

О фонологических единицах распознавания английских СМЫЧНЫХ

В. Н. Белоозеров

Выделены акустические сегменты субфонемной длительности в процессы реализации английских смычных (взрывных) согласных: Произведён дистрибутивный анализ возможности обобщения сегментов с учётом акустической близости и оптимальности использования в качестве единиц восприятия. Оказалось, что при этом сегменты объединяются иначе, чем это принято в традиционной фонологии. Противопоставление по глухости/звонкости согласных заменяется противопоставлением по консонантности/вокальности. Показано, что на уровне опознавания значащих единиц система фонем восприятия может преобразовываться к другой системе единиц, приближающейся к традиционному фонологическому описанию.

1. В акустическом процессе реализации начальных смычных согласных английского языка чётко выделяются участки, имеющие различную физическую природу и легко опознаваемые в потоке речи как слуховым анализом, так и приборными методами. Рассмотрим сначала слог типа ТА, где Т – смычный согласный, а А – любой гласный. Для определённости будем сначала рассматривать переднеязычные шумные смычные согласные. Для губных и заднеязычных согласных аналогичные рассуждения могут быть повторены без изменений.

Глухой смычный реализуется в указанной позиции последовательностью следующих процессов: t^* - взрывной шум в результате размыкания ротовой преграды напором воздуха из лёгких; h – фрикативный шум свободного истечения воздуха (Он имеет окраску тембра следующего гласного, но мы здесь этот фактор рассматривать не будем). В целом слог представляется записью $[t^* h A]$, где реализацию гласного А мы не рассматриваем. В абсолютном начале произнесения мы можем добавить в транскрипцию предшествующий звуку элемент молчания #: $[# t^* h A]$, а в безударных непервых слогах может отсутствовать придыхание: $[t^* A]$. (Внутри квадратных скобок записывается транскрипция звукового события как перечисление символов последовательно звучащих сегментов.)

Звонкий смычный в начальной позиции начинается обычно с такого же сегмента t^* , с которого начинается реализация глухого смычного. Факультативно этот сегмент может быть слабо озвончен; в таком случае мы обозначим его символом d^* . Но никакого звукового шлейфа, предшествующего взрыву при стандартном английском произношении не наблюдается. Голосовые связки вступают в работу обычно сразу после взрыва, когда ртовый канал перестраивается от смычного уклада к артикуляции гласного.

Результирующий вокальный звуковой процесс обозначим символом $d<$. Весь слог со звонким смычным представляется записью $[t^* d< A]$, или $[# t^* d< A]$.

Смычный глухой в исходе слога может быть реализован той же последовательностью $[t^* h]$, но обычно фрикативный сегмент h отсутствует. При этом между гласным и взрывом

реализуется краткая пауза звука #, во время которой ртовый канал перестраивается от гласного к смычке. Закрытый слог с глухим смычным может быть представлен записями [A # t*] или [A # t* h].

В случае звонкого согласного, закрывающего слог, перестройка ртовых органов происходит во время работы голосовых связок, что вызывает краткий вокальный глайд с характерным движением формант к локусу смычного звука. Обозначим этот участок символом d> . В результате смыкания ртовой преграды следует прекращение вокального звука и краткая пауза. Завершаться слог может размыканием преграды с соответствующим глухим звуком взрыва t*, но чаще в этой позиции взрыв не наблюдается. Впрочем в интервокальном положении и в слогах, закрытых сочетанием звонких согласных, наблюдается озвончённый взрыв d*, который факультативно может быть реализован и в абсолютном исходе, а в интервокальном положении появляется с необходимостью. В целом слог представляется следующими вариантами записи: [A d>], [A d> # t*], [A d> d*].

Таким образом на месте традиционных фонем [t], [d] мы имеем следующие физически различаемые процессы, по которым можно распознавать речевые знаки: t* , d* , d< , d> , h , # . Рассмотрение губных и заднеязычных гласных даёт нам соответствующие сегменты p* , b* , b< , b> и k* , g* , g< , g> . В целом типы слогов со смычными представляются следующей таблицей.

