Sukhada
Indian Institute of Technology
(BHU) Varanasi, Uttar Pradesh,
India

Soma Paul
International Institute of
Information Technology (IIIT),
Hyderabad, Telangana, India

Theory of *sāmarthya* in Indian grammatical tradition: The foundation of universal semantic representation

Sukhada and Soma Paul

Abstract
Semantic Representations become useful resources for various multilingual NLP applications such as Machine Translation, Multilingual Generation, cross Lingual QA, to name a few. Universal Semantic Representation (USR) is a recently developed semantic representation system that is based on Indian Grammatical Tradition and Paninian Grammar. This paper critically examines the notion of *sāmarthya* from Indian Grammatical Tradition, proposes that the principles of *sāmarthya* can account for the well-formedness of integrated word-sentential-discourse structure, and finally explains how the theory of *sāmarthya* motivates the multi-layered information encoding strategy of USR.

Keyword: Sāmarthya, padavidhi, saṃgati, vākya, mahāvākya, Indian Grammatical Tradition, Universal Semantic Representation

1. Introduction
Indian Grammatical Tradition (IGT henceforth) views language as a holistic phenomenon. Words are not derived as isolated units in Pāṇini's grammar, but as units that are semantically connected with other words in the sentence (Raster 2015). Thus, every word in a sentence contributes meaning in two ways: (i) the concept it represents and (ii) the compositional participatory role it plays in the sentence in relation to other words. The compositionality is expressed by certain linguistic cues, such as nominal inflection or the prepositional or postpositional markers attached. In natural languages, information encoding is not limited to the word or sentential or propositional level, but it goes beyond the sentential level as sentences are connected across the discourse. This is explicitly recognized by Pāṇini's rule (A 2.1.1): *samarthaḥ padavidhiḥ* [1] which is discussed in detail in section 2. The examination of this rule is very significant in this paper because we have adopted it as one of the primary guiding principles for our newly introduced IGT-based semantic representation, the Universal Semantic Representation (USR).

The fundamental component of the content specified in the semantic representation of texts is argument structure – who did what to whom, where, when, why, i.e., events, their participants and relations between them (Abend and Rapparport, 2017). However, in *communication*, speakers express how they view the situation which the mere argument structure of events can never capture. For example, let us consider the verb *jā* 'go'. The argument structure of this verb allows a goer, the destination and the mode/means of 'going'. So, we can say *hari apanī gādī se goā gayā* 'Hari went to Goa in his car'. Now, the speaker can decide to negate the event or emphasize the going of Hari or going in his own car and so on. Accordingly, the speaker will form different sentences which do not merely represent the argument structure of the verb. For example, those sentences can be *hari apanī gādī se hī goā gayā* 'It is in his own car that Hari went to Goa', *hari śāyada apanī gādī se goā gayā* 'Hari might have gone to Goa in his own car', *hari goā nahīṃ gayā.*

Corresponding Author:
Sukhada
Indian Institute of Technology
(BHU) Varanasi, Uttar Pradesh,
India

[1] An operation on words (takes effect only) when the words are semantically connected.

'Hari did not go to Goa'. Thus what is expressed in communication is *vivakṣā*: the intention of the speaker about the meaning to be conveyed by the words (Garg *et al.*, 2023; Sukhada *et al.*, 2023) [9, 18]. One interesting study is to examine the interplay of *samarthaḥ padavidhiḥ* and *vivakṣā* - the former constrains or licenses verbal Expressions and the latter belongs to the domain of the speaker's thought/idea. In this regard, Bhartṛhari's observation is very insightful:

Citrasyaikasya rūpasya yathā bhedanidarśanaiḥ
nīlādibhiḥ samākhyānaṃ kriyate bhinnalakṣaṇaiḥ
tathaivaikasya vākyasya nirākāṅkṣasya sarvataḥ
śabdāntaraiḥ samākhyānaṃ sākāṅkṣairanugamyate
(vākyapadīyam, 2.8-9)

According to Bhartṛhari, the process of conveying an idea to a listener or reader involves three stages. He uses the analogy of a painter to describe these stages step by step. For example, when a painter paints a picture,
1. He has a vision of the whole picture he wants to paint.
2. He observes the different parts of the figure or the objects separately and individually.
3. He finally paints the picture part by part.

