

УДК 811.581,37+811.581,35

Сяо И

преподаватель

Белорусский государственный
университет иностранных языков
г. Минск, Беларусь**Xiao Yi**

Lecturer

Belarusian State University
of Foreign Languages
Minsk, Belarus
xiao.eric@yandex.by

СТРУКТУРНО-СЕМАНТИЧЕСКИЕ МОДЕЛИ
ОБРАЗОВАНИЯ КИТАЙСКИХ ИДЕОГРАММ, ИМЕНУЮЩИХ
АБСТРАКТНЫЕ ПОНЯТИЯ: К РАЗРАБОТКЕ БАЗЫ ЗНАНИЙ
ДЛЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

В данной статье с целью последующего создания базы знаний ИИ анализируются структурно-семантические модели образования 82 идеограмм, обозначающих абстрактные понятия в китайской письменности. Выделены три основные семантические категории идеограмм – абстрактных понятий: 1) государственно-общественные абстракции (73,1 %); 2) религиозно-ритуальные абстракции (15,9 %); 3) эмоционально-интеллектуальные абстракции (11,0 %). Определено, что идеограммы – абстрактные понятия в китайской письменности образуются путем соединения семантических компонентов со следующими значениями: 1) фигура человека и части его тела (人 ‘человек’, 心 ‘сердце’); 2) инструменты власти (矢 ‘стрела’, 戈 ‘алебарда’); 3) ритуальные объекты (祭 ‘алтарь’, 鼎 ‘треножник’). Работа вносит вклад в понимание механизмов знакообразования в иероглифике, подчеркивая роль культурно-исторического контекста в формировании идеограмм-абстрактных понятий. Она является основой для создания баз знаний для искусственного интеллекта.

Ключевые слова: идеограмма; субъект; иероглиф; смысловой компонент; структурно-семантическая модель; китайский язык.

STRUCTURAL-SEMANTIC MODELS OF CHINESE CHARACTERS
DENOTING ABSTRACT CONCEPTS: TOWARD THE DEVELOPMENT
OF KNOWLEDGE BASES FOR ARTIFICIAL INTELLIGENCE

This article analyzes structural-semantic models of 82 Chinese characters representing abstract concepts to inform the creation of AI knowledge bases. Three primary semantic categories of abstract-character characters are identified: 1) Socio-political abstractions (73,1 %); 2) Religious-ritual abstractions (15,9 %); 3) Emotional-intellectual abstractions (11,0 %). The study reveals that Chinese abstract characters are formed by combining semantic components with the following meanings: 1) Human figures and body parts (e.g., 人 ‘person’, 心 ‘heart’); 2) Instruments of authority (e.g., 矢 ‘arrow’, 戈 ‘halberd’); 3) Ritual objects (e.g., 祭 ‘altar’, 鼎 ‘tripod’). This work contributes to understanding hieroglyphic sign-formation mechanisms, emphasizing the role of cultural-historical context in shaping abstract-character semantics. It provides a foundation for developing AI knowledge bases.

Key words: ideogram; subject; logograph; semantic component; structural-semantic model; Chinese language.

Китайская иероглифика, будучи графической письменной системой с уникальной структурной сложностью, занимает особое место в мировой лингвистике как исследовательский объект. Ее ценность обусловлена многомерной репрезентацией не только языковых явлений, но и философских концептов, социально-исторических трансформаций и этнокультурных параметров эволюции древневосточной цивилизационной модели. Идеограмма – это иероглифический знак, который образуется через соединение только смысловых компонентов без включения какой-либо фонетической части. Идеограммы китайской иероглифической письменности формируются тысячелетиями в виде знаковой системы с внутренними особыми принципами организации. Они продолжают привлекать внимание не только лингвистов, но и ученых, работающих в области искусственного интеллекта. Понимание механизмов порождения китайских идеограмм, особенности такой сложной семантической области, как абстрактные понятия, позволит научить не только изучающих китайский язык, но и программное обеспечение правильно интерпретировать и, при необходимости, породить китайские знаки. Целью данной статьи, соответственно, является разработка структурно-семантических моделей соединения китайских иероглифов-идеограмм, которые обозначают абстрактные понятия различного типа.

