

## **К ПРОБЛЕМАТИКЕ ЧАСТОТНОГО МОРФЕМИАРИЯ БОЛГАРСКОГО ЯЗЫКА**

В современной лингвистике принято считать морфему элементарной нагруженной смыслом частицей языка, нерасчленимым носителем смысла. Полагают также, что многообразие морфем в языке невелико (мощность морфемного многообразия современного болгарского языка, если считать нерасчленимыми сегментами традиционно выделяемые корневые морфемы, аффиксы, окончания, едва ли превысит число 3000), что новые морфемы в истории языка либо вообще не возникают, либо возникают так редко, что трудно проиллюстрировать бесспорным примером это явление.

Следовало бы ожидать, что лингвисты должны заняться прежде всего составлением исчерпывающих перечней морфем для каждого языка вместо заведомо обреченного на незавершенность в рамках реальных сроков (даже при компьютерной обработке материала) труда по составлению полных словарей, исчерпывающих многообразие слов языка, особенно если автор стремится дать разностороннее описание каждой лексемы. Так, к примеру, И. Мельчук и А. Жолковский в своем «Толково-комбинаторном словаре русского языка» ([Мельчук 1984]) смогли описать, по словам И. Мельчука ([Мельчук 1995: 10]), на более чем девяти сотнях страниц лишь четверть процента русской лексики!

Этот разлад между теорией и практикой лингвистического описания языков остается фактом и по сей день. Лингвисты продолжают работать над все более обстоятельным и разноаспектным описанием лексики языка в словарях разных типов, хотя и считают азбучной истинной ограниченность и, следовательно, в принципе легкость составления списка, исчерпывающего все многообразие морфем языка. Все еще нет справочников, описывающих морфемное многообразие современных европейских языков. Можно указать, разумеется, на такие исследования, как труд З. Ф. Оливериуса «Морфемы русского языка» ([Оливериус 1976]) и «Словарь морфем русского языка» А. И. Кузнецовой и Т. Ф. Ефремовой ([Кузнецова 1986]), но они содержат сведения лишь о **части** морфем русского языка. Некоторые количественные оценки морфем приводятся лишь в словаре Оливериуса, но они в немалой степени представляют искаженную глоттометрическую картину употребления морфем в текстах, поскольку в исследование не включены наиболее часто появляющиеся

в языковой практике морфемы – морфемы окончаний; данные о частоте русских суффиксов и префиксов опубликованы А. И. Кузнецовой и О. А. Лавреновой ([Кузнецова 1975]). В словаре Кузнецовой и Ефремовой, содержащем богатейший материал об «окружениях» (контекстах) корневых морфем, также нет информации о «грамматических морфемах». Оба эти исследования (труд Оливеруса и словарь Кузнецовой и Ефремовой) по существу – улучшенная, обогащенная материалом разновидность «словообразовательных словарей», издающихся и в России, и в Болгарии, таких, как, например, широко известные словари З. А. Потихи ([Потиха 1981]), А. Н. Тихонова ([Тихонов 1978]) и последний изданный в Болгарии в 1999 г. словообразовательный словарь современного болгарского литературного языка под редакцией Й. Пенчева ([Пенчев 1999]).

**Частотные словари и частотные морфемиирии.** Глоттометрические характеристики включенных в морфемиирий морфем являются необходимой предпосылкой для реализации преимуществ, которые дает справочник типа морфемиирия лингвисту-теоретiku и лингвисту-методисту (специалисту по практическому обучению языку). Глоттометрические характеристики морфем дают более надежную информацию обо всей языковой материи исследуемого языка, чем данные, которые можно получить из частотных словарей.

Достаточно надежными статистическими оценками о встречаемости слов в языковой практике являются данные лишь о немногих словах, включенных в частотный словарь, и эти данные надежны только для того текстового массива, по которому делался словарь.

