

# Lawrence J. Raphael. Acoustic Cues to the Perception of Segmental Phonemes

В и о л а И в а н о в а

29.11.18.

О чём статья: какие акустические признаки являются важными для восприятия и распознавания звука?



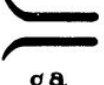

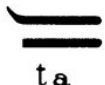
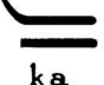



## В з р ы в н ы е с о г л а с н ы е

### С п о с о б о б р а з о в а н и я

Р а н н е е п о н и м а н и е : “stops as a class could be approximated satisfactorily by representing the sounds as vertical bars ..” (Liberman et al. 1952)

З а т е м : з а м е т и л и , ч т о п о с л е в з р ы в а ч т о - т о п р о и с х о д и т с г л а с н ы м и .  
В ы я с н и л о с ь , ч т о с о ч е т а н и е з в о н к и й в з р ы в н о й + г л а с н ы й м о ж е т б ы т ь с и н т е з и р о в а н о б е з в з р ы в а , е с л и ч а с т о т а н а ч а л ь н о г о у ч а с т к а F1 у в е л и ч и в а е т с я . П о к р а й н е й м е р е , э т о р а б о т а л о д л я г л а с н ы х с о т н о с и т е л ь н о в ы с о к о й F1 (Cooper et al. 1952, Liberman et al. 1954).

1. Schematic formant transition patterns varying in place and manner

	FRONTAL	MIDDLE	BACK
VOICED STOPS	 ba	 da	 ga
UNVOICED STOPS	 pa	 ta	 ka
NASALS	 ma	 na	 ŋa

Р и с . 1

И с т о ч н и к к а р т и н к и :

<https://www.phon.ucl.ac.uk/courses/spsci/acoustics/week2-5.pdf>

Это для звонких согласных. А что с глухими согласными?

Есть еще один акустический признак: silence of the closure interval.

Если между [s] и [l] в слове *slit* поместить кусочек пустоты, люди будут слышать *split*.

## Место образования

(Harris et al. 1958, Liberman et al. 1952, Liberman et al. 1954)