Similar to a painter, there exists a single unified idea/thought in the mind of the speaker or the author during verbal communication. He has a vision of how to communicate the entire idea to the listener or the reader - his *vivakṣā*. Accordingly, he utters or writes it part by part. These parts are not isolated; on the contrary, they are sewn by the principles of *sāmarthya* that contribute to a well-formed expression of the idea.

In this paper, we study, in detail, how the concept of *sāmarthya* integrates the representation of the meaning of the lexical, sentential and discourse levels in the USR.

Section 2 presents the theoretical understanding of *sāmarthya* as explained in IGT. Section 3 proposes the notion of *sāmarthya* motivating information encoding at lexical, syntactico-sem- antic and discourse levels. Section 4 explains how such information is represented in a stratified manner in USR making it a compact representation system of textual meaning.

2. Sāmarthya
Panini's sūtra 2.1.1 (samarthaḥ padavidhiḥ [2]) says that wherever a rule related to a *padavidhi* is called for a grammatical operation, it should be applicable only to the words that are *samartha*, i.e. to the words that are related to each other, the words that have direct mutual connection with each other. According to Pāṇini, no grammatical operation can take place, be it word formation [3], compound formation or sentence formation, unless they qualify the condition of being *samartha* [4] (Joshi, 1968; Sastri, 1957; Mahavir, 1968).

Thus, the concept of *sāmarthya* is a fundamental principle for any grammatical operation on a language string. The words *samartha* and *sāmarthya* are used interchangeably in Indian Grammatical Tradition. There are two types of *sāmarthyas*: 1.) *ekārthībhāva sāmarthya* (single integrated meaning), and 2.) *vyapekṣā sāmarthya* (meaning-interdependence).

2.1 *Ekārthībhāva sāmarthya* (Single Integrated Meaning)
When the words having direct semantic connection become one word, as in compounds and derivational morphology, they are called to have *ekārthībhāva sāmarthya*. The word *samartha* here means *saṃgatārthaḥ samarthaḥ* (capable in the sense of *saṃ* 'with, together with, together' (Apte, p. 1628), + *gata* 'to signify, denote, convey an idea or sense of' (Apte, p. 648)) and *saṃsṛṣṭārthaḥ samarthaḥ* (capable in the sense of *saṃ* + *sṛṣṭa* 'connected, joined' (Apte, p. 1701)) (Mahabhāṣya, 2.1.1). The objective of *ekārthībhāva sāmarthya* is to present the words, derived through one of the five vṛttis [5, 6] as one pada (eka-pada) or as a single unit. For example, the compositional compounds like *rājapuruṣaḥ* (king-man) is derived from *rājñaḥ* (king's) + *puruṣaḥ* (man), *yudhiṣṭhiraḥ* (one who is always stable in the battle) from *yudhi* (in battle) + *ṣṭhiraḥ* (stable), *bird cage* from *cage for birds*. Non-compositional compounds such as *kṛṣṇasarpaḥ* (cobra), blackboard form one unit and denote a single integrated meaning.

2.2 *Vyapekṣā sāmarthya* (Meaning-Interdependence)
In the *vyapekṣā sāmarthya*, the word *samartha* means seen together (*saṃprekṣitārthaḥ samarthaḥ*) and bound together (*saṃbaddhārthaḥ samarthaḥ*) (Mahābhaṣya 2.1.1). The word *vyapekṣā* literally means "special expectancy" (*viśiṣṭā apekṣā = vyapekṣā*) with other words. That is why Bhojaraja in his Sṛmgāraprakāśa, Kaiyata in his Mahābhāṣyapradīpa, Haradatta in his Padamaṃjarī and Jñānendra Sarasvatī in his Tattvabodhinī have further explained *vyapekṣā* as the relationship between the words in a sentence when they leave their individual meaning due to mutual interdependence [7]. For instance, the words denoting subject, verb, object, etc. are seen as mutually semantically bound together in a sentence to convey a unique meaning in the sentence [8, 9]. That is why, be

---

[2] yaḥ kaścidiha śāstre padavidhiḥ śrūyate sa samartho viditavyaḥ | vidhīyate iti vidhiḥ | padānāṃ vidhiḥ padavidhiḥ | sa punaḥ samāsādiḥ | (Kashika 2.1.1)

