

КРИТИКА И БИБЛИОГРАФИЯ

ОБЗОРЫ

В. С. ИВАНОВ

О НЕКОТОРЫХ НАПРАВЛЕНИЯХ В РАЗВИТИИ КОЛИЧЕСТВЕННЫХ МЕТОДОВ В ВЕНГЕРСКОМ ЯЗЫКОЗНАНИИ

Количественными методами изучения языков называется совокупность математических, статистических приемов, с помощью которых качественное описание языка дополняется количественными данными.

Хотя в настоящее время правомерность применения количественных методов в языкознании и не подвергается сомнению, следует сказать, что прежде на этот счет имелись различные точки зрения. Были исследователи, которые считали, что математические, статистические, количественные методы вряд ли принесут какую-либо пользу лингвистике (некоторые представители традиционной лингвистики) и были ученые, считавшие эти методы всемогущими (Г. Хердан, Дж. Уотмоу и А. Марков).

Однако большинство современных лингвистов рассматривает количественные методы как необходимость, помогающую дополнить качественное описание языка.

Развитию количественных методов способствовали прежде всего особенности структуры языка. Ведь любой язык состоит из единиц, легко поддающихся счету. Это звуки, буквы, слова, слоги, предложения и многие другие единицы. Для исследователя, описывающего язык, важно знать не только, как образуется то или иное явление, но и как часто оно употребляется. Специалистов не удовлетворяют общие замечания, что такое-то языковое явление встречается «часто», или «довольно часто», а такое-то «редко» или «довольно редко». Задача состоит в том, чтобы получить возможно более точные данные. Применению количественных методов в языкознании способствовало также развитие точных наук.

Использование количественных методов в языкознании объясняется и целым рядом других причин. Сюда относится необходимость определения стиля того или иного писателя, выяснения закономерностей стиха, установление авторства, дешифровка древних памятников, расшифровка секретных документов, перевод текста с одного языка на другой с помощью машин, решение серии задач по теории информации и др. Применение количественных методов диктуется также соображениями методики преподавания иностранных языков.

Венгерский язык с точки зрения применения количественных методов представляет особый интерес. Как язык, имеющий ярко выраженный агглютинативный характер, он обладает великим множеством исчисляемых единиц. С другой стороны, структура венгерского слова предельно ясна. Это корень, к которому в постпозиции, а иногда и в препозиции присоеди-

няются в определенном порядке различные форманты. Можно сказать, что структура венгерского слова сама наталкивает на мысль о применении количественных методов для изучения языка. Следует упомянуть и о том, что работа венгерских ученых в этой области в нашей научной печати почти не освещена. Между тем в исследованиях венгерских ученых немало интересного, важного и для изучения других языков.

Исходным пунктом развития количественных методов в венгерском языкознании следует считать 1896 г., когда известный венгерский языковед Ж. Шимони, занимаясь вопросом наиболее целесообразного расположения букв на клавиатуре венгерской пишущей машинки, установил, что какие-либо статистические сведения о связи букв и звуков в венгерском языке отсутствуют. Тем самым он обратил внимание языковедов на необходимость исследований именно в этой области.

За прошедшие с момента высказывания Ж. Шимони семьдесят с лишним лет венгерское языкознание добилося немалых успехов в области количественного изучения венгерского языка.

Количественные методы помогли выяснить вопрос о величине словарного запаса венгерского языка. Эту проблему, в частности, обстоятельно исследовал И. Келемен¹. В своей статье «Объем словарного запаса венгерского языка» он отмечает, что четырнадцатилетний ребенок знает 6000—7000 слов. Среди взрослых попадаются такие образованные люди, словарь которых насчитывает 50 000 слов. Из венгерских классиков наиболее богатым словарем обладал Я. Арань (23 000 слов). Далее И. Келемен пишет, что собрать воедино все слова венгерского языка чрезвычайно трудно. Правда, в 1868—1873 гг. М. Баллаги выпустил так называемый «Полный словарь венгерского языка»². В этом словаре 83 260 слов, но среди них много таких, значение которых в наше время никому неизвестно; многие из этих слов не употреблялись и во времена Баллаги. В 1899 г. на основе словаря Цуцора и Фогараши и материала шести других словарей Венгерская Академия наук составила и выпустила в свет «Перечень слов», в котором насчитывается 122 067 самостоятельных единиц. Наиболее ценным с точки зрения подхода к слову И. Келемен считает «Толковый словарь венгерского языка», изданный Институтом языкознания АН ВНР. В словаре собрано более 100 000 слов. На основании сравнения всех имеющихся в его распоряжении словарных источников с «Перечнем слов» 1899 г. И. Келемен составил ряд таблиц, убедительно показывающих расширение словарного запаса венгерского словаря.