По мере того, как составители частотных словарей приходили к использованию все более и более строгого математико-статистического аппарата обработки и показа результатов исследования частоты слов по текстам, они все лучше понимали, что для лингвистики (и теоретической, и прикладной, в том числе и методической) важно выяснить принадлежность того или иного слова к определенному классу слов по признаку встречаемости, установить наиболее целесообразную с точки зрения теории и / или практики классификацию слов по частотному признаку. Но задача составления так называемого частотно-рангового списка слов оказалась не только практически неосуществимой, но и теоретически неверно поставленной – теоретически (с точки зрения математической статистики) невозможно за данным словом, будь оно даже очень часто встречающимся во всех стилях языка, закрепить строго определенный порядковый номер в частотно-ранговом списке, т. е. сказать,

например, что оно является определенно наиболее частым словом в болгарском языке или что оно является точно вторым, точно третьим, точно десятым и т. д. в частотно-ранговом списке.

Дело в том, что лингвистические феномены, называемые «словами», не поддаются строгому, с точки зрения логики, подсчету по двум причинам.

Во-первых, обычно частотные словари, информируя о частоте определенного «слова» (лексемы, лексической единицы), фактически сообщают о числе появлений определенной морфемы, являющейся лишь частью (основой) многообразия полиморфем (словоформ), каждая из которых содержит также, кроме основы, в принципе не учитываемые «грамматические» (флексионные) морфемы. Таким образом, если принять совокупность учтенных при составлении (лексемного) частотного словаря текстов за один текст, по сумме частот всех учтенных в нем лексических единиц в принципе невозможно получить информацию о всей языковой материи (обо всех морфемах) этого текста.

Составители лексемных частотных словарей либо полностью пренебрегают флексионными морфемами (т. е. из поля зрения лингвиста выпадают весьма весомые части массы текста), либо учитывают их выборочно, как «грамматические формы» (падежи, числа, лица, времена, наклонения), причем в принципе исчезает индивидуальность морфем-носителей грамматических значений в грамматических формах – сообщается, например, формы какого падежа встречаются чаще или реже, но не сообщается, каков удельный вес того или иного падежного окончания в числе появлений этого падежа.

Фактически лексемные частотные словари информируют не обо всей языковой материи, не обо всем ее многообразии, а о многообразии языковой материи некоего аморфного (аграмматического) трансформата, некоего «китаизированного» варианта описываемого языка («китаизированного» – в смысле традиционного мнения лингвистов не синологов о китайском языке как о типичном представителе не содержащих грамматических морфем «корневых» языков). А между тем удельный вес флексионных частей сообщения далеко не пренебрежимо мал.

Поэтому чем менее богато многообразие флективных морфем языка, тем менее заметной становится недостаточность лексемного частотного словаря. Поэтому, например, частотный словарь английского языка в меньшей мере искажает представление о количественных отношениях морфем в английском языке, чем лексемный словарь любого славянского языка, в том числе и болгарского (в болгарском

языке именная флексия относительно бедна, но многообразие глагольных флексий велико и не позволяет ограничиться составлением лексемного словаря).

Частотные словари, дающие информацию о частоте словоформ, на первый взгляд, как будто устраняют отмеченный выше недостаток лексемных частотных словарей. Но мы указали лишь на одну из двух причин, по которым «слова» как лингвистические феномены не поддаются строгому учету в частотных словарях. Вторая причина, по которой частотные словари в принципе не дают надежной информации о количественной стороне языковой материи, в частотных словарях словоформ не только не устраняется, но ее весомость даже усугубляется.

Естественно, что чем больше многообразие классов, по которым распределяются феномены, подсчитываемые при составлении справочника по количественной стороне языка, тем больше (и во много раз больше) должен быть текстовый массив (корпус), положенный в основу исследования. А многообразие словоформ значительно больше многообразия лексем.