[3] padasya vidhiḥ padavidhiḥ

[4] ekaśeṣanirdeśādvā || atha vā ekaśeṣanirdeśoyam | samarthasya ca samarthayośca samarthānāṃ ca samarthānāmiti || (ākṣepabhāṣyam) evamapi ṣaṭprabhṛtīnāmeva prāpnoti | ṣaṭprabhṛtiṣu hyekaśeṣaḥ parisamāpyate || (samādhānabhāṣyam) naiṣa doṣaḥ | pratyekaṃ vākyaparisamāptirdṛṣṭeti dvyekayorapi bhaviṣyati || (ākṣepabhāṣyam) evamapi vibhaktīnāṃ na prāpnoti - samarthātsamarthe padātpada iti | (samādhānabhāṣyam) evaṃ tarhi samarthapadayorayaṃ vidhiśabdena sarvavibhaktyantaḥ samāsaḥ -

---

samarthasya vidhiḥ samarthavidhiḥ, samarthayorvidhiḥ samarthavidhiḥ, samarthānāṃ vidhiḥ samarthavidhiḥ, samarthādvidhiḥ samarthavidhiḥ, samarthe vidhiḥ samarthavidhiḥ| padasya vidhiḥ padavidhiḥ, padayorvidhiḥ padavidhiḥ, padānāṃ vidhiḥ padavidhiḥ, padādvidhiḥ padavidhiḥ, pade vidhiḥ padavidhiḥ| samarthavidhiśca samarthavidhiśca samarthavidhiśca samarthavidhiśca samarthavidhiśca samarthavidhayaḥ| padavidhiśca padavidhiśca padavidhiśca padavidhiśca padavidhiśca padavidhayaḥ| sāmarthavidhayaśca padavidhayaśca samarthaḥ padavidhiḥ| pūrvaḥ samāsa uttarapadalopī| yādṛcchikī vibhaktiśca || (Patañjali 2.1.1)

[5] yaḥ kaścidiha śāstre padavidhiḥ śrūyate sa samartho viditavyaḥ | vidhīyate iti vidhiḥ | padānāṃ vidhiḥ padavidhiḥ | sa punaḥ samāsādiḥ | samarthaḥ śaktaḥ | vigrahavākyārthabhidhāne yaḥ śaktaḥ sa samartho vidhitavyaḥ (Kashika 2.1.1)

[6] `Samāsādiḥ` iti| ādiśabdena taddhitavṛtyādīnāṃ grahaṇam| (Nyasa 2.1.1)

[7] *Tatrodbhūtasambandhavyatirekāṇāṃ padārthānāmitaretarapratyākāṅkṣa vyapekṣā"* (śṛmgāraprakāśa 8); *"parasparākāṃkṣārūpā vyapekṣā"* (kaiyaṭa, mahābhāṣyapradīpa 2.1.1); *"svārthaparyavasāyināṃ padānām ākāṅkṣādivaśādyaḥ parasparasambandhaḥ sa vyapekṣā"* (jñānendra sarasvatī tatvabodhinī)

[8] Vibhaktividhāne'pi yāstāvat kārakavibhaktayastā yeṣveva kārakeṣu karmādiṣu vidhīyante teṣāṃ kriyayā sambandho'styeva|

it a vākya (sentence) or a mahāvākya (a collection of sentences), the vyapekṣā sāmarthya ensures that they have a meaning interdependence on each other [10] with regards to conveying a unified meaning.

2.3 *Saṃgati* (Semantic Compatibility in a Discourse)

saṃgati is the term used to denote the semantic connection between different components of a text. The text jaiminīya-nyāya-mālā-vistara [10] talks about three types of *saṃgatis* [11] (coherence) which lead to the comprehensive understanding of a discourse: i.) *śāstra-saṃgati* (subject level coherence), ii.) *adhyāya-saṃgati* (chapter/book level coherence) and iii.) *pāda-saṃgati* (section level coherence).

The next level saṃgati is the topic level saṃgati called *adhikaraṇa-saṃgati* (Chattopadhyay, 1992). The mīmāṃsakas [12] have discussed about six topic level saṅgatis which bind one topic with another (Bhattacarya, 1989; Brahmacari, 2008): i.) ākṣepa (objection), ii.) dṛṣṭānta (example), iii.) pratyudāharaṇa (counter-example), iv.) prasaṃga (corollary/incidental illustration), v.) upodghāta (prerequisite), and vi.) apavāda (exception).

