Indian Knowledge System

Foundational Concepts for Science and Technology

UNIT 2

By:

Mr. Hirdesh Sharma

JIMS, Greater Noida

Linguistics (भाषा विज्ञान)

- Linguistics is a branch of language research that provides a scientific study of a language.
- Language has been the most effective tool for our communication since time immemorial.
- The advancement of knowledge and collaborative working require a common method of communication.
- Language plays this role in a civilized society.

Components of a Language

- Language is a tool used by everyone in a community and it is very difficult to maintain it unchanged.
- Communication is a key to trade, science and technology and societal progress.
- The respective part of a language deals with the ability of an individual to receive language inputs from multiple sources.
- The productive part of a language is to transmit back to others for their consumption.

Components of a Language

- Receptive Skills
    i. Listening (sound)
    ii. Reading (script)
- Productive Skills
    i. Speaking (sound)
    ii. Writing (script)

Panini’s Work on Sanskrit Grammar

- Panini composed 3983 rules to accommodate all the patterns and variations in Sanskrit language.
- The basic approach of Panini and its distinguishing features make Sanskrit a powerful language and eternal in its appeal.
- One of the Vedangas known as Vyakarana focuses on linguistics and phonetics (भाषाविज्ञान और ध्वन्यात्मकता) aspects of Sanskrit language.
- Astadhyayi is considered a fine creation of human intelligence and the best available descriptive model of a language.

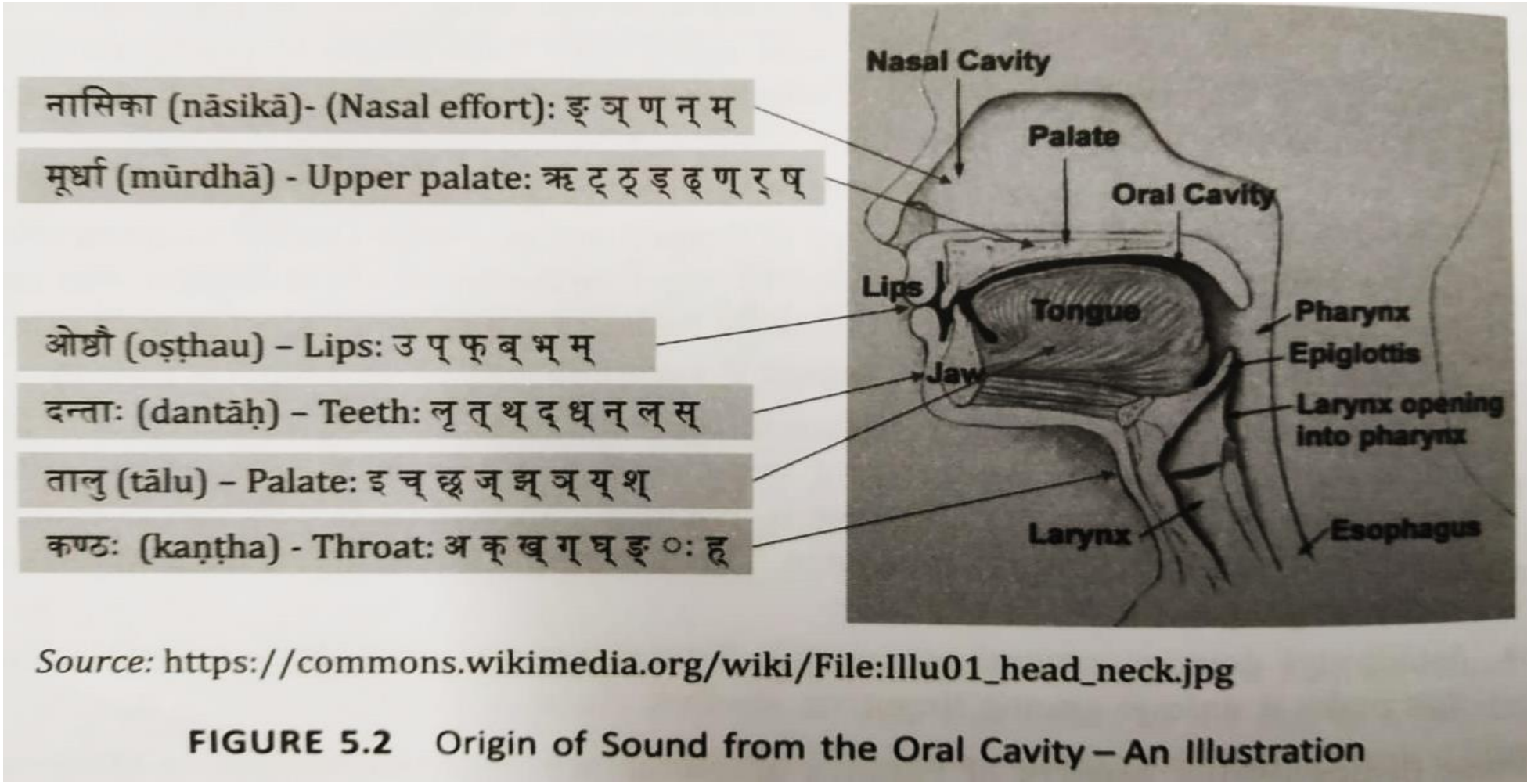
Basic approach of Panini and its features