Материалом исследования послужили 82 идеограммы, именующие абстрактные понятия, отобранные путем сплошной выборки из наиболее объемных и авторитетных словарей китайского языка: «汉语大字典» (Большой словарь китайских иероглифов) и «说文解字» (Объяснение простых и толкование сложных знаков) [1; 2]. В результате анализа отобранных идеограмм нами выделены три семантические группы знаков, которые создаются в китайском языке по типологии идеограмм. В каждой семантической группе идеограмм нами выявлены структурно-семантические модели их образования, т. е. те смысловые компоненты, которые входят в их состав, и как они комбинируются для создания того или иного знака.

Китайские идеограммы, обозначающие различные абстрактные понятия, представлены в виде трех семантических категорий знаков: 1) идеограммы, связанные с государством и обществом, составившие 73,1 % от общего числа отобранных единиц; 2) идеограммы, связанные с религией, составившие 15,9 %; 3) идеограммы, выражающие различные эмоции и мышление (11,0 %).

Наибольшую долю среди идеограмм, обозначающих абстрактные понятия, составили идеограммы, связанные с государством и обществом (73,1 %). Данные идеограммы создаются в китайской письменности по 13 структурно-семантическим моделям, количественная представленность которых в группе указана в таблице 1.

Т а б л и ц а 1

Количественная представленность структурно-семантических моделей идеограмм, связанных с государством и обществом

Структурно-семантическая модель	Количество, %
«человек /часть тела» + «государство и общество»	20,0 %
«производства и бытия» + «производства и бытия»	20,0 %
«человек /часть тела» + «предмет природы»	11,7 %
«человек /часть тела» + «производства и бытия»	10,0 %
«человек /часть тела» + «Человек /часть тела»	8,3 %
«человек /часть тела» + «абстрактная сущность»	6,7 %
«государство и общество» + «предмет природы»	6,7 %
«производства и бытия» +«зоонимы»	5,0 %
«человек /часть тела» +«зоонимы»	3,3 %
«производства и бытия»+«зоонимы»	3,3 %
«производства и бытия»+«абстрактная сущность»	1,7 %
«государство и общество» +«государство и общество»	1,7 %
«абстрактная сущность»+«абстрактная сущность»	1,7 %

Наиболее распространенная структурно-семантическая модель позволяет создавать идеограммы-абстракции государства и общества путем объединения, с одной стороны, семантических компонентов, обозначающих человека или его часть, с другой стороны, семантических компонентов со значением понятий, связанных с государством и обществом. Например, идеограмма 侯 ‘Хоу (второй из пяти рангов шанской и чжоуской аристократии)’ складывается из смысловых компонентов 亻 ‘человек’, 夂 ‘человек на высоте’ и 矢 ‘стрела’, так как ранг Хоу – высокий уровень аристократии, аристократы этого ранга владели территорией вдоль границы, которую обязаны были оборонять [2; 3].

Высокочастотна структурно-семантическая модель, заключающаяся в редупликация смысловых компонентов, выражающих объекты для производства и бытия. Например, идеограмма 則 ‘закон, принцип’ имеет первоначальную форму написания 𠄎, включая такие семантические конститuentы, как 鼎 ‘емкость-треножник или четырехножник’ и 刃 ‘нож’, поскольку древние китайцы вырезали тексты законов для утверждения [3; 5].

Третья модель позволяет создавать идеограммы путем объединения, с одной стороны, семантических компонентов, обозначающих человека или часть его тела, с другой стороны, семантического компонента со значением природных предметов и явлений. Например, идеограмма 姓 ‘фамилия

(по материнской линии) создана из смысловых компонентов 女 ‘женщина’ и 生 ‘появление жизни’, так как первые китайские фамилии произошли от древнего матрилинейного общества.