Обычно считают, что многообразие лексем болгарского языка – порядка 200 тысяч (ср. [Георгиев 1965: 120]). А о многообразии словоформ не приводятся даже приблизительные оценки. Но легко себе представить, что оно должно быть порядка, по меньшей мере, десяти миллионов и даже более. Лексемы существительных, наиболее бедные словоформами, реализуются четырьмя-шестью словоформами – четыре для большинства феминальных и неутрумных субстантивов (здесь и далее в примерах подчеркиванием отмечены ударные гласные), ср.: *жена* – *жената* – *женѝ* – *женците*, *стадо* – *стадото* – *стада* – *стадата*; пять – для большинства маскулинарных субстантивов, если не считать разными словоформами неполные и полные артиклевые формы, а при их различении – шесть, ср.: *град* – *грѣдат* и *града* – *градове* – *градовете* – *града*). У адъективных лексем – восемь или девять словоформ (аллолекс), напр.: *нов* – *новият* и *новия* – *нова* – *новата* – *ново* – *новото* – *нови* – *новите*. Если же добавить словоформы сравнительной и превосходной степеней (ср.: *по-нов*, *най-нов* и т. д.), количество аллолекс большой группы адъективных лексем возрастет втрое. Что касается словоформ глагольных лексем, им поистине несть числа ([Пашов 1976: 186]).

Составители частотного словаря «по словоформам» сталкиваются, кроме того, с затруднением, которое в менее явном виде существует и для составителя частотного лексемного словаря. Понятие «словоформа» для многих языков (в том числе и для болгарского) не всегда

совпадает в лингвистической традиции с основной счетной единицей частотных словарей – «графическим словом», определяемым как последовательность букв между двумя пробелами, без какого-либо «внутреннего» пробела между буквами. В особенности это относится к формам глагола. Болгарская лингвистическая традиция расценивает (в соответствии с аналогичной традицией в грамматике классических и некоторых современных языков) выражения типа *писал съм* или *цях да пиша* как сложные формы (в грамматической традиции классических языков – «перифрастические формы») спряжения глагола, несмотря на то, что они состоят из двух, из трех, иногда даже из большего количества (*цях да съм писал*) графических слов. Поскольку в частотном словаре «частицы» типа *да* и вспомогательные глагольные формы типа *съм* и *цях* учитываются отдельно, возникает проблема, как быть со «сложными словоформами» вроде *цях да пиша*. Если включить целиком *цях да пиша* в многообразие словоформ «спрягаемого глагола» (термин «спрягаемый глагол» противоречит строению не одной сложной формы: в *цях да пиша* спрягаются, т. е. изменяются по лицу и числу, и вспомогательный, и «спрягаемый» глагол, а в *писал съм* «спрягаемый глагол» не изменяется по лицу, т. е. спрягается только вспомогательный глагол), то не следует учитывать входящие в «сложную форму» частицы и вспомогательные глаголы как отдельные слова (словоформы). Если же не включать сложные глагольные формы в парадигму «спрягаемого» глагола, то останутся неучтенными все так называемые сложные глагольные формы. И в том и другом случае остается открытым вопрос об определении объема исследуемого массива текста: следует ли отказаться от удобного определения слова как последовательности букв между двумя пробелами, или независимо от учета выражений типа *цях да пиша* и как аллолекс (словоформ) не вспомогательного глагола (в данном случае *пиша*) подсчет количества слов в тексте вести по пробелам.

Из всего вышесказанного можно сделать вывод, что при составлении справочника о количественной стороне феноменов языка можно и нужно идти не по пути увеличения многообразия исследуемых феноменов, а в сторону уменьшения этого многообразия, но учитывая при этом все типы частиц языковой материи, все морфы – корневые, инфиксальные, аффиксальные, грамматические (флексионные). При таком решении можно надеяться, что даже относительно небольшой текст (объемом около ста тысяч традиционных графических словоформ) даст пока совсем отсутствующую информацию о вероятностном поведении действительно элементарных частиц ма-

терии болгарского языка, морфем – в бодуэновском смысле этого термина.

Переход к морфемной основе количественного описания материи болгарского языка не означает, что из морфемиария вообще невозможно получить информацию о частотном поведении «слов» и «словоформ». Феномены болгарского языка, о частотном поведении которых будет содержаться информация в морфемиарии, выбираются не по признаку простоты или сложности их морфемной структуры, а по строго вероятностным критериям. Морфемиарий дает информацию о наиболее часто появляющихся в болгарской языковой практике выражениях (полиморфемах) независимо от степени сложности их морфемного состава.

