

## The System

Sounds

### Two goals

- The first is to put **language sound structure in context**.
  - Why do human languages have a sound structure about which we need to say anything more than “vocal communication is based on noises made with the eating and breathing apparatus”?
  - What are the apparent “design requirements” for this system, and how are they fulfilled?
- The second is to provide a **concrete sense of what the language sound systems are like**.
  - We will discuss, in a certain amount of detail, a few aspects of the phonetics and phonology of English and Hindi. Along the way, some of the terminology and theory of phonetics and phonology will emerge.

### Sound structure of language

Key topics are

- the anatomy, physiology, and acoustics of the human vocal tract;
- the nomenclature for the vocal articulations and sounds used in speech, as represented by the International Phonetic Alphabet;
- hypotheses about the nature of a sound's features and their organization into segments, syllables and words;
- the way that features like tone align and spread relative to consonants and vowels;
- the often-extreme changes in sound of morphemes in different contexts;
- the way that knowledge of language sound structure unfolds as children learn to speak;
- the variation in sound structure across dialects and across time.

### Apparent design features of human spoken language

We can list a few characteristics of human spoken languages:

- Large vocabulary: 10,000-100,000 items
- Open vocabulary: new items are added easily
- Variation in space and time: different languages and “local accents”
- Messages are typically structured sequences of vocabulary items

### Phonetics and Phonology

Formally, these are the two sub-disciplines concerned with speech and are mutually dependent.

- Phonetics describes the concrete, physical form of sounds (how they are produced, heard and how they can be described)
- Phonology is concerned with the function of sounds that is with their status and inventory in any given language.

### Compare with the vocalization of other primates

Referential vocabulary is

- Small: < 35 items
- Closed vocabulary: new “names” or similar items (novel combinations for new messages) are not added
- System is fixed across space and time: widely separated populations use the same signals
- Messages are usually single items, perhaps with repetition

## Similarities with other primates

Some general characteristics of other primate vocalizations that are retained by human speech:

- Vocalizations communicate individual identity
  - Vocalizations communicate attitude and emotional state
- Some potential advantages of the human innovations:
- Easy naming of new people, groups, places, etc.
  - Signs for arbitrarily large inventory of abstract concepts
  - Language learning is a large investment in social identity

## What makes it work?

- In subsequent use, you (and those who listen to you speak) need to distinguish this one word accurately from tens of thousands of others.
- (The perceptual error rate for spoken word identification is less than one percent: In judgement tasks where words are chosen at random and spoken by arbitrary and previously-unknown speakers. In more normal and natural contexts, performance is even better!)
- Let's call this the **pronunciation learning problem**. If every word were an arbitrary pattern of sound, this problem would probably be impossible to solve.

## How can it work?

Experiments on vocabulary sizes at different ages suggest that children must learn an average of more than 10 items per day, day in and day out, over long periods of time.

- A sample calculation:
  - 40,000 items learned in 10 years
  - 10 x 365 = 3,650 days
  - 40,000 / 3,650 = 10.96
- Most of this learning is without explicit instruction, just from hearing the words used in meaningful contexts. Usually, a word is learned after hearing only a few dozen examples. Experiments have shown that young children can often learn a word (and retain it for at least a year) from hearing just one casual use.



"Keep out! Keep out! K-E-E-P O-U-T."

## Learning the sound

Leaving aside the question of how to figure out the meaning of a new word, and focusing on how to learn its sound:-

- You only get to hear the word a few times -- maybe only once.
- You have to cope with many sources of variation in pronunciation: individual, social and geographical, attitudinal and emotional.
- Any particular performance of a word simultaneously expresses the word, the identity of the speaker, the speaker's attitude and emotional state, the influence of the performance of adjacent words, and the structure of the message containing the word.
- Yet you have tease these factors apart so as to register the sound of the word in a way that will let you produce it yourself, and understand it as spoken by anyone else, in any style or state of mind or context of use.

## The Phonological Principle

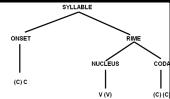
- In human spoken languages, the sound of a word is not defined directly (in terms of mouth gestures and noises). Instead, it is mediated by encoding in terms of a **phonological system**.
- A word's pronunciation is defined as a structured combination of a small set of elements
  - The available phonological elements and structures are the same for all words (though each word uses only some of them, *cat* vs *pit* vs *lenf*)
- The **phonological system** is defined in terms of patterns of mouth gestures and noises
  - This 'grounding' or "actualization" of the system is called *phonetic interpretation*
  - Phonetic interpretation is the same for all words

## How does the phonological principle help?

The pronunciation learning problem is split into two different problems, each one easier (individually) to solve.

