Phonemes and phonology

John Goldsmith

January 21, 2010

Phonemics

- Phonemics is a theory of categorization of sounds in a language. Each language has, or displays, a large set of sounds.

- These sounds are grouped into closely related families of sounds; each such family acts as a single unit from a phonemic point of view. We call these families "phonemes," and in some respects we are more aware of these phonemes than we are of the very sounds themselves. A word can be viewed as a sequence of sounds, or as a sequence of phonemes. Neither view is more correct; they are at different levels of abstraction.

- The different sounds that comprise (or realize) a phoneme do not always appear in the same "environment": the "environment" of a sound is the sound to its left and to its right (that's a first approximation of the definition).

- The goal is to minimize the number of phonemes, by showing (for example) that two sounds are part of the same phoneme, and we do that by showing that the two appear in *complementary distribution*:[1] that is, that there is no overlap in the contexts in which the two appear, and we can henceforth predict, given a location in an utterance, which of the two sounds *could* appear there.

  [1] Complementary distribution

- The difference between the two sounds, in this sense, is *predictable*: if you know the principles by which the sounds (called the *allophones*) of a phoneme are distributed, then you can simplify (or compress) your notational system: you can indicate just the name of (or symbol for) the phoneme.

- *Which* sound realizes the phoneme in a given word will be determined, or predicted, by the principles of the distribution of that phoneme's allophones.

Now, in more detail

A phonemic analysis is a process that takes as its input either (1) a set of utterances, transcribed phonetically or (2) a speaker of a language, and produces a set of symbols which represent distinct phonemes. There are 7 further conditions:

- The set must be minimal, in the sense that there may be no smaller set of phonemes that satisfies the conditions for phonemic analysis.

- It must be possible to represent every utterance of the language as a string of phonemes.

- It must be possible to represent every utterance of the language as a string of phones, where phones are the symbols used in the phonetic transcription.

- In the description of a particular utterance, there must be a one-to-one relationship between the phonemes in (2) and the phones in (1) (between the "tokens," not the types).

- It must be possible to establish rules of allophony: these rules specify the correspondences between the two levels of representation mentioned in (4). In particular, a rule of allophony says, "Phoneme M is realized as Phone P in context C", where context C says what sounds are present to the left or right (or both) of Phone P.

  A phonemic analysis always takes the form:

Phoneme	Phone	Context
/p/	[q]	A_B
	[r]	C_
	[s]	elsewhere

  where the underscore marks the focus of the environment we are considering.

- Uniqueness: A phonemic analysis must provide a unique phonemic representation for any given phonetic representation.

  Typographical convention: phonemes are placed inside slashes /phonemes/, and phones at the phonetic level are placed within square brackets: [phones].

  I-Raising: /ay/ is realized as [ʌy] when followed by a voiceless obstruent (*ptksf*) in the same word, otherwise it is realized as [ay].
  A rule of allophony always ends with the statement, "otherwise it is realized as..." - if only because that is the simplest way to state it (i.e., one realization can be stated without specifying the context).
  What does a symbol mean?

- Philosophy #1: A phonetic symbol is understood to represent a specific linguistic sound. A phonemic symbol has meaning only insofar as it relates to specific phones. The particular symbol used has essentially no other significance. The mean lies in the correspondence rules (rules of allophony).

- Philosophy #2: Human beings are extraordinarily good at discovering the phonemic relationships (allophony rules) that structure language, and the realization that is given in the "otherwise" formulation of the allophony rule is the (psychologically) real sound-image that comprises the phoneme. Phonemic analysis is primarily practical: it is meant to be practiced.

In practice, students are taught to look at all pairs of sounds that are similar and to test whether these two sounds are in an allophonic relationship (i.e., whether they might be realizations of the same phoneme). This requires one to learn what it means for two sounds to be similar (of course, one could simply consider all pairs of sounds....). Next, one (successively, iteratively) looks at each pair of sounds, and asks whether that pair might be realizations of the same phoneme.

The best test of whether two phones might be realizations of the same phoneme is the minimal pair test: if we can find two words that are different words and which are identical, except that one contains phone P, and the other contains phone Q in the same position, then the two constitute a minimal pair with regard to the pair P and Q, and P and Q cannot be allophones of the same phoneme.

- ɪ and ɛ in English: *pit* and *pet* are different words; hence these two sounds cannot be allophones of the same phoneme.

- Similarly, *i* and ɪ exist in English, and *pit* and *Pete* (or *peat*) are minimal pairs with these sounds (hence, they cannot be allophones in English).

- By contrast, these last two sounds occur in Canadian French, where no such minimal pairs exist (and they can be analyzed as allophones of a single phoneme).