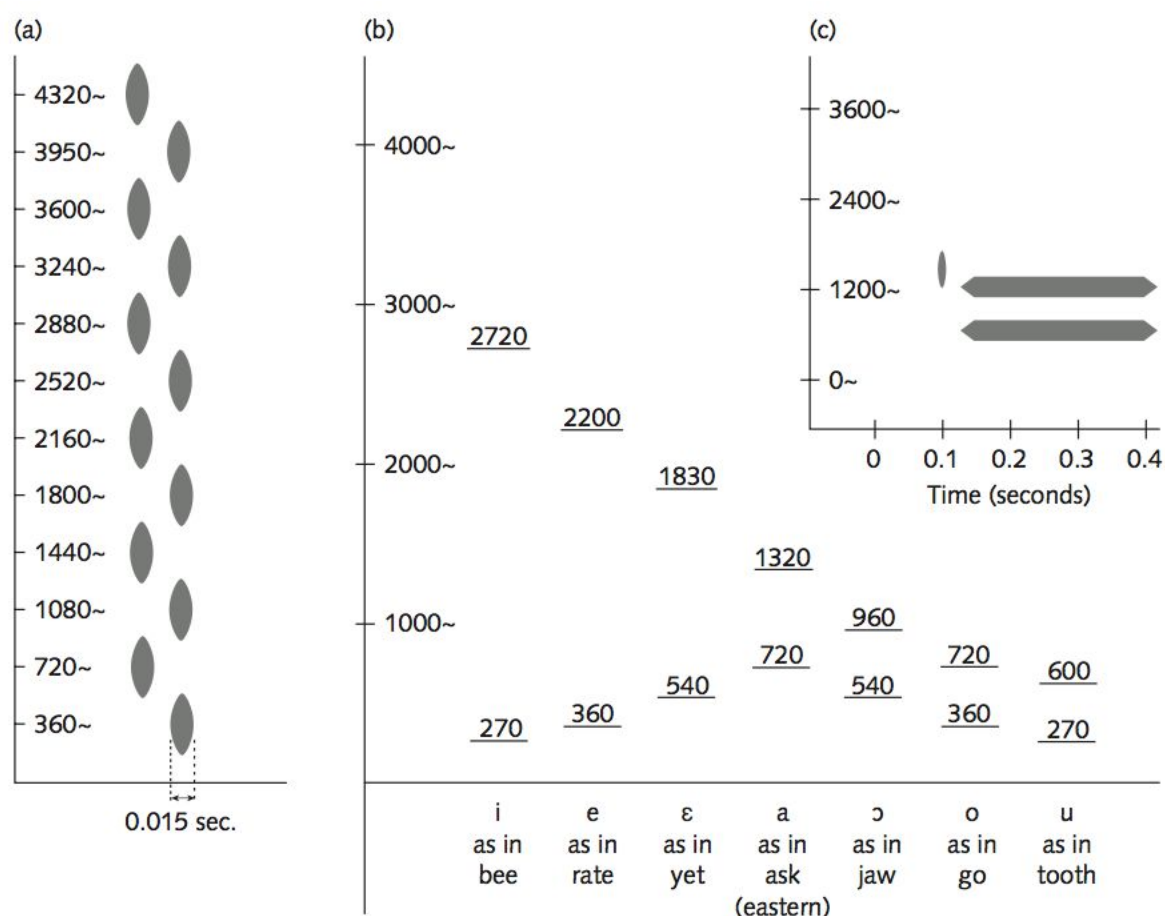


Рис. 2. Синтетические стимулы, которые образовывали комбинации взрыва определенной частоты + двухформантный гласный.

В этом эксперименте было показано, что восприятие взрывов — контекстно-зависимо: - взрыв 360 Гц воспринимался как [p]

- взрывы частотой выше 3000 Гц воспринимались как [t]
- взрывы частотой 1440 Гц воспринимались как [p] перед [i] и [u], но как [k] перед [a]

И опять нужно смотреть на форманты, но уже на F2. Другой эксперимент (Liberman et al. 1954)

Испытуемым давали послушать гласные с немного измененными переходными участками F2. Смычки перед гласным не было.