- Phonological representations are digital, i.e. made up of discrete elements in discrete structural relations.
  - Copying can be exact: members of a speech community can share identical phonological representations (abstract)
  - Within the performance of a given word on a particular occasion, the (small) amount of information relevant to the identity of the word is clearly defined.
- Phonetic interpretation is general, i.e. independent of word identity
  - Every performance of every word by every member of the speech community helps teach phonetic interpretation, because it applies to the phonological system as a whole, rather than to any particular word.

## A model system – Islish



### • Islish has

- three vowels -- /a/, /i/, /u/ -- and every Islish syllable must contain one of these.
- seven consonants that can start syllables -- /p/, /t/, /k/, /b/, /d/, /g/, /s/
- syllable may also lack an initial consonant and may optionally end with the consonant /n/
- Islish thus has 48 possible syllables:
  - the syllable onset has 8 options (/p/, /t/, /k/, /b/, /d/, /g/, /s/ or null),
  - the syllable nucleus has 3 options (/a/, /i/, /u/)
  - the syllable coda has 2 options (/n/ or nothing)

(and  $8 \times 3 \times 2 = 48$ )

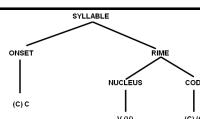
## Duality of patterning

- The linguist Charles Hockett (1960) used this phrase to describe this two-stage encoding of the basic elements of a communication system, in which a large set of elements that refer to concepts or to things in the world (e.g. words) are created as combinations of a small set of discrete elements (e.g. phonemes), which are meaningless except for their role in connecting words to concepts.
- André Martinet (1965) used the phrase "double articulation" for the same idea. A more contemporary description might say that the lexicon of human languages is *digitally encoded*.

## Words in Islish

- Islish words are made up of 1 to 4 syllables.
- There are **5,421,360** possible Islish words --
 
$$(48 \times 48 \times 48) + (48 \times 48 \times 48) + (48 \times 48) + (48) = 5,421,360.$$
- Thus the phonological elements of Islish, as we have described them, are /I/, /a/, /U/, /p/, /t/, /k/, /b/, /d/, /g/, /s/, /n/. The phonological structures of Islish include the notions of syllable, onset, nucleus, and coda.
- Example of Islish words: /kanpiuta/ "computer", /kaa/ "class", /pi/ "raging", /bata/ "dal and vegetables", /pikin/ "speech"
- What you won't find in Islish: /stind/, /spug/ etc.

## Syllable



- Human speech, like many animal vocalizations, tends to involve repetitive cycles of opening and closing the vocal tract.
- In human speech, we call these cycles *syllables*.
- A syllable typically begins with the vocal tract in a relatively closed position -- the syllable *onset* -- and proceeds through a relatively open *nucleus*.
- We have robust intuitions about syllables –
  - Hippopotamus (hi-po-po-to-mus, \*hip-op-o-tam-us)
  - floccinaucinihilipilification (flok-si-nok-si-ni-hi-li-pi-li-fi-ca-shan)

## Syllable structure of words

- In giving the phonological encoding of these words, we've omitted the structure, because it is unambiguously recoverable from the string of the elements present.
- For instance, /kanpiuta/ must be
  - a four-syllable word
  - The first syllable contains the onset /k/, the nucleus /a/, and the coda /n/, etc.
- kan-pi-u-ta

## Real languages

- Have more complex phonological systems than our made-up language does.
- However, it remains true that phonological structures are recoverable from strings of phonological elements, and therefore can be omitted for convenience in writing (transcribing).
- In this way of writing down phonological representations as strings of letter-like phonological elements, the “letters” are called **phonemes** (**abstract** representations of sounds).



## From phonemes to mouth gestures and noises (and back again)

- We've exemplified half of the situation: the “Islish” example explains what kind of thing a phonological system is, and how the pronunciation of words can be specified by “spelling” them in phonological terms (phonemes).
- What about the **phonetic interpretation** of words, that is, the interpretation of phonemic strings in terms of mouth gestures and the accompanying noises? How does that work?



