



INFORMATION TECHNOLOGY
UNIVERSITY

Carnegie Mellon University

Speech-based Social Media for Low-Literate Masses

Agha Ali Raza

SLT4D

Situations where information and connectivity are the missing links to Development

Use Speech and Language Technologies to:

- Disseminate information to low-literate masses
- Gather information from masses
- Develop linguistic resources for resource-scarce languages



Examples include speech-based:

- Marketplace (Craig's List), Citizen Journalism, message boards/blogs
- Health, agriculture, jobs, education,...
- Information gathering through surveys and polls,...
- Gathering Speech corpora

How do you reach your target users?

A REACHABILITY problem



You build it... but *they*
don't come!



How do you get your target
audience to use your
application?

An UPTAKE problem



Users come, but they
don't stay!

A RETENTION problem



Reachability

Challenges:

- PCs are not feasible
- Smartphones are not always feasible
- TV, radio are non-interactive
- SMS assumes literacy

Speech over simple phones is a viable way



Hurdles

User Interface hurdle: Even simple speech interfaces confuse low-literate, non-tech-savvy users

Motivation hurdle: Users are not motivated enough to change their ways

Uptake and spread hurdle: How do you spread your services to poorly connected masses

Our Strategy:

Entertainment



INFORMATION TECHNOLOGY
UNIVERSITY

Carnegie Mellon University

For Entertainment:

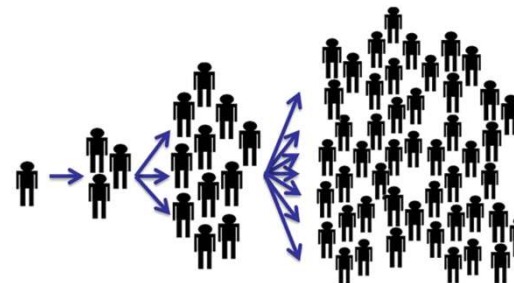
1. Users would overcome UI hurdles (Smyth et al. 2010)



2. Users don't need any convincing



3. Users may spread the services to others



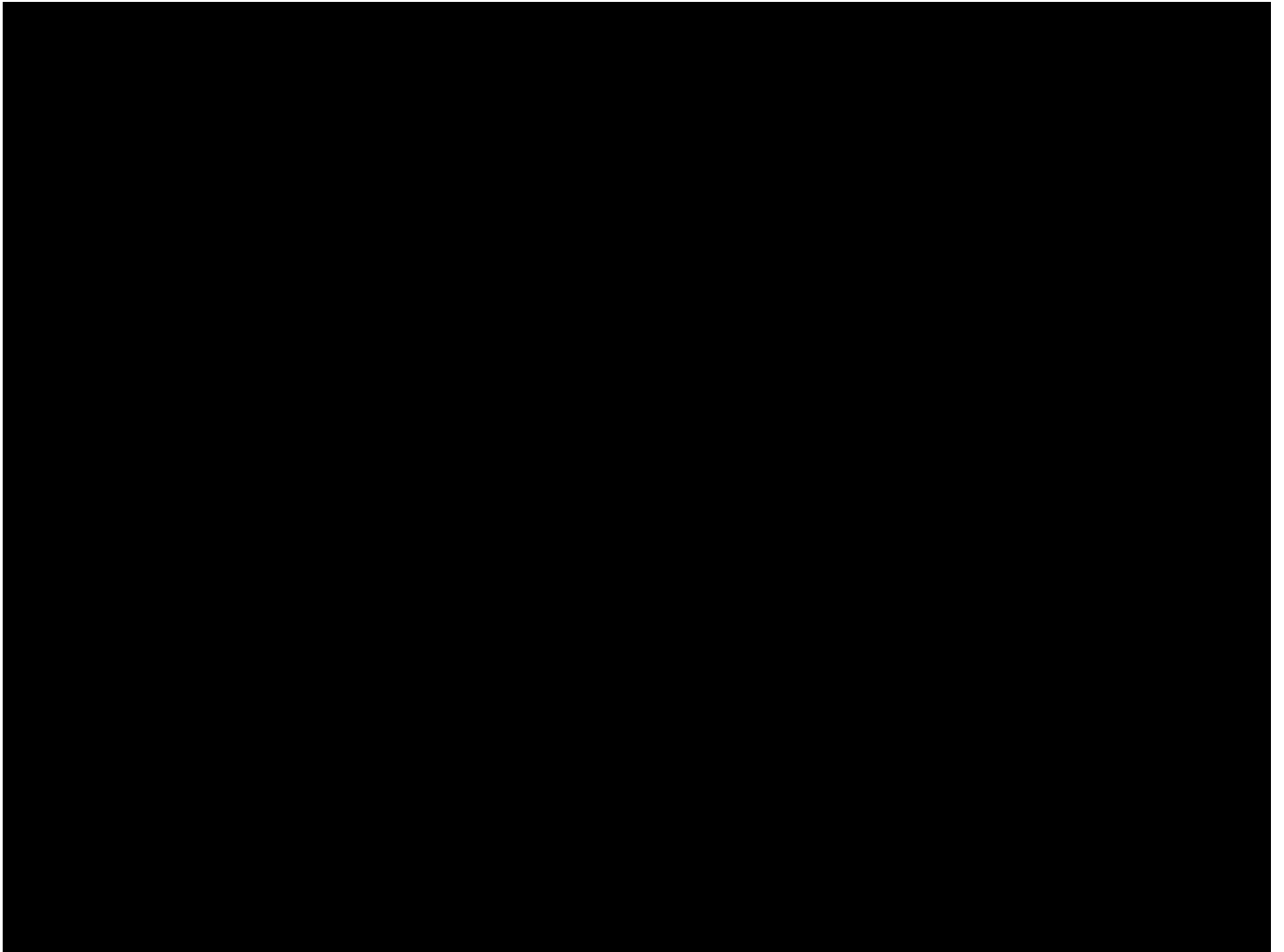
***Viral Entertainment* as a vehicle
for disseminating
Development related services**