Четвертая структурно-семантическая модель составляется из семантических компонентов, выражающих понятия человека или части его тела и семантического компонента, обозначающего предмет производства или бытия. Например, 守 ‘должность чиновника’ – идеограмма, в состав которой включаются смысловые компоненты 寸 ‘точка биения пульса, цунь (мера длины)’ и 宀 ‘крыша здания’. В понимании древних китайцев компонент 寸 ‘точка биения пульса, цунь (мера длины)’ представляет собой репрезентант инструмента для выражения понятий, связанных с правилами или порядками.

Пятая структурно-семантическая модель, по которой создаются идеограммы, связанные с государством и обществом, это редупликация смысловых компонентов, выражающих человека и части его тела. Например, 位 ‘позиция во властях’ – идеограмма структуре которой содержатся следующие смысловые компоненты: 亻 ‘человек’, 立 ‘смотрящий человек, видеть’. Сочетание данных компонентов дало древним китайцам понятия «позиция во властях», так как письменный знак 亻 ‘человек’ первоначально мог выражать социальное понятие «гражданин», т. е. человек с политическими правами [2; 4].

Шестая выделенная нами структурно-семантическая модель построения идеограмм представляет собой объединение семантического компонента со значением человека и части его тела и семантического компонента, выражающего количество. Например, идеограмма 丈 ‘чжан (китайская сажень, равна 3,33 метра)’ создана из смысловых компонентов 又 ‘правая рука’ и 十 ‘десять’, так как по древним правилам Китая единица «чжан» – это длина 10 рук [4, с. 165].

Седьмая разработанная нами структурно-семантическая модель позволяет образовать идеограмму с помощью таких смысловых компонентов, как «обозначение государства и общества» и «обозначение природного предмета». Например, 域 ‘государственная территория, пределы страны’ составлена из смысловых компонентов 或 ‘страна, государство’ и 土 ‘земля’.

Наименьшее распространение имеют следующие модели идеограмм, обозначающих абстрактные понятия, связанные с государственными делами и общественно-социальной жизнью: 8) «производство и бытие» + «зооним», например, 寶 ‘драгоценность’ = 宀 ‘крыша здания’ + 王(玉) ‘нефрит’ + 缶 ‘амфора’ + 貝 ‘ракушка’; 9) «человек/часть тела» + «зооним», например, 夔 ‘правила, совещание’ = 又 ‘правая рука’ + 萑 ‘сова’; 10) «производство и бытие» + «предмет природы», например, 災 ‘катастрофа’ = 火 ‘огонь’ + 巛 ‘река’; 11) «производство и бытие» + «абстрактная сущность» например, 罪 ‘преступление, нарушение закона’ = 罒 ‘сетка’ + 非 ‘вопреки, ошибка’;

12) «государство и общество» + «государство и общество» например, 族 ‘род, клан’ = 旂 ‘флаг’ + 矢 ‘стрела’; «абстрактная сущность» + «абстрактная сущность» например, 章 ‘часть музыкального произведения’ = 音 ‘музыка’ + 十 ‘десять’.

Не менее широкую по объему группу составили идеограммы, которые обозначают абстрактные понятия, связанные с религией (15,9 % от общего числа идеограмм-абстрактных понятий). Идеограммы этой семантической категории знаков созданы по 7 семантическим моделям, количественная представленность которых указана в таблице 2.

Т а б л и ц а 2

Количественная представленность структурно-семантических моделей идеограмм со значением абстрактных понятий, связанных с религией

Структурно-семантическая модель	Количество, %
«государство и общество» + «государство и общество»	23,1 %
«государство и обществом + «предмет природы»	23,1 %
«человек / часть тела» + «Человек / часть тела»	15,4 %
«человек / часть тела» + «предмет природы»	15,4 %
«человек / часть тела» + «производства и бытия»	7,7 %
«государство и общество» + «производства и бытия»	7,7 %
«предмет природы» + «предмет природы»	7,7 %

Наиболее распространенная модель позволяет образовать идеограмму, связанную с религией, путем редупликации семантических компонентов, выражающих понятия, связанные с государством и обществом. Например, 禮 ‘ритуал, этикет для уважения божеств и предков’ представляет собой идеограмму, куда включаются смысловые компоненты 礻 ‘алтарь’ и 豊 ‘посуда для жертвоприношения’. Письменный знак 礻 ‘алтарь’ в понимании древних китайцев может представлять собой инструмент для обозначения понятий, связанных с верованием и ритуалом [3; 4].

