Syllable Contact Law in Persian

Afshin Rahimi

Language and Linguistics Department, Division for Computational Linguistics, Sharif University of Technology
a_rahimi@ce.sharif.edu

Moharram Eslami

Language and Linguistics Department, Division for Computational Linguistics, Sharif University of Technology
m_eslami@sharif.edu

Bahram Vazirnezhad

Language and Linguistics Department, Division for Computational Linguistics, Sharif University of Technology
bahram@sharif.edu

1. INTRODUCTION

This is a universal tendency among languages that sonority of consonant clusters in the boundary of syllables has a falling pattern. In accordance with this universal tendency, structures with falling sonority difference are more unmarked structures. The source of this well-formedness has been widely discussed in literature to be either phonetic based or perceptual based. Here we investigate the patterns of sonority in boundary of syllables in Persian.

2. METHOD

This study investigates sonority change in the boundary of Persian syllables. "FLexicon" Persian lexicon was used as the main source of data for analysis. FLexicon contains more than 50000 lexemes with phonemic transcription. Syllable structure of Persian is in the form of CV(C)(C). Syllable structures were extracted from the phonemic transcription. Consonant clusters in the boundary of syllables are extracted. Consonants of Persian are assigned a sonority level from 5(highest sonority) to 1(lowest sonority). Sonority levels are defined as: *(Liquids=5 >Nasals=4 >Fricatives=3 >Affricates=2 >Plosives=1)*. The marginal probabilities for each consonant occurring in the boundary of syllables were computed. As an example the number of consonant /t/ occurring in onset position is divided to total number of syllable boundaries to achieve marginal probability of the consonant in onset position. The marginal probabilities of all consonants occurring in onset and coda positions were computed. The product of the marginal probabilities for the two consonants in syllable boundary consonant cluster was considered as the expected occurrence of a consonant cluster. The joint probability of a consonant cluster was computed by counting the number of occurrence in syllable boundary divided by total number of syllable boundaries. The division of joint probability as the observed probability to the product of marginal probabilities as the expected probability can be an estimation of how various consonants prefer to co-occur in the boundary of the syllable. This observed/expected measure is called PMI. Positive PMI values indicate a preference for co-occurrence of consonants and negative values indicate a preference for consonants not to co-occur. This relation between the sonority pattern (falling or rising pattern) in the boundary of syllable and preference to co-occur is investigated in this study. The probabilities are computed once using type frequency within lexicon and once using token frequencies using corpus to investigate the impact level (lexicon or corpus) of markedness constraints.

3. RESULTS

The results show that falling sonority patterns occur more in the boundary of syllables within lexicon. The same trend for falling sonorities occurs in corpus.

The PMI indicates no direct relation between falling/rising sonority patterns and preference for co-occurrence of consonants in the boundary of syllables. The results are the same for type probabilities within lexicon and token probabilities in corpus.

Figure 1- The vertical axis is the frequency of occurrence within lexicon. The horizontal axis is the sonority difference between last consonant in coda position and the first consonant in onset position in a A.B consonant cluster in the boundary of syllables. Negative sonority difference indicates a falling sonority pattern and Positive sonority difference indicates a rising pattern in the boundary of syllables. As can it be seen Negative sonority difference syllables have more frequency within lexicon.
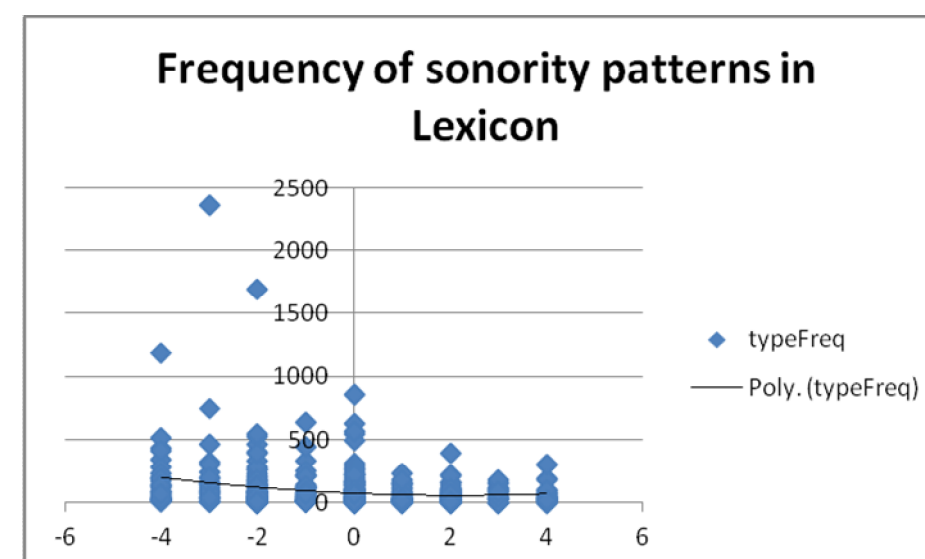