Таблица 1. Типы слогов со смычными согласными			
тип согласной	Губная	переднеязычная	заднеязычная
Смычная глухая	#p*hA, p*A	#t*hA, t*A	#k*hA, k*A
	A#p*, A#p*h	A#t*, A#t*h	A#k*, A#k*h
Смычная звонкая	#p*b<A, b*b<A	#t*d<A, d*d<A	#k*g<A, g*g<A
	Ab>, Ab>b*, Ab>#p*	Ad>, Ad>d*, Ad>#t*	Ag>, Ag>g*, Ag>#k*

В одной клетке таблицы здесь и далее записываются варианты произношения слога, которые не могут различать смысла высказываний в английском языке и следовательно не образуют позиций противопоставления входящих в них элементов.

2. В ходе восприятия речи языковой механизм должен преобразовать входной сигнал к форме внутреннего представления знаков и по совпадению формы принять решение о получении того или иного знака во входящем звуковом потоке. Воспринятые различаемые акустическим детектором звуковые процессы должны послужить материалом для решения о том, какие единицы внутреннего представления знаков языковой механизм должен зафиксировать в принимаемом сигнале. Очевидно, что шести распознаваемых элементов слишком много для представления сигналов, когда мы знаем, что для этого достаточно всего двух классических фонем. Для оптимизации своей работы решающее устройство языкового механизма может уменьшить избыточность сигнала, сгруппировав воспринимаемые звуковые сегменты так, чтобы представить сигналы с затратой минимальных ресурсов на запись и хранение знаков. После группировки мы получим систему действующих (виртуальных) фонем восприятия речи, где одной фонеме может соответствовать несколько разных звуковых сегментов, наблюдаемых как по отдельности, так и в последовательности друг с другом.

Соображения технической или биологической реализации языкового механизма требуют группировки в одну единицу распознавания (фонему) акустических процессов, имеющих сходную физическую природу. В нашем случае имеются следующие типы процессов: вокальные d> и d< , шумовые t* , d* и h , а также пауза # . Вокальные сегменты d> , d< и шумовые t* , d* находятся попарно в дополнительной дистрибуции, поэтому они

безусловно объединяются в одну функциональную единицу $d' = \{ d>, d< \}$ и $t'' = \{ t^*, d^* \}$ на правах альтернативных аллофонов, поскольку они акустически (перцептивно) близки, а отождествление дополнительно распределённых единиц не нарушает различимости ни одной пары знаков. (Здесь и далее фигурные скобки обозначают множество элементов, перечисленных через запятую внутри скобок)

Обобщённому сегменту t'' факультативно предшествует пауза # без изменения значения сигнала, поэтому эту последовательность можно воспринимать как единый звуковой образ: $t' = [\# t''] = \{ [\# t^*], d^* \}$.

В слогах типа АТ сегменту d^* всегда предшествует $d>$, но считать их функционально связанным нельзя, поскольку d^* теперь входит в состав фонемы t' , а $d>$ - в состав фонемы d' , которые противопоставлены как различающие значение речевых сигналов.

Соотношение сегментов t^* и h , которые в нашем корпусе данных распределены дополнительно, требует специального рассмотрения.

3. В новых обозначениях полный список слогов типа ТА и АТ может быть представлен таблицей 2. Здесь смысл обозначений p' , b' и k' , g' аналогичен t' , d' : В таблице показаны также слоги с глухими проточными согласными (акустическую реализацию которых мы здесь рассматривать не будем, но учёт которых необходим для определения фонематической роли сегмента h).

По этой таблице видно, что сегмент h дополнительно распределён по отношению к трём звуковыми единицами – p' , t' , k' , но объединить его в одну фонему можно только с одним из них. Спектральные характеристики шума h наиболее близки к k^* (им свойственно сосредоточение шумовой энергии в области низких частот), поэтому можно рассматривать гипотезу о формировании фонемы $k' = \{ h, k' \} = \{ h, \#k^* \}$, для которой взрывной заднеязычный и проточный ларингал были бы позиционными аллофонами, а взрывной аллофон был бы реализацией геминированной фонемы. Такое решение интуитивно представляется не корректным. Действительно, нужно учесть, что звук h может быть произнесён с разной степенью интенсивности и длительности вплоть до полного исчезновения (в диалектной и небрежной речи). hA должен иметь в речи также вариант hhA , с большей длительностью фрикативного сегмента. Тогда слоги hhA и $k'hA$ оказываются противопоставленными, а дополнительная дистрибуция h относительно k' , t' , p' нарушается. Поэтому гипотеза об объединении h с каким-либо взрывным сегментом в одну фонему отклоняется.