- The entire vocabulary of the Sanskrit language could be created using the 3983 rules.
- Language processing and word generation are strictly rule-based and derivative in nature.
- The entire scheme for word generation follows a highly modular approach.
- The derivation of words using the rules could be done using step by step process.
- The vocabulary is not fixed or static.

Phonetics (स्वर-विज्ञान) in Sanskrit

- Phonetics is the study of sounds in a language, particularly the production of sound in a language and how it communicates the language corresponding to the scripts of the language.
- It also addresses the issue of how the sound is perceived in the language.
- Phonetics in the Sanskrit language has been addressed in some details since this is vital because the ancient Indian knowledge tradition is oral.

Phonetics (स्वर-विज्ञान) in Sanskrit



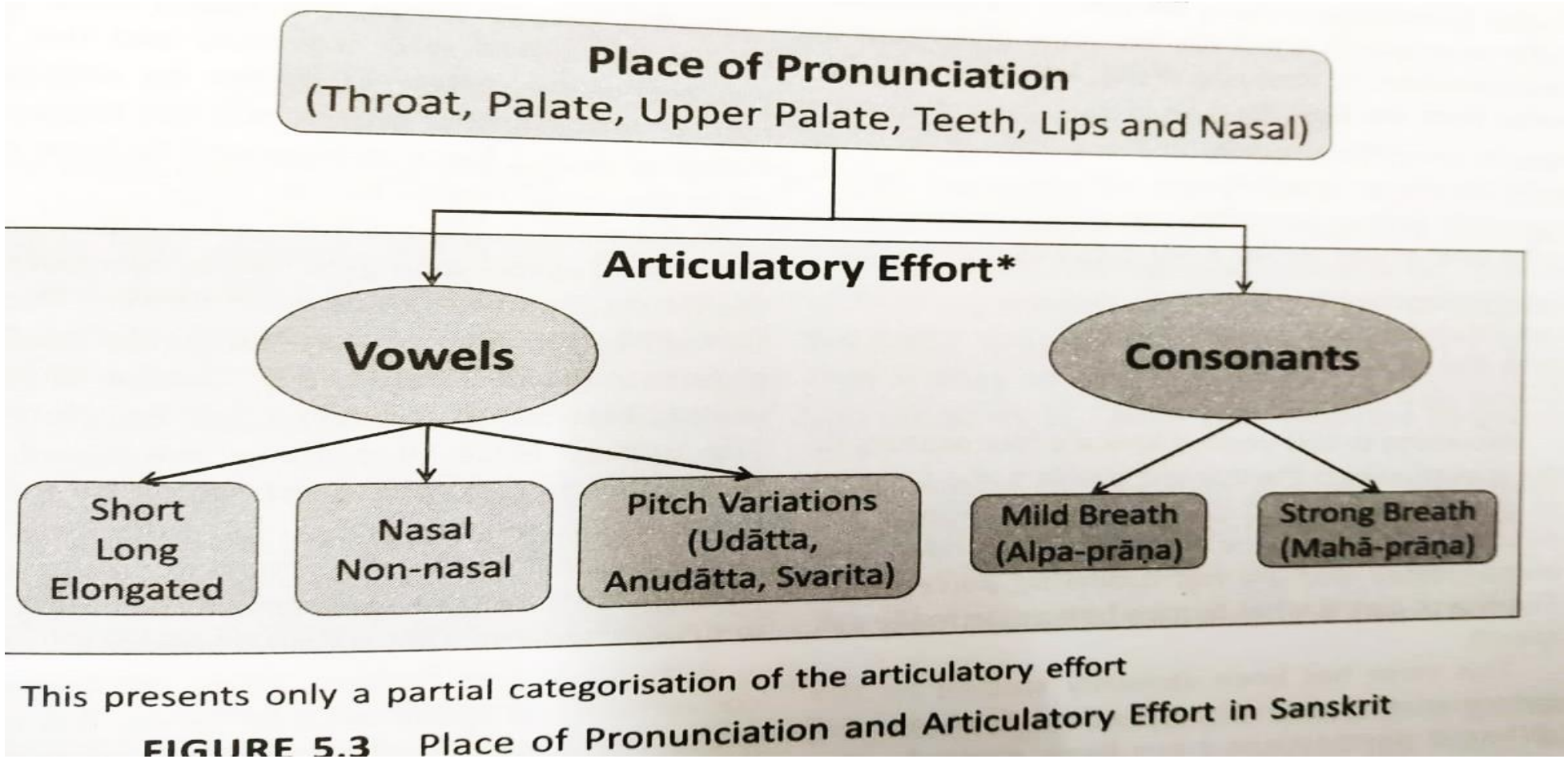
Source: https://commons.wikimedia.org/wiki/File:Illu01_head_neck.jpg