The naiyāyikas (logicians) have also classified the topic level saṃgati into six different types: a.) prasaṃga (corollary), b.) upodghāta (prerequisite), c.) hetūtā (causal dependence), d.) avasara (opportunity for further inquiry), e.) nirvāhakaikya (common connection between adjacent sections), f.) Kāryaikya - (common effect (kārya) connecting the adjacent sections based on joint causal factors), (Sastri, 1916; Das M 2016).

Thus the relationship that exists between two distinct sentences, paragraphs, sections, or chapters in a text or idea is known as saṃgati. The theoretical framework of saṃgati guarantees the compatibility of meaning among the various components of a text or idea. Thus, the concept of saṃgati ensures that various elements of a linguistic expression have *sāmarthya* (cohesion) as explained by Patañjali *saṃgatārthaḥ samarthaḥ* (mingled together).

One might argue what Patañjali meant by the vigraha-vākya [13] (paraphrase) *saṃgatārthaḥ samarthaḥ* (mingled together) of the term *samartha* was the mutual connection between the components of compounds and he does not explicitly talk about sāmarthya in the context of mutual connection between the elements of a discourse. Nevertheless, we observe that the etymological root of the terms *saṃgati* and *saṃgata-artha (saṃgatārtha)* is the same: "*saṃ + gam*". Apte in his Sanskrit

dictionary also describes the meaning of the word *saṅgata* as "joined or united with, come together, associated with" (Apte, p. 1605) and the word *saṅgati* as "fitness, appropriateness, applicability, consistent relation" (Apte, p. 1605). Thus similar to the meaning of the word *saṃgata*, the word *saṃgati* also contributes to the *sāmarthya* (interrelation, mutual connection, interconnection) among various components of a text/speech. These components have been presented in the next section.

3. Sāmarthya licensing meaning structure
Based on the above discussion on sāmarthya in IGT, we propose that the theory of sāmarthya can license the well-formedness of a text and it can be seen operating at three different yet connected levels:
1. Lexico-conceptual Level
2. Syntatico-semantic (Propositional) Level
3. Discourse Level

Lexico-conceptual level
This includes the formation of words from *dhātus* (verbal roots) and *prātipadikas* using affixation (*pratyaya-vidhi)* and other word-internal grammatical operations.

Syntactico-Semantic Level
This level includes *vākya* and Multiword expressions (MWE). A *vākya* [14,15] (sentence) is a collection of words (*padas)* possessing three conditions: a.) *yogyatā* (Compatibility), b.) ākāṃkṣā (expectancy and c.) Āsatti/sannidhi (proximity):
1. Yogyatā [16] (compatibility or fitness of words): The padas in a vākya are mutually related to each other in such a way that they do not have any absurdity in terms of meaning.
2. Ākāṃkṣā [16] (expectancy): The words in a vākya have an expectancy of another word (s) with regard to sense completion.
3. Āsatti [17]/sannidhi [18] (proximity): For comprehension of what is said in a vākya, it must have an absence of interruption.

For example, a sentence is composed of a dhātu (the head of the sentence) and its participants which are syntactically licensed and semantically compatible with the dhātu. The relationships among the dhātu and its participants (the prātipadikas) are marked by the *tiṅ pratyayas* (verbal inflections) and the *sup pratyayas* (nominal inflections). The *tiṅ pratyayas* majorly mark the role of either a *kartā kāraka* (doer/agent) or the *karma kāraka* (theme/patient) of an action. The attachment of the *sup pratyayas* to the *prātipadikas* that are participants in accomplishment of that particular action is subject to whether the roles denoted by the *prātipadikas* are already expressed or not by the verbal inflection, the *tiṅ*. In case the roles of the *prātipadikas* remain unexpressed [19] (Panini, 2.3.1) by the verbal inflections, then only the *sup*

---

upapadavibhaktiṣvapi sahayukte'pradhāne 2.3.19 ityevamādiṣu yuktagrahaṇādīni santi, tatrāpi sāmarthyamastyeva | (Nyasa 2.1.1)

[9] Samarthānāmiti ko'rthaḥ? saṃbadhdārthānāṃ saṃsṛṣṭārthānāṃ vetyarthaḥ | tatra vākye saṃbadhdārthatā | vyapekṣā hi tatra sāmarthyam | anyo'nyāpekṣā u vyapekṣā | ākāṃkṣāsannidhiyogyatveṣu satsuḥ yaḥ parasparasambandhaḥ sā vyapekṣā | (Padamanjari 2.1.1)