The voice is articulated by the lips and the tongue. . . . Man speaks by means of the air which he inhales into his entire body and particularly into the body cavities. When the air is expelled through the empty space it produces a sound, because of the resonances in the skull. The tongue articulates by its strokes; it gathers the air in the throat and pushes it against the palate and the teeth, thereby giving the sound a definite shape. If the tongue would not articulate each time, by means of its strokes, man would not speak clearly and would only be able to produce a few simple sounds.

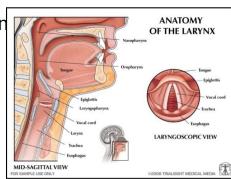
**HIPPOCRATES** (460–377 B.C.E.)

	অ আ ই ঈ উ ঊ	এ ঐ ও ঔ		
ক ক	ল	ং অং		
ক	খ	গ	ঘ	ঙ
চ	ছ	জ	ঝ	ঞ
ট	ঠ	ড	ঢ	ণ
ত	থ	দ	ধ	ন
প	ফ	ব	ভ	ম
য	ৰ	ল	ৱ	
শ	ষ	স	হ	

The preceding table lists eight groups of letters:  
(1) Vowel, (2) Kshemavart, (3) Moudri, (4) Vaish, (5) Vimala, (6) Jayini, (7) Sarvesvar, and (8) Kali or Kaulini. These are respectively the Prasanta Sudra of the following eight groups of letters.

## Basic sound production in the vocal tract: buzz, hiss and pop

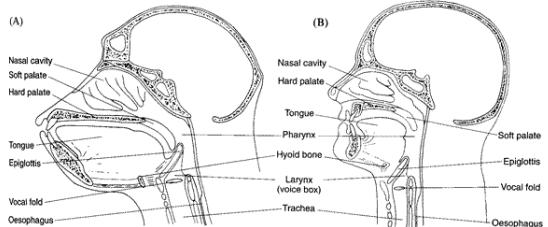
- There are three basic modes of sound production in the human vocal tract that play a role in speech:
- the buzz of vibrating vocal cords
- the hiss of air pushed past a constriction
- the pop of a closure released.



## Mechanism of this sound production

This very simple and general:

- As the air encounters the folds, the stream of air creates pressure. As this pressure builds up, it causes the vocal folds to blow apart.
- As the vocal folds blow apart, the air traveling through them speeds up.
- The acceleration of air in turn causes the pressure to drop at the **glottis** (or the space between the vocal folds).
- Because of the low air pressure, the vocal cords are sucked together and approximate/close again. (Bernoulli's principle)
- The folds stay together, of course, only until the pressure builds up again to blow the vocal folds back apart.
- The cycle repeats, as air pressure again forces an opening.
- In many such sounds, the pattern of opening and closing is irregular, producing a burp-like sound without a clear pitch. If the circumstances are right, a regular oscillation can be set up, giving a periodic sound that we perceive as having a pitch.
- Many animals have developed their larynges so as to be able to produce particularly loud sounds, often with a clear pitch that they are able to vary for expressive purposes.



## The larynx

### Vibration of the vocal folds

Mélanie Canault  
Olivier Rastello

Coordination : Patrice Thiriet  
ISTR – Lyon1 University

Rhône-Alpes

creative commons



## Laryngeal buzz

- The *larynx* is a rather complex little structure of cartilage, muscle and connective tissue, sitting on top of the *trachea*. It is what lies behind your "Adam's apple". The original role of the larynx is to seal off the airway, in order to prevent aspiration of food or liquid, and also to permit the thorax to be pressurized to provide a more rigid framework for heavy lifting and pushing.
- Part of the airway-sealing system in the larynx is a pair of muscular flaps, the *vocal cords* or *vocal folds*, which can be brought together to form a seal, or moved apart to permit free motion of air in and out of the lungs. When any elastic seal is not quite strong enough to resist the pressurized air it restricts, the result is an erratic release of the pressure through the seal, creating a sound.