FIGURE 5.2 Origin of Sound from the Oral Cavity – An Illustration

Phonetics (स्वर-विज्ञान) in Sanskrit

- The entire transmission of the Vedas from time immemorial has been possible on account of a well-developed science of phonetics.
- Vowels have a temporal factor in the production of the sound.
- Three variations (short: hrasva, long: dirgha and prolate: pulta) have been specified for the pitch of the vowel sound.

Phonetics (स्वर-विज्ञान) in Sanskrit



This presents only a partial categorisation of the articulatory effort

FIGURE 5.3 Place of Pronunciation and Articulatory Effort in Sanskrit

Patterns in Sanskrit Vocabulary

- The ultimate building block of any language is the word.
- The words are the combined in several ways to communicate ideas and transact knowledge.
- The word in Sanskrit can be divied into two categories: Noun forms (Known as subanta) and verb forms (known as Tinanta).

Word= Base+ Suffix

- After adding a suffix to the base, relevant grammatical rules are invoked to generate the final word.

Patterns in Sanskrit Vocabulary

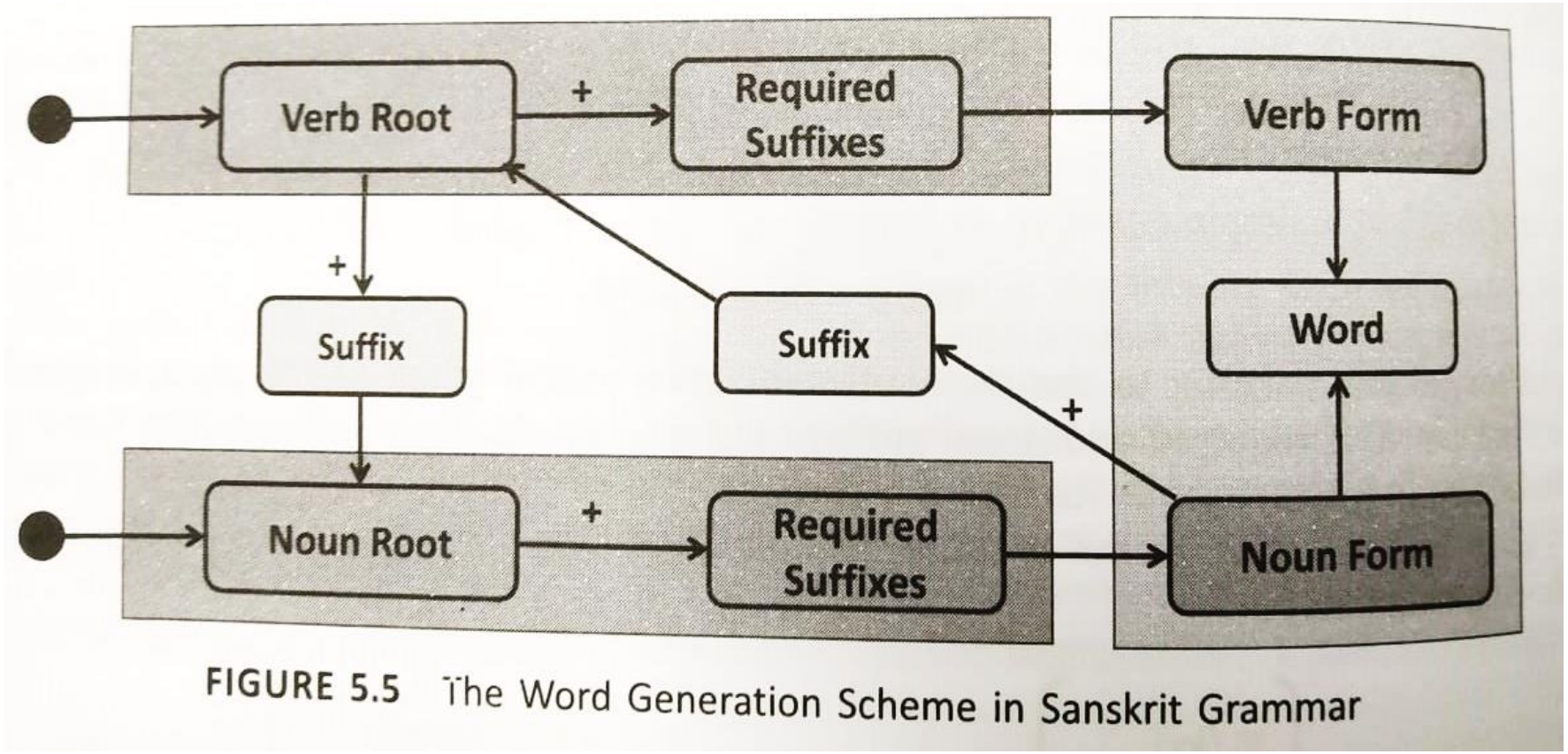


FIGURE 5.5 The Word Generation Scheme in Sanskrit Grammar

Computational concepts in Astadhyayi

- Sanskrit grammar has a robust mechanism to generate an infinite number of words.