[10] https://archive.org/details/in.ernet.dli.2015.487249/page/n22/mode/1up?view=theater

[11] śāstre'dhyāye tathā pāde nyāyasaṅgatayastridhā.śāstrādiviṣaye jñāte tattatsaṅgatirūhyatām. (jaiminīya-nyāya-mālā-vistara 1.5)

[12] Ūhitvā saṃgatīstisrastathā cāntarasaṃgatim.ūhed-ākṣepa-dṛṣṭānta-pratyudāharaṇādikam. (jaiminīya-nyāya-mālā-Vistara 1.23)…Yathaitatsaṃgaticayamūhitam tathā pūrbottarādhikaraṇayoḥ parasparamavāntarasaṃgatirūhanīyā.…SĀ cānekarūpā. Ākṣepasaṃgatir-dṛṣṭānasaṃgatiḥ pratyudāharaṇasaṃgatiḥ prāsaṅgikasaṃgatir-upodghātasaṃgatir apavādasaṃgatiścetyevamādirūpā. (jaiminīya-nyāya-mālā-vistara 1.23)

[13] vṛttyarthāvabodhakaṃ vākyaṃ vigrahaḥ (Laghusiddhantakaumudi 2.1.4)

[14] vākyaṃ syādyogyatākāṃkṣāsattiyuktaḥ padoccayaḥ (sahityadarpana 2.1)

[15] ekārthaḥ padasamūho vākyam (a vakya is a group of words that together express one thought) (kāśikā, 8.1.8; padamaṅjarī 8.1.8)

[16] ākāṃkṣā pratītiparyavasānavirahaḥ | sa ca śroturjijñāsārūpaḥ | (sahityadarpana 2.1)
 kā punariyāmākāṃkṣā nāma? pratipatturjijñāsā nāma | (śṛṃgāraprakāśa 8)

[17] āsattirbuddhyavicchedaḥ (sahityadarpana 2.1)

[18] atha sannidhiḥ kaḥ? yogyasyākāṃkṣitasya yadānantaryam | (śṛṃgāraprakāśa 8)

[19] *anabhihite (A. 2.3.1)*

pratyayas are attached to the *prātipadikas* to specify their kāraka roles in an action.

The important aspect to be noticed here is that the process of word formation according to Panini is not an isolated one; rather word formation is deeply interlinked with sentence formation (Mahavir, 1968). For example, let us take the sentences (1) and (2):

1. Rāmaḥ granthaṃ paṭhati
   Rama-Nom.3rd.Sg book-Acc.Sg read-kartṛvācya [20]. 3rd Sg
2. Rama reads a book. rāmeṇa granthaḥ paṭhyate Rama-Nom.3rd.Sg book-Acc.Sg read-karmavācya [21].3rd.Sg

A book is read by rama

The sentence (1) is in *kartṛvācya* (active voice). In *kartṛvācya*, a verbal inflection on the verb (here '-*ti*' on 'paṭha' (to read)) marks [22] the *kartṛ-kāraka* (agent/doer); whereas in a passive voice, a verbal inflection marks the *karma-kāraka* (object/theme/patient) as shown in the sentence (2). Since the verbal inflection ('-*ti*') has already marked the *kartṛ* (doer) of the action as 'kartṛvācya.3rd.Sg', the 3rd case ending (tṛtīyā vibhakti) which is generally used to mark the *kartṛ-kāraka* of an action cannot be applied to the kartṛ (rāma) as the 3rd case ending rule can be applicable only if the role of the *kartṛ-kāraka* has not been already marked/expressed/specified otherwise [23]. Similarly, the 2nd case ending (dvitīyā vibhakti) marks the role of a karma-kāraka if not specified otherwise [24]. Since the *karma-kāraka* has not been marked by any other suffix attached to the other members of the sentence (1), dvitīyā vibhakti is applied to *grantha* (book) to denote its role of the *karma-kāraka* of the verb *paṭha* (to read) in the sentence. This illustration confirms that the derivation of the different components in a sentence is mutually semantically interdependent and suggests that sentences are not formed if their components do not satisfy the principles of *sāmarthya* [25].