Переход к морфемной основе означает лишь, что накопленный лексикографами опыт по выделению, например, тех словосочетаний (идиом, фразеологизмов и т. п.), о которых изучающий данный язык должен получить информацию – опыт, имплицитно опирающийся на встречаемость этих словосочетаний в языковой практике, т. е. на частоту не слов, составляющих словосочетание, а словосочетания в целом, – лишь в частотном морфемиарии может получить эксплицитно, явно математико-статистическую форму.

Однако составители морфемиариев сталкиваются с проблемами, о которых стоит рассказать подробнее. Трудности, связанные с решением этих проблем, объясняют отсутствие справочников типа морфемиариев, причину нежелания лингвистов заниматься их составлением. Проблемы, возникающие перед составителями, можно разделить на две группы:

1) проблемы, связанные с морфемным членением текста и с идентификацией морфем. По определению Б. де Куртенэ термином морфема обозначается кратчайшая единица языка, связываемая с каким-то смыслом ([Бодуэн 1963: 240, 290]). Оставаясь в рамках бодуэновского учения о морфеме, мы автоматически, добровольно принимаем, что не имеем права называть репрезентантом морфемы *a*) такой сегмент сообщения, который можно разделить на более мелкие части, и каждая из них при этом будет связана с частью смысла всего сегмента (в противном случае мы вошли бы в противоречие с постулатом элементарности, неразложимости бодуэновской морфемы); *б*) такой сегмент сообщения, о котором мы не в состоянии утверждать, что он является носителем смысла (в противном случае мы допустили бы противоречие с постулатом об обязательной смысловой нагрузке морфемы). Однако в практике морфемного анализа эти два условия лингвисты постоянно нарушают, объявляя элементарными сегменты

сообщения, которые они по традиции привыкли считать нерасчленимыми. Указанные две опасности становятся весьма весомым препятствием на пути лингвиста, как только он попытается исчерпывающе описать все морфемы, из которых состоит определенное, даже не очень большое (порядка нескольких десятков слов) сообщение.

2) проблемы, связанные со способами описания семантики морфем. Здесь очень трудно использовать опыт, накопленный многовековой практикой лексикографов. Ближе к нуждам морфемологии стоит практика грамматического описания, поскольку грамматическое описание по существу является описанием семантики наиболее часто встречающихся в языке морфем (схноморфем). К сожалению, традиция грамматического описания языка, как правило, не увязывает элементарные грамматические значения с их материальными носителями. Кроме того, очень часто грамматическое значение не описывается, а лишь называется определенным (не всегда удачным) завещанным традицией термином. Но проблема описания семантики морфем заслуживает обсуждения в отдельной статье. Предмет обсуждения в настоящей статье – морфемное членение болгарского сообщения и идентификация морфем болгарского языка.

**Проблематика морфемного членения болгарского сообщения и идентификации морфем.** Проблема членения объекта «до предела», до элементарной, в принципе не членимой далее частицы является одной из древнейших проблем философии. Бодуэновское понятие морфемы – аналог «абсолютного атома» Демокрита, поскольку у Бодуэна нет оговорки об относительности ее неделимости. Бодуэн не сомневался в абсолютной элементарности освященных традицией индоевропейскими корнями и аффиксами. И для современных лингвистов набор корней и аффиксов остается набором неизбежно элементарных, неделимых частиц. Однако понятие инфикса у индоевропейцев намекает на неполноту веры в неизбежность, непроницаемость по крайней мере корневых морфем. Да и гипотезы о смысловой нагрузке разных степеней аблаута (чередования корневых вокалов) можно считать провозвестниками разложения лингвистического атома – бодуэновской морфемы.

