

Инновационная роль компьютерных технологий в языковых исследованиях

В последние десятилетия в связи с бурным развитием автоматизированных систем, которые базируются на использовании компьютерных технологий, все шире поднимается проблема использования ЭВМ в лингвистике для выполнения как теоретических исследований, так и решения прикладных проблем. Разработке этих проблем посвящены работы Л.С.Выготского, Арапова М.В., Соколова А.А., Стахова А.П., Н.И.Жинкина, А.А.Леонтьева, Р.Г.Регушевского, К.Б.Бектаева, И.А.Мельчука и др. В настоящее время еще нет завершенных универсальных теорий, которые можно было бы положить в основу построения практических компьютерных программ. Однако имеются интересные гипотезы и фрагменты теорий, дающие объяснения отдельным процессам порождения и понимания речи.

Данная статья посвящена исследованию сочетаемостных возможностей основ в сложных двухкомпонентных словах на основе количественного и психолингвистического анализа.

Среди новообразований современного русского языка сложные слова составляют значительный пласт лексики, обладающий чёткими структурными и семантическими закономерностями. Предметом исследования являются композиты и юкстапозиты, выраженные именами существительными, на примере которых наиболее глубоко можно проследить эволюцию не только способов словосложения, но и языка в целом.

Впервые использована оригинальная методология исследования структурной организации сложных слов в современном русском языке – концепция “золотой” пропорции. Исследование гармонического построения композитов и юкстапозитов в концепции “золотой”

пропорции проводится на основе лингво-информационного подхода, суть которого заключается в объединении традиционных лингвистических методов и приёмов с методами и приёмами, заимствованными из теории искусственного интеллекта, теории информации, а также компьютерных технологий обработки лингвистической информации и т.д.

На фонетическом (акустическом) уровне указанный подход предусматривает, кроме упомянутого выше, использование специализированной информационно-измерительной системы звуковых сигналов на основе персонального компьютера. Целью исследований является выявление с помощью лингво-информационного подхода закономерностей гармонического построения сложных слов.

Поскольку предметом исследования выступают разные сложные единицы (компози́ты и юкстапозити), в частности двухкомпонентные существительные, которые содержат две части, то оценивание гармоничности построения таких слов будет означать оценивание соотношения между длиной слова и длиной его большого компонента.

Как основной оценивающий показатель при количественном анализе гармоничности построения двухкомпонентных компози́тов и юкстапозитов на структурном уровне следует использовать отношение P , которое определяет отношения длины двухкомпонентного компози́та (юкстапозита) к длине большого компонента. В методологическом аспекте для измерения длины слова или длины компонентов сложных слов можно использовать два подхода. Суть первого, наиболее простого, состоит в том, что за единицу измерения длины берётся непосредственно слог. Соответственно длина всего слова или его компонента будет равна количеству слогов в конкретном слове или компоненте. Этот подход является удобным для анализа сложных единиц, зафиксированных письменно, в статической форме, так как любые изменения их длины или длины их компонентов на структурном уровне принципиально

невозможны. Значение пропорции вычисляем с помощью отношения количества слогов в двухкомпонентном слове к количеству слогов в большом компоненте.

Второй подход предусматривает за единицу измерения длины композита (юкстапозита) или длины компонента выбрать единицу физической величины – времени, например, миллисекунду (мс). Это означает, что на фонетическом (акустическом) уровне номинативную единицу рассматриваем в динамике как отрезок потока речи, которая акустически характеризуется определенной продолжительностью. Длина (продолжительность) звучание композита (юкстапозита) в этом случае может не иметь постоянного значения. Такая продолжительность зависит от ряда факторов, в частности от количества согласных в слогах, их характера, темпа речи, который всегда определенной мерой индивидуальный у разных лиц, от количества и места слогов в слове, от интонации речи и т.д.

Если во время исследования сложных слов на структурном уровне граница деления композита (юкстапозита) на компоненты четко очерчена, то на фонетическом (акустическом) уровне определение временной границы деления на компоненты фактически совпадает с границей слогоделения между этими же компонентами, особенно у композитов.

Систематизированная таблица количественных значений Π для семей “золотых” *p-пропорций* на структурном уровне выглядит так [5]:

	Π					
	2/1,	4/2,	6/3,	8/4,	10/5,	12/6
	3/2,	5/3,	7/4,	8/5,	10/6,	11/7, 13/8
	4/3,	6/4,	9/6,	13/9,	...	
	5/4,	7/5,	10/7,	14/10	...	
	6/5,	8/6,	11/8,	...		

	7/6,	9/7,	12/9,	...
	8/7,	10/8,	13/10,	...

Схематично анализ пропорции построения композитов можно проиллюстрировать словом *авиа(3)техника(3)* $\rightarrow P = 6:3 = 2.0$, а юкстапозитов – словом *кают(2)-компания(4)* $\rightarrow P = 6:4 = 1.5$

На структурном уровне семью дихотомии (“золотой” 0-пропорции) представляют такие значения P композитов и юкстапозитов: *хлеб(1)-соль(1)* $\rightarrow P=2:1$, *зверо(2)ферма(2)* $\rightarrow P=4:2$,

гальвано(3)пластика(3) $\rightarrow P = 6:3$.