Венгерские ученые стремились также определить наличие корневых слов в венгерском языке. В. Тольнай в статье «Бессмертный венгерский язык»³ отмечает, что в языке XIX в. насчитывалось 55% корневых слов, 65% лексикологических единств и 88,4% повседневных слов венгерского происхождения. Интересовался В. Тольнай и численностью иностранных слов. По его расчетам в языке XX в. на 1000 корневых слов приходится 550 венгерских, 100 славянских, 95 германских, 90 тюркских, 80 латинских и 85 слов неизвестного происхождения. По расчетам на основании словников получились несколько иные данные: на 1000 слов приходится 650 венгерских, 100 славянских, 85 тюркских, 75 германских, 65 латинских, 25 слов из других языков и неизвестного происхождения. Однако данные словников ни в коей мере не учитывают частоту употребления слов. Поэтому В. Тольнай провел дополнительные расчеты, подвергнув с учетом употребительности анализу живую речь и язык прессы. У него полу-

¹ K e l e m e n J., A magyar szókincs nagysága, «Iskolai nyelvújvíelő», Budapest, 1959.

² B a l l a g i M., A magyar nyelv teljes szótára, I—II, Budapest, 1868—1873.

³ T o l n a i V., Halhatatlan magyar nyelv, «Magyar nyelv», 2, 1896.

чился следующий результат в процентах: 88% венгерских слов, 3% славянских, 3% латинских, 2% германских, 1% тюркских, 3% слов из других языков и неизвестного происхождения. В «Системе современного венгерского языка», изданной в 1961 г. в Будапеште, подсчитано количество слов в современном венгерском языке. Автор раздела о словарном запасе, Д. Сепе, указывает, что в зависимости от принципов подсчета число слов в венгерском языке колеблется от 800 000 до 1 065 000⁴.

Довоенные статистические работы о буквах, звуках и словах связаны с исследованиями венгерских стенографов, и прежде всего З. Немеша. Секретарь венгерского общества стенографистов З. Немеш в 1933 г. опубликовал в Сегаде брошюру «Наиболее употребительные слова венгерского парламентского языка»⁵. Эта работа заслуживает более подробного изложения. Известно, что работа стенографистки теснейшим образом связана с закономерностями языка. Когда стенографистка следит за мыслью и набрасывает на бумагу услышанное, она стремится максимально использовать особенности и правила данного языка, специфику его фонетики, словообразования, языковые аналогии, конструктивную связь суффиксов, показателей и окончаний, взаимозависимость единственного и множественного числа, сокращения и условные обозначения постоянно употребляемых выражений. Поэтому для стенографии особо важное значение имеет частота употребления слов, суффиксов, показателей, окончаний, слогов и других языковых элементов. Впервые в истории венгерского языкознания З. Немеш на основе анализа текста в 100 000 слов установил статистику суффиксов, показателей, окончаний и выделил наиболее употребляемые слова в парламентском языке.

Автор «Наиболее употребительных слов венгерского парламентского языка» изложил результаты исследования в трех таблицах, в которые он включил слова, попадавшие в тексте по крайней мере двадцать пять раз. В первой таблице, насчитывающей всего 504 единицы, слова располагаются по алфавиту; рядом с каждым словом помещается цифра, отражающая частотность употребления данного слова в тексте.

Вторая таблица занимается частотой употребления слов в несколько ином плане. Она выясняет, сколько раз слово употребляется самостоятельно, и сколько раз в составе производных слов. Эта таблица имеет также графу специальных дополнений и замечаний. Третья таблица посвящена словам и формантам, часто встречающимся в составе сложных слов.

Помимо упомянутой брошюры, З. Немеш опубликовал также «Статистику слова»⁶.

