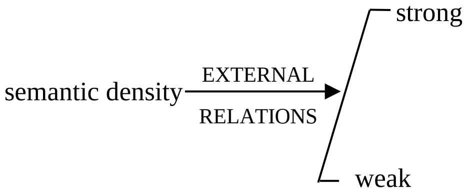


Figure 5.42 Network of EXTERNAL RELATION

The slant square bracket (∠) denotes a topological relation between the two disjunctive choices 'weak' and 'strong'. There can be indefinite strengths of semantic density between the two categories. This network can adequately describe the complexity of concepts presented in images, offering a tool for evaluating the technicality of the field construed by the images. According to Martin's (1992: 544) classification of field's technicality, images with relatively strong relationalities tend to represent uncommon-sense fields, and those with relatively weak relationalities typically present common-sense fields. The analysis of images' relationalities provides a more fundamental interpretation of the field's technicality represented in the images.

5.2.1.3 The full model for semantic density of images

Thus far, we have discussed the two factors contributing to images' semantic density: internal organization (the field-specific meanings realized by an image' internal structures) and external relations (the complexity of concepts presented in an image). However, it is not yet clear whether the two factors determine images' semantic density by the same scale, that is, whether they are given the same weight when used for analyzing images' semantic density. We argue that the complexity of concepts represented in an image should be privileged over the field categories construed by the image. Put differently, when examining an image's semantic density, we first compare the complexity of the concepts represented in the image and then the number of field-specific meanings realized. This is because the fundamental principle of assessing an image's semantic density is the overall complexity of meaning. As discussed above, external relations are concerned with the complexity of the concepts themselves presented in an image, while internal organization models the number and types of field categories construed in that image. If the concepts presented in an image embody relatively weak relationalities, it tends to locate in a relatively simple constellation of meanings no matter how many field-specific meanings the image construe. This means that images with relatively weak relationalities but construing multiple field-specific meanings embody weaker complexity than those with relatively strong relationalities but construing a single field-specific meaning. For example, the left image in Figure 5.43 shows two people running. It construes multiple field-specific meanings: an activity (running), a classification (men and women), and a composition (body and its parts). However, the concepts presented in the image exhibit relatively weak relationalities as it depicts a rather common-sense event. Although the image construes multiple field-specific meanings, its overall complexity is relatively weak due to its location in a relatively simple constellation of meanings. In contrast, as discussed previously (see the analysis of the left image in Figure 5.39), the right image in Figure 5.43 locates in a relatively complex constellation of meanings though it construes a single field-specific meaning. Therefore, the right image embodies a stronger semantic density than the left image in Figure 5.43.

Figure 5.43 Two people running (left) and a diagrammatic representation of a helium atom (right)

Based on the above analysis, we model the semantic density of images by arranging the two factors at two hierarchies, with external relations being privileged over internal organization, as shown in Table 5.1.

Table 5.1 A model for the semantic density of images

Semantic density (SD)	**External relations**	**Internal organization**
+	strong	multiple field-specific meanings
		single field-specific meaning
	weak	multiple field-specific meanings
-		single field-specific meaning

'+' and '-' in the table indicates stronger and weaker strengths of semantic density. The first hierarchy 'external relations' describes the relationalities of concepts presented in an image. Images involving concepts with stronger relationalities embody stronger semantic density. This hierarchy assesses the technicality of the field represented in the image. The second hierarchy 'internal organization' describes the field-specific meanings construed in an image. It involves two categories arranged on a continuum: [single field-specific meaning] and [multiple field-specific meanings]. Within [strong] and [weak], the more field-specific meanings an image involves, the stronger the semantic density of the image. This hierarchy evaluates the complexity of field categories construed by images. The model can be broadly interpreted as it first looks at whether an image construes technical fields and then the number of these field meanings.

As indicated by the arrow, all the categories in the table are topological, meaning that they sit in a continuum of strength: the strength [weak; single field-specific meaning] is weaker than [weak; multiple field-specific meanings], which is weaker than [strong; single field-specific meaning], which is, in turn, weaker than [strong; multiple field-specific meanings]. To more easily mark the difference in these categories' semantic density, we assign different scales of strength to them: [weak; single field-specific meaning] (SD− −), [weak; multiple field-specific meanings] (SD−), [strong; single field-specific meaning] (SD+), and [strong; multiple field-specific meanings] (SD++). These categories and scales of semantic density enable us to compare and trace images' semantic density in the next section.

5.2.2 The development of images in secondary school chemistry textbooks

Section 5.2.1 has established a model for the semantic density of images in secondary school chemistry textbooks. We are now in a position to enact this model to analyze how the images develop through schooling in terms of their semantic density to build the hierarchical knowledge structure of chemistry. In particular, we will focus on two issues in relation to the knowledge-building of chemistry: how images in secondary school chemistry textbooks develop through schooling to build knowledge that is about the same topic; (2) how the range of the semantic density of images used in each curriculum stage develops as learning progresses to higher schooling levels. The first issue concerns the roles of images in building specific related knowledge (Section 5.2.2.1), while the second focuses on the overall development of images across learning stages (Section 5.2.2.2). To address these issues, we will trace the images' development in terms of their semantic density. The analysis will reveal that to build the hierarchical knowledge structure of chemistry, the images' semantic density grows from weaker to stronger to, on the one hand, connect chemical knowledge to common-sense fields, and on the other hand, build increasingly complex uncommon-sense fields. To show this, we will first focus on knowledge-building through images that are about the same topic.

5.2.2.1 Building knowledge about the same topic through images

In our data, different images within and/or across year levels are used to represent knowledge related to the same topic, which provokes a significant question as to knowledge-building through images, i.e., what roles images play in building the knowledge structure of chemistry. This issue will be addressed in this section through a detailed analysis of images related to the same subject. It will enact the model for the semantic density of images and trace their development along the knowledge-building process. We will see that the model helps reveal the complexity of the field images construe and makes explicit what each image does in relation to the overall knowledge-building. Broadly, to build the hierarchical knowledge structure of chemistry, the secondary school chemistry textbooks use everyday images with relatively weak semantic density to link to phenomena in daily life (common-sense field) and power images with increasingly strong semantic density to develop chemical concepts underpinning the phenomena (uncommon-sense field). The images build the hierarchical knowledge structure of chemistry by developing increasingly complex theories while maintaining connections with the empirical phenomena. We will illustrate this through a group of images about the combustion of carbon.

Figures 5.44 to 5.46 show the combustion of carbon from the textbooks for Year 10 (Figures 5.44 and 5.45) and Year 11 (Figure 5.46). Figures 5.44 and 5.45 occur in two adjacent chapters in the Year 10 textbook.

Figure 5.44 A photo of a campfire (Zhang et al. 2015: 125)

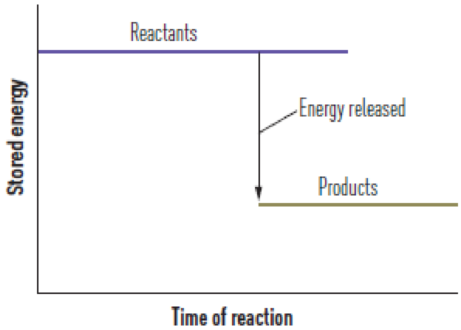


Figure 5.45 A line graph for the combustion of carbon (Zhang et al. 2015: 148)

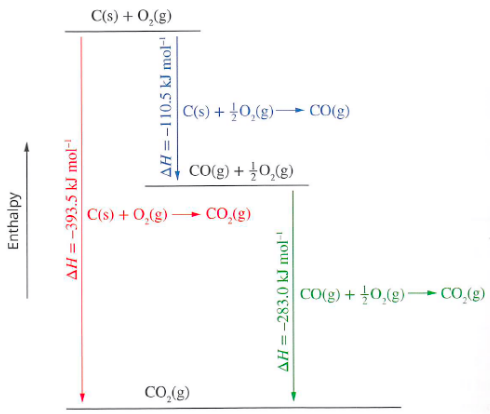


Figure 5.46 An energy level diagram for the formation of carbon dioxide from carbon and oxygen via carbon monoxide (Chan et al. 2018: 486)

Figure 5.44 occurs in Chapter three in the Year 10 textbook. It shows a campfire at a campsite, where the burning of wood typically provides light and warmth. In terms of external relations, this everyday image enjoys relatively weak relationalities to other meanings, suggesting that it construes a common-sense field. In terms of internal organization, it mainly construes an activity, i.e., the wood is burning. According to the model for the semantic density of images, Figure 5.44 locates at the bottom of the continuum [weak; single field-specific meaning] (SD− −), suggesting that the image has relatively weak semantic density. With this weak semantic density, the image depicts an everyday phenomenon of carbon combustion that can be easily understood.

Figure 5.45 occurs in the same chapter as Figure 5.44. It shows the change of energy brought about by the combustion of carbon. The vertical axis represents the energy level of chemical species involved in the combustion reaction and the horizontal axis indicates reaction time. The line graph indicates that as the reaction goes on, energy decreases from a higher level to a lower level. In contrast to Figure 5.44, this graph is a power image as it presents the technical concept ‘exothermic reaction’, which embodies relatively strong relationalities. The field construed by this image is relatively technical. In terms of internal organization, it construes a single field-specific meaning, i.e., an activity – the energy in chemical species changes from a higher level to a lower level. According to the model for the semantic density of images, Figure 5.45 locates at the bottom of the continuum within [strong] ([strong; single field-specific meaning]) (SD+), indicating that the image has a stronger semantic density than Figure 5.44. With this relatively strong semantic density, the image starts to introduce theoretical concepts to explain the everyday phenomenon of carbon combustion shown in Figure 5.44. The growth in semantic density builds the field’s technicality, shifting from the common-sense field presented in Figure 5.44 to the uncommon-sense field expressed in Figure 5.45.

The knowledge of energy change during combustion reaction is further developed through Figure 5.46 in Year 11. This image is an energy level diagram for the formation of carbon dioxide from carbon and oxygen via carbon monoxide. The diagram shows two pathways to the formation of carbon dioxide, in which the overall change of energy involved is the same. As analyzed in Section 5.2.1 (see the analysis of Figure 5.37), the diagram involves relatively strong relationalities and construes all of the four types of field-specific meanings. According to the model for the semantic density of images, the diagram locates at the top of the continuum within [strong] ([strong; multiple field-specific meanings]) (SD++), a strength that is stronger than Figure 5.45. With this relatively strong semantic density, Figure 5.46 expands the theoretical explanation of carbon combustion to a further level that involves much more chemical concepts, including enthalpy, exothermic reactions, incomplete and complete combustions, molecular composition, etc. The technicality of the field construed by this image grows even stronger than construed by Figure 5.45.

Overall, the increase of the images' semantic density from relatively weak through relatively strong to even stronger reveals that to build the knowledge of carbon combustion, chemistry needs everyday images to present empirical phenomena, thereby giving students a sense of what the objects they are learning look like in daily life, and power images with increasingly strong semantic density to build more and more complex technical understandings of the phenomena. The development of the images' semantic density attests that chemistry embodies a typical hierarchical knowledge structure where general theories are created from lower levels which in turn generalize knowledge from empirical phenomena (Bernstein 1999: 162). The images build this knowledge structure by developing increasingly complex theoretical explanations while maintaining connections to the empirical world. Interpreted from the perspective field, the knowledge-building through images involves shifting from common-sense fields to uncommon-sense fields and building increasingly strong technicality.

The above analysis indicates that to build knowledge about the same subject, the images' semantic density exhibit a consecutive progression. However, this does not mean that all images embody relatively weak semantic density in lower schooling levels and relatively strong semantic density in higher schooling levels. In fact, each level uses images embodying a range of semantic density, and everyday and power images are used throughout the levels. To more comprehensively account for the knowledge-building of chemistry through images across schooling levels, it is necessary to explore the range of semantic density across the curriculum stages, which is the issue the next section will focus on.

5.2.2.2 The rising range of images' semantic density in secondary school chemistry textbooks

As introduced in Section 3.3, the curriculum of secondary school chemistry in NSW, Australia includes six years from Year 7 to 12, categorized as three stages: Stage 4 (Years 7 and 8), Stage 5 (Years 9 and 10), and Stage 6 (Years 11 and 12). We focus on one topic – chemical reactions – that appears across all three stages. Images are used throughout the stages in these textbooks to represent chemical reactions, and so offer a means of seeing the development of their semantic density across the year levels. It will show that while images' semantic density increases across the stages, weaker semantic density remains through the stages. This means that the images exhibit a widening range of semantic density across stages. Through this range of semantic density, textbooks model students' ability to move between the more complex theoretical knowledge and the simpler everyday knowledge.

5.2.2.2.1 Stage 4

In Stage 4, Figures 5.47 and 5.48 show the simplest and most complex images about chemical reactions. They represent the range of the semantic density of images in this stage.

Figure 5.47 A photo of a rusted car (Zhang *et al*. 2014a: 213)

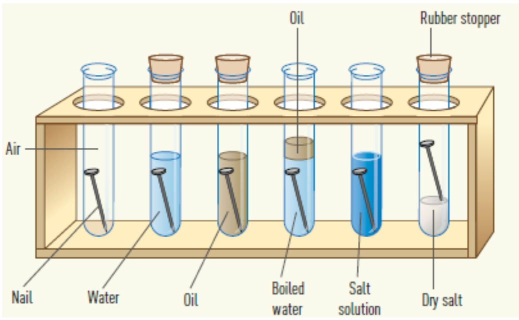


Figure 5.48 A diagram of a rusty nail experimental setup (Zhang *et al*. 2014a: 214)

Figure 5.47 shows a rusted car, illustrating a common issue for objects made of iron – rust. It depicts an example of the oxidation of iron in everyday life. In terms of the model for the semantic density of images, the image locates at the bottom of the continuum [weak; single field-specific meaning]) (SD– –), a relatively weak strength of semantic density. It construes a common sense field and does not involve technicality. Figure 5.48 shows six test tubes containing different environments: air, water, oil, boiled water, salt solution, and dry salt. In these test tubes, iron nails are placed to observe their respective rusting speed. As with the previous image, this diagram illustrates 'rusting' in some sense. However, it does so through a power image that locates at the bottom of the continuum of semantic density within [strong] ([strong; single field-specific meaning]) (SD+). The image involves a technical concept that enjoys relatively strong relationalities, i.e., oxidation of iron, which is shown by the six distinct environments with differing concentrations of air. This classification of reaction environments indicates that the key variable influencing the oxidation of iron is oxygen. The field construed by the diagram is relatively technical. With this range of semantic density from relatively weak to relatively strong, the images in Stage 4 connect to everyday phenomena and build relatively theoretical understandings underpinning them, though not that technical at this stage.

5.2.2.2.2 Stage 5

Stage 5 shows similarities to Stage 4 in that it also uses images with relatively weak semantic density that locates at the bottom of the continuum within [weak] ([weak; single field-specific meaning]) (SD– –). This is illustrated by Figure 5.49, which shows two segments of an orange.

Figure 5.49 Two segments of orange (Zhang *et al*.: 2014b: 134)

In terms of external relations, the image involves relatively weak relationalities as it depicts a kind of fruit we often eat in daily life, construing a common-sense field. In terms of internal organization, the image realizes a composition taxonomy between the overall orange and its peel and pulp. This image is used to illustrate a food that has weak acidity and hence is edible. Unlike Stage 4, however, Stage 5 includes images with semantic density locating between [strong; single field-specific meaning] and [strong; multiple field-specific meanings] in the continuum of the strength of semantic density (a scale between SD+ and SD++). Figure 5.50 is an example of such images.

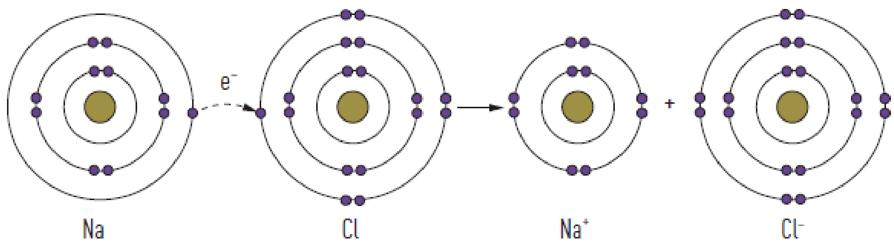


Figure 5.50 The formation of sodium chloride (Zhang et al.: 2014b: 109)

Figure 5.50 illustrates the formation of sodium chloride through an 'equation diagram', showing that a sodium chloride compound (Na^+ and Cl^-) is formed by a sodium atom (Na) donating an electron to a chlorine atom (Cl) (shown by the dashed arrow attached with an 'e^-'). This diagram involves several technical concepts, such as cations, anions, and atoms. The image thus involves relatively high relationalities and is a power image. In terms of internal organization, it construes two types of field-specific meanings: the activity that the sodium atom donating an electron to chlorine atoms to form sodium chloride and the compositional relation between atoms or ions (the whole) and nucleus and electrons (the parts). As the concepts in this image enjoy relatively strong relationalities, the above field categories are relatively technical. With more technical components enjoying relatively strong relationalities, Figure 5.50 exhibits stronger semantic density than Figure 5.49, pushing the strongest semantic density in Stage 5 to a higher level than in Stage 4. However, as shown by Figure 5.49, the weakest semantic density in Stage 5 remains at the bottom of the continuum of semantic density (SD– –), indicating that the images in Stage 5 show a wider range of semantic density than in Stage 4.

5.2.2.2.3 Stage 6

Stage 6 takes a further step than both Stages 4 and 5. In addition to once more using everyday images, for example, Figure 5.51, with relatively weak semantic density locating at the bottom of the continuum of semantic density ([weak; single field-specific meaning]) (SD– –), it also relies heavily on power images. In particular, it draws on particularly strong power images, as shown in Figure 5.52 for example.

Figure 5.51 Batteries (Chan et al. 2018: 390)

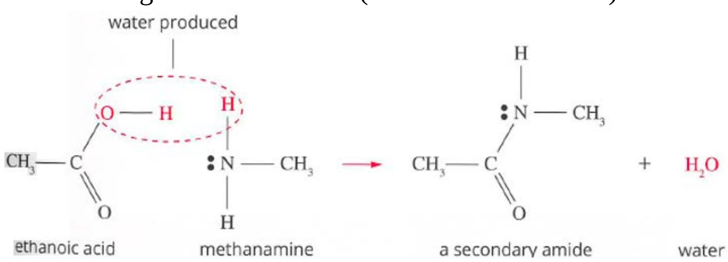


Figure 5.52 A diagram showing the formation of a secondary amide through a condensation reaction between ethanoic acid and methenamine (Chan et al. 2019: 359)

Figure 5.52 shows the formation of a secondary amide through a condensation reaction between ethanoic acid and methenamine. The concepts presented in this image are extremely technical and enjoy rather strong relationalities. For example, the structural formula of ethanoic acid is composed of two groups: methyl group (CH_3-) and carboxyl group (-COOH), which relate to two highly technical concepts in organic chemistry: alky groups and functional groups, which in turn relate to numerous other chemical meanings, such as atoms and covalent bonds. This indicates that the field construed by the image is highly technical. In terms of internal organization, the diagram construes multiple field-specific meanings. Firstly, it realizes an activity that ethanoic acid reacts with methenamine to form a secondary amide and water. Secondly, the structural formulas construe multiple levels of compositional taxonomies. For example, the structural formula representing ethanoic acid involves two levels of compositions: the molecule is composed of two groups, i.e., methyl group (CH_3-) and methenamine group (-COOH), which are in turn composed of atoms and covalent bonds. Thirdly, the diagram construes a classification - carboxyl group (-COOH) and amino group (-NH_2) are two subclasses of functional groups. Fourthly, the structural formulas also construe a property. As discussed in Section 4.1.3.1.2, the spatial arrangement of symbols construes spatial properties. With the four types of field-specific meanings that are all highly technical and enjoy rather strong relationalities, the diagram locates at the very top of the continuum of semantic density [strong; multiple field-specific meanings] (SD++), exhibiting extremely strong semantic density. Such images occur regularly in Stage 6 and push the semantic density of images to a much higher level than in Stage 5.

5.2.2.2.4 A widening range of semantic density of images through curriculum stages

The above analysis indicates that all stages use images with relatively weak semantic density to show everyday phenomena, but as the stage progresses, there is a steady expansion in the complexity of power images. This expands the semantic range of the images in the textbooks from lower to higher curriculum stages, as shown in Figure 5.53.