James P. Thomas, MD  
[voicedoctor.net](http://voicedoctor.net)

## The hiss of turbulent flow

- Another source of sound in the vocal tract -- for humans and for other animals - is the hiss generated when a volume of air is forced through a passage that is too small to permit it to flow smoothly. The result is turbulence, a complex pattern of swirls and eddies at a wide range of spatial and temporal scales. We hear this turbulent flow as some sort of hiss.
- In the **vocal tract**, turbulent flow can be created at many points of constrictions. For instance, the lower teeth can be pressed against the upper lip -- if air is forced past this constriction, it makes the sound associated with the letter (and IPA symbol) [f].
- When this kind of turbulent flow is used in speech, phoneticians call it *fricative*, and sounds that involve frication are called *fricatives*.
- Sip [s], zip [z], Bach [χ], shoe [ʃ], azure [ʒ], [v] as in Urdu words for 'sorrow'

## Sound shaping by the vocal tract: vowel colour and nasality

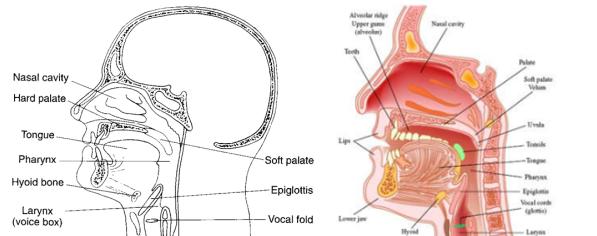
- Between the larynx and the world at large is about 15 centimetres of throat and mouth. This passageway acts as an acoustic resonator, enhancing some frequencies and attenuating others. The properties of this resonator depend on the position of the tongue and lips, and also on whether the *velum* or *soft palate* is lowered so as to open a side passage to the nasal cavity.

## The pop of closure and release

- When a **constriction** somewhere in the vocal tract is complete, so that air can't get past it as the speaker continues to breath out, pressure is built up behind the constriction. If the constriction is abruptly released, the sudden release of pressure creates a sort of a **pop**. When this kind of closure and release is used as a speech sound, phoneticians call it a *stop* or *occlusive* (focusing on the closure) or a *plosive* (focusing on the release).
- As with frication, a plosive constriction can be made anywhere along the vocal tract, from the lips to the larynx. However, it is difficult to make a firm enough seal in the pharyngeal region to make a stop, although a narrow fricative constriction in the pharynx is possible.

## Vowels

- Different positions of the tongue and lips make the difference between one vowel sound and another. As you can easily determine for yourself by experiment, you can combine any vowel sound with any pitch -- or with a whisper, which is a hiss created by turbulent flow at the vocal folds.



## Phonetic syllables: the scale and cycle of sonority

- Human speech, involves repetitive cycles of opening and closing the vocal tract or *syllables*
- The degree of vocal tract **openness** correlates with the loudness of the sound that can be made.
- Speech sounds differ on a scale of **sonority**, with vowels at one end (the most sonorous end!) and stop consonants at the other end. In between are fricatives, nasal consonants like [m] and [n], and so on.
- Languages tend to arrange their syllables so that the **least sonorous** sounds are restricted to the margins of the syllable - the onset in the simplest case - and the most sonorous sounds occur in the centre of the syllable.

## Sometimes syllabic position rules

- There are some cases where the same (or at least very similar) sounds can occur in several different syllabic roles.
- For example, the *glides* (sometimes called *approximants*) that begin syllables like "you" [ju] and "we" [wi] are almost exactly like vowels, except for their syllabic position.
- In fact, the mouth position and acoustic content of the "consonant" at the start of "you" and of the "vowel" at the end of "we" are just about exactly the same [j~-i and w~-u].

## The International Phonetic Alphabet

### • Bell's VISIBLE SPEECH

In the mid-19th century, Melville Bell invented a writing system that he called "Visible Speech". Bell was a teacher of the deaf, and he intended his writing system to be a teaching and learning tool for helping deaf students learn spoken language. However, *Visible Speech* was more than a pedagogical tool for deaf education - it was the first system for notating the sounds of speech independent of the choice of particular language or dialect. This was an extremely important step -- without this step, it is nearly impossible to study the sound systems of human languages in any sort of general way.

## [ju] [wi]

- In the International Phonetic Alphabet (IPA), the [j] stand for the sound we usually write as with the letter *y*. [u] refers to the vowel as in the word *pool*.
- The English word *we* would be written in the IPA as [wi], where the [w] is familiar from the spelling as well, and the [i] refers to the vowel found in *see, eat* etc.
- The articulation and sound of IPA [j] is like the articulation and sound of IPA [i], while the articulation and sound of IPA [w] is quite like that of IPA [u].
- What is different is the role in the syllabic cycle -- [j] and [w] are consonants, while [i] and [u] are vowels.