Таблица 2. Типы слогов с шумными согласными			
тип согласной	Губная	переднеязычная	Заднеязычная
Смычная глухая	$p'hA, p'A$	$t'hA, t'A$	$k'hA, k'A$
	$Ap', Ap'h$	$At', At'h$	$Ak', Ak'h$
Смычная звонкая	$p'b'A$	$t'd'A$	$k'g'A$
	$Ab', Ab'p'$	$Ad', Ad't'$	$Ag', Ag'k'$
Проточная глухая	fA	sA	hA
	Af	As	--

Примечание. К заднеязычным относим также ларингальный h , поскольку он акустически и артикуляционно максимально приближен к заднеязычным шумным и поскольку в английском языке отсутствуют другие согласные, которые бы могли образовать самостоятельный ларингальный ряд.

4. После этих объединений в исходном материале, представленном в фонемных обозначениях, не остаётся ни непротивопоставленных сегментов, ни связанных сочетаний

сегментов. В результате мы получили систему фонем, формально совпадающую с классической теорией: три смычных глухих p' , t' , k' и три смычных звонких b' , d' , g' . Но физическая реализация этих фонем включает в себя только часть звуковых сегментов, относимых к этим фонемам при традиционном рассмотрении. В качестве самостоятельных единиц восприятия выступают теперь экскурсия и рекурсия смычных согласных отдельно друг от друга, но не противопоставляясь друг другу, а выступая как позиционные аллофоны одной фонемы.

Из таблицы 3 видно, что сегмент h не несёт смыслоразличительной функции в положении, когда глухой смычный находится в абсолютном начале или абсолютном конце слога. Его можно считать вариацией взрывного сегмента и рассматривать их совместно как единый звуковой образ. Тогда глухие фонемы приобретут два аллофона – с придыханием и без придыхания. Однако сегмент h должен в других случаях опознаваться как самостоятельная фонема. Поэтому такое объединение с точки зрения распознающей системы бессмысленно, поскольку не ведёт к сокращению эталонов звуковых образов, подлежащих опознаванию в речи. В последовательностях типа $t'hA$, $At'h$ сегмент h так или иначе будет фиксироваться системой как присутствующий в произнесении звукотип. Поэтому при описании системы речевого звуковосприятия мы должны отвергнуть слияния таких сегментов в единый фонемный образ.

Тип согласной	Губная	Губная + переднеязычная	Переднеязычная	Заднеязычная	Задне- + переднеязычная
Смычная глухая	$p'hA, p'A$		$t'hA, t'A$	$k'hA, k'A$	
	$Ap', Ap'h$	$Ap't', Ap't'h$	$At', At'h$	$Ak', Ak'h$	$Ak't', Ak't'h$
Смычная звонкая	$p'b'A$		$t'd'A$	$k'g'A$	
	$Ab', Ab'p'$	$Ab't'$	$Ad', Ad't'$	$Ag', Ag'k'$	$Ag't'$
Проточная глухая	fA		sA	hA, hhA	
	Af		As	--	

Далее замечаем, что в непосредственном контакте с гласной «звонкие» фонемы b' , d' , g' факультативно сопровождаются гоморганным «глухим» элементом p' , t' , k' . Эти последовательности, не несущие смыслоразличительной функции, также можно считать аллофонами единой звонкой фонемы. Но согласно выше изложенным соображениям эти сегменты также не подлежат склеиванию на уровне речевого звуковосприятия.