Discourse Level

A *mahāvākya* [26] is a collection of sentences possessing the requisite qualities of a *vākya* i.e. *yogyatā* (compatibility), *ākāṃkṣā* (expectancy) and *āsatti/sannidhi* (proximity). The entire texts like rāmāyaṇa, mahābhārata etc are examples of *mahāvākyas*. Kumarila Bhatta in his Tantravārttika says: *svārthabodhe samāptānām aṅgāṅgitvavyapekṣayā Vākyānām ekavākyatvaṃ punaḥ saṃhatya jāyate* (Tantravārttika, 1.4.28) When the sentences are completed with regard to conveying their own meaning, then they have an *aṅgāṅgībhāva* (part-whole relationship) with each other, i.e. when put together, some sentences will become heads and some will become their dependents. Thus, the collection of such sentences put together is termed *mahāvākaya*. The equivalent term for *mahāvākaya* in modern linguistics is *discourse*.

saṃgati takes care of the compatibility of the meaning between phrases, clauses, sentences, paragraphs, sub-sections, sections and chapters of a text. The concept of *saṃgati* helps in distinguishing how the parts of a text/idea relate to each other. These relations allow the reader/listener to comprehend the deeper meaning of the text/idea.

4. Representation of USR

USR is a text-based data structure that is close to the Attribute Value matrix (AVM) representation. It is easier to read and write manually, as well as process computationally. As stated earlier, USR is modeled following the principles of *sāmarthya* and *vivakṣā*. The principles of *sāmarthya* guide the design of USR so that semantic compatibility at lexico-conceptual, syntactico-semantic and discourse levels are maintained in the representation. Before we illustrate in the sub-section 4.2 how this has been achieved we present one USR in the next sub-section.

4.1 An Example of a USR

Table-1 presents a USR that is represented in a tabular format. The first column is not part of the USR. But it defines the corresponding row. Currently, the following rows have been conceptualized:

- Concept Row
- Index row
- Semantic Category row
- Morpho-semantic row
- Dependency Relation row
- Discourse row
- Speaker's view row
- Scope row
- Sentence type row
- Construction

In Table-1, the first row represents concepts, the second index of each concept, the third semantic category of each concept, and so on.

USR_ID=1

Table 1: Example of a USR

Concept	hari	umā	koṭī	basa_1+aḍḍā_1/bus_1+stop_1	apanā/own	māṃ_1	choṭā_7/younger_1	bhāī_1/brother_1	nahīm_1/not	bāta+karā_1-yā_1/speak_1-past
Index	1	2	3	4	5	6	7	8	9	10
Semantic category	per/male	Per/Female	Place					Anim Male		
Morpho-semantic								pl		causative (ṇic)
Dep rel	10:pk1	10:pk1	4:r6	10:adhikaraṇa	6:dem	10:jk1	8:modifier	10:karma	10:neg	0:main
Discourse					1:coref 2:coref					
Speaker's view				def				certainty		
Scope	certainty[not[0:main]]									
Sent_type	negative									
CxN[27]	Conj:[1,2] compound:[3.2 rt 3.1]									

PK1: prayojaka karta (causer); jk1: prayojya karta (causee); dem: demonstrative; RT: Purpose.

---

[20] Active voice

[21] Passive voice

[22] laḥ karmaṇi ca bhāve cākarmakebhyaḥ (A. 3.4.69)

[23] Kartṛkaraṇayostṛtīyā (A. 2.3.18)

[24] karmaṇi dvitīyā (A. 2.3.2)

[25] Panini talks about rules related to euphonic changes that include sentence boundaries (Mahavir, 1968).

[26] vākyoccayo mahāvākyam| yogyatākāṃkṣāsattiyukta ityeva (sahityadarpana 2.2-3)

[27] CxN = Construction. Construction row represents a form-meaning pair when a complex form has relational information among its components or makes the compositionality explicit when one form is missing.

The USR shown in Table-1 will generate the Hindi sentence (3.a) and English sentence (3.b): (3).
a. Hari aura umā ne kotī ke basa aḍḍe para choṭe bhāīyoṃ ke sātha to apanī māṃ kī bāta nahīṃ karāī.
b. Hari and Uma did not certainly make their mother talk to the younger brothers at Koti bus stop.