## Contd.

- In the 1860's, Melville Bell's three sons -- Melville, Edward and Alexander -- went on a lecture tour of Scotland, demonstrating the *Visible Speech* system to appreciative audiences.
- In their show, one of the brothers would leave the auditorium, while the others brought volunteers from the audience to perform interesting bits of speech -- words or phrases in a foreign language, or in some non-standard dialect of English. These performances would be notated in *Visible Speech* on a blackboard on stage.
- When the absent brother returned, he would imitate the sounds produced by the volunteers from the audience, solely by reading the *Visible Speech* notations on the blackboard. In those days before the phonograph, radio or television, this was interesting enough that the Scots were apparently happy to pay money to see it!

## Sanskrit examples

- मधु + एतत् = मधेतत्
- अति + उत्रतम् = अत्युत्रतम्
- पितृ + आज्ञा = पित्राज्ञा
- लृ + आकृतिः = लाकृतिः

इति + आदि = इत्यादि  
 देवी + आवाहन = देव्यावाहन  
 सु + आगत = स्वागत  
 वधू + आगमन = वथ्वागमन  
 पितृ + आदेश = पित्रादेश

## Phonetic notation for elocution lessons - and for linguistic description



- After Melville Bell's invention, notations like *Visible Speech* were widely used in teaching students (from the provinces or from foreign countries) how to speak with a standard accent. This was one of the key goals of early phoneticians like Henry Sweet (said to have been the model for Henry Higgins, who teaches Eliza Doolittle to speak "properly" in Shaw's *Pygmalion* and its musical adaptation *My Fair Lady*).
- The International Phonetic Association (IPA) was founded in 1886 in Paris, and has been ever since the official keeper of the **International Phonetic Alphabet** (also IPA), the modern equivalent of Bell's *Visible Speech*. Although the IPA's emphasis has shifted to a more descriptive direction, there remains a lively tradition in Great Britain of teaching standard pronunciation using explicit training in the IPA.

The following are the Physiological Symbols for the English elements of Speech.

**CONSONANTS.**

p	p	pie	pea	lower-case p
t	t	tie	tea	lower-case t
k	k	kye	key	lower-case k
b	b	by	bee	lower-case b
d	d	dys	D	lower-case d
g	g	guy		lower-case g
m	m	my	me	lower-case m
n	n	nigh	knee	lower case n
ŋ	ŋ		ram	eng (or augma)
f	f	fe	fee	lower-case f
v	v	vie	V	lower-case v
θ	θ	thigh		theta
ð	ð	thy	thee	
s	s	sigh	sea	listen
z	z		Z	lower-case z
t (t̪)	t (t̪)	shy	she	mission
z (d̪)	z (d̪)			long z (or yugh)
l	l	lie	lee	lower-case l
w	w	why	we	lower-case w
r (r̪)	r (r̪)	rye	rye	lower-case r
j (y)	j (y)		ye	lower-case j
h	h	high	he	lower case h v

**VOWELS.**

ɛ ee in eel.	ɪ i in ill.	ʊ e in sheet.	ʌ a in shall.
ɔ oo in pool.	ə u in pull.	ɑ a in all.	ɒ o in doll.
ɑ a in father.	ʌ u in a sort.	ʊ u in curl.	ʊ u in cur.
			ʊ u in boy.

**GLIDES.**

ɹ w as in now.	ɹ r as in stir.	ɹ y as in may.	ɹ ʌ a as in near.
ɹ ʌ i in mine.	ɹ ʌ a in man.	ɹ ɔ əw in now.	ɹ ɔ əw in know.

**DIPHTHONGS.**

ɹ ʌ ɔ əw in now.	ɹ ɔ əw in boy.
------------------	----------------

**Illustration of the Physiological Alphabet.**

At the International Exhibition of 1876 in Philadelphia the Commissioners of the International Exhibition of 1876 have arranged an exhibition of the English Language, which is to be conducted by Professor A. Metzler, Prof. Dr. Brundum, and others.