- The following are some of the important common aspects seen in a computer language and Panini's rules:
  i. Vocabulary exclusively meant for his work
  ii. Abbreviated forms (mnemonics) for brevity and better retention of ideas
  iii. Exclusive syntax for Astadhyayi
  iv. An algorithmic approach to word generation
  v. Recursive logic

Computational concepts in Astadhyayi

TABLE 5.1 The Word Generation Scheme in Sanskrit Grammar – Examples

Base to Generate a Word	*Role of the Suffix (pratyaya)*	*Examples*	*Remarks*
Nominal Root	For generating singular, dual and plural of seven cases of nouns the relevant suffixes are applied.	*For the nominal root* 'राम' (Rāma), we can generate: रामः – रामौ – रामाः *(Rāmaḥ – Ramau – Rāmāḥ)* रामेण – रामाभ्यां – रामैः *(Rāmena – Rāmābhyām – Rāmaiḥ)*, etc.	Additional suffixes for generating feminine forms can be added.
Verb Root	For generating singular, dual and plural of 1st person, 2nd person and 3rd person of verbs the relevant suffixes are applied.	*For verbal root* 'पठ्' (paṭh) the present tense forms can be generated: पठति – पठतः – पठन्ति (paṭhati – paṭhataḥ – paṭhanti) पठसि – पठथः – पठथ (paṭhasi – paṭhathaḥ – paṭhatha) पठामि – पठावः – पठामः (paṭhāmi – paṭhāvaḥ – paṭhāmaḥ)	Relevant suffixes for generating 10 verb forms (6 tenses and 4 moods) can be added.