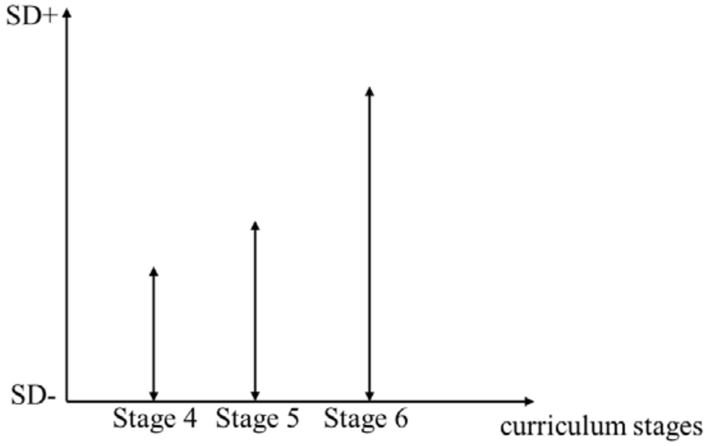


Figure 5.53 The widening range of the semantic density of images across the curriculum stages

The graph suggests that throughout the curriculum stages in secondary school chemistry, the knowledge expressed by the images in the textbooks maintains connections with the everyday empirical world while also reaching toward increasingly complex chemical theories. The textbook images thereby model a **growing range** of semantic density. In terms of field, the images in each stage construe common-sense fields and build increasingly technical fields across the stages. This once again shows how the images build the hierarchical knowledge structure of chemistry – in each year level, they develop increasingly complex fields involving more and more intricate and integrated chemical concepts while maintaining connections to the common-sense fields that present everyday phenomena.

This widening range of semantic density of images also provides insights into chemistry education research tradition's interpretation of the transitions between the different types of knowledge represented by images that are crucial for learning chemistry. In terms of Johnstone's (1991) chemistry triplet (macro, micro, and symbolic) (see Section 2.3.1 for a detailed introduction), across the three stages, the simplest images typically express macroscopic knowledge, i.e., what can be sensed, and the most complex images tend to express symbolic knowledge, i.e., symbols, formulas, and equations. The two images in Figures 5.51 and 5.52 are cases in point. In chemistry education research, it has been widely recognized that transitions between the different types of knowledge are critical for students to be successful in learning chemistry (Chittleborough et al. 2005; Gabel 1993; Johnstone 1991). However, chemistry education researchers seem to recognize only the transition of the referents in the images and remain unaware of the fundamental issue underlying chemistry learning - the change of meaning along with the transition. The widening range of images' semantic density in textbooks shown in this study suggests that the key to the transition from macroscopic knowledge to symbolic knowledge is that students are expected to be able to handle the images' increasing semantic density.

The above analysis of images in secondary school chemistry textbooks across curriculum stages demonstrates that students are expected to engage with images of different semantic density at each year level. This is often made explicit by a special type of image in chemistry – composite images, which often brings together images with different strengths of semantic density and allows for seeing how their meanings connect up. The next section will examine the knowledge-building of chemistry through these composite images.

5.2.2.2.5 Integrating semantic density through composite images

Most images in our data occur on their own. However, at times, the textbooks use composite images to integrate images with different degrees of semantic density, bridging knowledge with different levels of complexity. In terms of field, composite images integrate fields with varying degrees of technicality. For example, they can bridge everyday common sense and technical uncommon sense fields, as shown in Figure 5.54.

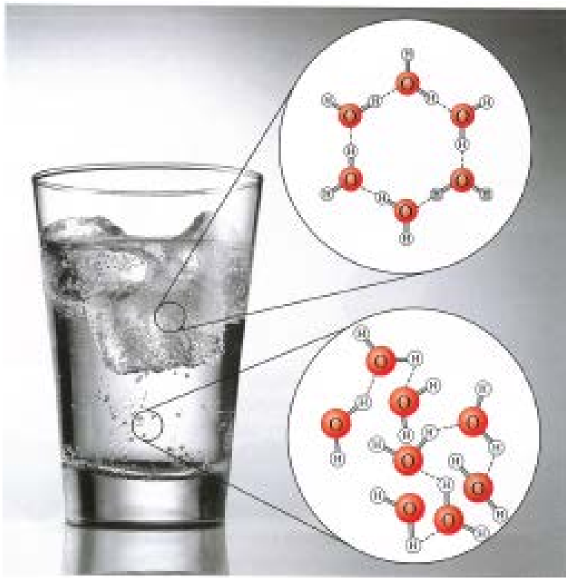


Figure 5.54 Ice is less dense than liquid water (Chan *et al*. 2018: 183)

Figure 5.54 brings together an image with relatively weak semantic density that shows a cup of water and ice cubes and two power images with a stronger semantic density that depicts the molecular organization of the ice cubes and water. Combining these images relates the relatively simple meanings of ice and water to the complex meanings situated within the field of chemistry. In this sense, the composite image offers a way of 'bridging' meanings with different degrees of complexity, potentially helping compound meanings from simpler into complex forms. In terms of field, the composite image allows for movement between the everyday common-sense field and the technical uncommon-sense field. Grammatically, the bridging is realized through intra-semiotic relations (O'Halloran 2005) between the different images. This intra-semiotic relation, as suggested by our data, can be modeled as '**elaboration**', a type of logico-semantic relations (Halliday and Matthiessen 2014). The photo shows the ice cubes and liquid water, while the diagrams elaborate these entities by restating them in terms of their molecular organization. It is this elaboration at the microscopic level that shifts the degree of semantic density from weaker to stronger.

Elaborating a macroscopic phenomenon at the microscopic level appears to be the key pattern of how composite images are organized to build chemistry knowledge. This relation also organizes composite images integrating images that both construe relatively technical fields. Figure 5.55 is an example of such composite images.

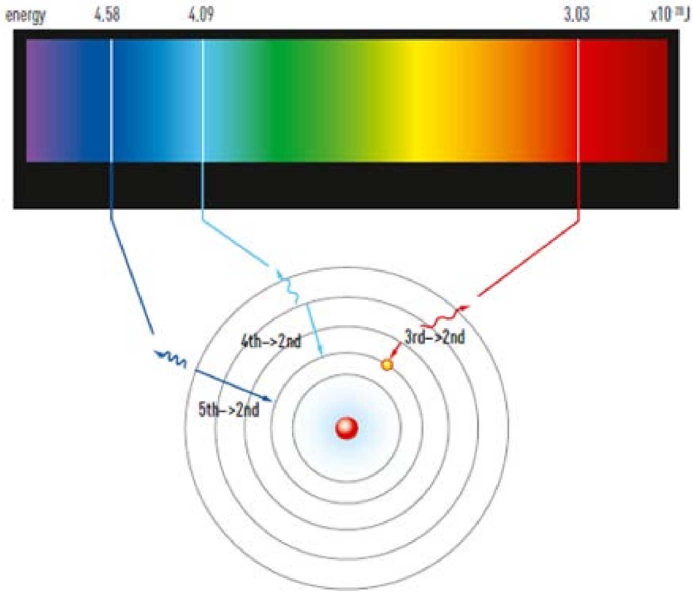


Figure 5.55 Electron jumps (Zhang et al. 2014b: 20)

Figure 5.55 illustrates energy released from electron jumps through a spectrum graph and a diagram showing electron jumps between electron orbits. The graph involves a concept that embodies relatively strong relationalities – electron jumps. All the different types of light colors represent different energy released from various electron jumps, which in turn relate to numerous technical concepts, such as electrons, electron orbits, and energy. In terms of field-specific meanings, the graph construes a classification taxonomy between the overall spectrum (the class) and the various light colors (the subclasses). In terms of the model for the semantic density of images, the graph locates at the bottom of the continuum within [strong] ([strong; single field-specific meaning]) (SD+). Like the graph, the diagram exhibits relatively strong relationalities as it presents several technical concepts, including electron jumps (denoted by ‘5th→2nd’, ‘4th→2nd’, and ‘3rd→2nd’), electron orbits (represented by the circles), nucleus (the red ball in the center), and electrons (the small yellow ball). In terms of field-specific meanings, the diagram construes three field categories: activities – the three electron jumps, classifications – the different types of electron jumps (‘5th→2nd’, ‘4th→2nd’, and ‘3rd→2nd’), and composition – an atom is composed of a nucleus and electrons. With all three field categories located in the technical field of chemistry, the diagram exhibits a stronger semantic density than the graph. The two images are integrated through the elaboration relation, whereby the diagram interprets the emitted spectrum (the three light colors) by explaining how they are produced at the microscopic level. The composite image thus builds the knowledge of electron jumps by enabling the movement between the weaker and stronger degrees of semantic density and relating the empirical phenomenon with the underpinning theoretical explanation.

The above analysis shows that composite images integrate different degrees of semantic density in a single snapshot and relate them by the elaboration relation, bridging either the everyday common-sense field and technical uncommon-sense field or uncommon-sense fields with different levels of technicality. This once again shows how images in chemistry textbooks build the hierarchical knowledge structure of chemistry: they develop increasingly uncommon-sense fields of chemistry that involve progressively intricate and integrated chemical concepts while maintaining connections to common-sense fields that present everyday phenomena.

5.3 Summary

This chapter has addressed the second research question of this study – how images build the knowledge of secondary school chemistry. This question includes two sub-questions: (1) what meaning images in the chemistry textbooks construe and how they realize the meaning; (2) how the images develop across schooling levels to build the hierarchical knowledge structure of chemistry in the textbooks. To explore the first sub-question, this chapter has examined the meanings construed by images in chemistry by establishing a model of these images in terms of field that links with the grammatical resources realizing the meanings. The finding shows that images in secondary school chemistry textbooks construe three broad types of field-specific meanings: item, activity, and property. Item is the static perspective of field. The images construe either singular item or taxonomic relations between multiple items. Particularly, they show a strong capacity for construing multiple levels of composition and classification taxonomies. In terms of composition, chemical images can establish multiple levels of part-whole relations between components that constitute a chemical matter. These multiple levels of composition are realized by a spatial analytical structure recursively embedded within another, allowing for construing as complex compositions of chemical matter as possible. In terms of classification, chemical images can set up a cline of type-subtype relation between different types of chemical matter. These classification relations can be realized by either multi-leveled overt taxonomy structures or multi-leveled covert taxonomy structures. The composition and classification taxonomies can occur simultaneously in a single chemical image.

Activity is the dynamic perspective of field. The images in the secondary school chemistry textbooks construe either unmomented or momented activities, with the former conveying the knowledge of a single chemical change while the latter typically showing how initial reactants reach final products. Grammatically, the unmomented activities tend to be realized by a single transformation process. Momented activities, on the other hand, involve further complications. They can be further described in terms of their cyclicality and the relation connecting them. In terms of cyclicality, they can be either cyclical or non-cyclical. Grammatically, the cyclical momented activities are realized by conversion processes, and the non-cyclical ones by complexes of transformation processes. In terms of the relation linking the activities, they can be related by either implication or expectancy. The expectancy sequences tend to be realized by complexes of transformation processes or conversion processes, and the implication sequences are typically realized by complexes of transaction processes and transformation processes.

Property describes the features of item and activity. The finding shows that the images in secondary school chemistry textbooks realize all the properties described in Doran and Martin's (2021) model and include a further distinction within the non-gauged arrayed property between categorical and non-categorical. These properties construe various aspects of chemical matter, including electronic polarities of electrons and protons, the spatial arrangement of these particles within atoms, pH values of substances, etc. When describing chemical reactions, these properties tend to show the amount of energy absorbed or released during reactions and the amount of time the reactions take. A significant distinction between the categorical and non-categorical properties is that the former tends to be typological while the latter is typically topological. Grammatically, these properties are not realized by any particular structures through which items and activities are realized, but by non-structural resources. The primary resource is symbolic labelings, including numbers, symbols (e.g. ‘+’ and ‘-‘), mathematical equations, etc. Another significant source for realizing the properties, especially non-gauged arrayed properties, is color gradations. In addition, the spatial arrangement of entities in images is also a key grammatical resource realizing the topological spatial property.

To investigate how the images develop to build the hierarchical knowledge structure of chemistry, this chapter has studied the development of images in secondary school chemistry textbooks in terms of their semantic density. To capture the development, this chapter has established a model for the semantic density of images in secondary school chemistry textbooks. The model builds upon two factors: internal organization and external relations. Internal organization concerns the number and types of field-specific meanings construed by an image, including composition, classification, activity, and property. External relations describe the technicality of these field-specific meanings. The model privileges external relations over internal organizations because the former determines whether a particular field is situated in a complex constellation of meaning.

With this model, this chapter has analyzed the knowledge-building through images in secondary school chemistry textbooks by tracing their development in terms of semantic density. Specifically, it has examined the development of images used for building knowledge about the same subject and the development of the range of semantic density images exhibit across the curriculum stages. The former concerns the roles of images in building specific related knowledge, while the latter focuses more on the overall development of images’ semantic density. The finding shows that when building knowledge about the same subject, the images’ semantic density exhibits a consecutive progression: from relatively weak through relatively strong to even stronger. This suggests that to build certain knowledge of chemistry, the textbook needs to use everyday images to present empirical phenomena, thereby giving students a sense of what the objects they are learning look like in daily life, and power images with increasingly strong semantic density to build more and more complex technical understandings of the phenomena. The development of the images' semantic density attests that chemistry embodies a typical hierarchical knowledge structure where general theories are created from lower levels which in turn generalize knowledge from empirical data. The images build this knowledge structure by developing increasingly complex theoretical explanations while maintaining connections to the empirical world. Interpreted from the perspective field, the knowledge-building through images involves shifting from common-sense fields to uncommon-sense fields and building increasingly strong technicality.

The analysis of the development of the range of semantic density of images shows that throughout the curriculum stages in secondary school chemistry, the knowledge expressed by the textbook images maintains connections with the everyday empirical world while also reaching toward increasingly complex chemical theories. The textbook images thereby model a growing range of semantic density. In terms of field, the images in each stage construe common-sense fields and build increasingly technical fields across the stages. This once again shows how the images build the hierarchical knowledge structure of chemistry – in each year level, they develop increasingly complex fields involving more and more intricate and integrated chemical concepts while maintaining connections to the common-sense fields that present everyday phenomena.

The above findings indicate that to learn chemistry knowledge through images, students are expected to engage with images embodying different degrees of semantic density at each stage. This is often made explicit by a special type of image in chemistry – composite images. The analysis shows that composite images bring together images with different strengths of semantic density and allow students to move between the different levels of complexity. In terms of field, composite images integrate fields with varying degrees of technicality. Our data indicate that they either bring together everyday common-sense field and technical uncommon-sense field or integrate uncommon-sense fields with different levels of technicality. Grammatically, the different fields construed by the component images within the composite are linked through the logic-semantic relation of elaboration in SFL, through which the technical field provides theoretical explanations at the microscopic level for the empirical phenomena.

Chapter 6 Knowledge-building through Multisemiosis

Secondary school chemistry textbooks employ language, chemical formalisms, and images to build chemistry knowledge. This thesis has thus far explored the meaning-making of chemical formalisms and images individually and how they develop across schooling levels in the textbooks to build chemistry knowledge. This final chapter will first add in language and then bring together the different semiotic resources to investigate how secondary school chemistry is built through multisemosis. It will first establish a multisemotic picture of chemistry knowledge in terms of field (Section 6.1) and then examine how the various semiotic resources interact in chemistry texts across schooling levels to build knowledge (Section 6.2). Section 6.1 first briefly analyzes the field-specific meanings construed by the English language in the NSW secondary chemistry textbooks (Section 6.1.1) and then establishes a multisemiotic picture of secondary school chemistry in terms of the field meanings each resource realizes (Section 6.1.2). Following this, Section 6.2 explores how the resources interact in chemistry texts across schooling levels to build chemistry knowledge. Specifically, it focuses on the interaction between language and images (Section 6.2.1) and between language and chemical formalisms (Section 6.2.2). Finally, Section 6.3 summarizes the findings in this chapter.

6.1 Multisemiosis and the knowledge of chemistry

Secondary school chemistry textbooks use language, chemical formalisms, and images to construe chemistry knowledge. Chapters 4 and 5 have described the knowledge realized by chemical formalisms and images in the textbooks from the perspective of field. This section will first briefly describe the chemistry knowledge construed by language in terms of field and then bring together all the different semiotic resources to present a multisemiotic picture of chemistry.

6.1.1 Language in chemistry and field

Chemical language is a typical scientific language. It shares many similarities to scientific language in construing knowledge as documented by numerous studies

(Doran 2018c; Halliday and Martin 1993; Hao 2020; Martin 1989). In terms of field, it organizes the knowledge of chemistry through construing deep taxonomies, intricate activity sequences, and various properties.

To organize the knowledge of chemical matter, chemical language tends to develop multiple levels of taxonomic relations between items - both compositional and classificational. For instance, Example (6:1) construes the structure of atoms by setting up two levels of composition.

(6:1) *All atoms are made up of a small, positively charged nucleus surrounded by a much larger cloud of negatively charged electrons, as shown in Figure 3.1.1. The nucleus is made up of two types of subatomic particles – protons and neutrons.* (Chan et al. 2018: 146)

The first level of composition is between ‘atom’ (the whole) and ‘nucleus’ and ‘electron’ (the parts) (realized by the clause complex *all atoms are made up of a small, positively charged nucleus surrounded by a much larger cloud of negatively charged electrons*). The second level is between ‘nucleus’ (the whole) and ‘proton’ and ‘neutron’ (the parts) (realized by the clause *the nucleus is made up of two types of subatomic particles – protons and neutrons*). The two levels of compositional taxonomies enable chemistry to organize the relatively complex knowledge of atomic structure.

In addition, chemical language also construes the knowledge of chemical matter by establishing multiple levels of classificational taxonomies between items. In particular, the multiple levels of classifications allow chemistry to organize intricate classifications of different chemical substances. For instance, Example (6:2) construes different types of substances by setting up three levels of classificational taxonomies.

(6:2) *It is very important to be able to identify substances as pure or mixture. A substance is considered to be pure if all the particles that make it up are the same. There are two different types of pure substances. An element is a pure substance because every atom within the substance is the same. Compounds can be pure as well. A compound is formed when different atoms are chemically bonded together, like water. There are also two different types of compounds – molecular compounds and ionic compounds.* (Zhang et al. 2013: 178)

The first level of classification is between ‘substances’ (superordinate) and ‘pure substances’ and ‘mixture substances’ (subordinate), and the second level is between ‘pure substances’ (superordinate) and ‘elements’ and ‘compounds’ (subordinate). At the third level, ‘compounds’ itself includes two subtypes – ‘molecular compounds’ and ‘ionic compounds’. The three levels of classifications build relatively deep classificational taxonomies among the various chemical substances.

In addition to the knowledge of chemical matter, language in chemistry also organizes the knowledge of chemical reactions by construing different activities. It can express a chemical reaction by construing an unmomented activity, such as Example (6:3). This example employs a material clause to construe the decomposition of calcium carbonate.

(6:3) *Calcium carbonate decomposes into calcium oxide and carbon dioxide*.

(Zhang et al. 2015: 154)

Unmomented activities can also be itemized, known as “**itemized activities**” (Doran and Martin 2021). These activities are typically realized by nominalizations, as illustrated in Example (6:4). ‘oxidation’ is an itemized activity transformed from ‘something is oxidized’.

(6:4) *Reactions in which oxygen was a reactant were described as <u>oxidation</u>*.

(Chan et al. 2018: 373)

Chemical language can also construe activity sequences, or momented activities in Doran and Martin’s (2021) term, to organize the knowledge of chemical reactions. These activities can be related through either expectancy or implication (Doran 2018c). When describing chemical reactions, chemical language tends to realize momented activities related by expectancy, typically a temporal or procedural relation. For instance, Example (6:5) describes three activities in an experiment related by expectancy. That is, the activities of a copper metal being placed in the solution, the copper being coated with a black deposit, and the solution turning a pale blue color are related by purely a temporal relation – one is expected to happen after another.

(6:5) *When a coil of copper metal is placed in a solution of silver nitrate, the copper becomes coated with a black deposit. After a while, the solution turns a pale blue color.* (Chan et al. 2018: 357)

In contrast, when explaining chemical reactions, chemical language tends to construe momented activities related through implication, which typically explains natural phenomena by specifying the cause-and-effect relations between activities (Martin 1993, Wignell et al 1989). For instance, Example (6:6) explains why hydrocarbon burns with a yell, smoky or sooty flame by realizing two activities related by implication, i.e., the presence of glowing carbon particles causes the hydrocarbon burning with a yellow, smoky or sooty flame.

(6:6) *The hydrocarbon burns with a yellow, smoky or sooty flame, due to the presence of glowing carbon particles*. (Chan et al. 2018: 323)

Importantly, for momented activities related by expectancy, they can construe different types of chemical reactions in terms of the overall change they undergo. For instance, Example (6:7) construes two activities – 'magnesium reacting with hydrochloric acid' and 'produce magnesium chloride and hydrogen gas' – unfolding linearly. The **linear** momented activities construe an irreversible chemical reaction.