Подводя итог, выпишем аллофонный состав новых смычных фонем, который представляется следующим образом. «Глухие» фонемы реализуются аллофонами ауслатного взрыва: $p' = \{ [\# p^*], b^* \}$, $t' = \{ [\# t^*], d^* \}$, $k' = \{ [\# k^*], g^* \}$. «Звонкие» фонемы включают вокальные сегменты формантных переходов: $b' = \{ b>, b< \}$, $d' = \{ d>, d< \}$, $g' = \{ g>, g< \}$. Условия реализации того или иного аллофона задаются таблицами 4 и 5.

Фонемы	Условие реализации.	непосредственно после гласного сегмента	непосредственно перед гласным сегментом
b'		$b>$	$b<$
d'		$d>$	$d<$

g'	g>	g<
----	----	----

Таблица 5. Аллофоны «глухих» смычных		
Фонемы	Условие реализации.	
	В контакте со звонким	Вне контакта со звонким
p'	b*	# p*
t'	d*	# t*
k'	g*	# k*

5. Склеивание сегментов, отвергнутое в предыдущем пункте, может иметь место на следующем этапе речевосприятия, на этапе сравнения воспринятых звуковых последовательностей с хранящимися в памяти эталонами звучания речевых знаков. Здесь действительно наличие сегмента *h* после глухих смычных не существенно для опознавания знака. Также не существенно сопровождение звонкой фонемы гоморганной глухой. Склеивая здесь последовательности воспринятых фонем в тех случаях, когда они не противопоставлены единичной фонеме, мы упрощаем запись эталонов звучания знаков, сокращаем перебор вариантов и облегчаем тем самым опознавание воспринятой последовательности фонем как того или иного знака (словоформы). Таким образом получаем новую систему фонем:

$$p = \{ b^* , [\# p^*] , [\# p^* h] \}$$

$$t = \{ d^* , [\# t^*] , [\# t^* h] \}$$

$$k = \{ g^* , [\# k^*] , [\# k^* h] \}$$

$$b = \{ [b > b^*] , [b > \# p^*] , [b^* b <] , [\# p^* b <] \}$$

$$d = \{ [d > d^*] , [d > \# t^*] , [d^* d <] , [\# t^* d <] \}$$

$$g = \{ [g > g^*] , [g > \# k^*] , [g^* g <] , [\# k^* g <] \}.$$

В представлении этими единицами рассматриваемые слоги получают привычную двухфонемную запись: *At* , *Ad* , *tA* , *dA* и др. Единственной особенностью изложенного подхода остаётся представление звонких смычных сегментов в абсолютном исходе сочетания «глухой» фонемой: *Agt* , *Abt*.

6. Настоящее рассмотрение показывает, что на разных этапах обработки речевого сигнала в воспринимающем механизме могут функционировать различные системы функциональных звуковых единиц (фонем). В целом процесс восприятия звукового облика речевых сигналов имеет дело с единицами, близкими к фонемам (звукотипам), выявляемым дескриптивистской фонологией и теорией Ленинградской фонологической школы. Можно показать, что дальнейшее восхождение при восприятии речи от распознавания входного сигнала к распознаванию осмысленных языковых знаков связано с использованием системы гипотетических решений о глубинном представлении знаков, лингвистической теорией которых может быть понятие гиперфонемы П. С. Кузнецова. Само глубинное представление звуковой структуры знака задаётся морфофонемами и фонемами Московской фонологической школы. В канале речеобразования глубинная структура знака преобразуется в последовательность звукотипов через управляющие сигналы, описание которых можно найти в теориях архифонем Пражского лингвистического кружка, слабых фонем Р. И. Аванесова, синтагмофонем М. В. Панова. Настоящее обсуждение конкретного материала английских смычных является иллюстрацией общей фонологической концепции автора, изложенной в ранее опубликованных работах [1, 2].

Литература

1. Белоозеров В. Н. Моделирование фонологической системы языкового механизма. Диссертация ... кандидата филологических наук. – М.: МГУ, 1974. – 283 с. ([http:// systemling.narod.ru/phonolog.html](http://systemling.narod.ru/phonolog.html))
2. Белоозеров В. Н. Фонологическое моделирование языкового механизма // Количественные методы в гуманитарных науках. – М.: Изд-во МГУ, 1981. – С. 153 – 158.