Figure 2- The vertical axis is the frequency of occurrence in corpus. The horizontal axis is the sonority difference between last consonant in coda position and the first consonant in onset position in a A.B consonant cluster in the boundary of syllables. Negative sonority difference indicates a falling sonority pattern and Positive sonority difference indicates a rising pattern.
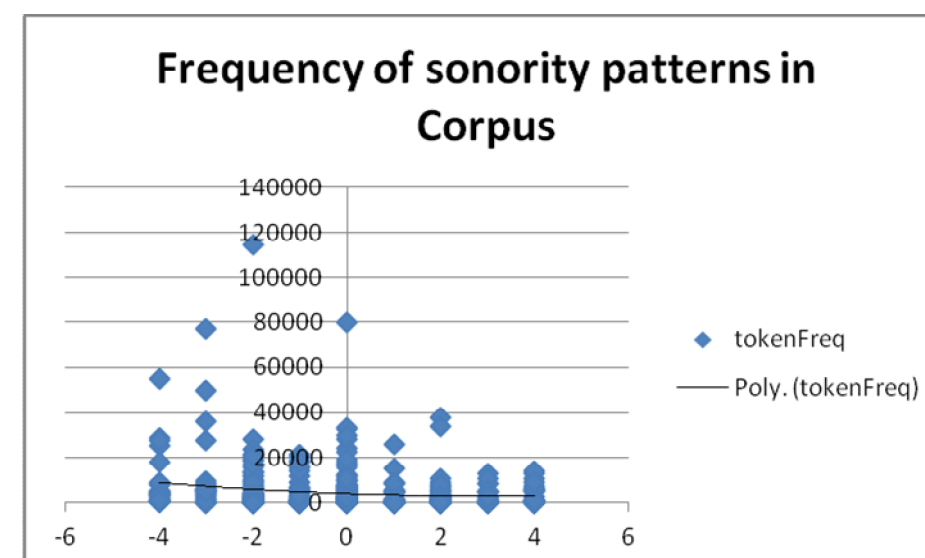

Figure 3- The vertical axis is the PMI of co-occurrence of A and B in a A.B consonant cluster in the boundary of syllables within lexicon. The horizontal axis is the sonority difference between last consonant in coda position and the first consonant in onset position in a A.B consonant cluster in the boundary of syllables. Negative sonority difference indicates a falling sonority pattern and Positive sonority difference indicates a rising pattern in the boundary of syllables. As it can be seen there is no direct relation between Lexicon PMI and sonority pattern in Persian.
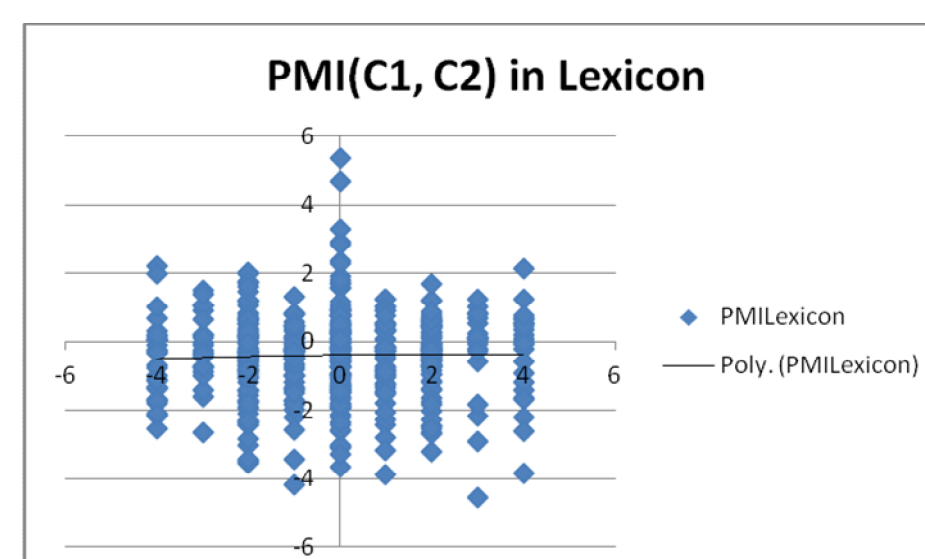