If we cannot find any minimal pairs distinguishing two phones P and Q, then we can look for a principle that will specify a context in which one of them is used (with the other used "elsewhere"). If we can find such a principle, then we have established that the two are allophones of a single phoneme. (Convince yourself that this is techically always possible if the corpus is finite.)

- There is another possibility that phonemic theory allows which we have not discussed: two phones P and Q which are realizations of the same phoneme M may be in free variation (in some context, or in all contexts): a word containing the sound P may be changed by replacing P by Q, and the result is another acceptable pronunciation of the same word.

- Notice that both the statement of free variation and the definition of minimal pair requires being able to know whether two words are the same or different.

- Bottom-up philosophy: phonemics is deeply rooted in a conception of analysis beginning with the phonetic, followed by analysis at the phonemic, followed by analysis of morphemes, and so on. This bottom-up philosophy can be interpreted as scientific methodology or as psychological theory about human speakers - or both. If it is taken as a theory about language users, it is a theory of language hearers rather than speakers. (Why?)

- I said above that with a finite corpus and no minimal pairs for sounds P and Q, it is always possible in principle to establish a phoneme that is realized as P and Q. But it is generally understood that the principles of allophony must have some phonological simplicity or naturalness to them.

- This was generally understood to mean, in addition, that the phones P and Q must not be too different phonetically, but that phrase was never successfully defined. Phonemicists all knew that some such clause was necessary to prevent an analysis in which [h] and the velar nasal are allophones of the same phoneme. (Why?)

Bottom-up, hearer-oriented analysis

Hearers have access to the string of phones representing an utterance. Do they have access to the presence of word-boundaries separating these words? Is this an empirical question? - that is, is it possible that some languages provide phonetic cues to word-boundaries, and others do not? This is important, because if allophony does not have access to word-boundaries, what do we do with the phonetic realization of the phrase *high time*? Which realization of the vowel of *high* do we find? Why not the other?

Generative phonology: speaker-oriented analysis

The speaker knows where the word-boundaries are, and what phonemes are being uttered, so let the conditioning on phoneme-realization be dependent on what is known to the speaker, not just on what the hearer will be able to hear. A cognitive analysis of what an adult, competent speaker does is not modeled by a linguist's discovery procedures (though a child's acquisition may be).

Ride and write

Phoneme	Phone	Examples		Phone	Examples	
/aj/	[aj]	bide	[bajd]	[ʌj]	bite	[bʌjt]
		guide	[gajd]		kite	[kʌjt]
		tribe	[trajb]		tripe	[trʌjp]
		aisle	[ajl]		type	[tʌjp]
		smile	[smajl]		like	[lʌjk]
		mime	[majm]		Mike	[mʌjk]
		mine	[majn]		knife	[nʌjf]
		buy	[baj]		knife	[nʌjf]
		try	[traj]		nice	[nʌjs]
					slice	[slʌjs]

Phoneme	Phone	Context
/aj/	[ʌj]	_[voiceless obstruent]
	[aj]	elsewhere

Sam and sang

Phoneme	Phone	Examples	
/æ/	[æ]	cap	[kæp]
		cat	[kæp]
		pack	[pæk]
		nap	[næp]
		Nat	[næt]
		knack	[næk]
		gap	[gæp]
		bat	[bæt]
		back	[bæk]
		bang	[bæŋ]
		sang	[sæŋ]
		gang	[gæŋ]
	[eᵊ]	Sam	[seᵊm]
		tan	[teᵊn]
		Nam	[neᵊm]
		Nan	[neᵊn]
		dam	[deᵊm]
		Dan	[deᵊn]

Phoneme	Phone	Context
/æ/	[eᵊ]	_{m,n}
	[æ]	elsewhere

French vowels

French oral vowels

Height	Vowel	example	Vowel	example	Vowel	example
	Front unrounded		Front rounded		Back	
High	i	vie	y	du	u	tout
Mid: tense	e	blé	ö	peu	o	mot
Mid: lax	ɛ	tête	œ	peur	ɔ	donne
Low:					a	plat

Height	Vowel	example	Vowel	example	Vowel	example
	Front unrounded		Front rounded		Back	
Mid: lax	ɛ̃	plein	œ̃	brun*	ɔ̃	bon
Low:					ã	dans

French consonants

	labial	alveolar	alveo-palatal	palatal	velar	uvular	laryngeal
Voiceless stop	p	t			k		
Voiced stop	b	d			g		
Voiceless fricative	f	s	ʃ				
Voiced fricative	v	z	ʒ			ʁ	
Nasal	m	n		ɲ	ŋ		
Liquid		l					
Glide	w			j ɥ			