(6:7) *Magnesium reacts with hydrochloric acid to produce magnesium chloride and hydrogen gas.* (Zhang et al. 2015: 121)

In contrast to the linear momented activities, some other momented activities can be **cyclical**. For instance, Example (6:8) shows two activities forming a cycle, i.e., it starts from nitrogen and hydrogen gas and ends at the same chemical substances. On the whole, the momented activities constitute a cycle. The cyclical momented activities construe a reversible chemical reaction. The distinction between linear and cyclical for momented activities complements Doran and Martin's (2021) distinction between linear and cyclical for a single activity.

(6:8) *Nitrogen and hydrogen gas molecules form ammonia in the forward reaction. Ammonia molecules decompose into nitrogen and hydrogen gas molecules in the reverse reaction.* (Chan et al. 2019: 60).

The above has shown that language in chemistry organizes the knowledge of chemical matter and chemical reactions by construing both deep taxonomies and intricate activities in terms of field. Complementing this, chemical language also construes various properties to organize the knowledge of both chemical matter and chemical reactions. To describe the qualities of chemical matter, it construes both qualitative and spatial properties. Example (6:9), for instance, construes the qualitative properties of charge polarity that protons, neutrons and electrons bear (shown in <u>*positively charge*</u> *protons*, <u>*neutral*</u> *neutrons*, and <u>*negatively*</u> *charged electrons*). In addition, it also construes a spatial property of the relative locations of nucleus and electrons (shown in the <u>*central*</u> *nucleus* and *electrons* <u>*spin around*</u> *the nucleus*).

(6:9) *Positively charged protons and neutral neutrons make up the central nucleus of the atom while negatively charged electrons spin around the nucleus in layered electron shells*. (Zhang et al. 2013: 142)

To describe chemical reactions, chemical language can also construe temporal properties, as illustrated in Example (6:10).

(6:10) *Nitrogen gas reacts with hydrogen gas* *rapidly* *under heating*. (Chan et al. 2019: 59)

All these properties can be graded and potentially ordered into arrays (Doran and Martin 2021). This allows chemical language to construe gradable properties of chemical matter and chemical reactions. For instance, Example (6:11) shows that the energy of electrons is gradable and electrons at the outer shell have the highest energy.

(6:11) *Electrons in the outer shell have the* *highest* *energy* (Chan et al. 2018: 98)

These gradable qualities can also be measured or quantified, known as "gauged properties" (Doran and Martin 2021), as illustrated in Example (6:12).

(6:12) *An electron in the first energy level has -13.6 eV of energy.* (Chan et al. 2019: 99)

Gauged properties can also describe chemical reactions. For instance, Example (6:13) shows the exact time for the reaction to reach an equilibrium state.

(6:13) *The reaction between nitrogen gas and hydrogen gas reaches an equilibrium state by six seconds*. (Chan et al: 2019:63)

This section has shown that language in chemistry organizes the knowledge of chemical matter and chemical reactions by construing deep taxonomies, intricate activities, and various properties. With the description of language completed, we can now bring together language, chemical formalisms, and images to present a multisemiotic picture of chemistry knowledge from the perspective of field in the next section.

6.1.2 A multisemiotic picture of chemistry knowledge in terms of field

The secondary school chemistry textbooks employ language, chemical formalisms, and images to construe knowledge. In terms of field, each of the semiotic systems realizes certain types of field meanings that constitute the knowledge of secondary school chemistry. To have an overall multisemiotic picture of the chemistry knowledge, we need to bring together the different semiotic resources and compare the similarities and differences in the field-specific meanings they construe. This allows us to see the functionalities of each semiotic resource in organizing chemistry knowledge and how they complement each other to expound the field of chemistry.

As noted in Chapter 4, chemical formalisms in secondary school chemistry include three major types: chemical formulas, structural formulas, and chemical equations.

Both chemical formulas and structural formulas specialize in construing the breadth of composition: chemical species can comprise as many types of elements as possible. The potential for realizing indefinitely broad composition taxonomies is afforded by the univariate structures in the formulas (iterations of chemical symbols) (see Section 4.1.3). In contrast to the broad composition taxonomies, formulas are rather restricted in developing the taxonomies' depth. They typically construe only one level of composition, for example, a molecule (the whole) and its component atoms (the parts). In addition to the composition taxonomies, structural formulas realize spatial properties, allowing them to show both what atoms occur in a molecule and their spatial arrangement. The properties reflect, to some extent, how chemistry conceives the molecule's actual structure. As discussed in Section 4.1.3, the compositional taxonomies and spatial properties can be viewed as the formulas' disciplinary affordance, which meets chemistry's needs of representing the technical knowledge of chemical substances' composition and internal structure.

These field-specific meanings about items are complemented by activities realized by chemical equations. Chemical equations in secondary school chemistry specialize in construing two types of activities: unmomented activities and momented activities, realizing irreversible reactions and reversible reactions. The momented activities involve only two activities (forward reactions and reverse reactions) related through expectancy (see Section 4.1.3.1.3). The description of these activities is enriched by the two field properties: a gauged property of coefficients and a qualitative property of physical states. The former enables chemical equations to describe the quantity relations of a reaction at the microscopic level (the molecule level), while the latter allows for construing chemical reactions at the macroscopic level (the material world that can be sensed). Working in tandem with these properties, the activities construed by chemical equations meet chemistry's needs of representing the technical knowledge of the changes that chemical substances undergo during reactions.

The above shows that chemical formalisms specialize in construing a relatively small range of field-specific meanings concerning compositions and internal structure of chemical matter and chemical changes these substances undergo. Chemical formalisms are thus designed for representing relatively narrow but highly technical fields. Images in secondary school chemistry, in contrast, cover more field-specific meanings situated in both everyday common-sense and technical uncommon-sense fields. As discussed in Chapter 5, images show a strong capacity for construing multiple levels of classification and composition taxonomies in a single snapshot, which builds the taxonomies' depth. However, the images seem not to be as specialized in developing the breadth of the compositional taxonomies. The interplay between chemical formulas and images thus forms a complementarity. For instance, Example (6:14) is the chemical formula for amidephrine molecules. It construes only one level of compositional taxonomy between the entire molecule (the whole) and the different atoms including carbon (C), hydrogen (H), nitrogen (N), oxygen (O), and sulfur (S) (the parts). However, the compositional taxonomy is relatively broad as it includes five components at the same level.

(6:14) $C_{10}H_{16}N_2O_3S$

Chemical images, in contrast, tend to develop deeper but much less broad compositional taxonomies. For example, Figure 6.1 shows the atomic structure of lithium atoms. It construes two levels of compositional taxonomies – the first level is between the lithium atom (the whole) and the nucleus (the entirety of the red and green balls in the center) and electrons (the blue ball), and the second level is between the nucleus (the whole) and the protons (the red balls) and neutrons (the green balls). Each level includes only two components, construing relatively narrow compositional taxonomies. Therefore, the relatively broad but much less deep compositional taxonomies construed by chemical formulas complement the relatively deep but much less broad compositional taxonomies construed by chemical images.

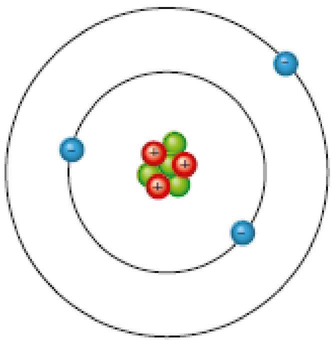


Figure 6.1 A diagrammatic representation of the atomic structure of lithium atoms

Another special feature of the compositional taxonomies construed by chemical images is that they tend to be accompanied by spatial properties since the images typically show the relative spatial locations of the components. In other words, chemical images construe compositional taxonomies and spatial properties simultaneously. Figure 6.1, for example, shows not only the compositions of a lithium atom but also relative positions of the components in the atom – the nucleus (the entirety of red and green balls) is in the center and electrons circle around it (the blue balls). The compositional taxonomies and spatial properties are presented in the image simultaneously. This is somewhat similar to structural formulas as they construe both the two field-specific meanings at the same time. Nonetheless, the spatial properties realized by images are more topological as they construe meanings of gradation, continuous change, and relations of relative nearness or connectedness (Lemke 1998: 87). Structural formulas thus can be viewed as sitting between image and symbolic modes such as chemical formulas and language.

In terms of activities, the momented activities construed by images in secondary school chemistry textbooks are more diverse than those construed by chemical equations. The momented activities realized by chemical equations involve only two activities (forward reactions and reverse reactions) related through expectancy, forming cyclical momented activities. In contrast, the momented activities construed by chemical images can include as many activities as possible related through expectancy or implication, which form either cyclical or non-cyclical momented activities (see Section 5.1.2). These momented activities tend to illustrate the procedures of yielding the final products from the initial reactants in chemical reactions or the causes leading to reactions. However, although images in chemistry construe a wider range of activities, they fall short of describing the quantitative relations of reactions (the molar ratios in which chemical species react and are formed) and the physical states of the chemical species involved in the reactions as represented by chemical equations (see Sections 4.1.1 and 5.1.2). This suggests that images in chemistry and chemical equations specialize in construing different aspects of the knowledge of chemical reactions, with the former focusing on procedures involved in chemical reactions or causes leading to the reactions and the latter on types of chemical species involved in the reactions and their quantitative relations and physical states.

Images in chemistry also construe various properties, including qualitative, spatial, and temporal. A special feature of these properties is that they can be topologically arrayed, which is afforded by images' capacity for displaying topological meaning (Lemke 1998). This allows the properties to be ordered along an array with infinitely small degrees of gradation, for example, the pH value and reaction time discussed in Section 5.1.3. Such topological arrays are beyond the construal power of both chemical formalisms and language.

The above analysis indicates that chemical formalisms and images realize most of the field-specific meanings of chemistry, covering classification and composition taxonomies, activity sequences, and properties. As shown in Section 6.1.1, most of these types of field-specific meanings can be realized by language. This then leads to the question as to why chemistry needs all these different semiotic resources to construe its knowledge. The most proper answer to this, we would argue, is that they form a complementarity. Each semiotic system has its functionalities for organizing chemistry knowledge, and they complement each other to constitute the complete picture of chemistry. In terms of the construal of the knowledge of chemical matter, although language can show elaborate compositional taxonomies, it falls short of showing the relative locations of the components topologically as shown by structural formulas and images. Although language can describe the relative locations, the meaning expressed is typological (Lemke 1998) and does not display the components' accurate positions. In field's terms, language is thus inadequate for construing topological spatial properties. Therefore, when chemistry needs to show this, it can turn to structural formulas or images.

In construing knowledge of chemical reactions, language shows more elaborated implication relations between activities, which complements chemical equations and images as the former involves only expectancy and the latter is relatively limited in construing implication sequences. For example, the clause complex in Example (6:15) construes two activities related through implication, i.e., the presence of glowing caron particles causes the hydrogen carbon burning with a yellow, smoky or sooty flame. This cause-and-effect relation is not shown in the corresponding chemical equation for the incomplete combustion of methane shown in Example (6:16).

(6:15) *The hydrocarbon burns with a yellow, smoky or sooty flame, due to the presence of glowing carbon particles*. (Chan et al. 2018: 324)

(6:16) $2CH_4(g) + 3O_2(g) \rightarrow 2CO(g) + 4H_2O(l)$

Nor does the image that presents this reaction, shown in Figure 6.2, indicate the cause-and-effect relation. It only depicts the different flames produced by incomplete and complete combustion. Therefore, when chemistry requires more elaborated implication sequences, it can turn to language.

Figure 6.2 Incomplete combustion and complete combustion of methane

When construing quantitative relations between reactants and products, chemical equations show a unique advantage over language and images since they show a balanced relation between the quantities of atoms before and after the reactions. In terms of field, chemical equations construe a special gauged property that is not afforded by language and images. In construing the properties of chemical matter and chemical reactions, images' capacity for displaying topological meaning enables them to construe topological arrays that involve indefinitely small gradations. They have stronger construal power over language and chemical formalisms in expressing the knowledge involving graded properties such as pH values and reaction speed.

The above shows that when one semiotic resource falls short, another will take over. Working together, they complement each other to form the full picture of chemistry knowledge. Table 6.1 summarizes the field-based affordances of each resource in chemistry.

Table 6.1 Field-based affordance of language, chemical formalisms, and images in chemistry

Field			**Language**	**Chemical formalisms**	**Images**
Taxonomy	Classification	Breadth	Yes	No	Yes
		Depth	Yes	No	Yes
	Composition	Breadth	Yes	Yes	No
		Depth	Yes	No	Yes
Activity	Unmomented	Linear	Yes	Yes	Yes
		Cyclical	Yes	No	Yes
	Momented	Expectancy	Yes	Yes	Yes
		Implication	Yes	No	Yes (restricted)
Property	Qualitative	Arrayed	Yes (typological)	No	Yes (topological)
		Non-arrayed	Yes	Yes (including a unique gauged property)	Yes
	Spatial-temporal	Arrayed	Yes (typological)	No	Yes (topological)
		Non-arrayed	Yes (typological)	Yes (spatial only)	Yes (topological)

In this section, we have discussed the functionality of each semiotic resource in organizing chemistry knowledge and how they complement each other to form a multisemiotic picture of chemistry. However, this is only a holistic view of the different resources' affordance in construing chemistry knowledge, and it is not yet clear how they interact in chemistry texts to build the knowledge of chemistry. We will thus turn to this issue in the next section.

6.2 Building chemistry knowledge through interaction between semiotic resources in chemistry texts

Language is the most common semiotic resource in secondary school chemistry textbooks for building knowledge. It is the first resource used to develop the technicality of chemistry by establishing deep compositional and classificational taxonomies, organizing technical activities into sequences, and construing various technical properties. However, as chemistry is inherently multisemiotic, language typically works with chemical formalisms and images to build the hierarchical knowledge structure of chemistry by developing increasingly integrative and generalized chemical concepts and connecting to the empirical phenomena. This section will thus explore the interaction between language and image (Section 6.2.1) and between language and chemical formalisms (Section 6.2.2) in chemistry texts from our data to examine how the interaction builds the knowledge structure of chemistry. Note that this section will not have a sub-section exploring the interaction among the three broad resources in a single chemistry text because our data show that language and chemical formalisms are always integrated as a whole text and examining the interaction among the three resources in a text hardly bears any differences from probing image-text relations.

6.2.1 Interaction between language and image

Texts comprising language and images prevail in secondary school chemistry textbooks across schooling levels. They are a crucial part of the textbooks for building chemistry knowledge. However, it is not yet clear how they interact in texts to build chemistry knowledge and how the interaction develops as the knowledge being built gets increasingly complex. This section will examine these issues in terms of LCT’s Semantics (Maton 2014). It will show that at the lower schooling levels, images tend to show stronger semantic gravity but weaker semantic density than language in order to connect theoretical concepts to everyday empirical phenomena. At the higher schooling levels, images gradually foster relatively strong semantic density by consolidating meaning from language. The interaction at the higher schooling levels builds increasingly integrative and generalized concepts and thus extends chemistry knowledge. The following will show this through a series of texts across the schooling levels.

Year 7, the lowest level in secondary schooling, tends to employ texts involving interaction between language and photographs to build chemistry knowledge. Text 1 is a typical example.

Text 1 (Zhang et al. 2013: 165) (original emphasis in bold fonts)

> **Vaporization** is the process of liquids turning into gases. The particles in a liquid jostle around and between each other. As the water in Figure 4.39 is heated the particles gain more energy, move faster and take up more space. If there is enough heat energy the particles are able to break free of the forces that hold them together. At 100℃, the boiling point of water, all of the liquid water particles will be changing state to become gas—this is called **boiling**.

Figure 4.39 The water vapor is between the kettle spout and the steam cloud

As shown in Text 1, the language text introduces the chemical concept of “vaporization” and the image exemplifies it with water boiling in a kettle. In terms of Unsworth’s (2007) model of language-image relations, Text 1 involves an exemplification relation of “image instantiates text” (see Section 2.2.2.1.1 for a detailed introduction to this relation). That is, the image presenting a boiling kettle is an illustration of the concept of vaporization. However, this interpretation of the language-image relation appears inadequate for explicating how the interaction builds knowledge. An alternative approach to this that enables ‘seeing’ the knowledge-building is the LCT dimension of Semantics (Maton 2014). As a reminder, Semantics includes two critical conceptual tools: semantic density (the complexity of meaning) and semantic gravity (the degree of context-dependence of meaning). In terms of semantic density, the language text embodies stronger complexity than the image. As we can see, the former comprises nine clauses, most of which construe technical field-specific meanings about vaporization. In terms of field, the relational clause *vaporization is the process of liquids turning into gases* construes an itemized activity (Doran and Martin 2021) of “vaporization” and an activity of “liquids turning into gases”. The itemized activity is further explained at the microscopic level by the material clause *the particles in a liquid jostle around and between each other*, which construes an activity of the particles’ movement. The rest of the clauses illustrate vaporization with the example of boiling water. In terms of field, they realize three activities of water particles' movement under heating, one activity of water changing physical states, and a gauged property of water's boiling point. In contrast to the numerous field-specific meanings of items, activities, and properties realized by the language text, the image construes only an activity – water vapor coming out of a kettle. The significantly larger elaboration of meaning in language indicates that the language text embodies much stronger semantic density than the image. This means that the language text construes significantly broader field meanings of "vaporization". A question then arises – what is the role of the image? We may unravel this in terms of both semantic density and semantic gravity.

In terms of semantic density, by bringing together the image with the language, the text links a relatively simple everyday activity of water boiling in a kettle to a complex technical concept of vaporization. In terms of field, the image with weaker semantic density connects the uncommon-sense field with the common-sense field. That is, it enables readers to relate the everyday activity of water boiling to the technical concept of vaporization.

In terms of semantic gravity, the image links the theoretical concept of vaporization to a specific empirical phenomenon. As shown in the above analysis, the language text mainly provides a theoretical explanation of water vaporization at the microscopic level, i.e., water particles' movement under heating leads to water vaporization. Such a theoretical construct is generalized and abstracted away from specific instances of water vaporization. The image, in contrast, depicts a specific phenomenon of water vaporization. Such interaction between language and image is modeled by Bezemer and Kress (2008) as a form of transduction involving losses in specificity. There is indeed a loss in specificity from the language text to the image, but this does not adequately explain the role of the image and how the interaction between them builds the knowledge of water vaporization. We can approach this issue from the perspective of semantic gravity. The image can be viewed as embodying stronger semantic gravity than the language text because the latter is much less context-dependent, i.e., the theoretical explanation construed is applicable for a much wider range of contexts. It applies to different empirical phenomena of water vaporization, of which what the image displays is only an instance. Thus, the image's role is to connect the theoretical explanation construed by the language to empirical data. In LCT terms, the interaction between language and image embodies a relatively wide semantic range, whereby weaker semantic gravity meanings become contextually located (Maton 2014:

129). The image thus links theoretical explanations construed by the language to an empirical phenomenon in everyday life.

Overall, from the image to the language text, the semantic density increases but the semantic gravity decreases, building the theoretical concept of vaporization and connecting to the everyday empirical phenomenon. Through the interaction between the language text and the image, Text 1 builds the hierarchical knowledge structure of chemistry.

The above analysis of the interaction between language text and image from the perspective of Semantics indicates one possible intersemiotic relation: language text embodies stronger semantic density but weaker semantic gravity than image. This type of intersemiotic relation prevails in early schooling levels. The interaction between language and image at higher year levels, by contrast, involves distinct relations in terms of Semantics. One is that language and image share similar levels of semantic density and semantic gravity, through which image reconstrues language and garners technical meaning. This is illustrated by Text 2, a chemistry text comprising language and a line graph from Year 10 Stage 5.

Text 2 (Zhang et al. 2015: 148)

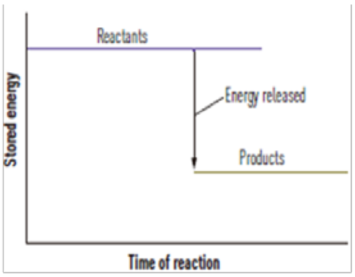


Figure 4.1 In an endothermic reaction, energy is released and the products have less stored energy than the reactants.

Reactions that release energy are called exothermic reactions. In an exothermic reaction, the product has less stored energy than the reactants at the start of the reaction (see Figure 4.1).