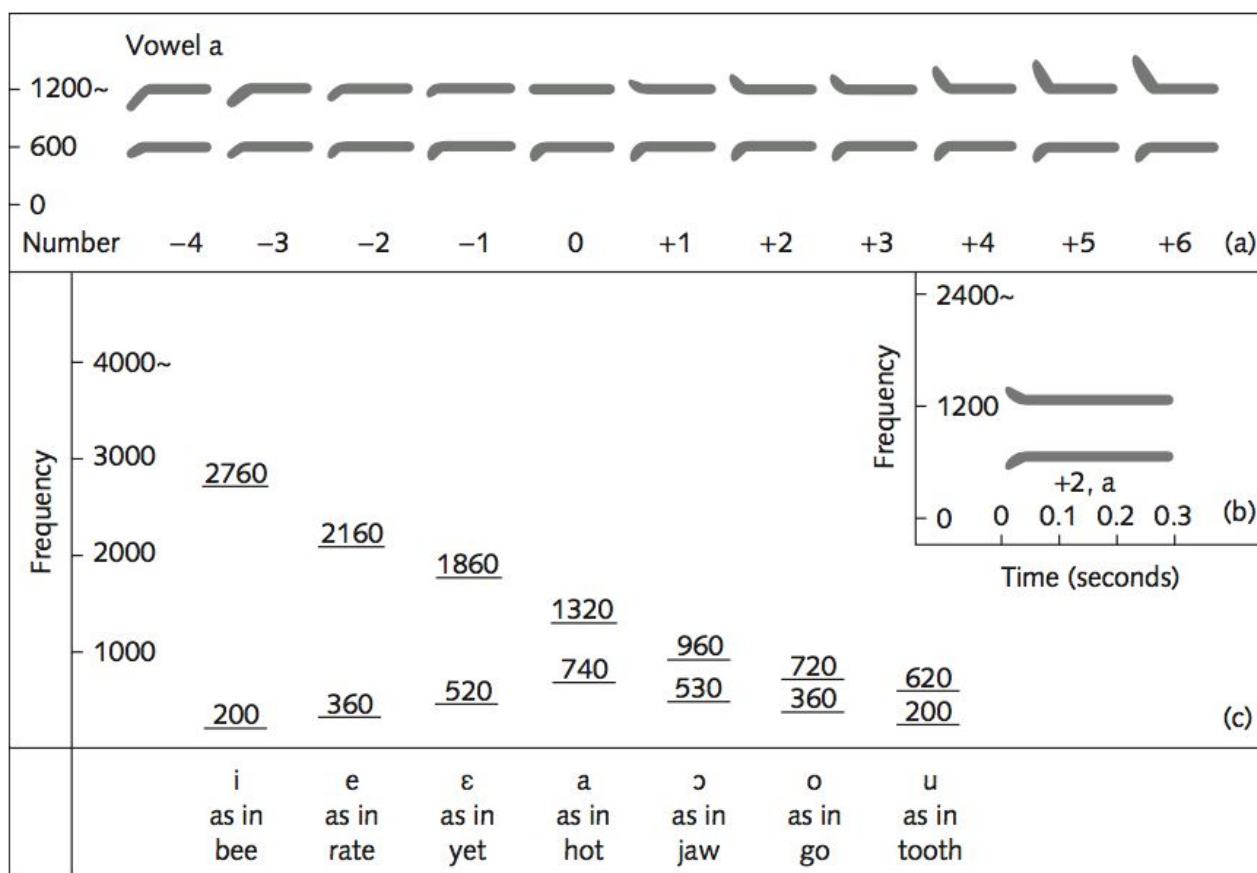


Рис. 3

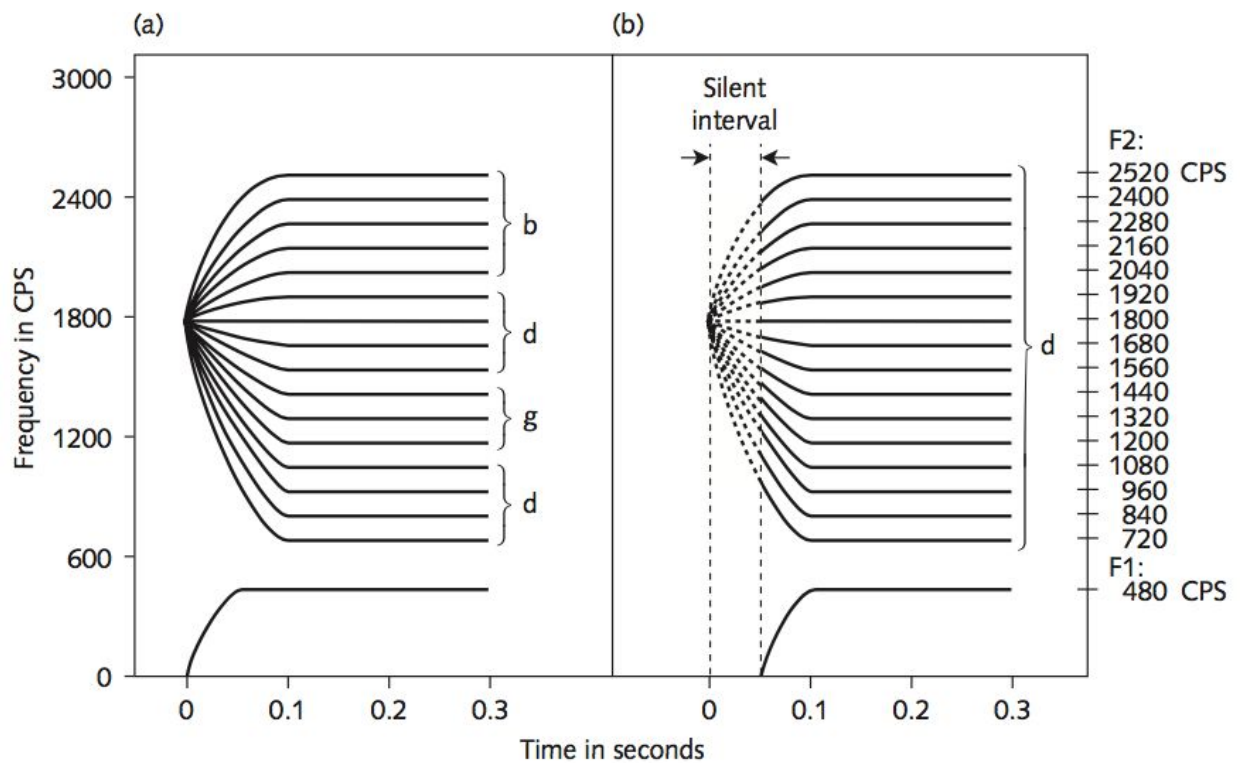
Что слышали испытуемые?