Уже упомянутая «Система современного венгерского языка» дает конкретные цифры числа слогов в венгерском слове. Примерно 22,5% всех слов состоит из одного слога, 60% — из двух, 15% — из трех, 2,3% — из четырех, 0,2% — из пяти и более слогов. Несколько иная картина получается при анализе связного текста, в котором слова берутся вместе с показателями и окончаниями. Здесь уже насчитывается 37% слов из одного слога, 33,8% — из двух слогов, 18,7% — из трех слогов, 8,5% — из четырех, 2% — из пяти и более слогов. Число односложных слов в связном тексте сильно увеличивается за счет часто встречающихся артиклей, частиц, указательных местоимений и других небольших слов.

Первый слог в слове примерно в 30% начинается с гласного звука, в 70% — с согласного. Последний слог в слове примерно в 20% всех случаев заканчивается гласным звуком, в 65% — одним согласным,

⁴ «A mai magyar nyelv rendszere», I kötet, Budapest, 1961.

⁵ N e m e s Z., A magyar parlamenti nyelv leggyakoribb szavai, Szeged, 1933.

⁶ N e m e s Z., Szóstatisztika, Budapest, 1939.

в 10% — двумя разными согласными. Стечение гласных звуков, открывающее слово, — чрезвычайно редкое явление. Оно встречается лишь в заимствованиях, составляющих около 1% всех слов. Стечение двух согласных в середине слова, напротив, довольно частое явление, встречается в 30% всех слов. Стечение трех и более согласных звуков не характерно для венгерского языка и встречается лишь в 2% слов. Однако стечение двух согласных в конце слова является обычным сочетанием и употребляется в 10% всех слов.

В венгерском языкознании количественные методы применялись также и для определения длины слова. В 1960 г. была опубликована работа И. Фонади «Длина слова в венгерской речи»⁷.

И. Фонади продельывает интересный опыт. Он берет устный рассказ одного человека, записывает его, исправляет и редактирует. Затем сопоставляет число слогов в словах устного и письменного рассказов. Число слогов И. Фонади подсчитывает в процентах. В результате получается таблица, наглядно показывающая разницу между письменной и устной речью (см. табл. 1).

В устной речи вместо длинных слов стараются употреблять их короткие варианты. Так, вместо слова *filmszínház* «кинотеатр» стали употреблять новое слово *mozi*, пущенное в ход писателем Й. Хелтай.

Венгерское языкознание имеет богатые традиции в области изучения количественных характеристик букв и звуков. Впервые попытался определить количественные характеристики венгерских букв Ж. Шимони, опубликовав в 1908 г. статью «Числовое соотношение букв в венгерском языке»⁸.

Оперируя скорее грамматическими, чем статистическими выкладками, Ж. Шимони устанавливает следующий порядок употребления согласных букв в конце слова: *k, t, l, n, m, s*. В соответствии с частотой употребления, гласные буквы также имеют свою последовательность: *a, e, i*. В начале слова буквы по степени встречаемости располагаются в следующем порядке: *k, f, m, sz, t, b, h, v, d, j, o, z(zs), u, ü*. Вторая и предпоследняя буквы слова в большинстве своем — гласные, а первая и последняя — согласные. Соотношение гласных и согласных звуков в венгерском языке составляет 43 : 57.

Две работы о буквах и звуках написал венгерский лингвист В. Тольнай. Первая из них, «Некоторые числовые данные о звуках и буквах», была опубликована в 1906 г.⁹ Автор статьи сопоставляет статистические данные о звуках и буквах в венгерском языке на основе анализа языка писателей П. Дюлаи, М. Йокаи, Ж. Бёти, Б. Тот и Ф. Херцег, поэтов Ш. Петефи, Верешмарти, Я. Араня, политических газетных статей и некоторых других сообщений. В. Тольнай берет в каждом случае по 2500 звуков и подсчитывает среднюю величину повторяемости каждого звука. Он устанавливает определенный порядок следования звуков в зависимости от частоты их употребления, а также сочетаемость гласных (*e + é, a + á, o + ó*). Вторая статья В. Тольнай, «О красоте языков», появилась в 1921 г.¹⁰ На этот раз в центре внимания автора — музыкальные элементы языка и частота их употребления. В. Тольнай приводит данные по употреблению гласных и согласных звуков по 14 языкам. Он стремится уточнить, какие согласные и их сочетания употребляются в каждом конкретном языке вообще и какие — только в начале или кон-

⁷ F o n a g y J., A szavak hossza a magyar beszédben, «Magyar nyelvőr», 3, 1960.