The information encoded in the USR given in Table-1 is as follows

▪ Every USR is given a unique id which can be referred within the USR to establish discourse connection between the USRs.
▪ The Concept, Sem (antic) cat (egory), Morpho-semantic and Speaker's view rows capture *lexico-conceptual* information of the concepts. For example,

Concept

▪ The Concept row represents *unique* concepts (not words) that refer to entities, events, quality, quantity and other properties of an entity or event. Hari, Uma, choṭā_7/younger_1, bhāī_1/brother_1, māṃ_1/mother_1, apanā/own, basa_1+aḍḍā_1/bus_1+stop_1, bāta+kara_1/ talk_1 are con- cepts. In the present version of USR nahīṃ_1 'not' is also been considered as a concept.
▪ Semantic category.
▪ Hari, Uma are per (sons).
▪ Younger brothers are anim (ate) entities.
▪ Hari is male and Uma is female.
▪ Morpho-semantic.
▪ Bāta+karā is a causative form of bāta+kara which is specified in this row.
▪ Plurality of younger brothers.
▪ Speaker's view.
▪ The bus stop is a def (inite) one that the speaker has in mind.
▪ The speaker is certain that Hari and Uma did not make their mother talk to their younger brothers (however they might have made her talk to someone else). nahīṃ_1 'not' is also been considered as a concept.
▪ Dep (endency) rel (ation) and Cxn row specify the syntactico-semantic level information in which the relation between the head and its dependents are specified. Dependents are modifiers of the head. It is the viśeṣya-viśeṣaṇa (head-modifier) relation.
▪ Dependency rel(ation)
▪ This row captures the kāraka and kāraketara (non-kāraka) relation between the *mukhya viśeṣya* (head) and its dependents within a proposition.
▪ Construction
▪ This row contains form-meaning pair/s. For example, a conjoined construction specifies that all entities involved enjoy equal status. A noun compound construction presents two or more nominal components with their internal relation underspecified.
▪ Discourse row captures inter-sentential relational information and also anaphoric rela- tions.
▪ Cohesion
▪ Pronominal co-referencing is specified for ensuring cohesion. For example, apanā/own corefers both Hari and Uma in this USR.
▪ Coherence
▪ *Saṃgati* or coherence in the text is established in this row through semantic connectives among USRs.
▪ The scope row captures scopal information if any.
▪ Reference to 0: Main in any other row than Dep Rel implies the verbs with all their dependents except neg, if

there is one. In other words, 0: Main represents the proposition.

4.2 sāmarthya: The Guiding Principle of USR

We will illustrate here some cases to show how *sāmarthya* is modeled in the representation of USR. The first example is the verbal concept. The concept row specifies the event *bāta+kara_1* 'talk'. The picture the speaker wants to draw is the negation of the idea that there is a motivator or causer or *prayojaka kartā* (here Hari and Uma together) who causes a conscient being (in this case the 'mother') to talk to younger brothers. This is the semantics of *hetu* [28] or cause. The causative meaning is represented in USR through *ṇic* pratyaya in the morpho-semantic row on the verbal concept *bāta+kara_1*. The principle of *sāmarthya* licenses the derivation of *bāta+karā* 'make (somebody) talk' from *bāta+kara_1+ṇic* 'talk+cause', which may not hold for another event, for example, in *rāma ne rāvaṇa ko mārā* 'Rama killed Ravana'. Even though Ravana is a conscious being, no human being, for that matter, has a choice for dying or not dying. Thus *mara* 'die' cannot be causativized (*asamartha*). Therefore, we will not derive *māra* 'kill' from *mara* 'die + ṇic' in USR.

At syntactico-semantic level, which is represented in the dependency row, the semantic and syntactic compatibility of the head and its dependents determines the sentential structure. For example, bāta+kara_1 is a communicative verb that requires two animate participants for communication. When we specify semantic roles for the participants in terms of kāraka relations, we consider the meaning of the verb, its *ākāṃkṣā* and *yogyatā* which in turn defines its *sāmarthya*. That is why, we assign *prayojya kartā* and *karma* relations to *māṃ_1* 'mother' and *chotā_1 bhāī_1* 'young brother' respectively for the event *bāta+kara_1+ṇic* in the USR relation given in Table-1.

Finally, the discourse row captures the cohesiveness and coherence of the text. In this case, for example, *apanā* co-refers to Hari and Uma. The semantics of discourse connectives are specified at this level which determines the well-formedness of the structure of the whole text.