Mahesvara Sutra

- The entire Sanskrit grammar of Panini rests on a fundamental set of sutras known as Mahesvara Sutras.
- These sutras, 14 in number, present letters of Sanskrit uniquely.
- The first four sutras covers the alphabets in the normal order.
- Sutras 5 to14 present the consonants in a somewhat obscure order than what are they normally used for.
- Each sutra end with a termination which is a consonant.

Use of Mnemonics

- In the previous discussion, we saw how the letters were jointly represented using mnemonics obtained out of combining set of letters specified through the Mahesvara sutra.
- Two set of suffixex for generating noun forms and verb forms; these two are concisely represented by two mnemonics sup and tin respectively.

Rule-based Grammatical Operations

- Panini's system of applying grammatical conditions to derive words exactly like a rule-based engine.
- Sutras from different locations of Astadhyayi under various heading come to operate where their conditions are satisfied.
- The final form remaining after all the operations are carried out is the word and it becomes eligible for use.

Rule-based Grammatical Operations

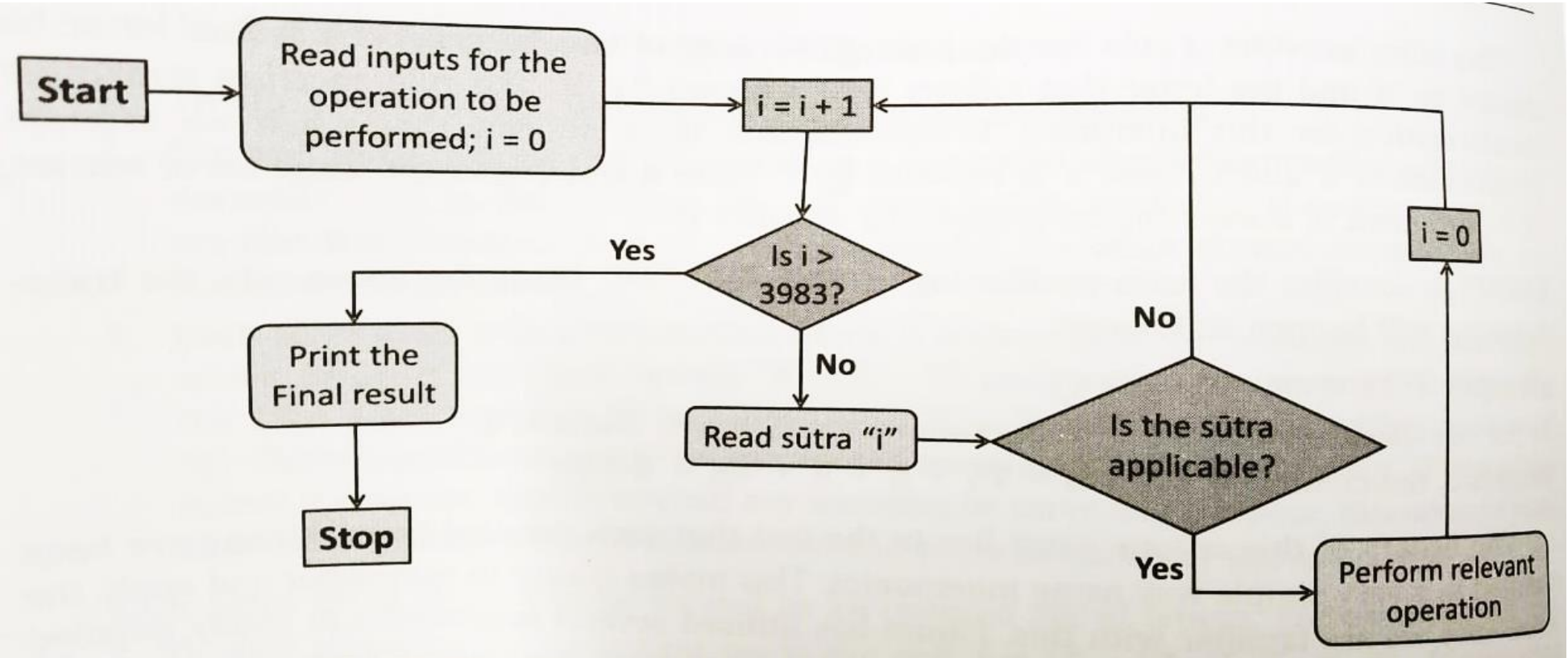


FIGURE 5.6 The Pāṇinian Algorithm for Grammatical Operations

Logic for Sentence Construction

- Words are to be grammatically correct in terms of construction.
- Karaka is a concept that helps to link the words in a sentence to the Kriya (action).
- Kriya and Karaka are the essential elements of any sentence.
- All other elements that are found in a sentence are woven around these two.
- Therefore, for a sentence to be complete there must be a verb, implicit or explicit, denoting an action.