Text 2 introduces the chemistry concept of “exothermic reaction”. In terms of Unsworth’s (2007) model of language-image relations, the text involves a relation of “exposition”, through which the image re-expresses what the language text construes. In terms of field, the body language text and the caption in Text 2 realize an activity (exothermic reactions release energy) and an arrayed qualitative property (the levels of energy stored in reactants and products). The line graph construes similar field-specific meanings. The vertical and horizontal axes represent stored energy and the time of reaction. The arrow leading ‘reactants’ to ‘products’ indicates the activity that energy is released during the reaction. The different values of ‘reactants’ and ‘products’ at the vertical axis represent stored energy in reactants and products. In terms of field, they also construe an arrayed qualitative property, which however is distinct from the arrayed property realized by the language since the former is topological while the latter is typological. This distinction is caused by the epistemological commitment of different semiotic modes (Bezemer and Kress 2008). The line graph expresses topological meaning, which inevitably shows the indefinitely small gradations in the levels of energy, whereas the language makes typological meaning and does not afford the indefinite gradations. Regardless of the distinction in the types of arrayed properties construed, the similar types and number of field-specific meanings realized by the language text and the line graph suggest that they embody similar semantic density. In terms of semantic gravity, both the language text and line graph are relatively context-independent because both focus on building the theoretical concept of ‘exothermic reactions’ and apply to a relatively wide range of contexts, i.e., they apply to different specific instances of exothermic reactions. Thus, they share similar semantic gravity.

The above intersemiotic relation in terms of Semantics indicates that the line graph reconstrues the language text. Through the interaction with language, the line graph garners technical field meanings and fosters relatively strong semantic density. A particular feature of the graph is that it can keep consolidating meanings through interacting with language at higher schooling levels and allow all the meanings to be related in a single snapshot. For example, Text 3, a chemistry text from Year 11 Stage 6, brings in the chemistry concept of “activation energy”.

Text 3 (Chan et al. 2018: 412) (original emphasis in bold fonts)

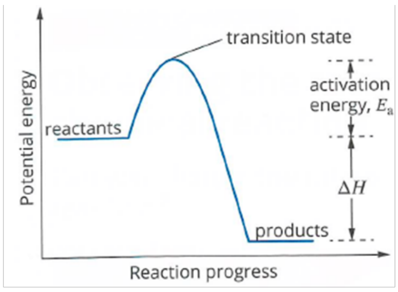


Figure 13.1.2. Energy profile diagram of an exothermic reaction

The minimum energy that a collision must possess for a reaction to occur is called the **activation energy**, *En*. When the energy of a collision is equal to or greater than the activation energy, a reaction can occur. The potential energy changes associated with a chemical reaction can be represented as an energy profile diagram. An energy profile diagram for an exothermic reaction is shown in Figure 13.1.2. On the energy profile diagram, the activation energy is measured from the energy of the reactants to the peak of the energy profile diagram.

In Text 3, the language text mainly explains the technical concept of “activation energy” and introduces its measurement. These technical meanings are encoded in the line graph, as shown by the linguistic labels of “transition state” and “activation energy”. In terms of field, this realizes a qualitative property that describes an activity – the collision between chemical substances causing chemical reactions. Through the interaction with the language text, the line graph garners this technical qualitative property. In addition to this newly encoded meaning, the line graph carries the knowledge built in Year 10 we have just analyzed, i.e., the field-specific meanings of an itemized activity (exothermic reaction) and arrayed qualitative properties (levels of stored energy in reactants and products), which are not specified by the language text in Text 3. To interpret this, we can analogize the interaction between the line graph and the language text to a snowball effect that gains weight as it rolls. The graph garners meaning from language through the year levels and allows them to be related in a single snapshot. The aggregation of technical meanings in the line graph across schooling levels displays a synoptic glimpse of the overall semantic density of the knowledge of exothermic reactions. The language texts in different year levels, however, are like different parts of a jigsaw, each construing one segment of the knowledge constellations. The line graph thus can be viewed as embodying stronger semantic density than the language in Text 3.

In terms of semantic gravity, the language text and the line graph share similar strengths of semantic gravity as both construe abstract theoretical concepts and can be applied to a wide range of contexts. This constitutes the third language-image relation in terms of Semantics found in our data − image embodying stronger semantic density than and similar semantic gravity as language text. This intersemiotic relation prevails in higher schooling levels, especially Years 11 and 12. The reason might be that knowledge at these year levels becomes rather complex, and images can present broad swathes of technical field meanings in a single snapshot while allowing them to be more easily interconnected. The interaction between language and images at higher school levels thus fosters relatively strong condensation, building increasingly integrative and generalized chemistry concepts.

The above three texts are the representative texts involving language and image in secondary school chemistry textbooks. In terms of Unsworth’s (2007) model of language-image relations, they involve primarily ‘exemplification’ and ‘exposition’. Our analysis of the interaction between language and image in terms of Semantics provides a new interpretation of language-image relations, revealing how the interaction builds chemistry knowledge. The exploration of intersemiotic relations between language text and image in terms of Semantics suggests three types of relations: (1) language text embodies stronger semantic density and weaker semantic gravity than images; (2) language text embodies similar semantic density and semantic gravity as images; (3) language text embodies waker semantic density than and similar semantic gravity as images. These relations tend to occur at different stages, revealing the characteristics of the knowledge-building at each stage. The interpretation of the intersemiotic relations between language and image in terms of Semantics adds another dimension to Unsworth's model of text-image relations, as shown in Figure 6.3.

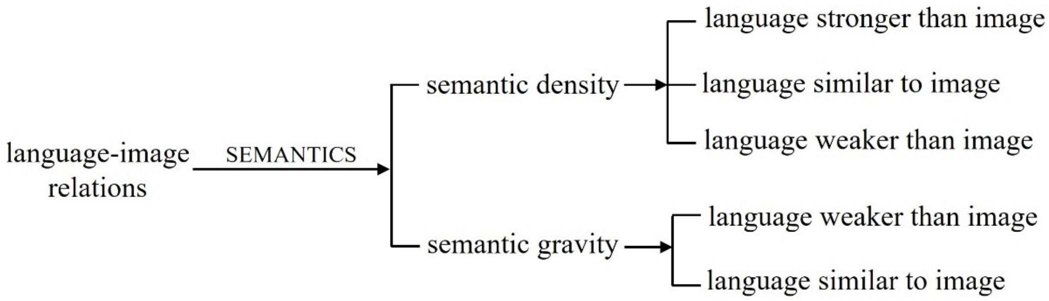


Figure 6.3 The Semantics dimension of language-image relations

The multiple intersemiotic relations between language and image in terms of Semantics enable chemistry texts to build increasingly integrative and generalized chemistry concepts and connect to everyday empirical phenomena. In terms of field, through the interaction with language, images can either gravitate chemistry concepts to link to everyday common-sense fields or condense meaning to build technical uncommon-sense fields. Yet, it is worth noting that this condensation of meaning through the interaction between language and image is only one way of aggregating meaning. Another is through the interaction between language and chemical formalisms, which will be explored in the next section.

6.2.2 Interaction between language and chemical formalisms

In secondary school chemistry textbooks, the interaction between language and chemical formalisms is crucial for building chemistry knowledge. Language can interact with all the subtypes of chemical formalisms, including chemical formulas, structural formulas, and chemical equations. One significant function of these interactions is that they allow chemical formalisms to condense meaning from language and build integrative and generalized chemistry concepts about chemical matter and reactions. This section will show this by tracing the interaction between language and chemical formalisms across schooling levels. It will first examine the interaction between language and chemical formulas and structural formulas and then between language and chemical equations. It will show that the formulas and equations keep consolidating meaning from language and amplifying chemistry's constellation of meaning.

6.2.2.1 Interaction between language and chemical formulas and structural formulas

Chemical formulas appear throughout secondary school year levels, while structural formulas are used only in higher year levels, i.e., Years 11 and 12. A feature of the interaction between language and the formulas is that the latter keeps condensing meanings from the former and aggregates them in a single formula. Interpreted from the perspective of field, the interaction enables formulas to garner the compositional taxonomies from language at the lower year levels and spatial properties at the higher year levels. This will be shown in the following through two texts about the chemical matter of oxygen and ozone gases.

Text 4, a text from Year 8 Stage 4 involving language and chemical formulas, introduces oxygen and ozone gases.

Text 4 (Zhang et al. 2014a: 179)

> Oxygen has the chemical symbol O. Oxygen can form diatomic or triatomic (three atoms) gases. The diatomic form, which has the chemical formula O_2, while ozone has the chemical formula O_3. The small '2' and '3' after the O symbol tell you how many atoms are involved in each molecule.

Since chemical formulas are a symbolic mode, they are always integrated with language as part of the text. In terms of field, the language construes the compositional taxonomies of oxygen and ozone gases. This field-specific meaning is reconstrued by the two chemical formulas O_2 and O_3. In some sense, the meaning construed by language is translated into the formulas, indicating that they share similar semantic density and semantic gravity. These compositional taxonomies constitute the basis on which structural formulas consolidate further meaning from language at higher year levels. Text 5, for example, is extracted from the Year 11 Stage 6 textbook. It introduces oxygen and ozone gases' molecular structures through the interaction between language and the two formulas.

Text 5 (Chan et al. 2018: 169)

> Oxygen gas consists of diatomic molecules with the formula O_2. Each oxygen atom in this arrangement is bound to one other oxygen atom. Ozone is another molecule containing only oxygen. Ozone molecules have the formula O_3 and consist of a central oxygen atom bound to two other oxygen atoms. Figure 5.3.16 shows the structure of these two molecules.

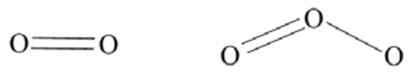


Figure 5.3.16 Oxygen and ozone are two molecules that contain only oxygen atoms

As shown in Text 5, the chemical formulas O_2 and O_3 are used directly in this stage, and the language does not describe their compositions anymore but introduces the general arrangement of atoms in the oxygen and ozone gas molecules, as described by the clauses "each oxygen atom in this arrangement is bound to one other oxygen atom" and "a central atom bound to two other oxygen atoms". In terms of field, they realize a spatial property, which is reconstrued by the structural formulas. However, the spatial property realized by the structural formulas is somewhat different from the language. The structural formulas present specific relative locations of oxygen atoms, whereas the language provides only a rather general description of the oxygen atoms' arrangement. For example, in the language text, "ozone molecules have the formula O_3 and consist of a central oxygen atom bound to two other oxygen atoms" shows that one oxygen atom is in the center and does not depict the specific relative positions of the three oxygen atoms as presented in the structural formula of ozone molecules shown in Figure 5.3.16. Such difference in the construal of spatial property by language and structural formulas mirrors the difference in the two semiotic modes' epistemological commitment (Kress 2003). In contrast to language as a typological oriented mode that makes categorical distinctions, structural formulas share some imagic nature of images and tend to construe topological meanings of "complex topological relations of relative nearness or connectedness" (Lemke 1998: 87). The topological meanings in structural formulas arise from the spatial arrangements of chemical symbols on a page, which allows their relative locations to be used as a resource for making meaning. The language, in contrast, does not afford this meaning potential.

In addition to the difference in the construal of spatial property, the structural formulas include meanings not construed in the language text. The single and double lines in the structural formulas represent different covalent bonds: single and double bonds, which in terms of field realize a classificational taxonomy. The knowledge of covalent bond classifications is not construed in the language in Text 5, but it has been built through language in earlier chapters in the Year 11 textbook. The structural formula thus aggregates meanings across schooling levels and condenses all of them in a single formula, imbuing it with stronger semantic density than the language text it interacts with.

The above analysis indicates that through the interaction, the formulas condense meaning from language and enable the meaning to be further condensed based on previous formulas across schooling levels efficiently. Working together, the language and the formulas develop increasingly complex meanings and extend the boundary of the knowledge of chemical matter.

6.2.2.2 Interaction between language and chemical equations

The interaction between language and chemical equations follows a similar pattern as the interaction between language and formulas in chemistry texts. The equations condense meaning from language, and its semantic density grows stronger as schooling levels progress. Working together, they expand the field constellation of chemical reactions. The following will illustrate this knowledge-building process through a series of texts from lower to higher year levels in secondary school chemistry.

Year 8 Stage 4 introduces the basic meaning of chemical equations through texts that involve interaction between language and the simplest chemical equations. Text 6 is a typical example.

Text 6 (Zhange et al. 2014: 211)

> You may see chemical reactions written in the basic form of the equation below.
>
> $$A + B \rightarrow C + D$$
>
> This means that the reactants A and B chemically react to form the products C and D. This is known as the general form of a chemical reaction. The arrow indicates that a reaction has taken place.

As shown in Text 6, the general chemical equation (the equation in which the chemical species are represented by the capital letters) is reconstrued by the language, which realizes an activity in field's term. The equation garners this field-specific meaning through interacting with the language. It forms the basic meaning of chemical equations, upon which other field-specific meanings of chemical reactions are developed in higher year levels. Year 10 Stage 5, for example, introduces the concept of "coefficient" to chemical equations, as shown in Text 7.

Text 7 (Zhang et al. 2015: 115)

> To balance the equation for the reaction between hydrogen and oxygen molecules, the hydrogen molecule in the reactants and the water molecule in the products each have a coefficient of 2. The oxygen molecule has a coefficient of 1, but a coefficient of 1 doesn't need to be shown.
>
> $$2H_2 + O_2 \rightarrow 2H_2O$$

In Text 7, the language introduces the coefficient each chemical species has, which construes a gauged property in terms of field. It describes the relative number of chemical species involved in the reaction. Once again, the chemical equation garners this field-specific meaning from the language text. The field meanings of chemical reactions can be expanded further still at higher year levels. Year 11 Stage 6, for instance, pushes the construal power of chemical equations a step further by introducing the concept of "physical states", as shown in Text 8.

Text 8 (Chan et al. 2018: 207)

> A fully balanced chemical equation includes the physical states or phases of the reactants. The four physical states of matter used in equations are:
>
> - liquid (l)
> - gas (g)
> - solid (s)
> - aqueous (aq), which refers to the solution formed when a substance dissolves in water.
>
> The state symbols of (l), (g), (s) or (aq) are written in brackets immediately after the element or compound in a chemical equation. For example, take the equation Lavoisier conducted:
>
> $$2Hg(l) + O_2(g) \rightarrow 2HgO(s)$$
>
> You can read this reaction as 'two parts of mercury liquid reacts with one part of oxygen gas to form two parts of mercury (II) oxide'.

The language in Text 8 introduces four physical states of chemical species involved in chemical reactions: liquid, gas, solid, and aqueous, represented by "(l)", "(g)", "(s)", and "(aq)". In terms of field, they construe a qualitative property. This field-specific meaning is condensed in the chemical equation through its interaction with the language text. Up to this stage, we can see that the chemical equation aggregates three field-specific meanings from language across the schooling levels: an activity, a gauged property, and a qualitative property. Each aggregation is based on previous equations, enabling chemistry to condense meaning efficiently and build the increasingly technical field of chemical reactions.

The above analysis of the interaction between language and chemical formalisms shows that through the interaction, the formalisms consolidate meaning from language and enable the meaning to be further condensed based on their previous forms at earlier schooling levels. The formalisms succinctly formalize knowledge that takes much longer to construe in language. The interaction thus allows chemistry texts to efficiently build increasingly integrative and generalized knowledge of chemical matter and reactions across schooling levels. This condensation of knowledge echoes the interaction between language and image analyzed in Section 6.2.1. However, the condensation arising from the interaction between language and chemical formalisms focuses on a relatively smaller range of field-specific meanings as noted in Section 6.2.2. It mainly concerns the compositions and spatial properties of chemical matter and the activities of chemical reactions. The language-image interaction, on the other hand, covers a wider range of fields, but it is not as specialized as the interaction between language and chemical formalisms in building the knowledge of chemical matter's compositions and structures and chemical reactions' quantitative relations and physical states. Therefore, the two interactions can complement each other to build the comprehensive field of chemistry. Another difference between the two interactions is that in the interaction between language and chemical formalisms, the formalisms never connect theoretical concepts to empirical phenomena. This is because chemical formalisms are designed purely for developing the technicality of chemistry. This can be complemented by the language-image interaction discussed in Section 6.2.1, which can connect chemical theories to empirical phenomena.

The above analysis shows that working together, the interaction among language, image, and chemical formalisms can foster strong condensation to build increasingly complex knowledge of chemistry and gravitate the technical meanings to connect to empirical phenomena. The division of labor across the semiotic resources and their complementarity build the hierarchical knowledge structure of chemistry by establishing increasingly integrative and generalized chemical theories and linking them to empirical data. Through these resources and their interaction, chemistry broadens its boundaries of knowledge and maintains connections to the material world.

6.3 Summary

This chapter has examined the third research question: how do language, image, and chemical formalisms work together to build chemistry knowledge in secondary school chemistry textbooks? To address this issue, this chapter has first established a multisemiotic picture of chemistry knowledge in terms of field-specific meanings realized by the resources and then studied how they interact to build chemistry knowledge in chemistry texts across schooling levels. The findings show that each semiotic resource has its functionality for organizing chemistry knowledge, and they complement each other to form the complete picture of chemistry. As far as the knowledge of chemical matter is concerned, language can construe its composition and classification taxonomies, but it cannot realize topological spatial properties as structural formulas and images do. Therefore, when representing chemical matter's structures, chemistry turns to structural formulas or images. In the construal of chemical matter's qualities, language and chemical formalisms fall short of representing the topological arrayed properties afforded by images. Thus, when construing qualities involving indefinite small gradations, images will take over. As for the knowledge of chemical reactions, chemical equations cannot express implication activity sequences as language and images do (though the implication sequences construed by images are less elaborated than those realized by language). Therefore, when chemistry requires more elaborated sequences, it will turn to language. Chemical equations, however, have unique advantages in expressing quantitative relations over language and images. They can show a balanced relation between the quantities of atoms before and after the reactions. In terms of field, chemical equations construe a special gauged property not afforded by language and images. Thus, when chemistry needs to show this property, it will turn to chemical equations. The above suggests that when one resource falls short, another will take over. The interplay between them forms a complementarity that expounds and expands the knowledge of chemistry.

To investigate how the semiotic resources interact in chemistry texts across schooling levels to build chemistry knowledge, we have examined the interaction between the resources in terms of LCT's Semantics. The finding shows that in terms of Semantics, the intersemiotic relation between language and image involves three types: (1) language embodies stronger semantic density and weaker semantic gravity than images; (2) language embodies similar semantic density and semantic gravity as images; (3) language embodies weaker semantic density than and similar semantic gravity as images. The first type tends to prevail at earlier schooling levels, while the second and third types typically occur at higher schooling levels. In terms of field, through the interaction with language, images can either gravitate chemistry concepts to link to everyday common-sense fields or condense meaning to build increasingly technical uncommon-sense fields. The intersemiotic relation between language and chemical formalisms, in contrast, exhibits relatively limited variations in terms of Semantics - the formalisms typically show similar semantic density as or stronger semantic density than language. The interaction allows chemical formalisms to consolidate meaning from language and enables the meaning to be further condensed based on their previous forms at earlier schooling levels. With such intersemiotic relation, chemistry texts efficiently develop increasingly integrative and generalized chemistry concepts.

However, the interaction does not connect to everyday empirical phenomena as the interaction between language and image does. Therefore, the two interactions complement each other to foster strong condensation to build increasingly technical fields of chemistry and gravitate the technical meanings to link to everyday empirical phenomena. The division of labor across the semiotic resources and their complementarity build the hierarchical knowledge structure of chemistry.

Chapter 7 Conclusion

This study aims to investigate how the multisemiotic resources including chemical formalisms, images, and language build chemistry knowledge in secondary school chemistry textbooks. To achieve this research objective, this study has explored the following research questions: (1) how do chemical formalisms build the knowledge of secondary school chemistry? (2) how do images build the knowledge of secondary school chemistry? (3) how do chemical formalisms, images, and language work together to build chemistry knowledge in the textbooks? Chapters 4 to 6 have addressed these questions respectively. This section will first summarize major findings to these questions (Section 7.1) and then discuss the contributions of this study (Section 7.2). Following this, it will ponder the research findings' implications (Section 7.3). Finally, this thesis will end with a discussion of the limitations of this study and directions for future research (Section 7.4).

7.1 Summary of findings

This section will review the findings to the research questions. The key findings relevant to each question are summarized as follows.

