

Role of Semantics on Hindi inflection

Report

Nehchal Jindal

Mentor: Amitabha Mukerjee
{nehchal, amit}@cse.iitk.ac.in

IIT Kanpur

18th April, 2013

SE367: Introduction to Cognitive Science

Abstract

Within inflection morphology, two approaches have been highly debated to understand mental process of generation of inflected forms, namely dual-mechanism and single mechanism connectionist model. Single-mechanism claims that semantics has a certain effect on the process. We test this hypothesis, by observing the effect of imageability on causative inflection of Hindi language verbs. The imageability ratings for verbs were collected. Response times of subjects were measured for inflection. While few results complied with the hypothesis, overall the hypothesis could not be successfully confirmed. The scale of experiment was too small jump to any reliable conclusions.

Introduction

Inflection is an important part of inflection and we use it very often in our language. An adequate model of inflection should explain how mental process for both regular and irregular inflection. Within the field of inflectional morphology, two models have been highly debated. Dual model [Pinker 1999; Pinker and Prince 1988] claims that regular verbs are inflected by rule and exception verbs are stored in lexicon. The rule is applied by default but if stored inflected form of verb is retrieved, rule application is blocked.

Propagators of connectionist model, discarding dual model, claim that there doesn't exist two separate mechanisms for regular and exception verbs. Single mechanism model is trained over phonology and semantics of word and same trained model works for regular as well as irregular verbs.

Dual model says there is no link between semantics and inflection. Connectionist model infers that semantics has certain effect on human mind inflection process.

In this experiment, effect of imageability (part of semantics) has been analyzed on causative inflection for Hindi language verbs.

Related Works

[Butler 2012] looked into role of semantics for English past tense inflection and found relation there is significant relationship between semantics and inflection response time and thus favored connectionist model

Causitive Form

The change in morphological structure of the word to form new related work is called inflection. E.g. देख → दिखा, दिखवा, दिखाइए, देखूँगा . Verbs are inflected for aspect, mood, tense, and agreement feature of gender, number and person.

Verbs have two causative forms, first causative and second causative.

	First causative	Second causative
गिर fall	गिरा cause x to fall	गिरवा cause y to cause x to move
चल move	चला cause x to move	चलवा cause y to cause x to move
सुन listen	सुना cause x to listen	सुनवा cause y to cause x to listen

Regular verbs form causative by adding 'आ' or 'वा' to root word.

कर → करवा

लिख → लिखवा

पढ़ → पढ़ा

Irregular (exception) verbs don't follow any particular rule.

फेंक → फिंकवा

खा → खिला

जोड़ → जुड़वा

Some verbs don't have causative form. E.g. सह (to bear).

We studied 150 verbs commonly used in Hindi selected from Hindi WordNet. Regular and exception verbs constituted 40% and 33% of 150 verbs respectively. Remaining verbs do not have causative form.

Imageability

Imageability is defined as the ease with which a word gives rise to a mental image (i.e., a mental picture or sound, or other sensory experience). Imageability is a part of semantics.

Highly imageable verbs can be imaged quickly and easily. Eg. peena. Less imageable arouse mental image with difficulty. E.g. उलझना

Methodology

First we rate the verbs for their imageability. Then we study response time for verb inflection by 4 participants. Finally, we analyze possible relation between imageability and mean response times.

Imageability Ratings Study

Imageability ratings were collected by online survey for 150 words. Each verbs had to be rated on scale of 1 to 7.

- 1 - least imageable
- 7 - most imageable

In survey, online form was designed such that each person gets random words to rate. 8 responses on average were recorded for each word.

Experiment

On the basis of imageability ratings 10 words were selected for each category:

1. *low imageability regular verbs*
2. *high imageability regular verbs*
3. *low imageability irregular verbs*
4. *high imageability irregular verbs*

The list of words selected is given in Table1.

Regular		Exception	
High	Low	High	Low
चढ़ना	भरना	पीटना	सीखना
गिरना	समझाना	उछलना	छापना
चिपकाना	बढ़ना	रोना	देखना
उड़ना	घटाना	खाना	भेजना
हँसना	पहुँचना	नाचना	माँगना
टपकना	बदलना	नहाना	छोड़ना
भागना	रखना	पीना	बाँटना
चलना	ठहरना	बैठना	कहना
लिखना	बनाना	तोलना	डालना

Table 1. The final list of words used for experiment.

Before actual experiment, participants underwent practice test. Practice test contained 10 words: चखना, सोना, खेलना, चढ़ना, पकाना, धोना, उतरना, लेटना, भेजना, पीना.

All 40 verbs were presented in randomized order to 4 subjects. Subjects were presented with the root form of verb on a computer screen one by one. In the experiment subject had to speak causative form of verb. Response time (RT) was measured for each subject, for each verb. FLXLab software was put to use for this purpose.



Figure 1: Subject responds to stimuli (root word) during experiment

Results

The mean response times are recorded in Table 2 and visualized in Figure 2.

	High	Low
Regular	1893 ms	2117 ms
Irregular	1937 ms	1846 ms

Table 2: Mean Response times (RTs)

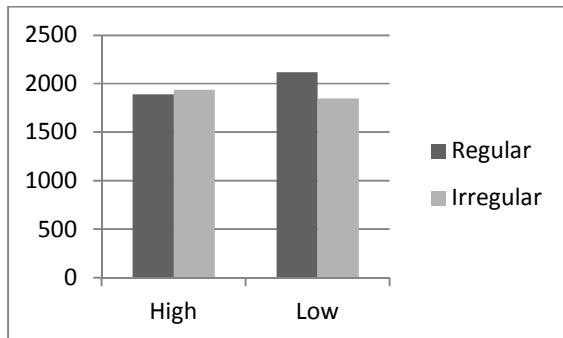


Figure 2: Mean Response times (RTs)

Discussion

In case of regular verbs, mean RT of low imageable words is higher than that of high imageable words, by 224ms. But for irregular verbs, mean RT of low imageable words is lower than that of high imageable words.

No concrete results can be inferred from the results of this experiment. According to connectionist model, regular inflection is

most common and so learnt probabilistic model would be more strongly encoded for regular transformations and hence take less time for inflection of regular verbs. In this experiment, this is true for highly imageable words, but not for low imageable words. Even for high imageable words, this difference of 44 ms, is relatively small. This may be due to the fact that difference between number of regular and irregular verbs in Hindi (40% and 33%) is small compared English (86% and 14% for past-tense inflection).

The connectionist model also claims that regular verbs are more dependent on phonology and semantics has lesser role while vice-versa is true to irregular verbs. Further, more imageability means greater semantic information, so high imageable words should be easily inflected as compared to less imageable words since single-mechanism model is trained over semantics (and phonology). The results of this experiment comply for regular verbs but not for irregular verbs.

Conclusion

This experiment has failed to verify the hypothesis of single mechanism connectionist model. But, it won't be fair to discard the hypothesis of connectionist model on basis of results of this experiment, considering the fact that subjects were very less in number, and tested verb database is too small to jump upon any reliable conclusion.

References

1. Butler, Rebecca, Karalyn Patterson, and Anna M. Woollams. "In search of meaning: Semantic effects on past-tense inflection. *The Quarterly Journal of Experimental Psychology*" 65.8 (2012): 1633-1656.

2. "Hindi Wordnet", IIT Bombay. 2008
<<http://www.cfilt.iitb.ac.in/wordnet/webhwn>
>