Polly (میاں مٹھو)

Polly is a **telephone-based, voice-based** service which allows users to make a short recording of their voice, **modify it** and **send the modified version to friends**.



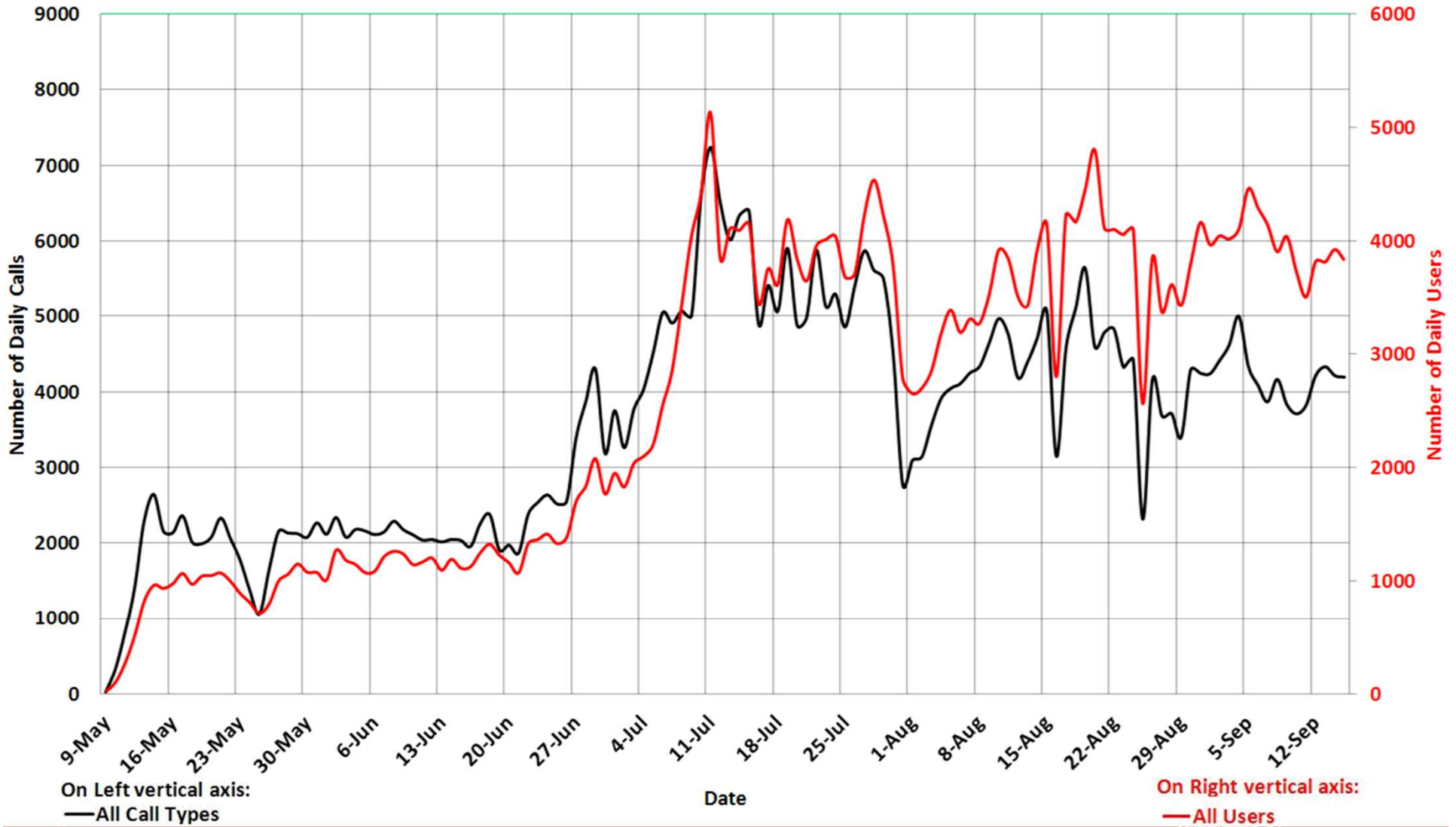


Deployment in Lahore

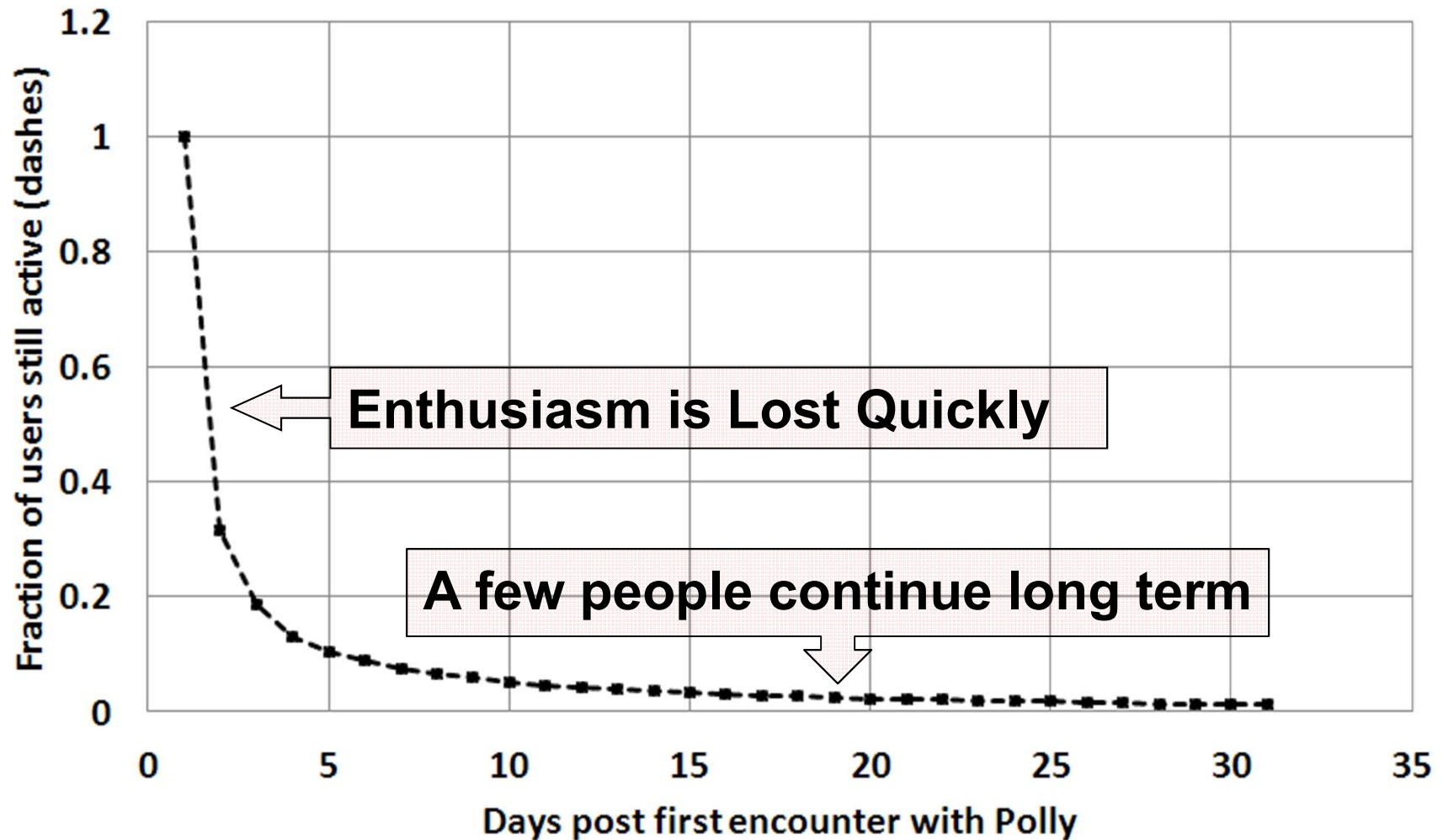
Seeded with **5 users**

- Within a year:
 - **636,000** calls
 - **165,000** users
 - Spreading to **1,000 new people daily**
- **34,000** people used the job search service
- listened **386,199** times to **728** job ads
- and **19,000** users forwarded them to their friends.