7.1.1 Knowledge-building through chemical formalisms

The first research question of this study is how chemical formalisms in secondary school chemistry textbooks build the knowledge of chemistry. This question includes two sub-questions: (1) how are chemical formalisms organized to make meaning and (2) how do the formalisms develop across schooling levels to build the hierarchical knowledge structure of chemistry in the textbooks. To address the first sub-question, this thesis has investigated the grammars of chemical formalisms in Section 4.1. Chemical formalisms include three major types: chemical equations, chemical formulas, and structural formulas, each of which is a distinct semiotic system. This thesis has described the grammar of these formalisms and generated their mefatunctionality based on the theoretical primitive of axis – system and structure. The description shows that the grammar of chemical equations includes four ranks: equation, term, formula, and symbol. Since chemical formulas are themselves a part of chemical equations, the formula and symbol ranks also constitute the grammar of chemical formulas. The grammatical systems at the four ranks organize primarily the experiential and logical metafunctions. The experiential metafunction pervades every rank of the grammar, and the logical metafunction occurs at all the ranks below the equation rank. The grammar also realizes a textual component but only at the term rank. The description of structural formulas shows that the grammar includes two ranks: formula and symbol. Similar to chemical formulas and chemical equations, the grammatical systems of structural formulas also realize primarily the experiential and logical metafunctions. In terms of grammatical analysis, this study finds two special structures. One is the **subjacency structure**, a relation of non-iterative nuclear dependency. This structure tends to occur in both chemical formulas and structural formulas. The other is the **ideationalized prosodic structures**, typically occurring in structural formulas and chemical equations to construe ideational meanings of bonding relations between atoms, molecular polarities, and quantitative relations of chemical species involved in chemical reactions.

To capture the similarities and differences between the three chemical formalisms in making meaning, Section 4.1.3 has developed a functional semiotic typology for the formalisms in terms of their grammatical organization and field-specific meanings realized. The finding shows that chemical formulas and structural formulas are similar in that both of them are dominated by univariate structures, each capable of iterating chemical symbols multiple times. However, in contrast to chemical formulas, structural formulas can iterate in a two-dimensional space, which affords a reading of spatial layout. This leads to a distinction in the field-specific meanings they realize − chemical formulas realize only a compositional taxonomy, whereas the two-dimensional organization of the structural formulas realizes both a compositional taxonomy and their spatial properties. The two formulas are also similar in that they both involve subjacency structures. However, they employ these structures to realize different field-specific meanings − a gauged property of the number of atoms for chemical formulas and a qualitative property of partial charge polarities for structural formulas. In contrast to chemical formulas and structural formulas dominated by univariate structures, chemical equations are organized largely around multivariate structures. In terms of field, these multivariate structures construe activities of different types of reactions at the equation rank and various properties and quantitative relations of reactions at the term rank. In addition to the discussion of similarities and differences in meaning-making in terms of univariate and multivariate structures, chemical formalisms can also be compared in terms of whether they involve **ideationalized prosodic structures**. The finding shows that both structural formulas and chemical equations employ prosodic structures to construe qualitative properties, in contrast to chemical formulas that do not involve such prosodic structures. The functional semiotic typology analysis of the chemical formalisms indicates that they complement each other to construe the chemistry knowledge of chemical matter and chemical reactions.

To address the second sub-question – how do chemical formalisms develop across schooling levels to build the hierarchical knowledge structure of chemistry in secondary school chemistry textbooks, Section 4.2 has traced the development of the formalisms' meaning accounted for in the grammatical descriptions in Section 4.1 across the three curriculum stages in the NSW secondary school chemistry textbooks. In specific, it has analyzed how formulas and chemical equations develop across the stages to build the two crucial components of chemistry knowledge – chemical matter and chemical reactions – in Sections 4.2.1 and 4.2.2.

Section 4.2.1 has first established a model for the semantic density of formulas used in secondary school chemistry in terms of the field-specific meanings they realize and then used the model to trace the formulas' development in terms of their semantic density. The finding shows that the semantic density of the various formulas used in the textbooks can be sequenced as five scales, with molecular formulas embodying the weakest strength and stereo formulas exhibiting the strongest strength. This continuum of scales constitutes a tool for tracing the development of the formulas' semantic density across curriculum stages in the NSW secondary school chemistry textbooks. Enacting the tool reveals that formulas with different degrees of semantic density are used in different curriculum stages. They exhibit a widening range of semantic density across the three curriculum stages, with the bottom of the range being the most basic molecular formulas construing the compositions of chemical matter and the top reaching increasingly strong semantic density that construes more and more complex fields of chemical matter. This indicates that to build the knowledge of chemical matter, the formulas used in each stage maintain connections to the simpler technical field meanings and establish increasingly technical fields of chemical matter. From the simpler to the more complex field, the meanings presented by formulas are organized into the hierarchical knowledge structure of chemistry.

Section 4.2.2 has examined how chemical equations develop across schooling levels to build the knowledge of chemical reactions. Like Section 4.2.1, it has first established a model for the semantic density of the various chemical equations used in secondary school chemistry in terms of field and enacted to model to analyze how the equations develop across schooling levels to build the knowledge of chemical reactions. The finding shows that the semantic density of the chemical equations can be sequenced as six scales along a continuum, with incomplete molecular equations embodying the weakest strength and reversible ionic equations exhibiting the strongest strength. This continuum of scales constitutes a tool for tracing the development of the equations' semantic density across curriculum stages in the NSW secondary school chemistry textbooks. Enacting the tool to analyze the development of the equations reveals that there is a consecutive increase in chemical equations' semantic density from Stage 4 to 5, expanding the field of chemical reactions from compositional taxonomies (compositions of chemical species) and an activity (a forward reaction) to compositional taxonomies, an activity, and two extra properties concerning chemical reactions (quantities of chemical species involved in the reactions and their physical states). Building upon the complete molecular equations used in Stage 5, Years 11 and 12 in Stage 6 exhibit a widening range of semantic density, with Year 11 introducing ionic equations and Year 12 bringing in reversible equations, construing increasingly complex fields of chemical reactions. The growth of chemical equations' semantic density across the schooling levels builds the hierarchical knowledge structure of chemistry.

7.1.2 Knowledge-building through images

The second research question of this study is how images build the knowledge of secondary school chemistry. It includes two sub-questions: (1) what meaning images in the chemistry textbooks construe and how they realize the meaning; (2) how the images develop across schooling levels to build the hierarchical knowledge structure of chemistry in the textbooks. To address the first sub-question, this study has explored the meaning construed by images in chemistry by establishing a model of these images in terms of field that links with the grammatical resources realizing the meaning. The finding shows that images in secondary school chemistry textbooks construe three broad types of field-specific meanings: item, activity, and property. Item is the static perspective of field. The images construe either singular item or taxonomic relations between multiple items. Particularly, they show a strong capacity for construing multiple levels of compositional and classificational taxonomies. In terms of composition, chemical images can establish multiple levels of part-whole relations between components that constitute a chemical matter. These multiple levels of composition are realized by a spatial analytical structure recursively embedded within another, allowing for construing as complex compositions of chemical matter as possible. In terms of classification, chemical images can set up a cline of type-subtype relation between different types of chemical matter. These classification relations can be realized by either multi-leveled overt taxonomy structures or **multi-leveled covert taxonomy structures**. The composition and classification taxonomies can occur simultaneously in a single chemical image.

Activity is the dynamic perspective of field. The images in the secondary school chemistry textbooks construe either unmomented or momented activities, with the former conveying the knowledge of a single chemical change while the latter typically showing how initial reactants reach final products. Grammatically, the unmomented activities tend to be realized by a single **transformation process**. Momented activities, on the other hand, involve further complications. They can be further described in terms of their cyclicality and the relation connecting them. In terms of cyclicality, they can be either cyclical or non-cyclical. Grammatically, the cyclical momented activities are realized by conversion processes, and the non-cyclical ones by complexes of transformation processes. In terms of the relation linking the activities, they can be related by either implication or expectancy. The expectancy sequences tend to be realized by complexes of transformation processes or conversion processes, and the implication sequences are typically realized by complexes of transaction processes and transformation processes.

Property describes the features of items and activities. The finding shows that the images in secondary school chemistry textbooks realize all the properties described in Doran and Martin’s (2021) model and include a further distinction within the non-gauged arrayed property between **categorical** and **non-categorical**. These properties construe various aspects of chemical matter, including electronic polarities of electrons and protons, the spatial arrangement of these particles within atoms, pH values of substances, etc. When describing chemical reactions, these properties tend to show the amount of energy absorbed or released during reactions and the amount of time the reactions take. A significant distinction between the categorical and non-categorical properties is that the former tends to be **typological** while the latter is typically

topological. Grammatically, these properties are not realized by any particular structures through which items and activities are realized, but by non-structural resources. The primary resource is symbolic labelings, including numbers, symbols (e.g. '+' and '-'), mathematical equations, etc. Another significant source for realizing the properties, especially non-gauged arrayed properties, is color gradations. In addition, the spatial arrangement of entities in images is also a key grammatical resource realizing the **topological spatial property**.

To investigate the second sub-question - how the images develop to build the hierarchical knowledge structure of chemistry, this study has studied the development of images in secondary school chemistry textbooks in terms of their semantic density. To capture the development, it has established a model for the semantic density of images in secondary school chemistry textbooks. The model builds upon two factors: internal organization and external relations. Internal organizations concern with the number and types of field-specific meanings construed by an image, including composition, classification, activity, and property. External relations describe the technicality of these field-specific meanings. The model privileges external relations over internal organizations because the former determines whether a particular field is situated in a complex constellation of meanings.

With this model, this study has analyzed the knowledge-building through images in secondary school chemistry textbooks by tracing their development in terms of semantic density. Specifically, it has examined the development of images used for building knowledge about the same subject and the development of the range of semantic density images exhibit across the curriculum stages. The former concerns the roles of images in building specific related knowledge, while the latter focuses more on the overall development of images' semantic density. The finding shows that when building knowledge about the same subject, the images' semantic density exhibits a consecutive progression: from relatively weak through relatively strong to even stronger. This suggests that to build certain knowledge of chemistry, the textbook needs to use everyday images to present empirical phenomena, thereby giving students a sense of what the objects they are learning look like in daily life, and **power images** with increasingly strong semantic density to build more and more complex technical understandings of the phenomena. The development of the images' semantic density attests that chemistry embodies a typical hierarchical knowledge structure where general theories are created from lower levels which in turn generalize knowledge from empirical data. The images build this knowledge structure by developing increasingly complex theoretical explanations while maintaining connections to the empirical world. Interpreted from the perspective of field, the knowledge-building through images involves shifting from common-sense fields to uncommon-sense fields and building increasingly strong technicality.

The analysis of the development of the range of semantic density of images shows that throughout the curriculum stages in secondary school chemistry, the knowledge expressed by the textbook images maintains connections with the everyday empirical world while also reaching toward increasingly complex chemical theories. The textbook images thereby model a **growing** range of semantic density. In terms of field, the images in each stage construe common-sense fields and build increasingly technical fields across the stages. This once again shows how the images build the hierarchical knowledge structure of chemistry – in each year level, they develop increasingly complex fields involving more and more intricate and integrated chemical concepts while maintaining connections to the common-sense fields that present everyday phenomena.

The above findings indicate that to learn chemistry knowledge through images, students are expected to engage with images embodying different degrees of semantic density at each stage. This is often made explicit by a special type of image in chemistry – composite images. The analysis shows that composite images bring together images with different strengths of semantic density and allow students to move between the different levels of complexity. In terms of field, composite images integrate fields with varying degrees of technicality. Our data indicate that they either bring together everyday common-sense field and technical uncommon-sense field or integrate uncommon-sense fields with different levels of technicality. Grammatically, the different fields construed by the component images within the composite are linked through the logic-semantic relation of elaboration in SFL, through which the technical field provides theoretical explanations at the microscopic level for the empirical phenomena.

7.1.3 Knowledge-building through multisemiosis

To address the last research question - how do chemical formalisms, images, and language work together to build chemistry knowledge in secondary school chemistry textbooks, this thesis has first briefly described language in chemistry in terms of field and then established a multisemiotic picture of chemistry knowledge by bringing together language, chemical formalisms, and images in terms of their field affordance in chemistry. Following this, it has studied how they interact in chemistry texts to build chemistry knowledge.

The finding shows that language, chemical formalisms, and images have their own functionalities of organizing chemistry knowledge and complement each other to constitute the full picture of chemistry. As far as the knowledge of chemical matter is concerned, language can construe its compositional and classificational taxonomies, but it cannot realize topological spatial properties as structural formulas and images do. Therefore, when representing chemical matter's structures graphically, chemistry can turn to structural formulas and images. In the construal of chemical matter's qualities, language and chemical formalisms fall short of representing the topological arrayed properties afforded by images. Thus, when construing qualities involving indefinite small gradations, images will take over. As for the knowledge of chemical reactions, chemical equations cannot express implication activity sequences as language and images do (though the implication sequences construed by images are less elaborated than those realized by language). Therefore, when chemistry requires more elaborated sequences, it will turn to language. Chemical equations, however, have unique advantages in expressing quantitative relations over language and images. They can show a balanced relation between the quantities of atoms before and after the reactions. In terms of field, chemical equations construe a special gauged property that is not afforded by language and images. Thus, when chemistry needs to show this property, it can turn to chemical equations. The above suggests that when one resource falls short, another will take over. The interplay between them forms a complementarity that expounds and expands the knowledge of chemistry.

To investigate how the semiotic resources interact in chemistry texts across schooling levels to build chemistry knowledge, this study has examined the interaction between the resources in terms of LCT's Semantics. The finding shows that in terms of Semantics, the intersemiotic relation between language and image involves three types: (1) language embodies stronger semantic density and weaker semantic gravity than images; (2) language embodies similar semantic density and semantic gravity as images; (3) language embodies weaker semantic density than and similar semantic gravity as images. The first type tends to prevail at earlier schooling levels while the second and third types typically occur at higher schooling levels. In terms of field, through the interaction with language, images can either gravitate chemical concepts to link to everyday common-sense fields or condense meaning to build increasingly technical uncommon-sense fields. The intersemiotic relation between language and chemical formalisms, in contrast, exhibits relatively limited variations in terms of Semantics: the formalisms typically show similar semantic density as or stronger semantic density than language. The interaction allows chemical formalisms to consolidate meaning from language and enables the meaning to be further condensed based on their previous forms at earlier schooling levels. With these intersemiotic relations, chemistry texts efficiently develop increasingly integrative and generalized chemistry concepts. However, the interaction does not connect to everyday empirical phenomena as the interaction between language and image does. Therefore, the two interactions complement each other to foster strong condensation to build increasingly technical fields of chemistry and gravitate the technical meanings to link to everyday empirical phenomena. The division of labor across the semiotic resources and their complementarity build the hierarchical knowledge structure of chemistry.

7.2 Contributions of this study

The present study has explored the knowledge-building of secondary school chemistry from a multisemiotic perspective. It has made several contributions to the field of multimodal discourse analysis and knowledge-building studies.

In terms of multimodal discourse analysis, this study has made contributions regarding chemical formalisms, images, and intersemiotic relations. The following will discuss the contributions in turn.

In terms of chemical formalisms, this study has made the following contributions. Firstly, it has expanded the disciplinary map of multimodal discourse analysis to chemistry and modeled the grammars of chemical formalisms, including chemical formulas, chemical equations, and structural formulas. In contrast to scholars in systemic-functional multimodal discourse analysis (SF-MDA), this study does not assume that the three metafunctions developed from the description of English (Halliday 1985) will unproblematically transfer to chemical formalisms and argues that assuming these categories will homogenize description and water down the specific functionalities of chemical formalisms. Using the more fundamental theoretical primitive of **axis** from which systemic functional macro-theoretical categories such as metafunction, rank, and stratum are derived (Martin 2013b, 2015a), this study has described chemical formalisms in terms of their own grammatical organizations and discovered their own functional components and grammatical hierarchies. This work can be viewed as an effort to step out of what Mitchell (1986) has referred to as '**linguistic imperialism**'. That is, we should be cautious of simply blindly imposing the model of language onto new semiotic systems, which could conceal their very nature.

Another contribution of the description is its emphasis on **system**. This study has established exhaustive grammatical systems for chemical equations, chemical formulas, and structural formulas, covering all data in secondary school chemistry. With such a comprehensive data set, the description is able to generate the metafunctions of the different chemical formalisms. The emphasis on system in semiotic description is another major difference from SF-MDA. Starting from O'Tool's (1994) modeling of images, the SF-MDA tradition seems to have been focusing more on structure and "lacks explicit system/structure cycles and is based more on Halliday's earlier scale and category modeling" (Martin 2015b: 45). It should be noted that with no emphasis on system, the grammar developed is more of a functional one than a systemic one (Zhang 2018). The description developed in this study thus sets up an example for future studies towards a more systemic description of semiotic systems other than language.

In addition, the establishment of the functional typology of chemical formalisms in this study contributes to the extension of functional language typology (Caffarel et al. 2004) to **functional semiotic typology** (Doran 2019). This extension is enabled by taking the more fundamental theoretical primitive of axis as the descriptive foundation, allowing the chemical formalisms to be described on their own terms. Once the description is done, we are able to compare the different chemical formalisms' intrinsic functionalities and their internal grammatical organization, upon which the functional semiotic typology is developed. This task of developing functional semiotic typology seems to be beyond the reach of the SF-MDA tradition since it assumes that non-linguistic semiotic resources share the same metafunctional organization as English, making it difficult to compare and contrast the similarities and differences between semiotic resources in a principled nature. In addition to the extension of functional semiotic typology to chemical formalisms, this study adds another dimension of comparing and contrasting academic formalisms discussed by Doran (2019), i.e., whether the formalisms involve **ideationalized prosodic structure**. Our analysis in Section 4.1.3 shows that chemical equations and structural formulas are similar in that they both involve prosodic structures ideationalized for construing chemistry knowledge.

In terms of the grammatical analysis of chemical formalisms, this study has made two contributions. One is that it discovers a unique structure called **ideationalized prosodic structure**. Prosodic structures are typically associated with interpersonal meaning in language, but in chemical formalisms they tend to be ideationalized. The ideationalized prosodic structures occur in both structural formulas and chemical equations, realizing the field-specific meaning of properties.

In terms of images, this study has also made several contributions. Firstly, it has established a model of chemical images from the perspective of field that links with the images' grammatical organization. Kress and van Leeuwen's (2006) model of images has offered a detailed description of images' grammar. However, there have been scant studies modeling what meaning images construe at a **more abstract level**. This study contributes to this by building a system network that captures the meaning realized by chemical images in terms of field and the grammatical resources realizing the meaning.

Secondly, the analysis of the grammatical resources realizing the field meanings discovers some structures not accounted for in Kress and van Leeuwen's description of images. For example, this study has proposed a non-agentive narrative process called a **transformation process** constituted by a **Source**, a Vector, and a **Product**. Kress and van Leeuwen's description of classificational structures distinguishes between single-leveled and multi-leveled within only overt taxonomy structures. This study has found that covert taxonomy structures also involve a distinction between single-leveled and **multi-leveled**. Based on this, this study has established a renovated model for classificational structures of images. The grammatical structures accounted for in this study have enriched Kress and van Leeuwen's description of images.

Thirdly, this study has established a model for the semantic density of images. Kress and van Leeuwen (2006: 108-109) noted that different processes can be embedded in one image to form a powerful and multidimensional structure. Doran (2018c) takes a step further and argues that the more structures an image includes, the stronger the semantic density of the image. However, these studies seem to have ignored the complexity of the concepts represented by an image when assessing its semantic density. The model for the semantic density of images proposed in this study has filled this gap by considering both the **complexity of concepts** represented in an image and the number and type of field-specific meanings realized by the multiple structures. The analysis of the knowledge-building of chemistry through images has proven the model's usefulness and practicality.

In terms of intersemiotic relations, this study has modeled language-image relations in terms of LCT's **Semantics**. Most previous studies on language-image relations focus on logical-semantic relations (e.g. Martinec and Salway 2005) or discourse semantic relations (e.g. Unsworth 2007) between language and image. However, these studies are relatively restricted in informing the knowledge-building of chemistry through the interaction between language and image. The model of language-image relations in terms of semantic density and semantic gravity proposed in this study has filled this gap by making explicit how the interaction builds increasingly technical chemical concepts and relates them to empirical phenomenon.

In terms of knowledge-building analysis, this study has made the following contributions. Firstly, it has provided a semiotic perspective to the knowledge-building of chemistry. In chemistry education, studies on chemistry learning have long been dominated by constructivisms, i.e., knowledge-building of chemistry is viewed as the development of metal structures within learners' minds (Bodner 1986; Herron 1975; Shiland 1999; Taber 2000; Wink 2014). This approach to knowledge-building of chemistry explores 'knowing' (state of consciousness in learners' mental models) rather than the knowledge itself, which falls in what Maton (2014) criticized as '**knowledge-blindness**' – knowledge as an object of study is obscured. By focusing on the development of mental structures in learners, this research tradition ignores the forms taken by chemistry knowledge and how the knowledge is structured as a discipline. It thus fails to show what forms of chemistry knowledge should be taught to students and how the knowledge progresses as the learning proceeds. The present study has attempted to fill this gap by adopting a semiotic perspective to the knowledge-building of chemistry. Specifically, it has focused on how the multisemiotic resources, including chemical formalisms, images, and language, construe knowledge and how they develop across schooling levels to build knowledge in secondary school chemistry textbooks. This semiotic perspective has revealed the content knowledge (field meanings) built at each curriculum stage and how the knowledge progresses across the stages, which has made explicit how the multisemiotic chemistry discourse builds the hierarchical knowledge structure of chemistry.