- губные согласные – когда переходный участок F2 возрастал
- альвеолярные – когда переходный участок F2 был ровным или немного убывал
- велярные – когда переходный участок F2 резко убывал.

Но вообще всё это опять зависело от того, каким был гласный: у гласных переднего ряда высокая F2, поэтому она не убывала и не была плоской, а всегда возрастала. Moreover, the amount of frequency change, even when

its direction fell within the scope of the general description, displayed considerable variability from one vowel context to another. (?)

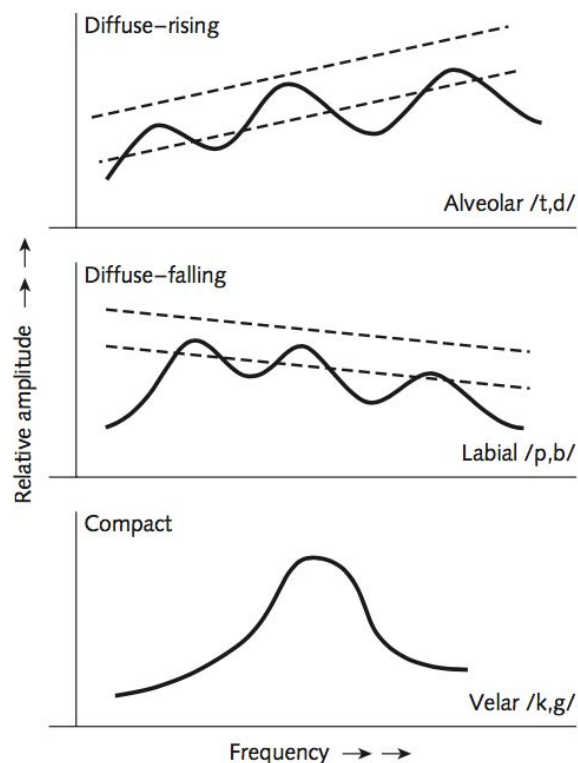
П о я в л я е т с я и д е я л о к у с а (Delattre, Liberman, & Cooper 1955)



Р и с . 4

(Dorman and Raphael 1977): р а з м ы к а н и е с м ы ч к и и п е р е х о д н ы е  
у ч а с т к и г л а с н о г о – ф у н к ц и о н а л ь н о  
э к в и в а л е н т н ы е п р и з н а к и , и г о в о р я щ и е  
и с п о л ь з у ю т в с е д о с т у п н ы е п р и з н а к и , а н е  
т о л ь к о к а к о й - т о о д и н .

(Steven and Blumstein 1978): а ч т о е с л и с п е к т р у ч а с т к а 10-20 м с  
п о с л е р а з м ы к а н и я с м ы ч к и – э т о т о ж е  
а к у с т и ч е с к и й п р и з н а к ?



Р и с . 5.

