

References: Phonetics



[Phonetics](#) is a branch of linguistics for the study of sounds of human speech: physical properties, production, acoustics, articulation, etc.

Phoneme

In any given language, a [phoneme](#) is the smallest sound segment that can be used to distinguish one word from another. For example "bat" and "chat" have only one sound different but this changes the word. The phonemes in question are "B" and "CH". What exactly these are and how many exist varies a bit and may be influenced by accents included. Generally, US English consists of 39 to 44 phonemes. See ARPAbet below for more phoneme examples.

Grapheme

The definition of a [grapheme](#) is somewhat inconsistent in the literature. In our context, a grapheme is the smallest symbol that distinguishes one written word from another. For example, "bat" and "chat" have a difference of two graphemes, even though "CH" is considered to be a single phoneme. In US English, 26 letters and a space combine for 27 possible graphemes.

Lexicon

A lexicon for speech recognition is a lookup file for converting speech parts to words. An example of this is [cmudict](#), the Carnegie Mellon tool for speech recognition compatible with the open source [Sphinx](#) project. Here's a short excerpt:

```

AARDVARK      AA R D V AA R K
AARON         EH R AH N
AARON'S       EH R AH N Z
AARONS        EH R AH N Z
AARONSON      EH R AH N S AH N
AARONSON'S    EH R AH N S AH N Z
AARONSON'S(2) AA R AH N S AH N Z
AARONSON(2)   AA R AH N S AH N
...

```

ARPAbet

A set of phonemes developed by the Advanced Research Projects Agency(ARPA) for the Speech Understanding Project (1970's).

[ARPAnet on Wikipedia](#) [ARPAnet dictionary at CMU](#):

Phoneme	Example	Translation
AA	odd	AA D
AE	at	AE T
AH	hut	HH AH T
AO	ought	AO T
AW	cow	K AW
AY	hide	HH AY D
B	be	B IY
CH	cheese	CH IY Z
D	dee	D IY
DH	thee	DH IY
EH	Ed	EH D
ER	hurt	HH ER T
EY	ate	EY T
F	fee	F IY

Phoneme	Example	Translation
G	green	G R IY N
HH	he	HH IY
IH	it	IH T
IY	eat	IY T
JH	gee	JH IY
K	key	K IY
L	lee	L IY
M	me	M IY
N	knee	N IY
NG	ping	P IH NG
OW	oat	OW T
OY	toy	T OY
P	pee	P IY
R	read	R IY D
S	sea	S IY
SH	she	SH IY
T	tea	T IY
TH	theta	TH EY T AH
UH	hood	HH UH D
UW	two	T UW
V	vee	V IY
W	we	W IY
Y	yield	Y IY L D
Z	zee	Z IY
ZH	seizure	S IY ZH ER