Высокочастотны также объединения, с одной стороны, семантических компонентов, обозначающих государство и общество, с другой стороны, семантического компонента со значением природных предметов и появлений. Например, идеограмма 社 ‘божество земли’ включает такие семантические конститuenty, как 礻 ‘алтарь’ и 土 ‘земля’.

Третья модель позволяет создавать идеограммы путем редупликации семантических компонентов, выражающих понятия, связанные с человеком и

частью его тела. Например, 禘 ‘молитвенная церемония с участием жреца или представителя божества’ – идеограмма, в составе которой содержатся смысловые компоненты 巫 ‘жрец’, 口 ‘два рта, гомон’, и 手 ‘пара рук’. Компонент 巫 ‘жрец’ представляет собой репрезентант субъекта производимого действия, а компоненты 口 ‘два рта, гомон’ и 手 ‘пара рук’ – инструменты.

Четвертая структурно-семантическая модель составляется из семантических компонентов, выражающих понятие человека или части его тела, и семантического компонента предмета природы. Например, идеограмма 仙 ‘небожитель, бессмертный’ создана из смысловых компонентов 人 ‘человек’ и 山 ‘гора’, так как по китайской древней мифологии небожитель и бессмертный человек живут на высокой горе.

Наименьшее распространение имеют следующие модели идеограмм, обозначающих религиозные понятия: 5) «человек / часть тела» + «производство и бытие», например, 魁 ‘оборотень’ = 鬼 ‘дух’ + 多 ‘рисунок’; 6) «государство и общество» + «производство и бытие», например, 福 ‘счастье, везение’ = 礻 ‘алтарь’ + 畀 ‘наполненная емкость для урожая или вина’; 7) «предмет природы» + «предмет природы», например, 坤 ‘энергия земли’ = 土 ‘земля’ + 申 ‘гром и молния’.

Наименьшую долю составили идеограммы, обозначающие различные эмоции и мышление человека (11,0 % от общего числа идеограмм, именуемых абстрактные понятия). Идеограммы этого семантического типа создаются в китайской письменности согласно 6 структурно-семантическим моделям. Количественное распределение единиц, которые могут быть отнесены к той или иной модели, представлено в таблице 3.

Т а б л и ц а 3

Количественная представленность структурно-семантических моделей, по которым образуются идеограммы, обозначающие различные эмоции и мышление человека в китайском языке

Структурно-семантическая модель	Количество, %
«Человек / часть тела» + «человек / часть тела»	22,2 %
«Человек / часть тела» + «абстрактная сущность»	22,2 %
«Человек / часть тела» + «предмет природы»	22,2 %
«Производства и бытия» + «предмет природы»	11,1 %
«Предмет природы» + «предмет природы»	11,1 %
«Человек / часть тела» + «государство и общество»	11,1 %

Структурно-семантическая модель, по которой чаще образуются знаки подобного типа, представляет собой редупликацию семантических компонентов, выражающих значение человека или его частей. Например, идеограмма 色 ‘настроение; выражение лица’ образована в результате соединения смысловых компонентов 亻 ‘человек’ и 巴 ‘человек на коленях’.

Высокочастнота структурно-семантическая модель, представляющая собой объединение, с одной стороны, семантических компонентов, обозначающих человека или часть его тела, с другой стороны, семантического компонента со значением абстрактных понятий. Например, идеограмма 意 ‘мысль, воля’ включает такие семантические конститuentы, как 心 ‘сердце’ и 音 ‘голос, музыка’, т. к. в понимании древних китайцев сердце — это место, где душа находится, является главным органом ментальной деятельности [2; 4].