(Ohde et al. 1995) повторили эксперимент выше на взрослых и детях (5, 6, 7, 9 и 11 лет) и пришли к выводу, что спектр участка после взрыва содержит достаточно информации для определения места образования согласного, а что ни одна группа испытуемых не полагалась на переходные участки формант.

(Lahiri, Gewirth, Blumstein 1984): в малаялам, английском и французском этот признак не справился с различением [b] и [d]. Альтернативный подход: change in the distribution of spectral energy from the burst onset to the onset of voicing. И с таким подходом всё получилось. Авторы приходят к выводу, что переходные участки формант не являются акустическими признаками для различения места образования.

(Dorman and Loizou 1996) повторили эксперимент (Lahiri et al. 1984), но с использованием естественной речи, и пришли к противоположному выводу: их испытуемые классифицировали звуки (губные и альвеолярные) по переходным участкам формант.

Авторы предостерегают: в естественной речи могут содержаться такие признаки, которых нет в искусственных стимулах, но которые могут быть важны для восприятия, в то время как способы классифицировать звуки в искусственных условиях могут не работать на восприятии естественной речи.

Таким образом, можно выделить несколько акустических признаков для определения места образования согласного, но:

- их иерархия непонятна, потому что всё зависит от контекста
- все исследователи соглашались в том, что слушающие полагаются на все доступные признаки. Как они это делают? Возможно, в зависимости от контекста выбирается более заметный (salient) признак.

### Признак глухости-звонкости

Тут важно различать позиции согласных:

- 1) в начале слога, в конце слога, интервокальный
- 2) в начале ударного слога, в начале безударного слога.

В английском:

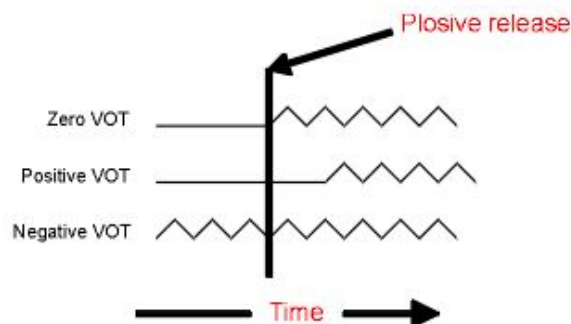
[p, t, k] в начале ударного слога – глухие и придыхательные.

[b, d, g] – звонкие и непридыхательные.

(Lotz et al. 1960, Reeds & Wang, 1961): в словах *spill, still, skill* у брали начальный [s], и испытуемые в урезанных словах слышали *bill, dill, gill*. Вывод: придыхательность более важный признак в английском, чем глухость-звонкость.

Дальше начинаются измерения VOT – время (мс) между разрывом смычки и началом фонации (голоса).

(Lisker & Abramson 1964, 1967, 1970): если на отрезке VOT есть фонация и F1 гласного поднимается, испытуемые слышат звонкие [b, d, g]. Если убрать фонацию и начальный фрагмент F1, испытуемые слышат [p, t, k].



Р и с . 6. VOT

И с т о ч н и к : <https://www.phon.ucl.ac.uk/home/johnm/siphtra/plostut2/plostut2-2.htm>

(Lisker and Abramson 1970): носители разных языков по-разному классифицируют смычные согласные. Носители английского: 2 группы согласных (короткий +VOT, длинный +VOT) Носители испанского: 2 группы согласных (-VOT, +VOT) Носители тайского: 3 группы согласных (-VOT, короткий +VOT, длинный +VOT)

F0 – считается вторичным признаком, но ее эффект также обнаружен в экспериментах.

(Lehiste & Peterson 1961, Ohde 1984): если после размыкания смычки F0 находится на том же уровне или растет – это звонкий согласный, если уменьшается – глухой.

В позиции конца слога: зависит от контекста. Считается, что наиболее заметный признак в большинстве контекстов – наличие/отсутствие фонации.

Возможно, имеет значение также переходный участок F1 перед конечным согласным, но не все с этим согласны.

Кто-то считает, что дело в длине гласного.

Кто-то считает, что оба этих признака важны для различения глухих и звонких в позиции конца слога.