⁸ S i m o n y i Zs., A betűk számaránya a magyarban, «Magyar nyelvőr», 2, 1908.

⁹ T o l n a i V., Egy néhány számadat a hangokról es betűkről, «Magyar nyelvőr», 4, 1906.

¹⁰ T o l n a i V., A nyelvek szépségéről, «Magyar nyelv», 1, 1921.

це слова. В. Тольнай считает, что в начале венгерских слов нет стечений согласных, два согласных в начале характерны лишь для заимствованных слов.

Статистикой гласных звуков подробно интересовался Т. Гарноци, опубликовавший две работы. Первая из них, «Распределение венгерских гласных», вышла в свет в 1942 г.¹¹, вторая, «Замечания и дополнения к вопросу о распределении гласных в венгерском языке», — в 1943 г.¹²

В 1953 г. большое статистическое исследование о фонетическом составе венгерского языка опубликовала Э. Вертеш¹³. Сопоставляя данные Т. Гарноци о языке венгерского поэта Э. Ади с количественными данными о языке современного венгерского писателя П. Вереша, Э. Вертеш отделяет звуковые особенности языка от особенностей стиля писателей, и приходит к выводу о том, что выяснить частоту употребления отдельных звуков речи было бы затруднительным делом. Можно лишь установить частоту употребления отдельных звуковых сочетаний. Э. Вертеш изучила возможности сочетаний гласных и согласных звуков в языке Э. Ади и П. Вереша и установила, что выявить зафиксированные сочетания можно лишь с учетом стиля писателя. В своей работе Э. Вертеш рассматривает как с точки зрения качества, так и с точки зрения количества особенности слогов в венгерском слове. Э. Вертеш занималась также вопросом о том, насколько те или иные статистические данные применимы к заимствованным словам. Она установила, что имеется определенная разница между употреблением сочетаний согласных звуков в венгерских и заимствованных словах. Стечения согласных звуков в начале слова наблюдаются в славянских и германских заимствованиях, стечения трех согласных в середине слова — в германских заимствованиях. Стечения гласных звуков в середине слова или в конце употребляются лишь в латинских, неолатинских или в псевдолатинских заимствованиях. Происхождение слов накладывает отпечаток и на количество слогов.

Подсчетом удельного веса отдельных гласных и согласных звуков в венгерском языке занимался Л. Деме в статье «Звуковой облик нашего языка», появившейся в 1966 г.¹⁴. Автор статьи выводит ряд закономерностей частоты употребления звуков. Так, например, звук *e*, по данным Л. Деме, составляет 10,6% всех звуков языка и 25% всех гласных звуков. Не менее широко распространен звук *a*, который составляет 9,6% всех звуков и 23,6% всех гласных звуков. Гласные составляют 41,2% всех звуков, а согласные — 58,3%. Соотношение гласных переднего и заднего ряда почти одинаково. Частота употребления гласных заднего ряда лишь на 2% отстает от частоты употребления гласных переднего ряда. Согласных звонких звуков в венгерском языке на 12% больше, чем незвонких.

Языковая статистика играет немалую роль при исследовании диалектов. Определенный интерес в этом отношении представляет очерк В. Фаркаша «Статистические проблемы фонем в истории типов диалектов», защищенный в качестве диссертации на соискание ученой степени кандидата филологических наук¹⁵. С целью дать возможно более точную качественную характеристику фонем в диалектах В. Фаркаш привлекает для исследования венгерские средневековые кодексы и анализирует их.

¹¹ T a r n ó c z y T., A magyar magánhangzók eloszlása, «Magyar nyelv», 4, 1942.

¹² T a r n ó c z y T., Megjegyzések és kiegészítések, a magyar magánhangzók eloszlásának kérdéséhez, «Magyar nyelv», 4, 1943.

¹³ Vertes E., Statistische Untersuchungen über den phonetischen Aufbau der ungarischen Sprache, «Acta ling. hung.», 3—4, 1953—1954.

¹⁴ D e m e L., Nyelvünk hangtani arculata, «Magyar nyelvhelyesség», Budapest, 1966.