Figure 4- The vertical axis is the PMI of co-occurrence of A and B in a A.B consonant cluster in the boundary of syllables in corpus. The horizontal axis is the sonority difference between last consonant in coda position and the first consonant in onset position in a A.B consonant cluster in the boundary of syllables. Negative sonority difference indicates a falling sonority pattern and Positive sonority difference indicates a rising pattern in the boundary of syllables. As it can be seen there is no direct relation between Corpus PMI and sonority pattern in Persian.
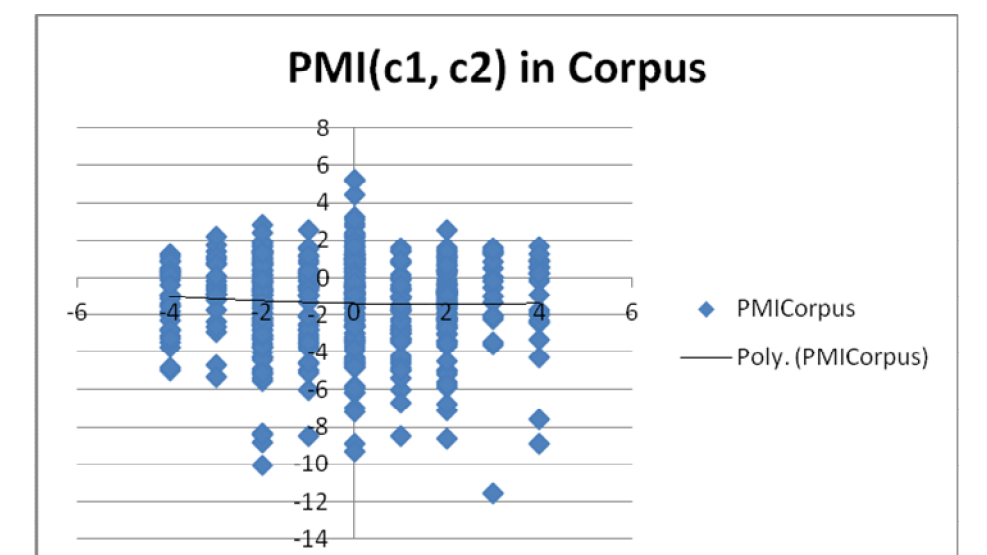
4. Discussion

Higher frequency of occurrence for falling sonority patterns in the boundary of syllables support the idea that falling sonorities are of phonetic and perceptual base. The phonetic and perceptual basis of this preference causes diachronic and synchronic phonological change in the language. This change demonstrates itself in the frequency of falling sonority patterns within syllables in Flexicon and in corpus. It means that there are more falling sonority pattern consonant clusters within lexicon, and those words that have these patterns occur more in corpus. The preference is not very forced everywhere as the slope of the fitting line is low. This means that there is no cut-off categorical point that makes rising patterns ungrammatical.

PMI analysis indicates that in Persian, unlike some other languages like Kirkiz and Kazakh, the amount of observed occurrence of patterns comply with the amount of expected ones. The patterns support a hypothesis against SCL. The higher frequencies can be the effect of local conjunction of high sonority coda constraint and low sonority onset constraint. The Split Margin hypothesis as another explanation for syllable boundary sonority patterns considers the relation of complex onsets and the tendency for falling sonority patterns. The Persian language forbids onset clusters, so in a V1(C0)(C1)C2V2 consonant cluster there is no ambiguity in recognizing the syllable boundary position. The syllable boundary position is always one consonant before the vowel V2. So no other factor like falling sonority is needed to make syllable boundary perceptual cues more salient.