Со времени Б. де Куртенэ лингвисты научились «разлагать» звуки (фонемы) на акустические дифференциальные элементы. Оказалось возможным связать их с работой определенных мышц речевого аппарата. Каждая акция той или иной мышцы речевого аппарата человека порождает определенный акустический эффект, который можно связать с некоторым смыслом. Таким образом, каждая акция речевой

мышцы (иннервация или дезиннервация мышцы) через производимый ею акустический эффект сообщает нечто, передает некий смысл, т. е. реализует предельно краткий, далее неразложимый сегмент языковой материи. Именно такие сегменты авторы «Грамматики болгарского языка для владеющих русским языком» (далее – ГБР) назвали мономорфемами. В ГБР описаны наиболее частые мономорфемы – акции речевых мышц, определенно связанные с некоторым смыслом. Естественно, мономорфемы представляют собой самые частые морфологические феномены языка, появляющиеся в звуковом потоке речи с огромной частотой в разных типах контекстов (в разных морфемосочетаниях, полиморфемах), и потому их инвариантная семантика характеризуется высочайшей степенью абстрактности. Например, описанная в ГБР морфема, реализуемая акцией подбородочно-язычной мышцы, появляется в болгарском речевом потоке в составе более сложных феноменов, которые принято обозначать на письме буквами *a, e, u, ь* и общая частота которых в буквенной записи любого болгарского текста составляет не менее 30%. Эта морфема информирует об очень абстрактной семантике, названной авторами «агрегативностью».

Проблематику, связанную с выделением морфем (и полиморфем, т. е. сочетаний мономорфем, и самих мономорфем), следует соотносить с требованиями, предъявляемыми логикой к исследованиям в области всех современных наук. Но даже терминологический аппарат лингвистики, который обычно называют аллоэтическим и считают в наибольшей степени удовлетворяющим требованиям логической строгости описания объектов лингвистики, не в состоянии предложить лингвистам достаточно удобный набор терминов, с помощью которых стало бы возможным логически строгое описание многообразия морфем.

Список болгарских мономорфем (в ГБР показаны лишь самые частые), разумеется, с развитием исследований по физиологии речи и глоттометрической морфематики неизбежно будет дополнен и уточнен, но во всяком случае он будет весьма невелик. Однако каждая включенная в список мономорфема должна получить обстоятельную характеристику в посвященном ей разделе – в нем должны быть показаны многочисленные типы разных контекстов мономорфемы, из которых извлекается ее инвариантная семантика. А в настоящее время лингвистика даже не обладает необходимым терминологическим аппаратом и способами графической записи мономорфем. И лингвисты, и обычные носители языка привыкли работать с буквенной записью звукового потока речи. Поэтому составлять морфемиа-

рий «истинных» мономорфем в настоящее время не просто невозможно, но и нецелесообразно. Следовательно, в болгарском частотном морфемиарии будут представлены заведомо расчленимые морфемы – именно такие морфемы, которые выделял Б. де Куртенэ в своем анализе русских примеров, считая их элементарными: корневые морфемы, аффиксы, окончания, а также сихноморфемы, записываемые «отдельными словами», т. е. междометия, предлоги, союзы. Но при описании каждой такой полиморфемы будет показана ее структура – входящие в ее состав мономорфемы и (насколько это окажется возможным) их соотношения.

Вторая заслуживающая рассмотрения проблема – проблема границ морфемы. Дело в том, что мы представляем себе сообщение на естественном человеческом языке как последовательность непроницаемых (т. е. не проникающих друг в друга) сигналов (знаков). Эта линейная модель, модель абсолютной сукцессии сигналов, не адекватна феноменам естественных человеческих языков, но, поскольку она существует как имплицитная, неосознанная основа лингвистической работы, она мешает лингвистам, которые пытаются преодолеть постоянно возникающие трудности логического порядка путем локально действующих коррекций, не затрагивающих неудовлетворительного постулата о непроницаемости сигнала. По этой причине не отпадает и проблематика, связанная с определением границ морфем. Установить, с чего начинается и чем кончается морфема, оказывается далеко не так легко. Дело в том, что впечатление о большей или меньшей степени некоего качества лингвистического микрофеномена можно рассматривать как воздействие на него другого феномена, воздействие в смысле «проникновение в него другого феномена».