Secondly, this study contributes to the knowledge-building analysis of chemistry by using both SFL and LCT. The educational linguistics approach to knowledge-building informed by SFL (Christie and Martin 1997, 2007; Rose and Martin 2012) focuses on the discourse that organizes disciplinary knowledge. However, it pays much less attention to the disciplinary knowledge's structuring principles. LCT, in contrast, focuses on the organizing principles underpinning the development of knowledge, but it does not probe how the discourse construes the knowledge. The two approaches to knowledge-building analysis thus complement each other. By incorporating LCT with SFL, this study has managed to unveil what knowledge the multisemiotic chemistry discourse construes and how the knowledge develops across schooling levels. This study thus contributes to the current ongoing and fruitful dialogue and collaboration between SFL and LCT on knowledge-building (Christie and Maton 2011; Martin et al. 2019; Maton et al. 2021).

7.3 Implications of this study

This study has explored the meaning-making of the multisemiotic resources in secondary school chemistry textbooks and the knowledge-building of chemistry through these resources, which holds both theoretical and pedagogical implications. The most prominent theoretical implication of this study is that it provokes rethinking the status of categories such as metafunctions developed from the description of English in semiotic descriptions by SF-MDA. The present study has shown that not assuming the three metafunctions but taking the more fundamental theoretical primitive of axis enables this study to reveal chemical formalisms' intrinsic functionalities and grammatical organizations. The description in this study is heuristic for future work, especially for any studies that intend to describe and compare semiotic systems in a more principled nature and, even more ambitiously, to develop a functional semiotic typology of these semiotic systems.

This study also holds several significant pedagogical implications for chemistry education. Firstly, as noted in Chapter 4, secondary school students face difficulties in understanding chemical formalisms. To support them better understand the formalisms, explicit literacy pedagogy of the formalisms is needed. However, it has long been recognized that to develop a literacy pedagogy for a particular type of discourse, it is necessary to first understand how that discourse works (Rose and Martin 2012). The description of the grammars of the chemical formalisms in secondary school chemistry in Chapter 4 offers a comprehensive account of how the formalisms make meaning, which is potentially rather helpful for developing the literacy pedagogy. Similarly, the model of chemical images from the perspective of field that links with the images' grammatical organization in Chapter 5 accounts for what field meanings the images construe and how these meanings are realized, which could facilitate the development of literacy pedagogy for teaching chemistry with chemical images.

Secondly, the analysis of knowledge-building through the different semiotic resources is potentially helpful for teachers to use appropriate chemical formalisms or images at certain schooling levels when teaching chemistry and for students to better engage with and understand the discourse more efficiently. Chapter 4 has shown that different chemical formalisms embody different levels of semantic density, and hence they are used in different schooling levels. This offers teachers a guideline for using appropriate chemical formalisms at the proper learning levels. Chapter 5 has proposed a model for the semantic density of images and analyzed the development of the images' semantic density across the curriculum stages in secondary school chemistry. The model could be a tool for teachers to identify images' semantic density, and the analysis of the images' development across schooling levels is also helpful for teachers to avoid misusing images at certain learning stages. For example, with the model and the findings of the knowledge-building analysis, teachers can identify an image's semantic density and use it at the places where it best fits, supporting students better engage with the images and learn the knowledge more efficiently.

Thirdly, the analysis of knowledge-building through multisemiotic texts offers teachers guidance on using texts comprising of semiotic resources that complement each other in terms of their semantic density and semantic gravity for teaching chemistry. For example, Chapter 6 has proposed three types of language-image relations in chemistry texts in terms of semantic density and semantic gravity and has shown which schooling levels the different types of intersemiotic relations tend to occur. This would help teachers use texts involving appropriate language-image relations in terms of Semantics for teaching chemistry at certain learning stages. For instance, the finding shows that at the lower schooling levels, language tends to embody stronger semantic density but weaker semantic gravity than image. Teachers thus can use texts that incorporate everyday images connecting to empirical phenomena to illustrate the relatively technical language text at the lower learning stages.

7.4 Limitations of this study and directions for future studies

As an exploratory account of knowledge-building of secondary school chemistry from the perspective of semiotics, this study is by no means without limitations.

Firstly, since the object of study is the knowledge of secondary school chemistry, the grammatical description of chemical formalisms in this thesis considered data from secondary school chemistry textbooks. It is not yet clear if the grammars work for chemical formalisms used in higher education. Future studies based on the axis principle are warranted to investigate if the chemical formalisms at the university level exhibit similar or different grammars.

Secondly, due to the limitation of space, this study has established only the model for the semantic density of images and left out the model for the images' semantic gravity. It should be noted that the semantic gravity of images is also crucial for knowledge-building analysis. Therefore, accounting for the semantic gravity of images is worthy of investigation by future studies.

Thirdly, this study focuses on the semiotic resources used in chemistry textbooks, which are only part of the resources involved in the knowledge-building of secondary school chemistry. As a typical scientific discipline, chemistry is characterized by 'modal diversity' (Prain and Waldrip 2010). It uses numerous other semiotic resources to build chemistry knowledge, for example, ball-stick models for molecular structures and animations. These resources are also key components of the multimodal literacies of chemistry and thus worth exploration in the future.

In a nutshell, this study has taken a modest step towards the knowledge-building of chemistry through multisemiotic resources in secondary school chemistry textbooks. Further research is needed to include a wider range of data and expand the object of study to richer modalities.

Acknowledgements

The completion of this thesis is impossible without the help and support from numerous people. I would like to express my sincerest appreciation to them for their generosity and kindness.

Thank you Prof. Delu Zhang, my supervisor at Tongji University. Thanks for taking me as your PhD student. Doing PhD research under your supervision itself is an honor for me because every time when I tell people that you are my supervisor, their faces are full of admiration. Your profound understanding of SFL and multimodality have been a constant inspiration for me during my PhD research. It is through your work that I first learned about multimodal discourse analysis and chose the multisemiotic chemistry discourse as my research object. Thanks for your patience and insightful suggestions on numerous issues in my thesis.

Thank you Prof. Karl Maton, one of my supervisors at University of Sydney. Thank you for taking me as your visiting PhD student at the LCT Center. It was really a privilege for me to learn LCT from the creator, which I believe is the dream of many PhD students. You have taught me more than I thought I could know about LCT. Thanks for your fortnight meetings with me discussing my PhD project and offering me the chance to present my work at the internationally renowned LCT Roundtables and the Third LCT International Conference. This not only improved my understanding of LCT and the ability of using it to solve problems in my PhD project but also made me understand how to deliver academic ideas through conferences and seminars more effectively. I am also very grateful for your suggestions on and editing of my book chapter. I cannot thank your more, Karl!

Thank you Dr. Yaegan Doran, the other supervisor of mine at University of Sydney. I am not sure if there is anyone who is fortunate enough as me to have you as their PhD supervisor – the world leading scholar in combing SFL and LCT. I am so indebted to you for your constant caring and patient supervisions in the last two years. It is you who taught me systemic functional description and helped me solved all the issues regarding the grammatical description of chemical formalisms. It is you who taught me how to do academic presentations in a way that makes it easier for audiences to follow. It is you who patiently edited my thesis chapters, book chapters and research articles. I am also very grateful for your supervisions after I returned to Tongji University. You have done all this for free and I really wish one day I can repay your unselfish help and support. But now just let me say this – thank you Yaegan!

Thank you Prof. James Martin. Thanks for your informal supervisions on my description of chemical formalisms and my analysis of Semantics of images. Your suggestions have made my thesis much better. I would also like to thank you for offering me opportunities of presenting my work at the Tuesday PhD Student Seminars and the Sydney SFL Seminars and sending me your literature that are very helpful to my PhD research. Thank you Prof. Louise Ravelli. It was really a pleasure to know a nice friend and teacher like you in Sydney. Thanks for attending my presentations and offering me your insightful advices on my PhD project. Thank you Dr. Dongbing Zhang and Dr. Pin Wang for teaching me more about systemic functional description. I really enjoyed our discussion about the description of chemical formalisms in Sydney. Your suggestions have made my description much better. Thank you Prof. Susan Hood. Thanks for attending my presentations and offering me your valuable comments and suggestions. I would also like to thank Prof. Yongsheng Zhu, Prof. Yanning Yang, Prof. Xueyan Yang, Prof. Jing Dai and Associate Prof. Lihe Huang for your valuable comments and suggestions on my thesis at the pre-defense meeting. Your advices have made my thesis much better.

Thank you to my PhD mates in Sydney – both the SFL PhD mates including Xiaoqin, Yufei, Qingxin, Yuan, Wen, Rurong, Georgia, Gaga, Natalia and the LCT PhD mates including Matthew, Mauricio, Kirstin, Sharon and Patrick. Our meetings and gathering were one of the happiest memories for me in Sydney. Without your companion, it would be very difficult to reach the end of my PhD. I would also like to thank Dr. Chengfang Song for taking care of me like a big brother in Sydney.

Thank you to my PhD mates at Tongji University – Yanfang, Li, Jing, Ruiyun, Juan, and Ke. Thanks for your companion for the last half of year, which is the most difficult time for me. I am grateful for that you always share my anxiety and comfort me.

Thank you to my girlfriend Miss Gao. Thanks for constantly comforting me when I am frustrated because of my PhD work. Thanks for your love and always being there for me.

Last but not least, I thank my families – my parents and sisters. Thanks for always being so supportive and encouraging. Your unfailing love has been the power source for me to complete my PhD project.

References

Airey, J., Eriksson, U., Fredlund, T. and Linder, C. (2014). *The concept of disciplinary affordance*. Paper presented at the 5th International 360° Conference: Encompassing the Multimodality of Knowledge. Aarhus University: Aarhus, 8-10 May.

Airey, J. and Linder, C. (2017). Social semiotics in university physics education. In D. Treagust, R. Duit and H.E. Fischer (eds), *Multiple Representations in Physics Education* (pp. 95–122). Dordrecht: Springer.

Arnold, M. (2011). Images, diagrams, and narratives: Charles S. Peirce's epistemological theory of mental diagrams. *Semiotica*, (186), 5–20.

Baldry, A. and Thibault, P. (2006). *Multimodal Transcription and Text Analysis*. London: Equinox.

Barthes, R. (1977[1964]). *Image Music Text*. London: Fontana Press.

Bateman, J.A. (2014). *Text and Image: A Critical Introduction to the Visual-verbal Divide*. London, New York: Routledge.

Bateman, J.A. (2019). Towards critical multimodal discourse analysis: a response to Ledin and Machine. *Critical Discourse Studies*, 15(5), 531–539.

Bennett, J. (2003). *Teaching and Learning Science*. London: Continuum.

Bernstein, B. (1999). Vertical and horizontal discourse: an essay. *British Journal of Sociology of Education*, 20(2), 157–173.

Bezemer, J. and Kress, G. (2008). Writing in multimodal texts: a social semiotic account of designs for learning. *Written Communication*, 25(2), 166–195.

Biglan, A. (1973). The characteristics of subject matter in different academic areas. *Journal of Applied Psychology*, 57(3), 195–203.

Board of Studies NSW. (2012). Science K-10 (incorporating science and technology K-6) Syllabus. https://educationstandards.nsw.edu.au/wps/portal/nesa/k-10/learning-areas/science/science-and-technology-k-6-new-syllabus/version-log (accessed 28th April 2021).

Bodner, G.M. (1986). Constructivism: a theory of knowledge. *Journal of Chemical Education*, (63), 873–879.

Brooke, M. (2017). Using ‘semantic waves’ to guide students through the research process: from adopting a stance to sound cohesive academic writing. *Asian Journal of the Scholarship of Teaching and Learning*, 7(1), 37–66.

Brooke, M. (2020). Guiding teacher talk in the Content and Language Integrated Learning classroom using Semantics from Legitimation Code Theory. *Teaching in Higher Education*, 25(7), 812–824.

Brown, T.L., Lemay, H.E., Bursten, B.R., Murphy, C.J. and Woodward, P.M. (2012). *Chemistry: The Central Science (twelfth edition)*. New York: Prentice Hall.

Bruice, P.Y. (2017). *Organic Chemistry (eight edition)*. New Jersey: Pearson.

Caffarel, A., Martin, J.R. and Matthiessen, C. (eds) (2004). *Language Typology: A Functional Perspective*. Amsterdam: John Benjamins.

Canac, S. and Kermen, I. (2016). Exploring the mastery of French students in using basic notions of the language of chemistry. *Chemistry Education Research and Practice*, 17(3), 452–473.

Chan, D., Commons, C., Hecker, R., Hillier, K., Hogendoorn, B. Lennard, L., Moylan, M., O'Shea, P., Porter, M., Sanders, P., Sturgiss, J. and Waldron, P. (2018). *Pearson Chemistry 11 New South Wales Student Book*. Melbourne: Pearson Australia.

Cheng, M. and Gilbert, J.K. (2009). Towards a better utilization of diagrams in resesarch into the use of representative levels in chemical education. In J.K. Gilbert and D. Treagust (eds), *Multiple Representations in Chemical Education* (pp. 55–74). Dordrecht: Springer.

Chiappetta, E.L. and Koballa, T.R. (2002). *Science Instruction in the Middle and Secondary Schoools*. New York: Pearson Education.

Chittleborough, G., Treagust, D. and Mocerino, M. (2005). Non-major chemistry students' learning strategies: explaning their choice and examinng the implications for teaching and learning. *Science Education International*, 16(1), 5–21.

Chittleborough, G.D. (2004). The Role of Teaching Models and Chemical Representations in Developing Students' Mental Models of Chemical Phenomena. Unpublished Unpublished PhD Thesis, Curtin University of Technology.

Christie, F. and Martin, J.R. (eds) (1997). *Genre and Institutions: Social Processes in the Workplace and School*. London: Continuum.

Christie, F. and Martin, J.R. (eds) (2007). *Language, Knowledge and Pedagogy: Functional Linguistics and Sociological Perspectives*. London and New York: Continuum.

Christie, F. and Maton, K. (eds) (2011). *Disciplinarity: Functional Linguistic and Sociological Perspectives*. London and New York: Continuum.

Clarence, S. (2017a). A relational approach to building knowledge through academic writing: facilitating and reflecting on peer writing tutorials. In S. Clarence and L. Dison (eds), *Writing Centres in Higher Education: Working in and across the Disciplines* (pp. 49–65). Stellenbosch: SUN Press.

Clarence, S. (2017b). Surfing the waves of learning: enacting a Semantics analysis of teaching in a first-year Law course. *Higher Education Research & Development*, 36(5), 920–933.

Davidowitz, B. and Chittleborough, G.D. (2009). Linking the macroscopic and sub-microscopic levels: diagrams. In J.K. Gilbert and D. Treagust (eds), *Multiple Representations in Chemical Education* (pp. 169–191). Dordrecht: Springer.

Dimopoulos, K., Koulaidis, V. and Sklaveniti, S. (2003). Towards an analysis of visual images in school science textbooks and press articles about science and technology. *Research in Science Education*, 32(2), 189–216.

Djonov, E. and van Leeuwen, T. (2011). The semiotics of texture: From tactile to visual. *Visual Communication*, 10(4), 541–564.

Doran, Y.J. (2017). The role of mathematics in physics: Building knowledge and describing the empirical world. *Onomázein*, Special Issue II, 209–226.

Doran, Y.J. (2018a). Building knowledge through images in physics. *Visual Communication*, 18(2), 251–277.

Doran, Y.J. (2018b). Intrinsic functionality of mathematics, metafunctions in Systemic

Functional Semiotics. *Semiotica*, 2018(225), 457–487.
Doran, Y.J. (2018c). *The Discourse of Physics: Building Knowledge through Language, Mathematics and Image*. New York: Routledge.
Doran, Y.J. (2019). Academic formalisms: towards a semiotic typology. In J.R. Martin, Y.J. Doran and G. Figueredo (eds), *Systemic Functional Language Description: Making Meaning Matter* (pp. 331–358). New York: Routledge.
Doran, Y.J. and Martin, J.R. (2021). Field relations: understanding scientific explanation. In K. Maton, J.R. Martin and Y.J. Doran (eds), *Studying Science: Knowledge, Language, Pedagogy* (pp. 105–133). New York: Routledge.
Ebbing, D.D. and Gammon, S.D. (2008). *General Chemistry (eigth edition)*. New York: Houghton Mifflin Company.
Firth, J.R. (1957). A synopsis of linguistic theory, 1930-1955. In *Studies in Linguistic Analysis (Special volume of the Philological Soceity)* (pp. 1–32). Oxford: Basil Blackwell.
Gabel, D. (1993). Use of the particulate nature of matter in developing conceptual understanding. *Journal of Chemistry Education*, 70(3), 193–194.
Gabel, D. (1998). The complexity of chemistry and its implications for teaching. In B.J. Fraser and K.G. Tobin (eds), *International Handbook of Science Education (Vol. 1)* (pp. 223–248). London: Kluwer.
Georgiou, H. (2016). Putting physics knowledge in the hot seat: the semantics of student understanding of thermodynamics. In K. Maton, S. Hood and S. Shay (eds), *Knowledge-building: Educational Studies in Legitimation Code Theory* (pp. 176–192). New York: Routledge.
Gilbert, J.K. (2005). Visualization: a meta-cognitive skill in science and science education. In J.K. Gilbert (ed.), *Visualization in Science Education* (pp. 9–27). Dordrecht: Springer.
Gilbert, J.K. and Treagust, D. (2009). Introduction: macro, submicro and symbolic representations and the relationship between them: key models in chemical education. In J.K. Gilbert and D. Treagust (eds), *Multiple Representations in Chemical Education* (pp. 1–8). Dordrecht: Springer.
Gkitzia, V., Salta, K. and Tzougraki, C. (2011). Development and application of suitable criteira for the evaluation of chemical representations in school textbooks. *Chemistry Education Research and Practice,* (12), 5–14.
Goodwin, W.M. (2008). Structural formulas and explanation in organic chemistry. *Foundations of Chemistry*, 10(2), 117–127.
Halliday, M.A.K. (1961). Categories of the Theory of Grammar. *Word;*, 17(2), 241–292.
Halliday, M.A.K. (1963a). Class in relation to the axes of chain and choice in language. *Linguistics*, (2), 5–15.
Halliday, M.A.K. (1963b). The tones of English, 15(1), 1–28.
Halliday, M.A.K. (1966). Some notes on "deep" grammar. *Journal of Linguistics*, 2(1), 57–67.
Halliday, M.A.K. (1967a). Notes on transitivity and theme in English: part 1. *Journal of Linguistics*, 3(1), 37–81.
Halliday, M.A.K. (1967b). Notes on transitivity and theme in English: part 2. *Journal*

of Linguistics, 3(2), 199–244.

Halliday, M.A.K. (1968). Notes on transitivity and theme in English: part 3. *Journal of Linguistics*, 4(2), 179–215.

Halliday, M.A.K. (1969). Options and functions in the English clause. *Brno Studies in English*, 8, 81–88.

Halliday, M.A.K. (1970). Functional diversity in language as seen from a consideration of Modality and Mood in English. *Foundations of Language*, 6(3), 322–361.

Halliday, M.A.K. (1978). *Language as a Social Semiotic: The Social Interpretation of Language and Meaning*. London: Edward Arnold.

Halliday, M.A.K. (1979). Modes of meaning and modes of expression: types of grammatical structure and their determination by different semantic functions. In D.J. Allerton, E. Carney and D. Holdcroft (eds), *Function and Context in Linguistics Analysis: A Festshrift for William Hass* (pp. 57–79). Cambridge: Cambridge University Press.

Halliday, M.A.K. (1981[1965]). Types of structure. In M.A.K. Halliday and J.R. Martin (eds), *Readings in Systemic Linguistics* (pp. 29–41). Bristol: Batford Academic and Educational Ltd.

Halliday, M.A.K. (1985). *An Introduction to Functional Grammar*. London: Edward Arnold.

Halliday, M.A.K. (1993a). On the language of physical science. In M.A.K. Halliday and J.R. Martin (eds), *Writing Science: Literacy and Discursive Power* (pp. 59–75). London: The Falmer Press.

Halliday, M.A.K. (1993b). Towards a language-based theory of learning. *Linguistics and Education*, 5(2), 93–116.

Halliday, M.A.K. (1994). *An Introduction to Functional Grammar*. London: E. Arnold.