Logic for Sentence Construction

- On the other hand, a verb alone cannot make a meaningful sentence.
- The Sanskrit language uses a concept called karaka to provide in-built mechanisms for constructing unambiguous and grammatically correct sentence.
- A participant involved in the action in some manner is called Karaka.

Logic for Sentence Construction

TABLE 5.2 Issues in Sentence Formation – An Illustration

Sl. No.	*Sentence in English*	*Sentence in Sanskrit*
1	The **fat boy** eats the *tasty food* with the hand	स्थूलः बालकः स्वादु भोजनं हस्तेन खादति । **sthūlaḥ bālakaḥ** *svādu bhojanaṃ* hastena khādati
2	The **fat** hand eats the *tasty food* with the **boy**	स्थूलः हस्तेन खादति स्वादु भोजनं बालकः । **sthūlaḥ** hastena khādati *svādu bhojanaṃ* **bālakaḥ**
3	The **fat** *food* eats the tasty hand with the **boy**	स्थूलः भोजनं खादति स्वादु हस्तेन बालकः । **sthūlaḥ** *bhojanaṃ* khādati *svādu* hastena **bālakaḥ**
4	The *food tasty* eats the **fat** hand with the **boy**	स्वादु भोजनं खादति स्थूलः हस्तेन बालकः । *svādu bhojanaṃ* khādati **sthūlaḥ** hastena **bālakaḥ**
5	The *tasty* **boy** eats the **fat** *food* with the hand	स्वादु बालकः खादति स्थूलः भोजनं हस्तेन । *svādu* **bālakaḥ** khādati **sthūlaḥ** *bhojanaṃ* hastena

Logic for Sentence Construction

- In order to under this, let us consider a sentence,
- “The technician removes the machine form the office in the morning with a truck.”
- This sentence can be written in Sanskrit as, “yantrakarakah pratah kale yantram vahanena karyalayat apakaroti.”
- The kriya in this sentence is ‘removing’; therefore, every karaka will create a direct link to the kriya.

Logic for Sentence Construction

- The six Karaka are related to vibhakits (cases) in Sanskrit. The six karaka (and the corresponding cases) are as follows:
  i. Kartr- doer: one in whom the cause of action is resident. (first case)
  ii. Karma- the focus of the result of an action (second case)
  iii. Karana- instrument: That which aids in the attainment of the action. (third case)
  iv. Sampradana- receiver: That with which the karma desires to get associated. (fourth case)
  v. Apadana- reference point in separation: That which has ability to create division. (fifth case)
  vi. Adhikarana- the locus of the action: That which provides the substratum, context or references to performed through kartr or karma. (sixth case)