Вот один из наиболее очевидных примеров – корневая морфема существительного *рука*. Всем ясно, что в форме род. пад. *руки* букву *к* произносят иначе, чем в форме им. пад. *рука*. Говорят при этом об изменении в произношении *к* под воздействием следующего звука, т. е. считают, что звук *к* остается целиком частью корневой морфемы. Это так называемое комбинаторное изменение звуков, которое выделяется теперь как объект специальной области лингвистики – морфонологии (вне фонетики и вне морфологии).

Но можно рассуждать иначе. Можно считать, что те составляющие произношения буквы *к*, которые меняются (чередуются), не являются частями (составляющими, компонентами) корневой морфемы, а представляют собой части следующей за ней морфемы (флексионной), т. е. можно понимать воздействие следующей морфемы на

предыдущую не как действие на расстоянии, а как проникновение в нее. Если флексия – окончание род. пад. *и*, то проникающая в произношение *к* часть окончания – это дифференциальный элемент «дизность» (по терминологии Якобсона, Фанта и Халле).

Согласно привычной трактовке, дизность в прочтении *к* является частью корневой морфемы *рук*, хотя известно, что она по происхождению связана с флексийной морфемой *и*. Предполагается, следовательно, что между корневой морфемой и флексией есть какая-то четкая граница, которой в речи фактически нет – предположение основано на графике. Его принятие означает, что корневая морфема **непроницаема**, что дизность в произношении окончания *и* – это одна дизность, а дизность в произношении *к* – другая дизность, дизность, порожденная дизностью окончания, но все же другая, дизность **другой** морфемы.

Согласно принятой в ГБР трактовке морфем, которую можно назвать «волновой», дизность в произношении *к* является частью как корня, так и флексии, т. е. материя двух морфем накладывается одна на другую, две морфемы «звучат» одновременно, так же, как два волнообразных движения могут накладываться одно на другое, совершаться одновременно. О проблематике проницаемости морфем, морфемных границах, пределе морфемного членения и волновом характере морфем см. также в опубликованных работах [Янакиев 1977]; [Котова 1978; 1981; 1984]; [ГБР 2001]. Волновая трактовка морфем, предполагающая, что, подобно волне, морфема не имеет определенных границ в сообщении, в то же время учитывает, что волна морфемы имеет центр, вершину и затухает в соседних с вершиной зонах сообщения.

Итак, при волновой трактовке понятия *морфема* решение проблем, связанных с границами морфемы, получает несколько неожиданную форму: границ нет. Есть постепенное затухание, причем волна морфемы тем быстрее становится пренебрежимо малой, чем чаще встречается морфема в тексте. Мы же привыкли определять в сообщении границы морфем, например, говорим, что в словоформе *чтение* корневая морфема *чет* начинается с *ч* и кончается буквой *т*, однако при этом не уточняем, что именно начинается с *ч* и кончается *т*. Фонетика морфемы действительно может укладываться в рамки произнесения буквосочетания *чт*, но о семантике этого сказать нельзя. Понять, что *чт* в *чтение* несет информацию, называемую семантикой корневой морфемы *чет*, можно лишь после того, как за *чт* последует *е*, а за ним и *н* (одного *е* не хватает). В *почте* тоже есть *чт* с последующим *е*, но в *почте* нет корня *чет*. Только в буквосочетании

четании *чтен чт* может быть истолковано как «корневая морфема *чет*». И можно утверждать, что семантическая волна этой морфемы практически полностью загухает в *н* и что в сообщении, содержащем последовательность букв *чтен*, независимо от более отдаленных его частей, есть корневая морфема *чет* и центром ее фонетической сигнализации является *чт*.