В интервокальной позиции: звонкие короче, глухие длиннее.

Но: (Lisker 1957, 1981)

В парах типа rupee – ruby, rapid – rabid изменяли длительность второго согласного. Короткие по длительности (около 60 мс) оценивались как глухие, а длинные (около 130 мс) – как звонкие.

## Фрикативные и аффрикаты

### Способ образования

Основной акустический признак фрикативных – шум.

Аффрикаты: сочетают признаки смычных и фрикативных:

- длина пустых сегментов (тишина)
- длина шума

Что отличает аффрикату от фрикативного?

- наличие пустоты перед шумом
- шум короче

### Место образования

Для фрикативных:

- спектр шума: [s, z] vs [ʃ, ʒ] vs [ð, θ, f, v]
- переходные участки формант (особенно F2).

Для аффрикат: комбинация признаков для смычных и для фрикативных.

### Различение глухих и звонких

Для фрикативных:

для позиции начала слога и интервокальной:  
наличие фонации  
в конце слога: длительность шума и длина предшествующего гласного.

## Носовые

Способ образования: nasal resonance

Место образования: всё сложно. Все сходятся на том, что слушающие полагаются на переходные участки формант окружающих гласных.



## Глайды [w, j] и плавные [r, l]

### Способ образования

В отличие от других согласных:

- минимальные преграды для потока воздуха
- имеют свою формантную структуру

Как отличить гласный + глайд от дифтонга?

(O'Connor et al. 1957): у глайда в позиции перед гласным есть ровный, постоянный (steady-state) сегмент длиной 30-50 мс, чего нет в начале дифтонга.

(Lieberman et al. 1956): A second contrast concerns the rate of formant frequency change, which is faster for semivowels than for diphthongs but not as fast as the formant transitions preceding or following stops.

### Место образования

[w] и [j] отличаются также, как [u] и [i].

[r] и [l] различаются в третьей форманте: для [l] она более стабильна, для [r] – изменчивая.

## Гласные

Никто никогда не сомневался, что основные признаки для восприятия гласных – форманты. Но:

- 1) Частоты формант одного и того же гласного варьируются (в зависимости от длины речевого тракта носителя, скорости речи, интонации): *как люди распознают разные форманты как один и тот же гласный?*

Люди умеют «нормализовывать» разные варианты частот.

- 2) В структуре гласных больше двух формант: *влияют ли как-то третья форманта и выше на восприятие? F0?*

Да. F0 и форманты выше F2 являются более важными признаками при шепоте.

- 3) Гласные в связанной речи влияют не только на ближайшее окружение, но и на соседние слоги: *как это влияет на способность носителей распознавать гласные?*

(Ladefoged and Broadbent 1957): в эксперименте давали послушать фразу «Please say what this word is: bVt», где гласный V варьировался (всего было 4 гласных). Люди определяли гласный в целевом слоге в зависимости от того, какие формантные частоты были в предшествующем предложении.

Люди умеют «нормализовывать» разные варианты частот.

- 4) Длительность гласных в реальной жизни короче, чем в лабораторных условиях фонетистов, и гласные в разных контекстах различаются по длине. *Полагаются ли носители на длину гласного?*

(Gay 1970): один и тот же синтезированный звук воспринимался как монофтонг, когда был коротким по звучанию, и как дифтонг, когда его растягивали.

(Verbrugge and Shankweiler 1977): один и тот же слог [pVp], произнесенный быстро, хорошо воспринимается либо в таком же быстром предложении, либо изолированно, а помещенный в медленное предложение воспринимается плохо.

(Strange, Jenkins, and Johnson 1983): были записаны слоги типа [bVb]. Испытуемым давали слушать и определять два типа гласных: гласные, которые содержали только переходные сегменты, и стабильные извлеченные участки формант. При изменении длины гласного, испытуемые чаще ошибались в стабильных гласных, а в гласных без середины (только с переходными участками) ошибались минимально.