High Uptake



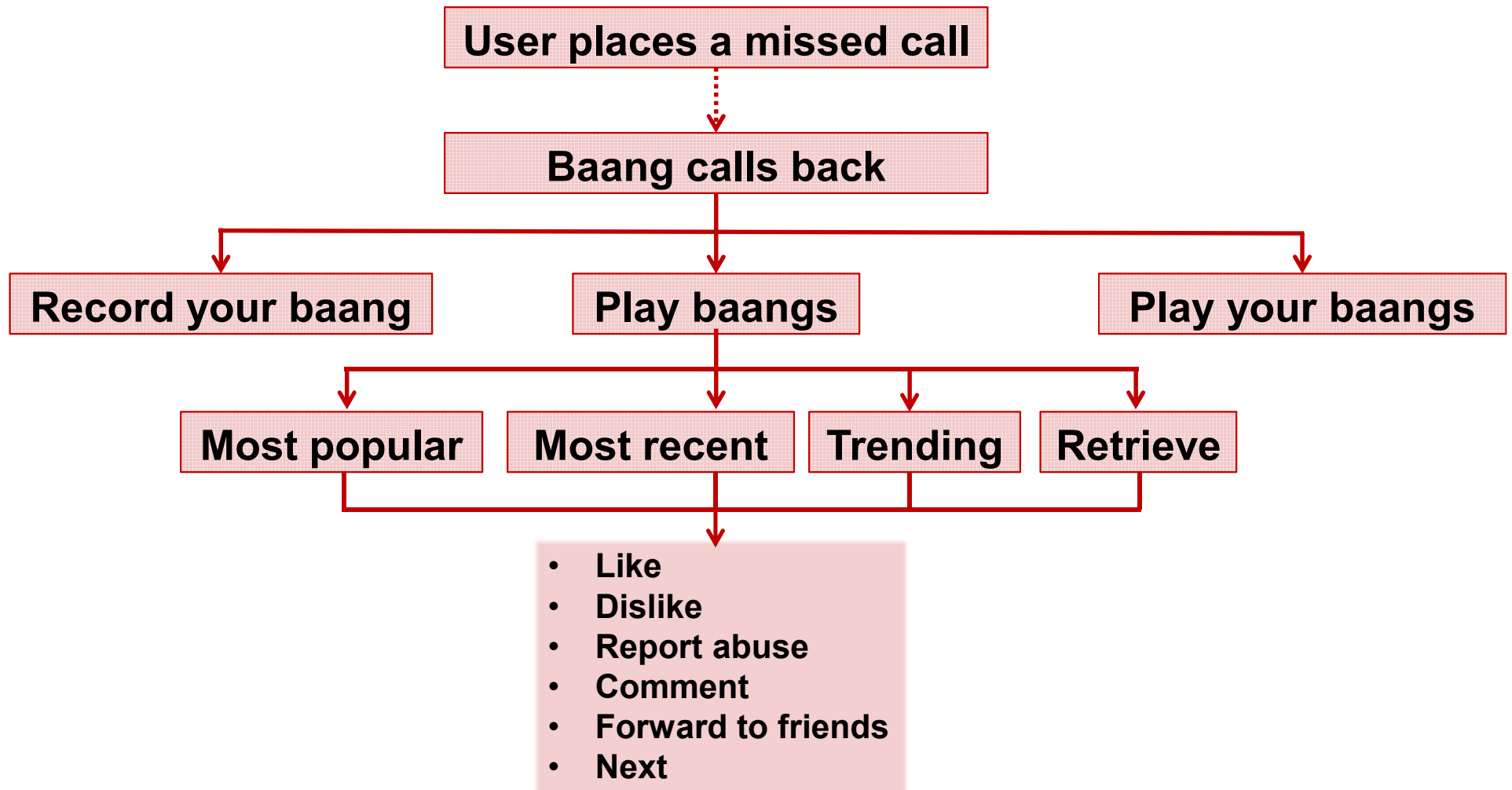
Low Retention



Speech-based Social Media

- **Baang¹**: A voice-based, telephone-based community forum (Pakistan; 2015-16)
- A voice-based “Reddit” for low-literates
- Users can:
 - Post short voice messages
 - Play messages recorded by others
 - Like/Dislike/Report/Comment/Forward
- Messages could be sorted by popularity, recency or recent-popularity (trending)

Baang (بانگ)



Baang (بانگ)

Within 71 days of its launch *Baang* accumulated:

- **42,500** phone calls (by **1,550** users)
- **8,469** posts (by **888** users)
- Posts were played **393,448** times (by **1,524** users)
- **106 hours** of speech data
 - Average recording duration: **45 seconds**
- Interactions included:
 - **29,312** up votes (by **1,110** users)
 - **14,097** down votes (by **734** users)
 - **7,714** report abuse votes (by **564** users)
 - **10,965** comments (by **603** users)

Content of the Posts

Out of 1,718 recordings:

- Actual content: 1,713
- Unclear/noise etc.: 6

Out of 1,585 recordings:

- Civilized: 1,483
- Uncivilized: 102

Gender and Languages

Out of 1,718 recordings:

- Male: 1,552
- Female: 166

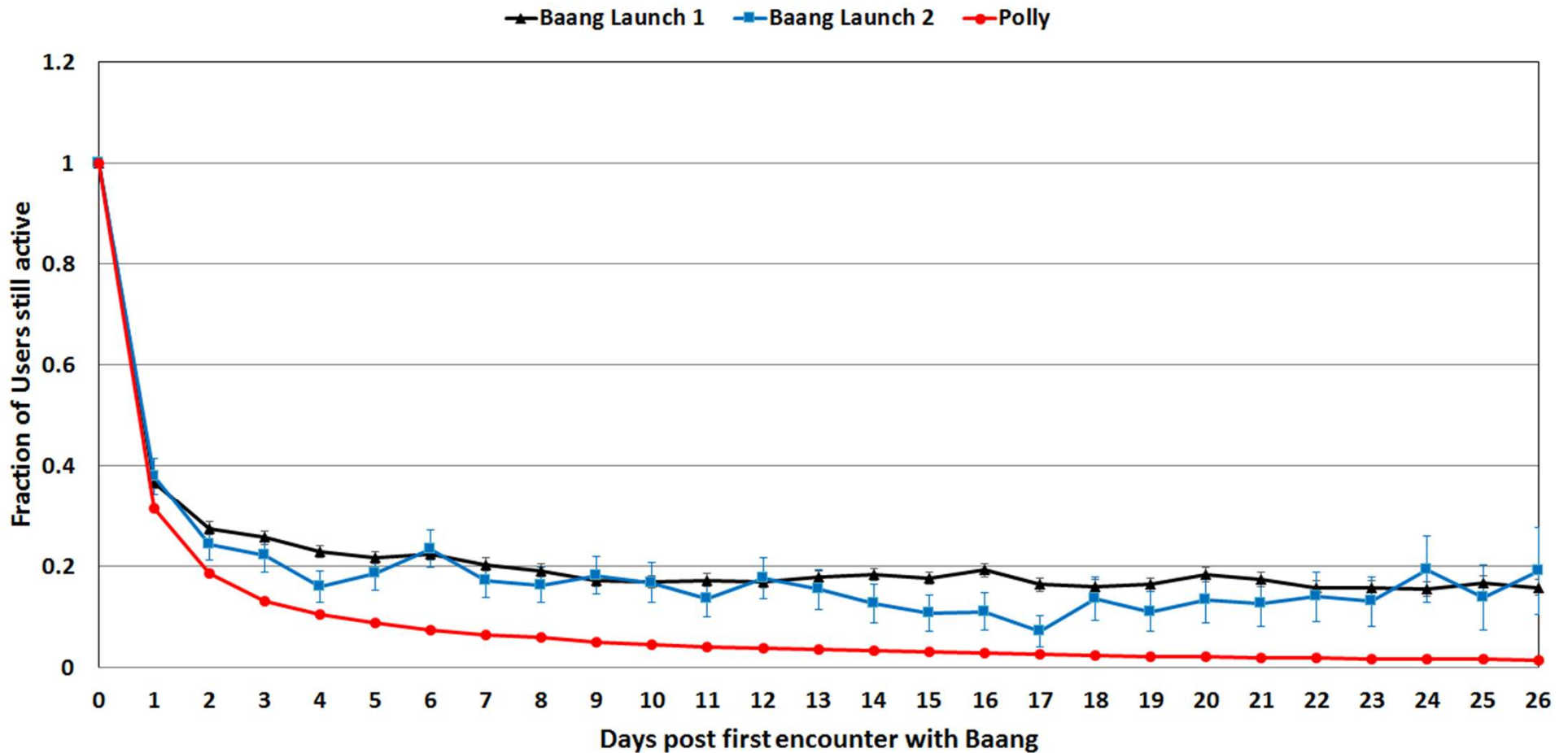
Out of 1,638 recordings:

- Urdu: 1,403
- Pushto: 82
- Arabic: 55
- Punjabi: 45
- Saraiki 26
- English 14
- Sindhi 12
- Farsi 1

Among 100 Top Voted Posts

- All 100 files contained actual content
- Female: 41 / Male: 59
- Civilized: 98 / Uncivilized: 2
- **Content Type:**
 - Naat: 34
 - Argument: 13
 - Poetry: 12
 - Dua, Hadees, Hamd: 10
 - Songs: 9
 - Famous quotes: 6
 - Quran recitation: 4
 - Random facts: 4
 - Advertisement: 2
 - Rude Language: 2
 - Requests for votes: 2
 - Personal introduction: 1
 - News updates: 1

User Retention



Significantly higher, sustained retention as compared to Polly

For more details please visit

<http://aghaaliraza.com>

<http://www.cs.cmu.edu/~araza/>



Thank you!