॥	॥	॥
p	pie	pea
t	tie	tea
k	kye	key
b	by	bee
d	dys	D
g	guy	
m	my	me
n	nigh	knee
ŋ		ram
f	fe	fee
v	vie	V
θ	thigh	
ð	thy	thee
s	sigh	sea
z		listen
t (t̪)	shy	she
z (d̪)		Z
l	lie	lee
w	why	we
r (r̪)	rye	rye
j (y)		ye
h	high	he

Note also the following:

tʃ (t̪ʃ) chime

dʒ (d̪ʒ) jive

lower-case p

lower-case t

lower-case k

lower-case b

lower-case d

lower-case g

lower-case m

lower case n

eng (or augma)

lower-case f

lower-case v

theta

eth

lower-case z

lower-case ʐ

ah (or long ə)

long ʐ (or yugh)

lower-case l

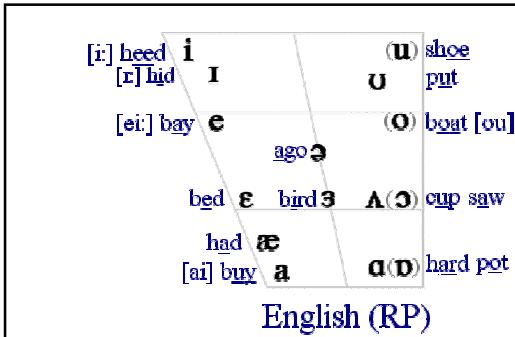
lower-case w

lower-case r

lower-case j

lower case h v

অ	আ	ই	ঝ	ত	ঞ
এ	ো	ো	ো	ো	ো
ক	খ	গ	ঘ	ঢ	ঙ
চ	ছ	জ	ঝ	ঢ	ঙ
ট	ঠ	ঢ	ঠ	ঢ	ঙ
ত	থ	দ	ধ	ন	
প	ফ	ব	ভ	ম	
য	ৱ	ল	ৱ		
ৱ	ষ	স	হ		



	Bilabial	Labio-dental	Dental	Alveolar	Post-alveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Epi-glottal	Glottal
Nasal	m	nj		n	ɳ	ɳ	ɲ	ŋ	ɳ	ɳ	ɳ	ɳ
Positive	p	b		t	d	t̪	d̪	c	j	k	g	g̪
Fricative	ɸ	β	f	v	θ	ð	s	z	ʂ	ʐ	χ	χ̪
Approximant			ɹ		ɹ̪		ɹ̪				h	h̪
Trill	B			Γ				R				
Tap or flap		V		ɾ		ɾ̪						
Lateral fricative			ɬ									
Lateral approximant				ɻ								
Lateral flap				ɺ								

I	ি	ি	ে	ু	ূ	া	়	ঽ	ষ	ঽ	ঽ	ঽ
i	heed	he	bead	heat	keyed	lower-case i						
t	hid	bid	bit	kid	kid	small capital I						
ɛɪ	hayed	hay	boyed	hate	Cade	lower-case e						
e	head	bad	bed	cad	epilon							
æ	had	bad	hat	cad	ash							
ɑ	hard	board	heart	card	script a							
ə	had	bad	hot	cod	(2) turned script a							
ɔ	hawed	haw	bowd	cawed	open o							
əʊ	hoed	hoe	boode	code	code							
u	who'd	whoed	hoot	coored	lower-case u							
ʌ	Hudd	bud	but	curd	turned v							
ɔɪ	herd	ber	bird	burt	reversed epsilon							
ai	hide	high	bide	height	lower case a (+ i'							
əʊ	how	bowed			(as noted above)							
ɔɪ	(əhooy	Boyd			(as noted above)							
ɛr	here	beard			(as noted above)							
ər	hair	bared			cared							
ən	hired	hire			(as noted above)							
ju	laud	hue	Bude	cued	(as noted above)							

अंत – Ant – End  
 पंख – Pankh – Feather  
 सुंदर – Sundar – Beautiful  
 पांच – Paanch – Five  
 हिन्दी – Hindi – Hindi  
 मैं – Main – I  
 आदमियों – Aadmiyon –  
 Men  
 महिलाओं – Mahilaon –  
 Women

हँसना – Hamsna – to laugh  
 हाँ – Haan – Yes  
 घडियां – Ghadiyan – Watches  
 चांद – Chand – Moon  
 कहाँ – Kahan – Where  
 पांव – Paany – Feet  
 थाइलैंड – Thailend – Thailand  
 भाषाएँ – Bhashayean – Languages  
 Nasalized sound for ऊ, ऊ, ण, न, म  
 अण्डा or अंडा – Anda – Egg  
 लम्बा or लंबा – Lamba – long, tall