Третья структурно-семантическая модель составляется из семантических компонентов, выражающих человека или части его тела и семантического компонента – предмета природы. Например, 厄 ‘переживание, мучение’ – идеограмма, в чей состав включаются смысловые компоненты 卩 ‘сидящий человек’ и 厂 ‘пещера на обрыве’.

Наименьшее распространение имеют следующие модели идеограмм, обозначающих абстрактные понятия, связанные с эмоциями и мышлением: 4) «производство и бытие» + «предмет природы», например, 夢 ‘сон’ = 苜 ‘подножный корм’ + 冫 ‘покрытие’ + 夕 ‘луна вечером’; 5) «предмет природы» + «предмет природы», например, 昊 ‘великое небо, небосвод’ = 日 ‘солнце’ + 天 ‘небо’; 6) «человек / часть тела» + «государство и общество», например, 詞 ‘слово, речь’ = 言 ‘рот говорящий’ + 司 ‘управление’.

Таким образом, анализ структурно-семантических моделей образования идеограмм, обозначающих абстрактные понятия различного типа в китайской письменности, показал, что графическое составление данных знаков обусловлено, прежде всего, включением в состав идеограмм семантических компонентов, которые выражают конкретные сущности, к которым логически могут быть отнесены описываемые идеограммами абстрактные понятия. Данными категориями предстают обозначения лиц или организма человека (亻 ‘человек’, 心 ‘сердце’, 女 ‘женщина’, 巴 ‘человек на коленях’, 口 ‘рот’ и др.) и объектов государственных и общественных дел (礻 (示) ‘алтарь’, 矢 ‘стрела’, 豐 ‘посуда для жертвоприношения’, 旂 ‘флаг’, 戈 ‘оружие’ и др.). Графемы, репрезентирующие эти категории модели мира, дополняются смысловыми компонентами, обозначающими важные для носителей китайского языка признаки, исторически связанные с целым рядом национально-культурных особенностей, широкая вариативность которых приводит к сосуществованию большого разнообразия знакообразовательных

моделей конструирования китайских идеограмм, описанных в данной работе. Среди дополнительных признаков наиболее часто используются смысловые компоненты, обозначающие объекты производства и бытия, объекты и явления живой природы.

ЛИТЕРАТУРА

1. 许慎. 说文解字 / 许慎, 张三夕, 刘果。 – 岳麓 : 岳麓书社, 2006. – 540页。 = Сюй, Шэнь. Объяснение простых и толкование сложных знаков (Шовэньцзецзы) / Сюй Шэнь, ЧжанСаньси, ЛюГо. Юэлу : Изд-во в Юэлу, 2006. 537 с.
2. 汉语大字典 / 汉语大字典编委会: 徐中舒、李格非、赵振铎等, 学术顾问组: 于省吾、王力等。 – 武汉 : 崇文书局, 2010. – 5727 页。 = Большой словарь китайских иероглифов / редкол. : СюйЧжоншу, Ли Гэфэй, ЧжаоЧжэньдо [и др.]. Ухань : Чунвэнь, 2010. 5727 с.
3. 黄天树. 说文解字通论。 – 北京: 北京大学出版社, 2014. – 361页。 = Хуан, Тяньшу. Общее суждение «Шовэньцзецзы» / Хуан Тяньшу. Пекин : Изд-во Пекинского ун-та, 2014. 361 с.
4. 李学勤. 字源。 – 天津: 天津古籍出版社, 2013. – 1420页。 = Ли, Сюэцин. Этимология иероглифов / Ли Сюэцин. Тяньцзинь : Старинные книги Тяньцзиня, 2013. 1420 с.
5. 谷衍奎. 汉字源流字典。 – 北京 : 语文出版社, 2008. – 2155页。 = Гу, Янькуй. Словарь этимологии китайских иероглифов / ГуЯнькуй. Пекин : Изд-во китайского языка и литературы, 2008. 2155 с.