Spelling	English	Québécois	Continental
petit	small	ptsi	pəti
tiroir	drawer	tsirwaʁ	tirwaʁ
diable	devil	dʑjab	djablə
Adèle	(name)	adɛl	adɛl
terre	earth	tɛʁ	tɛʁ
tâche	task	ta:ʃ(or tawʃ)	taʃ
tout	all	tʊt	tu
il dit	he says	i dʑi	i(l)di
elle dit	she says	adʑi	ɛldi
planter	to plant	plãte	plãte
torchon	dish towel	tɔʁʃõ	tɔʁʃõ
tiens	hold, take	tsjẽ	tjẽ
tuer	to kill	tsye	tye
dur	hard	dʑyr	dyr
diable	devil	dʑjab	djablə
dans	in	dã	dã
petite image	small picture	ptsitima:ʒ	pətitimaʒ

Spelling	English	Québécois	Continental
tout, toute	all	tʊt	tu, tut
vous, nous	you, us	vu, nu	vu, nu
riz	rice	ri	ri
du	of the, some	ʤy	dy
vite	quickly	vɪt	vit
parler	to speak	paʁle	paʁle
Colette	(name)	kɔlɛt	kɔlɛt

Cashinahua, Panoan (Peru, Brazil)

[paka]	bamboo	[mɨsu]	swollen hand
[taka]	liver	mɨʃu	dark, black
[kaka]	type of basket	[bɨtu]	spotted face
[baka]	fish	[bɨtsu]	to be squeezed
[daka]	to rest	[bisu]	face
[tsaka]	to kill	[kutʃa]	type of arrow
[tʃaka]	bad	[kuʃa]	to hit
[maka]	rat	[naka]	to chew
[tapa]	floor	[taβa]	washboard
[tama]	peanut	[tawa]	sugar cane
[bɨɾu]	eye	[kuja]	to have pus
[kana]	type of macaw	[hana]	type of bird
[isi]	unending	[isa]	bird (generic)
[isu]	spider monkey	[isɨ]	head painting
[dani]	body hair	[baɾi]	sun
[ba]	friend	[naβu]	people
[iwɨɾiwɨ]	bring quickly!	[dasiβi]	all
[daɾɨ]	medicinal plant	[bɨnu]	duck-like bird
[ʃana]	type of fruit	[sɨpi]	weaving design
[hɨpɨ]	type of palm	[janɨ]	quickly

Cashinahua consonants

	Labial	Alveolar	(Alveo-)palatal	Velar/Glottal
Plosive voiceless	p	t		k
Plosive voiced	b	d		g
Affricate		ts	tʃ	
Fricative	β	s	ʃ	h
Nasal	m	n		
Flap		ɾ		
Glide	w		j	

Table 2: Cashinahua consonants

8

	Front	Central	Back
Close	i	ɨ	u
Open		a	

Table 3: Cashinahua vowels

Distribution

Two analyses are possible:

Phoneme	Phone	Context
/b/	[b]	#_
	[β]	elsewhere

Phoneme	Phone	Context
/b/	[β]	V_V
	[b]	elsewhere

Likewise, here is one of two possible analyses:

Phoneme	Phone	Context
/d/	[d]	#_
	[ɾ]	elsewhere

Zoque

Phoneme	Phone	Context
/p/	[b]	[nasal]_
	[p]	elsewhere

Phoneme	Phone	Context
/t/	[d]	[nasal]_
	[t]	elsewhere

b	β	d	ɾ
_itu	ta_a	_aka	bɨ_u
_aka	na_u	_ani	ba_i
_itsu	dasi_ i	_ asibi	iwɨ_iwɨ
_isu		_aɾi	da_ i
_aɾi			
_a			
_ inu			

Table 4: Voiced obstruents in Cashinahua

	gloss		gloss
pata	mat	ŋgyunu	you fell
tatah	father	sis	meat
tʸitʸɨy	little	šohšahu	they cooked it
cima	calabash	kama	cornfield
cehcu	he cut it	nas	earth
kunu	he fell	ñanah	his mother
kenba	he sees	kaŋ	jaguar
myaŋdamu	you came	liŋba	he slashes
ʔɨŋdʸoʔpya	he is sleepy	win	face
ñǰehcu	you cut brush		

Table 5: Zoque words

Voiceless plosives	p	t,c	tʸ, č	k	
Voiced plosives	b	d, dz	dʸ, ǰ	g	
Fricatives		s	š		
Nasals	m	n	ñ	ŋ	
Liquids		l,r			
Glides	w		y		ʔ, h

Table 6: Zoque consonant inventory