К семье классической “золотой” 1-пропорции ($p=1$) относятся композиты и юкстапозиты с такими значениями P :

лейб(1)-медик(2) $\rightarrow P=3:2$, *авиа(3)модель(2)* $\rightarrow P=5:3$,

вагоно(3)строение(4) $\rightarrow P=7:4$.

Семья “золотой” 2-пропорции объединяет композиты и юкстапозиты с такими значениями P :

аэро(3)порт(1) $\rightarrow P=4:3$, *роман(2)-полемика(4)* $\rightarrow P=6:4$,

радио(3)прослушивание(6) $\rightarrow P = 9:6$

Семью “золотой” 3-пропорции представляют композиты и юкстапозиты с двумя значениями P :

рок(1)-панорама(4) $\rightarrow P=5:4$, *псевдо(2)публикация(5)* $\rightarrow P=7:5$,

вибро(2)диагностика (5) $\rightarrow P=7:5$.

К семье “золотой” 4-пропорции относятся композиты и юкстапозиты также с двумя значениями P :

пресс(1)-конференция(5) $\rightarrow P=6:5$, *фото(2)лаборатория(6)* $\rightarrow P=8:6$,

лор(1)отделение(5) $\rightarrow P=6:5$.

Семью “золотой” 5-пропорции представляют композиты и юкстапозиты со значениями P :

вода(2)регулирование(7) → П=9:7, фильм(1)-воспоминание(6) → П=7:6.

Заметим, если p равно или больше 3, то количество композитов и юкстапозитов уменьшается. Если же $p = 6, 7, 8$, то количество таких сложных слов резко падает.

Для определения пропорций построения исследуемых сложных единиц на фонетическом (акустическом) уровне необходимо прежде всего четко определить центр их деления на компоненты с помощью временной отметки. Далее за этой отметкой измеряем длину (продолжительность) большого компонента. Измеренные значения продолжительности звучания целого слова и большого компонента фиксируем. Количественное значение пропорции на этом уровне вычисляем из отношения:

$$P_{\phi} = \frac{\text{длительность звучания двухкомпонентного слова (мс)}}{\text{длительность звучания его большого компонента (мс)}}.$$

(индекс ϕ означает пропорцию на фонетическом (акустическом) уровне). Поскольку P_{ϕ} представляет собой относительную величину, то такой субъективный фактор, как неодинаковый темп речи у разных лиц не может существенным образом влиять на точность в оценивании этого показателя. Если в исследовании сложных слов на структурном уровне (измерение в слогах) удобно было использовать целые числа: 3/2; 5/3; 8/5 и т.д., то на фонетическом (акустическом) уровне для измерения временных интервалов необходимо использовать преимущественно дробные числа. Это связано с тем, что отношение вымеренных временных интервалов в двухкомпонентных словах, есть большей частью дробным числом. Учитывая все это, предлагается такая последовательность исследования пропорций построения композитов и юкстапозитов на фонетическом уровне. На первом этапе систематизируются отобранные сложные слова. Второй этап предусматривает процедуру “озвучивания” отобранных и систематизированных сложных слов. На третьем этапе проводим

фонетически-временной анализ этих слов на материале интонограмм. Итогом третьего этапа будет получение количественных значений продолжительности звучания всего сложного слова и большого компонента. Измерение указанных характеристик с помощью компьютерной интонограммы представлен на рисунке 1 на примере слова *автотранспорт*. Центр его деления (временная отметка – вертикальная прямая линия) приходится на значение 437 мс, общая же продолжительность (длина) составляет 1143 мс. Продолжительность (длина) большого компонента *транспорт* равняется 706 мс. На завершающем, четвертом, этапе определяются пропорции построения этого двухкомпонентного сложного слова, вычислив отношения соответственно приведенной основной формуле для $\Pi^*_ф$. По результатам этого квалификационного анализа исследуемое двухкомпонентное слово относится к соответствующей семье.

Пропорция построения слова *автотранспорт* вычисляется так:

$$\Pi_{ф} = (437 \text{ мс} + 706 \text{ мс}) : 706 \text{ мс} \approx 1.619.$$

Сравнив приведенные результаты с данными таблицы дробных показателей “золотых” p -пропорций, делаем вывод, что композит *автотранспорт* относится к семье классической “золотой” 1-пропорции.

I	II	II*
0	2.00	2.00
1	1.50; 1.67; 1.75; 1.60; 1.57; 1.625	1.62
2	1.33; 1.44	1.465
3	1.25; 1.40; 1.43	1.38
4	1.20; 1.38	1.32
5	1.17; 1.29	1.285
6	1.14; 1.30	1.26

Нами проведено фонетический анализ 80-ти композитов и юкстапозитов, определив их принадлежность к семьям p -пропорций [1]. На

основании проведенного исследования, которое основывается на принципах концепции «золотой» пропорции, сделаны такие выводы.