Halliday, M.A.K. (2002[1970]). Language structure and language function. In J. Webster (ed.), *On Grammar (Volum 1 in the Collected Works of M.A.K. Halliday)* (pp. 173–195). London: Bloomsbury Publishing.

Halliday, M.A.K. (2003[1973]). The functional basis of language. In J. Webster (ed.), *On Language and Linguistics (Volum 3 in the Collected Works of M.A.K. Halliday)* (pp. 298–322). New York: Continuum.

Halliday, M.A.K. (2004[1998]). Language and knowledge: the 'unpacking' of text. In J. Webster (ed.), *The Language of Science (Collected Works of M.A.K. Halliday)* (pp. 24–48). London: Continuum.

Halliday, M.A.K. (2007[1988]). Some basic concepts of educational linguistics. In J. Webster (ed.), *Language and Education (Collected Works of M.A.K. Halliday)* (pp. 341–353). London: Continuum.

Halliday, M.A.K. and Hasan, R. (1976). *Cohesion in English*. London: Longman.

Halliday, M.A.K. and Martin, J.R. (eds) (1993). *Writing Science: Literacy and Discursive Power*. London: The Falmer Press.

Halliday, M.A.K. and Matthiessen, C.M.I.M. (1999). *Construing Experience through Meaning: A Language-based Approach to Cognition*. London: Cassell.

Halliday, M.A.K. and Matthiessen, C.M.I.M. (2014). *Halliday's Introduction to Functional Grammar (fourth edition)*. London and New York: Routledge.

Hao, J. (2020). *Analysing Scientific Discourse From a Systemic Functional Linguistic Perspective: A Framework for Exploring Knowledge-building in Biology*. New York: Routledge.

He, Y.F. (2020). Animation as a Semiotic Mode: Construing Knowledge in Science Animated Videos. Unpublished PhD thesis, University of Sydney.

Herrington, D. and Daubenmire, P.L. (2016). No teacher is an island: bridging the gap between teachers' professional practice and research findings. *Journal of Chemical Education*, 93(8), 1371–1376.

Herron, D.J. (1975). Piaget for Chemists: explaining what "good" students cannot understand. *Journal of Chemical Education*, 52(3), 146–150.

Hjelmslev, L. (1969). *Prolegomena to a Theory of Language*. Madison, Milwaukee, and London: The University of Wisconsin Press.

Hodge, R. and Kress, G. (1988). *Social Semiotics*. New York: Cornell University Press.

Hoffmann, R. and Laszlo, P. (1991). Representation in chemistry. *Angewandte Chemie (International Edition in English)*, 30(1), 1–16.

Hood, S. (2016). Ethnographies on the move, stories on the rise: methods in the humanities. In K. Maton, S. Hood and S. Shay (eds), *Knowledge-building: Educational Studies in Legitimation Code Theory* (pp. 117–137). New York: Routledge.

Hood, S. (2017). Live lectures: the significance of presence in building disciplinary knowledge. *Onomázein*, special issue II, 179–208.

Huddleston, R.D. (1965). Rank and depth. *Language*, 41(4), 574–586.

International Association for the Evaluation of Educational Achievement. (2019). TIMSS 2019: International Results in Mathematics and Science. https://www.iea.nl/publications/study-reports/international-reports-iea-studies/timss-2019-international-report (accessed 28th April 2021).

Jacob, C. (2001). Analysis and synthesis: interdependent operations in chemical language and practice. *HYLE - International Journal for Philosophy of Chemistry*, 7(1), 31–50.

Jewitt, C. (2017). Different approaches to multimodality. In C. Jewitt (ed.), *The Routledge Handbook of Multimodal Analysis* (pp. 31–43). New York: Routledge.

Johnstone, A.H. (1982). Macro- and micro-chemistry. *School Science Review*, (64), 377–379.

Johnstone, A.H. (1991). Why is science difficult to learn? Things are seldom what they seem. *Journal of Computer Assisted Learning*, 7(2), 75–83.

Johnstone, A.H. (1993). The development of chemistry teaching: a changing response to changing demand. *The Forum*, 70(9), 701–705.

Johnstone, A.H. (2000). Teaching of chemistry - logical or psychological? *Chemistry Education Research and Practice*, 1(1), 9–15.

Jong, O. and Taber, K.S. (2014). The many faces of high school chemistry. In N. Lederman and S.K. Abell (eds), *Handbook of Research in Science Education Volume 5* (pp. 457–480). New York: Routledge.

Kirk, S. (2017). Waves of reflection: seeing knowledge(s) in academic writing. In *EAP in a Rapidly Changing Landscape: Issues, challenges and solutions* (pp. 109–117).

Reading: Garnet Education.
Kirk, S. (2018). Enacting the Curriculum in EAP: An LCT Analysis. Unpublished PhD thesis, Durhan University.
Klein, U. (2001). Berzelian formulas as paper tools in early nineteenth-century chemistry. *Foundations of Chemistry*, (3), 7–32.
Kozma, R.B. and Russell, J. (1997). Multimedia and understanding: expert and novice responses to different representations of chemical phenomena. *Journal of Research in Science Teaching*, 34(9), 949–968.
Kress, G. (1993). Against arbitrariness: the social production of the sign as a fundational issue in critical discours analysis. *Discourse and Society*, 4(2), 169–191.
Kress, G. (2003). *Literacy in the New Media Age*. London: Routledge.
Kress, G. and van Leeuwen, T. (2001). *Multimodal Discourse: The Modes and Media of Contemporary Communication*. London: Edward Arnold.
Kress, G. and van Leeuwen, T. (2006). *Reading images: The Grammar of Visual Design*. London and New York: Routledge.
Lambrinos, E.M. (2020). Building Ballet: Developing Dance and Dancers in Ballet. Unpublished PhD thesis, The University of Sydney.
Laugier, A. and Dumon, A. (2004). The equation of reaction: a cluster of obstacles which are difficult to overcome. *Chemistry Education Research and Practice*, 5(3), 327–342.
Lee, O. (2005). Science education with English language learners: synthesis and research agenda. *Review of Education Research*, 75(4), 491–530.
Lemke, J. (1998). Multiplying meaning: visual and verbal semiotics in scientific text. In J.R. Martin and R. Veel (eds), *Reading science: Critical and functional perspectives on discourses of science* (pp. 87–113). London: Routledge.
Liu, Y. (2011). Scientific Literacy in Secondary School Chemistry: A Multimodal Perspective. Unpublished PhD thesis, National University of Singapore.
Liu, Y. and O'Halloran, K.L. (2009). Intersemiotic texture: analyzing cohesive devices between language and images. *Social Semiotics*, 19(4), 367–388.
Machin, D. (2009). Multimodality and theories of the visual. In C. Jewitt (ed.), *The Routledge Handbook of Multimodal Analysis* (pp. 181–190). London and New York: Routledge.
Markic, S. and Childs, P.E. (2016). Language and the teaching and learning of chemistry. *Chemistry Education Research and Practice*, 17(3), 434–438.
Martin, J.R. (1989). Technicality and abstraction: language for the creation of speialized texts. In F. Christie (ed.), *Writing in Schools: Reader* (pp. 36–44). Geelong, Vic: Deakin University Press.
Martin, J.R. (1992). *English Text: System and Structure*. Philadelphia: John Benjamins Pub. Co.
Martin, J.R. (1993). Technicality and abstraction: language for the creation of specialized texts. In M.A.K. Halliday and J.R. Martin (eds), *Writing Science: Literacy and Discursive Power* (pp. 223–243). London: The Falmer Press.
Martin, J.R. (1996). Types of structure: deconstructing notions of constituency in clause and text. In E.H. Hovy and D.R. Scott (eds), *Computational and conversational*

discourse (pp. 39–66). Berlin: Springer.

Martin, J.R. (1997). Analysing genre: functional parameters. In F. Christie and J.R. Martin (eds), *Genre and institutions: Social processes in the workplace and school* (pp. 3–39). London: Continuum.

Martin, J.R. (2007). Construing knowledge: a functional linguistic perspective. In F. Christie and J.R. Martin (eds), *Language, knowledge and pedagogy: Functional linguistic and sociological perspectives: Functional linguistic and sociological perspectives* (pp. 34–64). London: Continuum.

Martin, J.R. (2013a). Embedded literacy: Knowledge as meaning. *Linguistics and Education*, 24(1), 23–37.

Martin, J.R. (2013b). *Systemic Functional Grammar: A Next Step into the Theory: Axial Relations*. Beijing: Higher Education Press.

Martin, J.R. (2015a). Halliday the grammarian: axial foundations. In M. Halliday and J. Webster (eds), *The Bloomsbury Companion to M.A.K. Halliday* (pp. 257–290). London: Bloomsbury Academic.

Martin, J.R. (2015b). Jim R. Martin. In M. Boeriis, E. Maagerø and E.S. Tønnessen (eds), *Social semiotics: Key figures, new directions* (pp. 42–68). London, New York: Routledge.

Martin, J.R. (2017). Revisiting field: specialized knowledge in secondary school science and humanities discourse. *Onomázein*, Special Issue III, 111–148.

Martin, J.R. and Maton, K. (2013). Cumulative knowledge-building in secondary schooling. *Linguistics and Education*, 24(1), 1–3.

Martin, J.R., Maton, K. and Doran, Y.J. (eds) (2019). *Accessing Academic Discourse: Systemic Functional Linguistics and Legitimation Code Theory*. New York: Routledge.

Martin, J.R. and Matruglio, E. (2019). Revisiting mode: context in/dependency in ancient history classroom discourse. In J.R. Martin, K. Maton and Y.J. Doran (eds), *Accessing Academic Discourse: Systemic Functional Linguistics and Legitimation Code Theory* (pp. 89–113). New York: Routledge.

Martin, J.R. and Rose, D. (2003). *Working with Discourse: Meaning beyond the Clause*. London and New York: Continuum.

Martin, J.R. and Rose, D. (2008). *Genre relations: Mapping culture*. London: Equinox Pub.

Martinec, R. (1998). Cohesion in action. *Semiotica*, 120(1/2), 243–268.

Martinec, R. (2000). Types of process in action. *Semiotica*, 135(3/4), 243–268.

Martinec, R. (2001). Interpersonal resources in action. *Semiotica*, 135(3/4), 117–145.

Martinec, R. and Salway, A. (2005). A system for image-text relations in new (and old) media. *Visual Communication*, 4(3), 337–371.

Maton, K. (2011). Theories and things: the semantics of disciplinarity. In F. Christie and K. Maton (eds), *Disciplinarity: Functional Linguistic and Sociological Perspectives* (pp. 62–84). London and New York: Continuum.

Maton, K. (2013). Making semantic waves: a key to cumulative knowledge-building. *Linguistics and Education*, 24(1), 8–22.

Maton, K. (2014). *Knowledge and Knowers: Towards a Realist Sociology of Education*.

New York: Routledge.

Maton, K. (2020). Semantic waves: context, complexity and academic discourse. In J.R. Martin, K. Maton and Y.J. Doran (eds), *Accessing Academic Discourse: Systemic Functional Linguistics and Legitimation Code Theory* (pp. 59–85). London: Routledge.

Maton, K. and Chen, R.T.H. (2016). LCT in qualitative research: creating a transation device for studying constructivist pedagogy. In K. Maton, S. Hood and S. Shay (eds), *Knowledge-building: Educational Studies in Legitimation Code Theory* (pp. 27–48). New York: Routledge.

Maton, K. and Doran, Y.J. (2017a). Condensation: a translation device for revealing complexity of knowledge practices in discourse, part 2-clausing and sequencing. *Onomázein*, Special Issue II, 77–110.

Maton, K. and Doran, Y.J. (2017b). Semantic density: a translation device for revealing complexity of knowledge practices in discourse, part 1-wording. *Onomázein*, Special Issue II, 46–76.

Maton, K. and Doran, Y.J. (2017c). Semantic density: a translation device for revealing complexity of knowledge practices in discourse, part 1 - wording. *Onomazein*, special issue II, 46–76.

Maton, K. and Doran, Y.J. (2017d). Systemic functional linguistics and code theory. In T. Bartlett and G. O'Grady (eds), *The Routledge Handbook of Systemic Functional Linguistics* (pp. 605–618). New York: Routledge.

Maton, K., Martin, J.R. and Doran, Y.J. (eds) (2021). *Teaching Science: Knowledge, language, Pedagogy*. New York: Routledge.

Matruglio, E., Maton, K. and Martin, J.R. (2013). Time travel: the role of temporality in enabling semantic waves in secondary school teaching. *Linguistics and Education*, 24(1), 38–49.

Matthiessen, C.M.I.M. and Halliday, M.A.K. (2009). *Systemic Functional Grammar: A First Step into the Theory*. Beijing: Higher Education Press.

McDonal, C.V. (2016). Evaluating junior secondary science textbook usage in Australian schools. *Research in Science Education*, 46(4), 481–509.

Mitchell, W.T. (1986). *Iconology: Image, Text, Ideology*. Chicago: University of Chicago Press.

Moje, E.B. (2008). Foregrounding the disciplines in secondary literacy teaching and learning: a call for change. *Journal of Adolescent and Adult Literacy*, 52(2), 96–107.

Moore, R. and Young, M. (2001). Knowledge and the curriculum in the sociology of education: towards a reconceptualisation. *British Journal of Sociology of Education*, 22(4), 445–461.

Mouton, M. and Archer, E. (2019). Legitimation code theory to facilitate transition from high school to first-year biology. *Journal of Biological Education*, 53(1), 2–20.

O'Halloran, K.L. (1999a). Interdependence, interaction and metaphor in multisemiotic texts. *Social Semiotics*, 9(3), 317–354.

O'Halloran, K.L. (1999b). Towards a systemic functional analysis of multisemiotic mathematics texts. *Semiotica*, 124(1/2), 1–29.

O'Halloran, K.L. (2005). *Mathematical Discourse: Language, Symbolism And Visual*

Images. New York: Continuum.

O'Halloran, K.L. (2011). Multimodal discourse analysis. In K. Hyland and B. Paltridge (eds), *Bloomsbury Companion to Discourse Analysis* (pp. 120–137). New York: A&C Black.

O'Tool, M. (1994). *The Language of Displayed Visual Art*. New Jersey: Associated University Press.

O'Toole, M. (1990). A systemic-functional semiotics of art. *Semiotica*, 82(3/4), 185–209.

Painter, C., Martin, J.R. and Unsworth, L. (2013). *Reading Visual Narratives: Image Analysis of Children's Picture Books*. Sheffield South Yorkshire, Oakville CT: Equinox Pub.

Parodi, G. (2012). University genres and multisemiotic features: accessing specialized knowledge through disciplinarity. *Forum Linguistico*, 9(4), 259–282.

Patton, M.Q. (2002). *Qualitative Evaluation and Research Methods*. Thousand Oaks, CA: Sage.

Prain, V. and Waldrip, B. (2010). Representing science literacies: an introduction. *Research in Science Education*, (40), 1–3.

Quadros, A., CarvalhoDa-Silva, D., César Silva, F., Pereira de Andrade, F., Gabriela Aleme, H., Cristina Tristão, J., Rodrigues Oliveira, S., José Santos, L. and DeFreitas-Silva, G. (2011). The knowledge of chemistry in secondary education: difficulties from the teachers' viewpoint. *Educación Química*, 22(3), 232–239.

Ravelli, L.J. and McMurtrie, R.J. (2016). *Multimodality in the Built Environment: Spatial Discourse Analysis*. New York, Abingdon: Routledge.

Rennie, L., Goodrum, D. and Hackling, M. (2001). Science teaching and learning in Australian schools: results of a national study. *Research in Science Education*, 31, 455–498.

Richardson, S. (2020). Teaching Jazz: A Study of Beliefs and Pedagogy Using Legitimation Code Theory. Unpublished PhD thesis, The University of Sydney.

Rose, D. and Martin, J.R. (2012). *Learning to Write, Reading to Learn: Genre, Knowledge and Pedagogy in the Sydney School*. London: Equinox.

Royce, T.D. (2002). Multimodality in the TESOL classroom: exploring visual-verbal synergy. *TESOL Quarterly*, 36(2), 191–205.

Royce, T.D. (2007). Intersemiotic complementarity: a framework for multimodal discourse analysis. In T.D. Royce and W.L. Bowcher (eds), *New Directions in the Analysis of Multimodal Discourse* (pp. 63–110). London: Lawrence Erlbaum Associates.

Saussure, F.D. (1959). *Course in General Linguistics*. New York: Philosophical Library.

Shiland, T.W. (1999). Constructivism: the implications for laboratory work. *Journal of Chemical Education*, 76(1), 107–109.

Simon, J.H. (1926). Chemical equations. *Journal of Chemical Education*, 3(11), 1305–1312.

Stenglin, M. (2009). Space odyssey: towards a grammar of three-dimensional space. *Visual Communication*, 8(1), 35–64.

Taber, K.S. (2000). Chemistry lessons for universities? a review of constructivist ideas.

University Chemistry Education, 4(2), 63–72.

Taber, K.S. (2009). Learning at the symbolic level. In E.J.K. Gilbert, J.K. Gilbert and D. Treagust (eds), *Multiple Representations in Chemical Education* (pp. 75–108). Dordrecht: Springer.

Talanquer, V. (2011). Macro, submicro, and symbolic: the many faces of the chemistry "triplet". *Journal of Science Education*, 33(2), 179–195.

Taskin, V. and Bernholt, S. (2014). Students' understanding of chemical formulas: a review of empirical research. *International Journal of Science Education*, 36(1), 157–185.

The New London Group (1996). A pedagogy of multiliteracies: designing social futures. *Harvard Educational Review*, 66(1), 60–93.

Treagust, D. and Chittleborough, G.D. (2001). Chemistry: a matter of understanding representations. In J. Brophy (ed.), *Subject-specific Instructional Methods and Activities* (pp. 239–267). Oxford: Elsevier Science Ltd.

Treagust, D., Chittleborough, G.D. and Mamiala, T. (2003). The role of submicroscopic and symbolic representations in chemical explanations. *International Journal of Science Education*, 25(11), 1353–1368.

Treagust, D., Duit, R. and Nieswandt, M. (2000). Sources of students' difficulties in learning chemsitry. *Educación Química*, 11(2), 228–235.

Unsworth, L. (2001). *Teaching Multiliteracies Across the Curriculum: Changing Contexts of Text and Image in Classroom Practice*. Buckingham: Open University Press.

Unsworth, L. (2007). Image/text relations and intersemiosis: towards multimodal text description for multileracies education. In L. Barbara and T.B. Sardinha (eds), *Proceedings of the 33rd International Systemic Functional Congress* (pp. 1165–1205).

Unsworth, L. (2020). Intermodal relations, mass and presence in school science explanation genres. In M. Zappavigna and S. Dreyfus (eds), *Discourses of Hope and Reconciliation: On J.R. Martin's contribution to Systemic Functional Linguistics* (pp. 131–153). London: Bloomsbury Academic.

Upahi, J.E. and Jimoh, M.A. (2015). Classification of end-of-chapter questions in senior school chemistry textbooks used in Nigeria. *Electronic Journal of Science Education*, 19(7), 1–16.

van Leeuwen, T. (1999). *Speech, Music, Sound*. London: Macmillan.

van Leeuwen, T. (2005). *Introducing Social Semiotics*. London: Routledge.

van Leeuwen, T. (2009). Parametric systems: The case of voice quality. In C. Jewitt (ed.), *The Routledge Handbook of Multimodal Analysis* (pp. 68–77). London and New York: Routledge.

Veel, R. (1997). Learning how to mean - scientifically speaking: Apprenticeship into scientific discourse in the secondary school. In F. Christie and J.R. Martin (eds), *Genre and Institutions: Social Processes in the Workplace and School* (pp. 161–195). London: Continuum.

Vidal, L. (2014). Theorizing Popular Education as a Knowledge Practice: the Case of Chile. Unpublished Unpublished PhD thesis, University of Technology, Sydney.

Walton, J. (2020). Making the Grade: Theorising Musical Performance Assessment. Unpublished PhD thesis, University of Sydney.

Weininger, S.J. (1998). Contemplating the finger: visuality and the semiotics of chemistry. *HYLE - An International Journal for the Philosophy of Chemistry*, 4(1), 3–27.

Wellington, J.J. and Osborne, J. (2001). *Language and Literacy in Science Education*. Buckingham: Open University Press.

Whorf, B.J. (1945). Grammatical categories. *Language*, 21(1), 1–11.

Wignell, P., Martin, J.R. and Eggins, S. (1993). The discourse of geography: ordering and explaning the experiential world. In M.A.K. Halliday and J.R. Martin (eds), *Writing Science: Literacy and Discursive Power* (pp. 115–182). London: The Falmer Press.