На первый взгляд кажется, что таким способом нельзя разграничить морфему и не имеющее смысла буквосочетание. О *чт* в *почте* нельзя сказать, что оно сигнализирует о какой-либо другой морфеме, можно сказать лишь, что оно является буквосочетанием, не представляющим собой центра корневой морфемы *чет*. Но этим утверждением не отвергается возможность его понимания как части иной морфемы (в данном случае корневой морфемы *почт*). Но и *почт* как целое недостаточно, чтобы видеть в нем независимо от контекста корневую морфему существительного *почта* – и *почт* как целое не сигнализирует о корневой морфеме существительного *почта*. Лишь следующая буква *а*, например, в *почта* или *почтамт* решает вопрос о принадлежности фонетического соответствия буквосочетанию *почт* к фонетике корневой морфемы существительного *почта*. А у буквосочетания *почт* вне какого-либо контекста можно лишь подозревать «значение»; даже если это сегмент сообщения на русском языке, мы знаем, что у него есть значение, но еще не знаем какое.

Волновой характер морфемы определяется вероятностными (глottomетрическими) характеристиками текстового массива, накопленного в памяти лингвиста. Каждый феномен языка, каждый сигнал существует в сознании лингвиста вместе с оценкой вероятности его связи с той или иной семантикой, с теми или иными ситуациями, явлениями вне языковой действительности. Например, лингвисту предложена буква *ч*, без контекста. Располагая надежной информацией о частоте встречаемости всех морфем, содержащих эту букву, лингвист может построить дизъюнкцию гипотез о принадлежности этой буквы (и ее фонетического соответствия) к графике (соответственно фонетике) разных морфем русского языка и каждой из этих гипотез приписать оценку достоверности, вероятности. Так, лингвист имеет право утверждать, что эту букву с наибольшей вероятностью надо расценивать как сигнал корневой морфемы вопросительного местоимения *что*. С гораздо меньшей вероятностью лингвист может утверждать, что *ч* является сигналом корневой морфемы *чет*. С еще меньшей вероятностью можно ожидать, что *ч* окажется связанным с корневой морфемой существительного *почта* и т. д. Если будет известна и следующая за *ч* буква сообщения и она

окажется *t*, то отпадут мало вероятные гипотезы о связи *ч* с какими-либо другими морфемами, кроме названных. Но отношения между вероятностями указанных трех гипотез не изменятся. Если же станет известна еще одна буква и это будет *e*, оценки правдоподобия трех гипотез значительно изменятся. Гипотеза о связи с корнем вопросительного местоимения окажется равной нулю, но станет значительно более правдоподобной гипотеза о принадлежности *ч* к корневой морфеме слова *чтение*, хотя гипотезу о принадлежности к корневой морфеме слова *почта* тоже нельзя считать невероятной.

Так фактически работает сознание лингвиста, когда он опознает морфему. Имплицитно он опирается на вероятностные критерии, но из-за того, что он не доводит своих рассуждений до чисел, многое в его работе окрашено субъективными помехами. Глоттометрия углубляет эти рассуждения лингвиста до уровня числовых оценок. С логической точки зрения такое углубление осуществить нетрудно – достаточно воспользоваться вероятностной трактовкой понятия «силы сцепления» между частями сообщения (когезии), чтобы дать количественное выражение рассуждениям лингвиста при анализе сообщения ([Янакиев 1977: 97–112]; [ГБР 2001: 25–26]). Трудно оказалось получить достаточно надежную информацию о частоте сегментов сообщения разной сложности. Морфемииарий имеет целью дать информацию о частотах сегментов небольшой сложности (наиболее частых мономорфем и наиболее частых, наименее сложных полиморфем). С точки зрения волновой трактовки морфем приводимые в морфемииарии частоты морфем надо понимать как частоты вершин (центров) морфем. Эти частоты дают также возможность оценивать степень затухания морфемы-волны в пределах полиморфем, для которых приведены глоттометрические данные в морфемииарии.

Надо отметить, что доведенная до числовой точности последовательность операций определения границ между морфемами (точнее, впадин между вершинами волн-морфем, вместе с оценкой глубины этих впадин) при составлении морфемииария приводит к результатам, в подавляющем большинстве случаев хорошо согласующимся с результатами добротного, но не глоттометрического, традиционного лингвистического анализа. В немногих случаях несовпадения причины следует искать путем усложнения математико-статистической обработки данных, стремясь более приблизить ясное своими числовыми операциями глоттометрическое описание феноменов языка к поведению лингвиста, не всегда отдающего себе отчет в том, как он получил свои результаты.