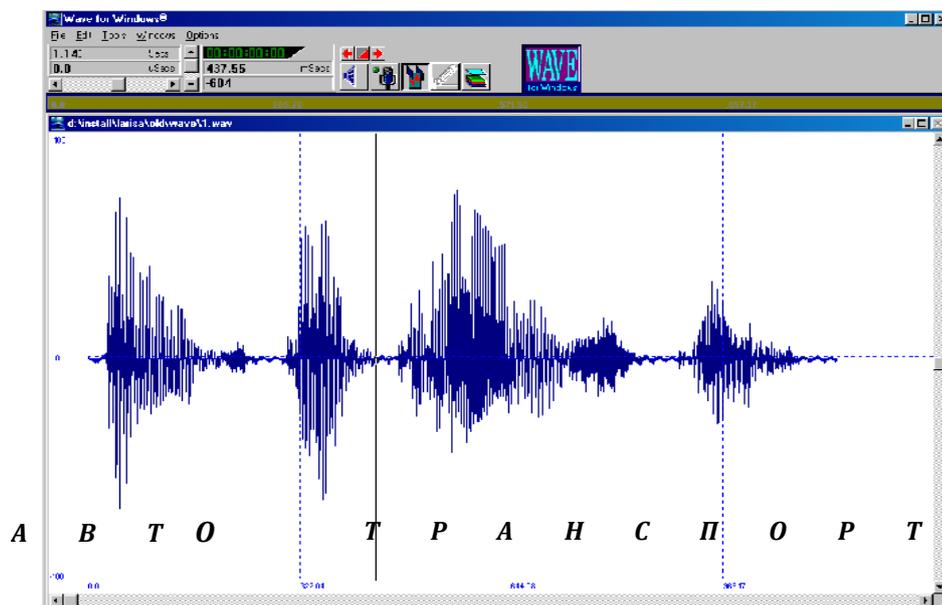


Рис. 1 – Интонограмма двухкомпонентного слова **автотранспорт**

Двухкомпонентные композиты и юкстапозиты, выраженные именем существительным, имеют гармоническое построение. В исследовании гармонического построения сложных слов в концепции “золотой” пропорции целесообразно использовать лингво-информационный подход как базисную форму изучения словообразования.

Сравнение количественных показателей на структурном и фонетическом уровнях показало, что доминантой двухкомпонентных композитов и юкстапозитов являются единицы, входящие в семью “золотой” 1-пропорции. Второе место занимают сложные слова из семьи дихотомии, третье – из семьи “золотой” 2-пропорции. Объемы семей “золотых” 3, 4, 5-пропорций резко падают.

У большей части слов пропорции на структурном и фонетическом уровнях совпадают, то есть они принадлежат к одним и тем же семьям, что подтверждает четкую закономерность в построении двухкомпонентных сложных слов. Например, юкстапозит *сон-трава* на структурном и

фонетическом уровнях относится к семье классической «золотой» пропорции. Пропорции построения некоторых слов во время произношения и на письме могут отличаться, то есть они входят в состав разных семей. Например, юкстапозит *люкс-секунда* на структурном уровне относится к семье «золотой» 2-пропорции, а на фонетическом – к семье «золотой» 1-пропорции. Не обнаружено ни одного сложного слова, в котором сочетаемостные возможности основ выходили бы за рамки указанных выше пропорций.

Наряду с выявлением структурных и количественных закономерностей в языкознании и, в частности в словообразовании, наиболее глубинным и «таинственным» остаётся вопрос: почему в человеческом языке существуют именно эти закономерности, а не какие-нибудь другие? Можно ли утверждать, например, что появляющиеся новые композиты и юкстапозиты пополняют указанные группы, или количественные соотношения для них будут другие? Ответ на этот вопрос может быть положительным, если вспомнить, что слова и язык вообще – это продукт умственной деятельности. Исследования физиологов позволили приоткрыть завесу «таинственности» в механизме работы человеческого мозга. Экспериментально установлено существование связи между частотами инвариантов ритмов мозга и «золотой» пропорцией. Вполне естественно предположить, что механизм работы человеческого мозга влияет определённым образом на процесс формирования и развития языка вообще и на количественные соотношения при построении сложных слов в частности. Поэтому на свет появляются именно те образования, пропорции которых укладываются в рамки «золотой» р-пропорции, чисел Фибоначчи. Количественный анализ целесообразно использовать в качестве одного из направлений в методологии изучения закономерностей построения сложных двухкомпонентных слов.

Использованная литература

1. Азарова Л.Є. Структура та фонетична побудова складних одиниць у концепції «золотої пропорції». Монографія. – Вінниця: ВЕГА. – 2001. – 284 с.
2. Арапов М.В. Квантитативная лингвистика. - М.: Наука, 1988. - 184с.
3. Леонтьев А.А. Язык, речь, речевая деятельность.- М.: Просвещение, 1969.
4. Соколов А.А. Тайны "золотого" сечения //Техника молодежи.- 1978. - №5.
5. Стахов А.П. Коды золотой пропорции. - М.: Радио и связь, 1984. - 151с.