

Unbalanced Numbers in Online Language Learning

Jayant Venkatanathan

Social Web 2009 Fall

Introduction

Livemocha is a social networking tool for language learning. Users who speak the language you learn critique on your exercises and you are expected to provide feedback for people learning languages you speak. The average user may be more inclined to provide feedback to those who speaks the language he is learning, increasing the chances of receiving feedback for his exercises. In such a scenario, users who speak a language not many are learning might be at a disadvantage. We identify such languages.

Ratio between number of users who speak A and learn B and the number of users who learn A and speak B.

Method

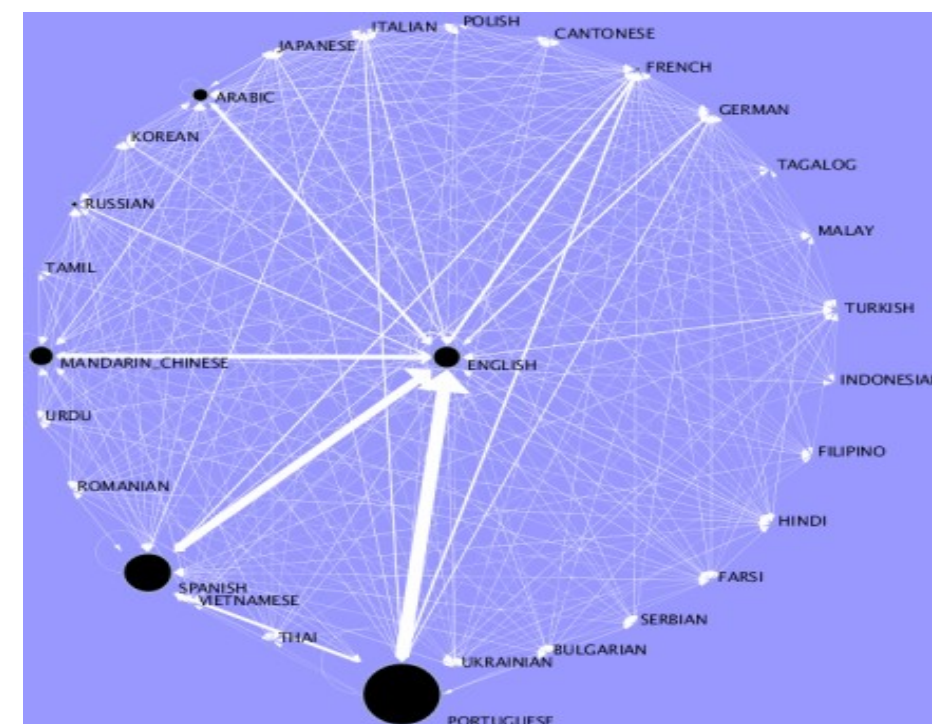
Each user has a profile-id which can be extracted from the url to his profile page. These ids are incremental, with each new user assigned a higher id. We randomly picked around 26,000 ids and extracted the languages the the user speaks and is learning.

Results

Number of Speakers and Learners of 5 Languages

Speakers-Number	Learners-Number
Portuguese-10675	English - 18804
English - 8158	French - 4920
Mandarin - 3731	Mandarin - 949
Arabic - 2564	Portuguese - 883
French -996	Arabic - 564

Language pair (A-B)	Learner-Speaker Ratio
Portuguese-Japanese	102
Arabic-English	7.8
Portuguese-English	21.78
English - French	4.04
Mandarin - Japanese	37



Network of Languages: Node size is proportional to number of speakers and edge with to number of learners of the target language

Discussion

The results show an imbalance in the distribution of speakers. For 10,675 Portuguese speakers, there are only 883 Portuguese learners, while for 4920 learners of French there are 996 speakers. The results are more striking for the ratio of language pairs. For every 102 Portuguese speakers who learn Japanese, there is one Japanese who learns Portuguese and for about every 8 Arabic speakers learning English there is one English speaker learning Arabic

Conclusion

The distribution of speakers and learners is highly unbalanced. This can put certain learners at a disadvantage. Such users can be easily identified and hence mechanisms to help them can be put in place.