Из всего сказанного выше ясно, что главная цель морфемиария – дать исчерпывающий перечень нагруженных смыслом частиц болгарской языковой материи, путем соединения которых образуются все возможные выражения болгарского языка. Вторая цель – показать закономерности соединения морфем друг с другом. Это не означает, разумеется, что при каждой морфеме будет дан полный список правил ее сочетаемости с другими морфемами. Но при каждой морфеме будут приведены доказанно наиболее частые, наиболее вероятные сочетания этой морфемы с другими морфемами болгарского языка (более сложные полиморфемы, в состав которых она входит).

Глоттометрическая информация в морфемиарии (данные о встречаемости наиболее частых в болгарской языковой практике морфем и статистические оценки надежности этих данных) позволяет обосновать отбор полиморфем для включения в морфемиарий. Иными словами, морфемиарий позволяет выявить то, что наиболее важно в болгарском языке в чисто синхронном плане. Пассивное ознакомление лингвистов-теоретиков, типологов, компаративистов с болгарским языком будет, можно полагать, наиболее эффективным с помощью морфемиария. Вместе с тем частотный морфемиарий, показывая статистический рельеф болгарской морфематики, будет полезен также составителям практических руководств для изучения болгарского языка как родного (в начальной и средней школе) и как иностранного.

#### Л и т е р а т у р а

- Бодуэн 1963 – *Бодуэн де Куртенэ И. А.* Избранные труды по общему языкознанию. Т. II. Москва, 1963.
- ГБР 2001 – *Котова Н., Янакиев М.* Грамматика болгарского языка для владеющих русским языком. М., 2001.
- Георгиев 1965 – *Георгиев В., Дуриданов И.* Езикознание. София, 1965.
- Котова 1978 – *Котова Н. В., Янакиев М.* О многообразии морфем в славянских языках // *Славянская филология.* Вып. X. М., 1978. С. 3–27.
- Котова 1981 – *Котова Н. В.* Морфонология и морфемный анализ // *Славянское и балканское языкознание. Проблемы морфонологии.* М., 1981. С. 329–339.
- Котова 1984 – *Котова Н. В.* Фонологические типы морфем и морфемный анализ. Исследования по славянскому языкознанию. М., 1984. С. 37–57.
- Кузнецова 1975 – *Кузнецова А. И., Лавренова О. А.* Частотный словарь суффиксов и префиксов русского языка // *Исследования по структурной и прикладной лингвистике.* Вып. 7. М., 1975. С. 92–99.
- Кузнецова 1986 – *Кузнецова А. И., Ефремова Т. Ф.* Словарь морфем русского языка. М., 1986.

*Славянский вестник. Вып. 2. М.: МАКС Пресс, 2004. 608 с.*

Мельчук 1984 – *Мельчук И. А., Жолковский А. К.* Толково-комбинаторный словарь русского языка. Опыт семантико-синтаксического описания русской лексики. Wien, 1984. (Wiener Slawistischer Almanach).

Мельчук 1995 – *Мельчук И. А.* Русский язык в модели «смысл – текст». М.; Вена, 1995.

Оливериус 1976 – *Оливериус З. Ф.* Морфемы русского языка. Прага, 1976.

Пашов 1976 – *Пашов П.* Българските глаголни времена (за основните им значения и онагледяването им със схеми и формули) // Помагало по българска морфология. Глагол / Съст. П. Пашов, Р. Ницолова. София, 1976.

Пенчев 1999 – *Пенчев Й.* Словообразователен речник на съвременния български книжовен език. София, 1999.

Потиха 1981 – *Потиха З. А.* Строение русского слова. М., 1981.

Тихонов 1978 – *Тихонов А. Н.* Школьный словообразовательный словарь русского языка. М., 1978.

Янакиев 1977 – *Янакиев М.* Стилистиката и езиковото обучение. София, 1977.