Wink, D.J. (2014). Constructivist frameworks in chemistry education and the problem of the "Thumb in the Eye". *Journal of Chemical Education*, 91(5), 617–622.

Zhang, D.B. (2020). Negotiating interpersonal meaning in Khorchin Mongolian: discourse and grammar. Unpublished PhD thesis, University of Sydney.

Zhang, D.L. (2018). A systemic-functional synthetic framework for multimodal discourse analysis. *Modern Foregin Languages*, 41(6), 731–743.

Zhang, H. and Hocking, G. (1996). Withdrawal of layered fluid through a line sink in a porous medium. *Journal of the Australian Mathematical Socienty*, 38(Part 2), 240–254.

Zhang, J., Alford, D., Hopley, S. and Tilley, C. (2014a). *Oxford Insight Science 8 Student Book*. Oxford University Press.

Zhang, J., Alford, D., Hopley, S. and Tilley, C. (2014b). *Oxford Insight Science 9 Student Book*. Melbourne: Oxford University Press.

Zhang, J., Alford, D., McGowan, D. and Tilley, C. (2013). *Oxford Insight Science 7 Student Book*: Oxford University Press.

Zhang, J., Filan, S.D., Hopley, S., Morante, R. and Tilley, C. (2015). *Oxford Insight Science 10 Student Book*: Oxford University Press.

Zhao, Q. (2012). Knowledge Building in Physics Textbooks in Primary and Secondary Schools. Unpublished PhD thesis, Xiamen University.

Zhao, S. (2010a). Intersemiotic relations as logogentic patterns: towards the restoration of the time dimension in hypertext description. In M. Bednarek and J.R. Martin (eds), *New Discourse on Language: Functional Perspectives on Multimodality, Identity and Affiliation* (pp. 195–218). London: Continuum.

Zhao, S. (2010b). Rank in visual grammar: some implications for multimodal discourse analysis. In A. Mahboob and N.K. Knight (eds), *Appliable Linguistics* (pp. 251–266). London: Continuum.

Zhao, W. (2020). Exploring college English teaching of rhetorical knowledge: A Legitimation Code Theory analysis. *Language Teaching Research*, 24(4), 1-21.

Appendix A Details of Data

Data used in this study

Stage	Year	Textbook
4	7	Oxford Insight Science 7 student book
	8	Oxford Insight Science 8 student book
5	9	Oxford Insight Science 9 student book
	10	Oxford Insight Science 10 student book
6	11	Pearson Chemistry 11 New South Wales Student Book
	12	Pearson Chemistry 12 New South Wales Student Book

Information of the textbooks

Stage 4

Zhang, J., Alford, D., McGowan, D. & Tilley, C. (2013). *Oxford Insight Science 7 Student Book*. Oxford University Press. (Year 7)

Zhang, J., Alford, D., Hopley, S. & Tilley, C. (2014a). *Oxford Insight Science 8 Student Book*. Oxford University Press. (Year 8)

Stage 5

Zhang, J., Alford, D., Morante, R. & Tilley, C. (2014b). *Oxford Insight Science 9 Student Book*. Oxford University Press. (Year 9)

Zhang, J., Filan, S., D., Hopley, S., Morante, R. & Tilley, C. (2015). *Oxford Insight Science 10 Student Book*. Oxford University Press. (Year 10)

Stage 6

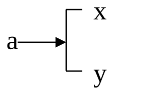
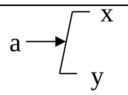
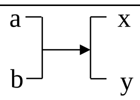
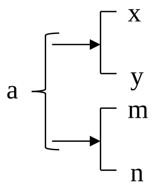
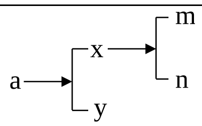
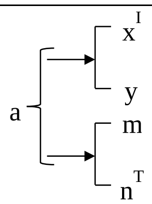
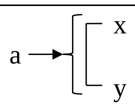
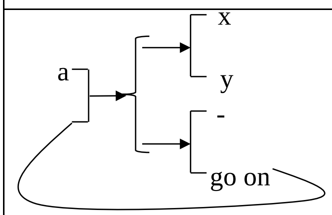
Chan, D., Commons, C., Hecker, R., Hillier, K., Hogendoorn, B., Lennard, L., Moylan, M., O'Shea, P., Porter, M., Sanders, P., Sturgiss, J. and Waldron, P. (2018). *Pearson Chemistry 11 New South Wales Student Book*. Melbourne: Pearson Australia. (Year 11)

Chan, D., Commons, C., Commons, P., Finlayson, E., Hillier, K., Hogendoorn, B., Johns, R., Lennard, L., Moylan, M., O'Shea, P., Porter, M., Sanders, P., Sturgiss, J. and Waldron, P. (2019). *Pearson Chemistry 12 New South Wales Student Book*. Melbourne: Pearson Australia. (Year 12)

Appendix B System Network Conventions

Systems are reproduced from Matthiessen and Halliday (2009) and realization statements are primarily reproduced from Martin (2013b) with some adaptations.

Systems

	**System:** If ‘a’, then ‘x’ or ‘y’, abbreviated as ‘a: x/y’
	**Clined system:** If ‘a’, then ‘x/y’ or any points along the cline
	**Disjunction in entry condition:** If ‘a/b’, then ‘x/y’
	**Simultaneity:** If ‘a’, then simultaneously ‘x/y’ and ‘m/n’
	**Delicacy ordering:** If ‘a’, then ‘x/y’; if ‘x’, then ‘m/n’
	**Conditional marking:** If ‘x’, then also ‘n’
	**And/or system:** If ‘a’, then ‘x/y’ or ‘x’ and ‘y’
	**Recursive system (logical):** If ‘a’, then ‘x/y’ and simultaneously option of entering and selecting from the same system again

Realization statements

A realization statement consists of an operator, such as ‘insert’ or ‘conflate’, and one or more operands, at least one of which is a grammatical function.

major type	operator	operand 1	operand 2	example
(i) structuring	Insert (+)	Function	-	+Subject
	Order (^)	Function	Function	Subject^Finite
(ii) layering	Conflate (/)	Function	Function	Subject/Agent
(iii) inter-rank realization	Preselect (:)	Function	feature(s)	Subject: nominal group

The three major types of realization statements are outlined in more detail below:

(i) Presence of Functions in the structure: the presence of a Function in a functional structure is specified by inserting the Function into the structure; the operation of insertion is symbolized by ‘+’, for example, +Subject.

Relative ordering of Functions and ordering relative to unit boundaries: two Functions may be ordered relative to one another in the functional structure and this relative ordering is symbolized by the notation ‘^’, for example, Subject^Finite.

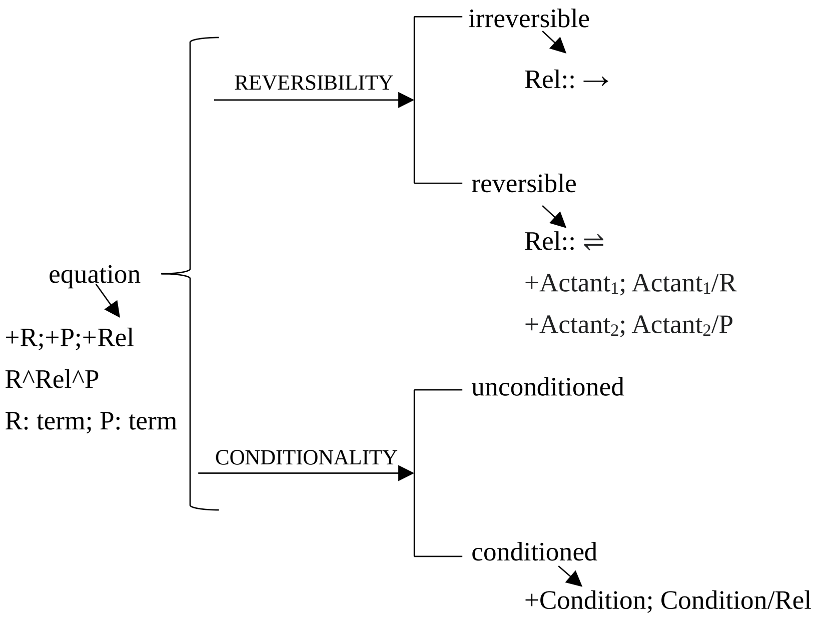
(ii) Conflation of one Function with another: one Function from one perspective is conflated with a Function from another perspective. That is, the two Functions are description of the same constitute but at different layers. Conflation is symbolized by ‘/’. For example, ‘Subject/Agent’ means that Subject (interpersonal) and Agent (ideational) apply to the same constituent.

(iii) Realization of a Function in terms of features from the rank below: the realization of a Function in a functional structure is stated by preselecting one of more features from the unit realizing it; preselection is symbolized by ‘:’, for example, Subject: nominal group.

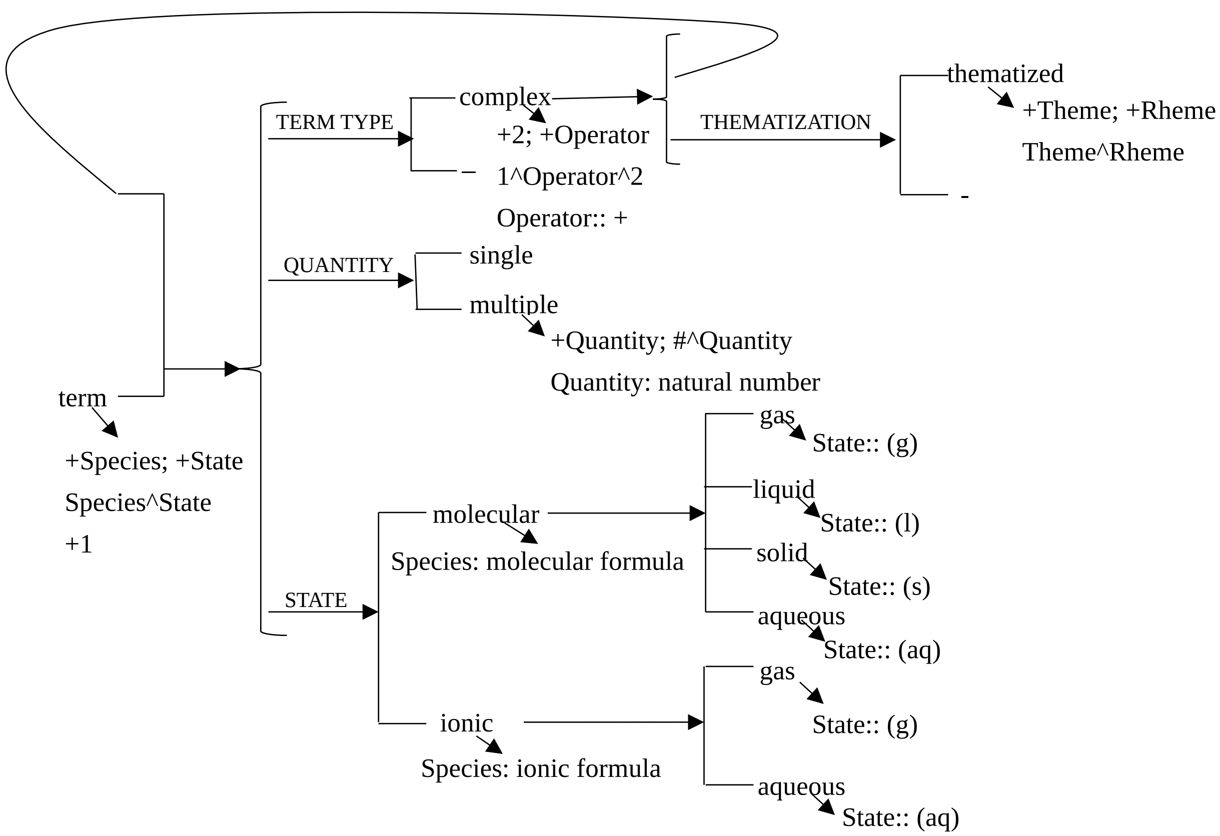
Appendix C Full System Networks for Chemical Formalisms

Chemical equations

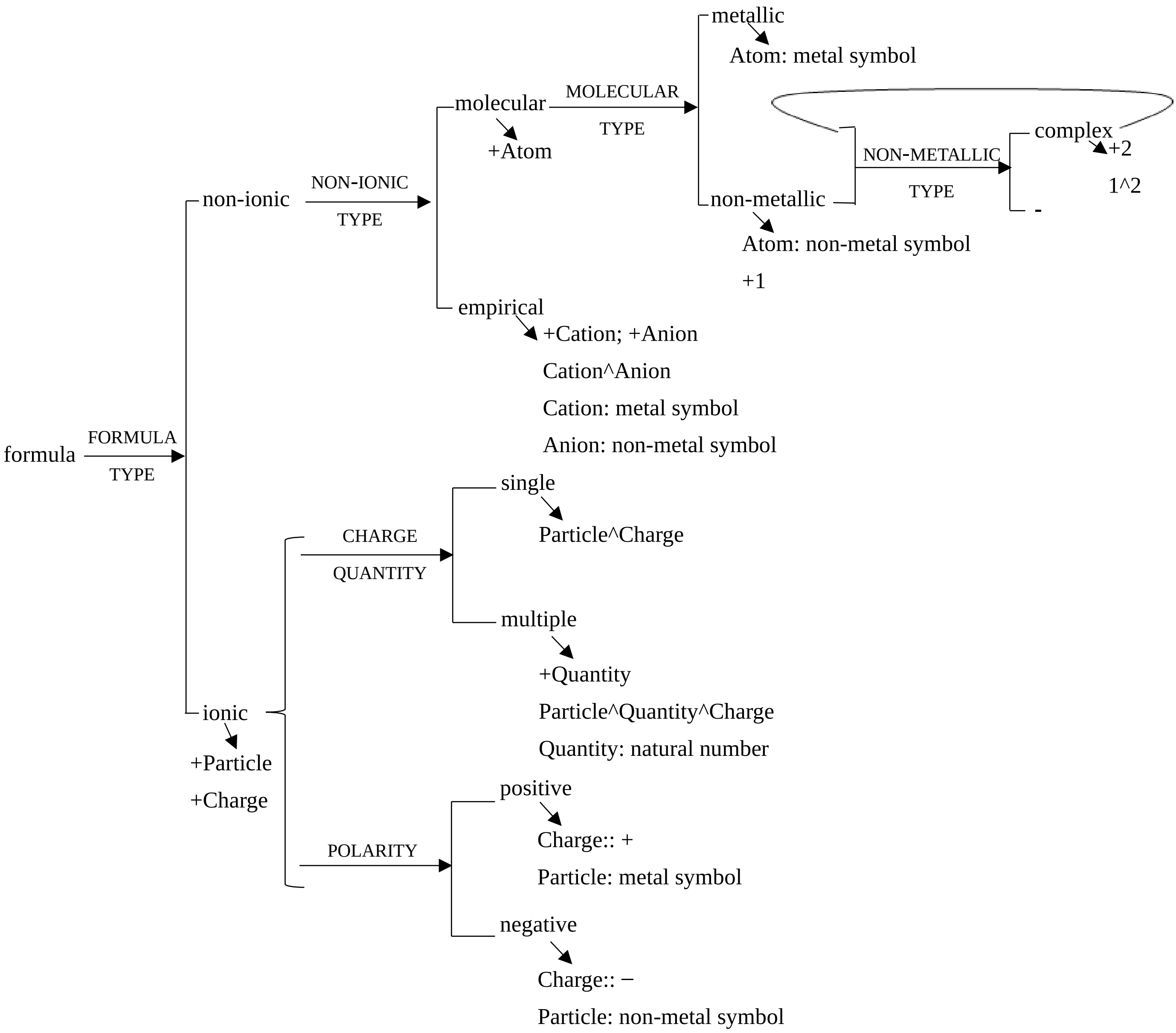
(i) The equation rank



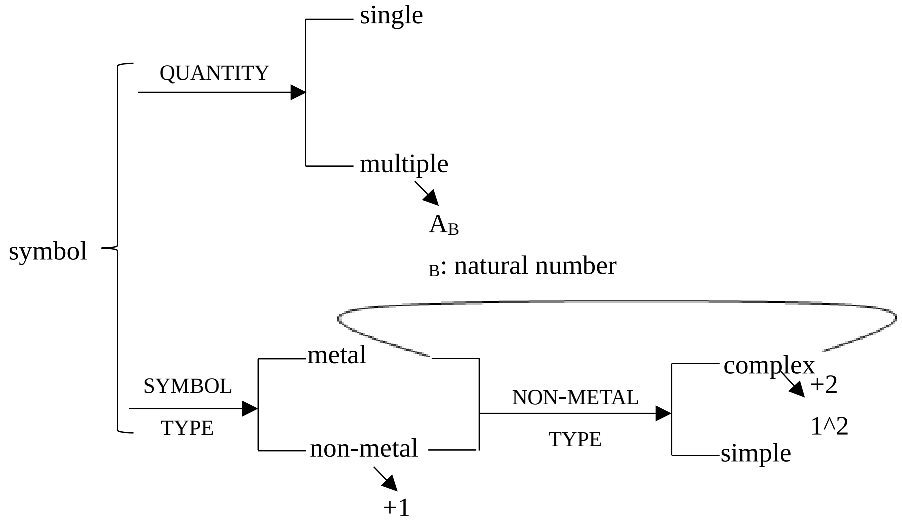
(ii) The term rank



(iii) The formula rank

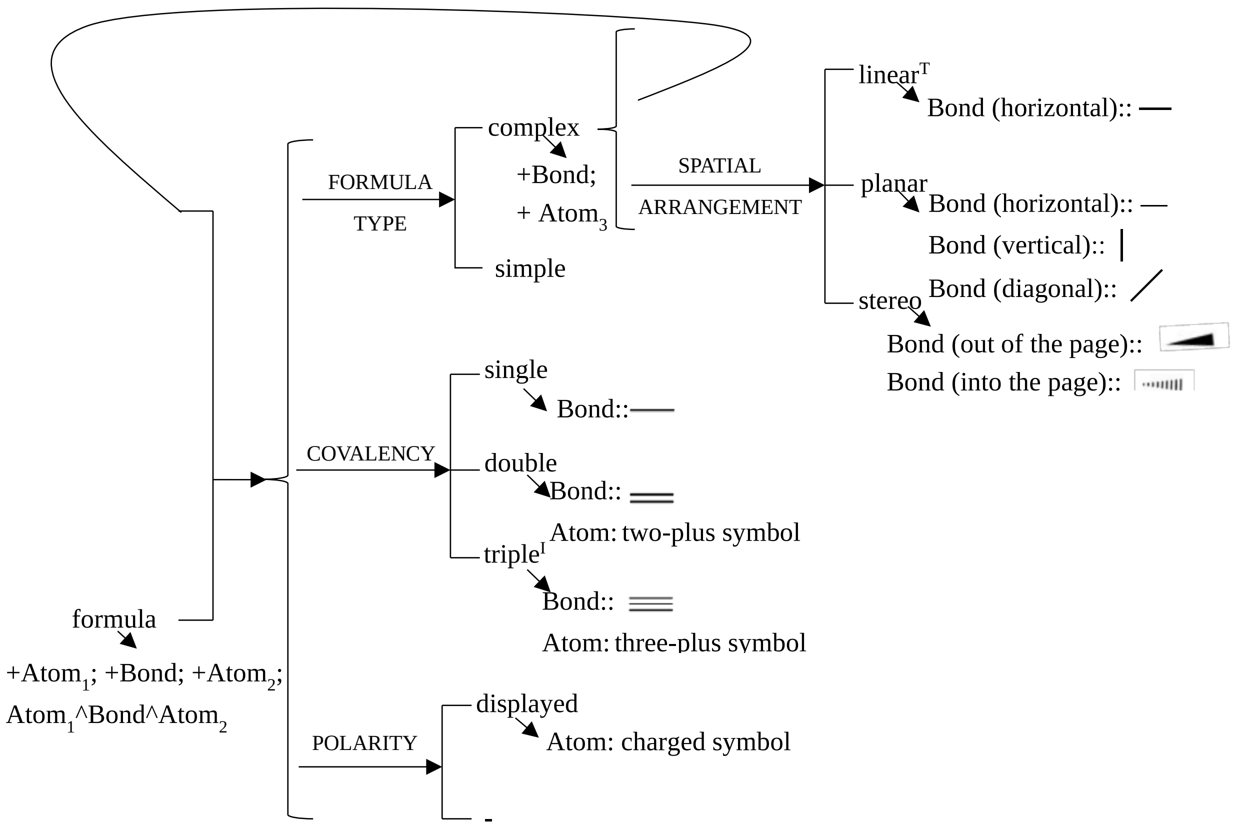


(iv) The symbol rank



Structural formulas

(i) The formula rank



(ii) The symbol rank