Logic for Sentence Construction

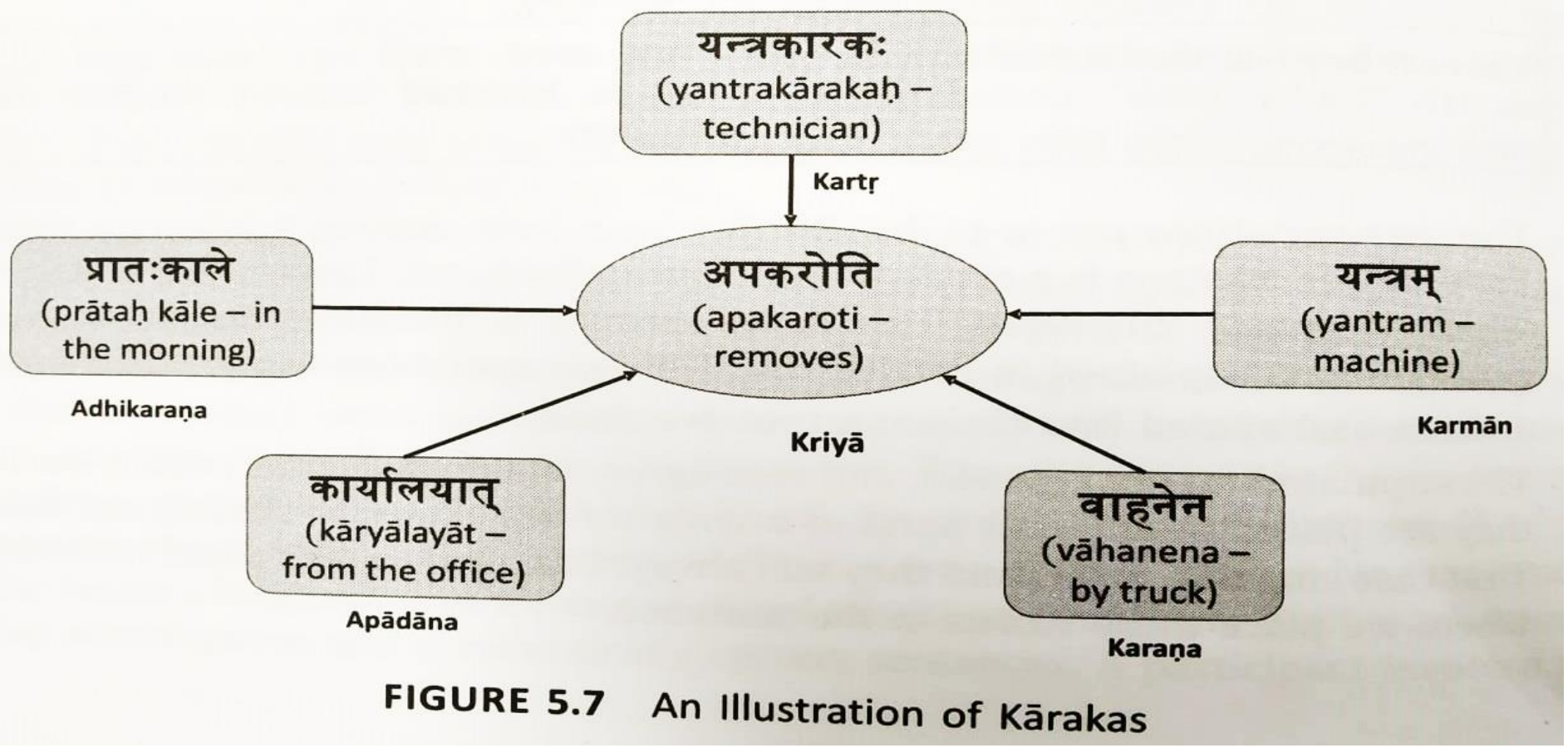


FIGURE 5.7 An Illustration of Kārakas

Importance of Verbs

- Language is required when we are in action. If no one is engaged in action, there is no need for language.
- The importance of verbs in further amplified by the fact that several of the noun roots are also derived from the verb roots only.
- Since most words (both verb forms and noun forms) originate from the verb roots (dhatus).
- In Sanskrit, we often find several synonyms for a word; each synonym for a word derived from dhatu.

Prefixes for Verb Forms

- Prefixes, known as upa-sargas, are appended to the verb forms in order to create additional words.
- There are 22 prefixes and one or more of these could be prefixed to a verb form.
- By adding the prefixes, it is possible to express the meaning in many ways.