- 5) В реальной речи частота формант не стабильна на всем отрезке: *могут ли носители различать гласные в таких условиях?*

(Jenkins, Strange, & Edman 1983, Strange et al. 1983, Strange 1989): да, могут. Для распознавания гласных в принципе не нужен стабильный участок формант. Люди умеют распознавать гласные по переходным участкам также хорошо, как и по целым гласным.

## З а к л ю ч е н и е

В восприятии важны многие акустические признаки, и, видимо, носители используют их все, в зависимости от контекста и ситуации.

Пока что не найдено простого (the simplest) способа связать акустические признаки и восприятие.

Таблица 1. Акустические признаки, важные для восприятия согласных.

	Взрывные согласные	Фрикативные	Аффрикаты	Носовые	Глайды и плавные
Способ	- наличие взрыва; - подъем на переходном участке F1 (для звонких); - F1 reduction (для глухих); - silence of the closure interval;	наличие шума	наличие тишины и шума (шум короче, чем у фрикативных)	носовой резонанс	- минимальные преграды для потока воздуха - имеют свою формантную структуру
Место	- частота взрыва;	- спектр шума - переходные	комбинация признаков взрыва	переходные участки формы	[w] и [j] отличаются так же,

	<ul style="list-style-type: none"> <li>- изменение F2;</li> <li>- интенсивность спектра после размыкания смычки</li> </ul>	участки формант (особенно F2).	ных и фрикативных	<ul style="list-style-type: none"> <li>нт</li> <li>о</li> <li>к</li> <li>р</li> <li>у</li> <li>ж</li> <li>а</li> <li>ю</li> <li>щ</li> <li>и</li> <li>х</li> <li>г</li> <li>л</li> <li>а</li> <li>с</li> <li>н</li> <li>ы</li> <li>х</li> <li>, н</li> <li>о</li> <li>в</li> <li>о</li> <li>о</li> <li>б</li> <li>щ</li> <li>е</li> <li>с</li> <li>о</li> <li>п</li> <li>р</li> <li>е</li> <li>д</li> <li>е</li> <li>л</li> <li>е</li> <li>н</li> <li>и</li> <li>е</li> <li>м</li> <li>е</li> <li>с</li> <li>т</li> <li>а</li> <li>в</li> <li>с</li> <li>ё</li> <li>о</li> <li>ч</li> <li>е</li> <li>н</li> <li>ь</li> <li>п</li> <li>л</li> <li>о</li> <li>х</li> <li>о</li> </ul>	<ul style="list-style-type: none"> <li>к</li> <li>а</li> <li>к</li> <li> [u]</li> <li>и</li> <li> [i].</li> <li> [r]</li> <li>и</li> <li> [l]</li> <li>р</li> <li>а</li> <li>з</li> <li>л</li> <li>и</li> <li>ч</li> <li>а</li> <li>ю</li> <li>т</li> <li>с</li> <li>я</li> <li>в</li> <li>F3:</li> <li>д</li> <li>л</li> <li>я</li> <li> [l]</li> <li>о</li> <li>н</li> <li>а</li> <li>б</li> <li>о</li> <li>л</li> <li>е</li> <li>е</li> <li>с</li> <li>т</li> <li>а</li> <li>б</li> <li>и</li> <li>л</li> <li>ь</li> <li>н</li> <li>а</li> <li>, д</li> <li>л</li> <li>я</li> <li> [r]</li> <li>–</li> <li>и</li> <li>з</li> <li>м</li> <li>е</li> <li>н</li> <li>ч</li> <li>и</li> <li>в</li> <li>а</li> <li>я</li> <li>.</li> </ul>
Глухость-звонкость	<ul style="list-style-type: none"> <li>- VOT</li> <li>- F0</li> <li>- длина предшествующего гласного</li> <li>- переходный участок F1 в последующем гласном</li> <li>- длина тишины</li> </ul>	<ul style="list-style-type: none"> <li>- наличие фонации</li> <li>- длительность шума и длина предшествующего гласного.</li> </ul>	комбинация признаков взрывных и фрикативных	*	*