¹⁵ F a r k a s V., Fonémastatisztikai problémák a nyelvjárástipustörténetben, «Nyelvtudományi értekezések», 55, 1966.

Все кодексы, учитывая употребление в них вариантов *ő*, делятся автором на четыре группы. В. Фаркаш постоянно оперирует цифрами, доказывая степень употребляемости той или иной фонемы в кодексе.

Вопросами применения количественных методов в изучении диалектов занимался также Л. Папп в книге «История диалекта и языковая статистика»¹⁶, получившей на конкурсе Венгерской Академии наук в 1961 г. первую премию.

Л. Папп считает, что при исследовании фонетических явлений следует обратить внимание на два вида статистики: на установление абсолютного числа примеров и контрпримеров и на абсолютное число частоты употребления анализируемого явления. Исследуемые явления следует сопоставлять даже и тогда, если они выражаются неодинаковыми единицами (морфемами) и встречаются в неодинаковом числе. Руководствуясь этим положением, Л. Папп на основе 400 текстов XVI в., содержащих 230 000 единиц, устанавливает соотношение *i*-кающего диалекта с языковой нормой. «Иканья» в венгерском языке представляет собой такое явление, при котором в некоторых венгерских диалектах вместо звука *é* употребляется *i*. Это явление в венгерском языкознании объясняется тем, что некогда в венгерском языке были два звука, т. е. *é* и *ê*, и появление *i* связано с историей этих звуков. Однако монографии о диалектах излагают этот вопрос суммарно, хотя они и отмечают разную степень *i*-канья в разных диалектах. Следует сказать, что венгерский классик Я. Арань в свое время указывал на необходимость систематического описания иканья. И. Сабо написал очерк, в котором дал полную картину иканья одного поселения. Л. Папп дал полное описание *i*-канья и *é*-канья в хоссупайском диалекте, пытаясь сопоставить языковые элементы *i* и *é/ê*. Помимо всего прочего, Л. Паппа эта проблема интересовала в плане истории звука. Еще до знакомства с исследованиями по этому вопросу Д. Пайжа, он весьма сомневался, что переход *é* в *i* может служить единственным объяснением *i*-канья. В своей работе Л. Папп с помощью количественных методов поднимает и исследует ряд проблем: анализирует *i*-кающие элементы архивов XVI в., выясняет частоту их употребления, статистику числа слогов и знаков, географию и социологию *i*-канья, а также обобщает свой метод. Л. Папп придерживается мнения, что определенных результатов исследования географических и общественных вопросов можно было бы добиться и не применяя количественных методов, без статистической обработки материала. Однако степень *i*-канья в отдельных письменных памятниках и у отдельных лиц можно определить лишь с помощью цифр. Без подсчетов и статистических сведений можно составить ложную картину на основе субъективных впечатлений.

Венгерские языковеды, учитывая опыт других ученых, пытались установить степень родства между финно-угорскими языками, оперируя определенными расчетами. Опираясь на идеи М. Сводеша, венгерский ученый И. Фодор вывел математическую формулу, устанавливающую примерное время расхождения родственных финно-угорских языков¹⁷.

Количественные методы сыграли большую роль в изучении венгерского языка. Именно с их помощью были установлены числовые характеристики звуков, букв, слогов, предложений и других исчисляемых языковых элементов. Был определен объем словарного запаса венгерского языка и величина словаря отдельных диалектов. Языковая статистика помогла найти правильное решение ряда вопросов истории диалектов

¹⁶ P a p p L., Nyelvjárástörténet és nyelvi statisztika, Budapest, 1963.

¹⁷ F o d o r J., A statisztikai módszer alkalmazásának néhány kérdése, «Magyar nyelvőr», 2, 1960.

Таблица 1

Число слогов в слове	1	2	3	4	5
В устной речи	53,82	32,18	8,72	3,50	1,76
На письме	36,36	29,54	23,86	8,00	2,27

венгерского языка. Было определено число слогов в словах устной и письменной речи, исследована частота структур предложений, обоснована статистика формантов. Проведены первые опыты по установлению примерной даты расхождения родственных финно-угорских языков. Появились исследования о длине слова. Возникли методы определения стиля того или иного писателя.

Основная задача, которая стоит в наши дни перед количественными методами, состоит в точном описании количественного облика венгерского языка.