Prefixes for Verb Forms

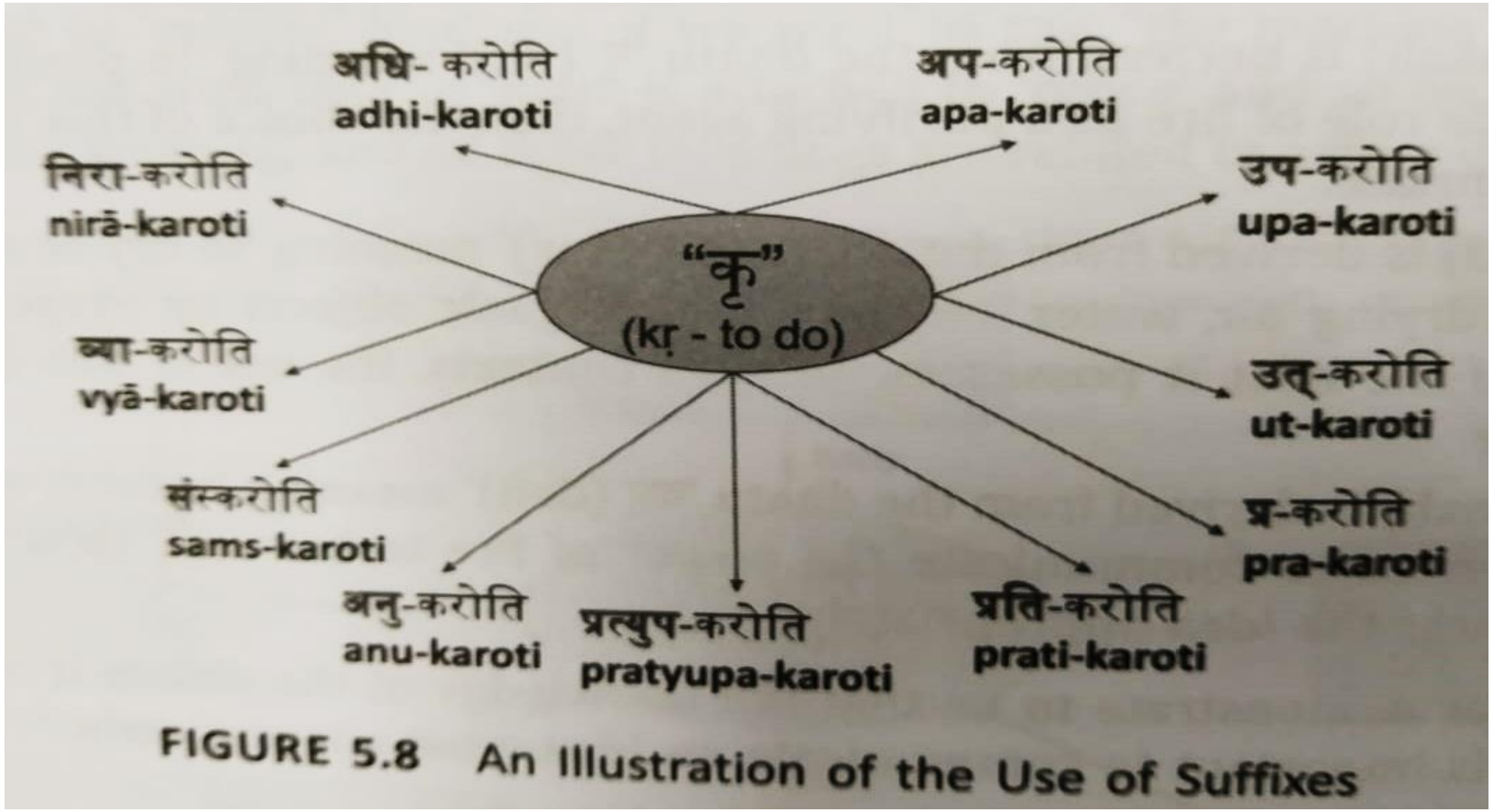


FIGURE 5.8 An Illustration of the Use of Suffixes

Role of Sanskrit in Natural Language Processing

- Natural Language Processing (NLP) is a branch of linguistics mainly concerned with processing of natural language data using computers and programming techniques.
- The principles of karaka and vibhkti will enable us to distill the components and the encoded information in a sentence.
- Indian linguists have described fourteen determiners to fix the meaning of a word in case of multiple of a word in case of multiple meaning.

Role of Sanskrit in Natural Language Processing

बाल	सुँ	वृक्ष	स्य	फल	अम्	खाद्	ति
bāla	Sũ	Vṛkṣa	Sya	phala	am	khād	Ti
boy	-	Tree	of	fruit	karma	To eat	Kartṛ
Noun base	1st case suffix	Noun base	6th case suffix	Noun base	2nd case suffix	Verb base	Present tense suffix

THANKS