5. Conclusion

We have presented, in this paper, the principles of *sāmarthya* (*saṃgatārthaḥ samarthaḥ*) as explained in Indian Grammatical Tradition. We have made an observation that the etymological derivation of *saṃgatārthaḥ* and *saṃgati* are the same. Thus, we conclude that *sāmarthya* accounts for the structural well-formedness of *pada*, *vākya* (sentence) as well as it establishes cohesion and coherence at the level of *mahāvākya* (Discourse). The structural well-formedness of USR the newly developed IGT-based Semantic Representation, is proposed to be based on the principle of *sāmarthya*. We have given some examples in section 4 to explain how the principle of *sāmarthya* motivates the design of USR in its current form. We envision this work as a pioneer step towards applying the principles and concepts from Indian Grammatical Tradition to build modern

---

[28] Tatprayojako hetuśca (A. 1.4.55).
kartuḥ prayojako hetusaṃjñaḥ kartṛsaṃjñaśca syāt
(Siddhāntakaumudī 2575)
That which is the mover thereof, i.e. of the independent source of action, is called hetu or cause, as well as kartā or agent (SC Vasu on A. 1.4.55)

knowledge-rich semantic resources which can be used for various NLP tasks. Currently, we have taken up the task of Multilingual Natural Language Generation (NLG) from the USRs. We engage ourselves in the experiments of introducing various semantic insights from IGT in USRs so that this becomes a useful resource for NLG tasks.

Acknowledgement

6. References
1. Abhyankar KV, Limaye VP. Mahâbhâsya-dipikâ of Bhartr-hari. Other Ed Visnu Prabhàkar Limaye. Poona: Bhandarkar Oriental Research Institute; c1967.
2. Abhyankar, Vasudev K, Śukla JM. A dictionary of Sanskrit grammar; c1986.
3. Apte, Shivaram V, *et al*. Prin. VS Apte's the practical Sanskrit-English dictionary; c1957.
4. Bharati A, *et al*. Natural language processing: a Paninian perspective; c1995.
5. Bhattacarya, Vidyasagar J. Śrī Mādhavācāryaviracita Jaiminīya Nyāyamālāvistaraḥ. Kṛṣṇadās Academy, Varanasi; c1989.
6. Caṭṭopādhyāẏa, Rāmapada, Chattopadhyay R. A Vaiṣṇava Interpretation of the Brahmasūtras: Vedānta and Theism. Brill; c1992. Vol. 3.
7. Das, Monali. Discourse Analysis of Sanskrit Texts: first attempt towards computational processing. Diss. unpublished Ph.D. thesis University of Hyderabad; c2016.
8. Fillmore, Charles J, Lee-Goldman R, Rhodes R. The framenet constructicon. Sign-based construction grammar. 2012;193:309-372.
9. Garg, Kirti, *et al*. Evaluation of Universal Semantic Representation (USR). Proceedings of the fourth international workshop on designing meaning Representations; c2023.
10. Joshi, Dattatray S. Patañjali's Vyākaraṇa-Mahābhāṣya. Samarthāhnika; c1972.
11. Kane, Vaman P. The Sahityadarpana. Motilal Banarsidass Publisher; c1997.
12. Kielhorn F. Ed. The Vyākaraṇa-Mahābhāṣya of Patañjali. No. 18-22. Government central book depot; c1909.
13. Mahavir. Samartha Theory of Pāṇini and Sentence Derivation. Munshiram Manoharlal Pvt. Ltd. New Delhi; c1984.
14. Mishra, Narayan. Kashika of Pt. Vamana and Jayaditya; c1996.
15. Ramkrishnamacharyulu KV. Annotating Sanskrit texts based on Śābdabodha systems. International Sanskrit Computational Linguistics Symposium. Berlin, Heidelberg: Springer Berlin Heidelberg; c2009.
16. Raster, Peter. The Indian grammatical tradition and the phenomenology of higher stages of language. Na; c2001.
17. Sastri PSS. Lectures on Patanjalis Mahabhasya, Chennai: Kuppuswami Sastri Research Institute; c2015. Vol 1.
18. Sukhada, Sirisipalli VH, Paul S. Generation of MRS Abstract Predicates from Paninian USR. HPSG 2023: 30th International conference on head-driven phrase structure grammar. University of Massachusetts Amherst, USA; c2023 Jul.
19. Sukhada. A Paninian perspective to information dynamics in language mapping structures between English and Hindi; c2016.
20. Tripathi J, Malaviya SL, Kashika S. A commentary on Panini's Grammar, Tara Printing Works, Varanasi; c1986.
21. The Ashtadhyayi of Panini. Satyajnan Chaterji; c1897. Vol. 6.