



# Analyzing Weibo, the Chinese Twitter

Ken Hu

President and Founder

Sentiment Analysis Symposium, October 30th 2012

# This Is Weibo



# This Is Chinese

你在天上願你的名兒被人尊  
敬願你的國圖降臨願你的旨  
意能彀成就在地如在天一樣  
我們需用的糧食求你今日給  
我求你免我的債照我免人債  
的樣兒不要由我入迷願你  
我出惡爲的是邦國權勢榮耀  
一概是你的直到世世代代這  
是我心所願的啊



# This is Chinese on Weibo

你以为我会告诉你们跟罗小弟半个月不见，今日，她讲了好几个猥琐笑话吗？并且很有笑果!!! @天气这么好 你只听了一个对不对? 😁 现在罗小弟正在直播新段子《就算打扫卫生也会是一场爱恋的开始》&《求知少女高调示爱高富帅地理老师》&《吃货们的奇葩生活》

9月30日 08:44 来自新浪微博 | 举报

转发(1) | 收藏 | 评论



# This is Chinese on Weibo

Numerous Sentences

Long Segments

你以为我会告诉你们跟罗小弟半个月不见，今日，她讲了好几个猥琐笑话吗？并且很有笑果!!! @天气这么好 你只听了一个对不对? 😊 现在罗小弟正在直播新段子《就算打扫卫生也会是一场爱恋的开始》&《求知少女高调示爱高富帅地理老师》&《吃货们的奇幻

新浪微博 | 举报

转发 | 评论

Homophones & Puns

No Capitalization

Slangs

Homographs



# This Is Us



# Hints

你以为我会告诉你们跟罗小弟半个月不见，今日，她讲了好几个猥琐笑话吗？并且很有笑果!!! @天气这么好 你只听了一个对不对？😁现在罗小弟正在直播新段子《就算打扫卫生也会是一场爱恋的开始》&《求知少女高调示爱高富帅地理老师》&《吃货们的奇葩生活》

9月30日 08:44 来自新浪微博 | 举报

转发(1) | 收藏 | 评论



# Hints

你以为我会告诉你们跟罗小弟半个月不见，今日，她讲了好几个猥琐笑话吗？并且很有笑果!!! @天气这么好 你只听了一个对不对? 😁 现在罗小弟正在直播新段子《就算打扫卫生也会是一场爱恋的开始》&《求知少女高调示爱高富帅地理老师》&《吃货们的奇幻

新浪微博 | 举报

转发(1) | 收藏 | 评论

Syntax





# Hints

你以为我会告诉你们跟罗小弟半个月不见，今日，她讲了好几个猥琐笑话吗？并且很有笑果!!! @天气这么好 你只听了一个对不对？😁现在罗小弟正在直播新段子《就算打扫卫生也会是一场爱恋的开始》&《求知少女高调示爱高富帅地理老师》&《吃货们的奇幻

新浪微博

转发(1) | 收藏 | 评论

Syntax

Stop Words



# Hints

Emoticons

你以为我会告诉你们跟罗小虎的绯闻吗？昨日，她讲了好几个猥琐笑话吗？并且很有笑果!!! @天气这么好 你只讲了一个对不对？😁现在罗小弟正在直播新段子《就算打扫卫生也会是一场爱恋的开始》&《求知少女高调示爱高富帅地理老师》&《吃货们的奇幻之旅》

新浪微博

转发(1) | 收藏 | 评论

Syntax

Stop Words



# Hints

Emoticons

你以为我会告诉你们跟罗小虎的绯闻吗？昨日，她讲了好几个猥琐笑话吗？并且很有笑果!!! @天气这么好 你只讲了一个对不对？😁现在罗小弟正在直播新段子《就算打扫卫生也会是一场爱恋的开始》&《求知少女高调示爱高富帅地理老师》&《吃货们的奇幻之旅》

新浪微博

1 | 收藏 | 评论

Syntax

Stop Words

Brackets



# Hints

Known Words

Emoticons

你以为我会告诉你们跟罗小虎过日子，她讲了好几个猥琐笑话吗？并且很有笑果!!! @天气这么好 你只讲了一个对不对？😁现在罗小弟正在直播新段子《就算打扫卫生也会是一场爱恋的开始》&《求知少女高调示爱高富帅地理老师》&《吃货们的奇幻之旅》

新浪微博

1 | 收藏 | 评论

Syntax

Stop Words

Brackets



# Take One

Negative

你以为我会告诉你们跟罗小弟半个月不见，今日，她讲了好几个猥琐笑话吗？并且很有笑果!!! @天气这么好 你只听了一个对不对? 😁 现在罗小弟正在直播新段子《就算打扫卫生也会是一场爱情的开始》&《...》&《...》&《吃我们的贵族生活》

9月30日 08:44 来自新浪微博 | 李

转发(1) | 收藏 | 评论

Positive



# Slangs & Puns

Pingyin

nǐ	yǐ	wéi	wǒ	huì	gào	su	nǐ	men	gēn	luō	xiǎo	dì	bàn	ge	yuè	bù	jiàn		jīn
你	以	为	我	会	告	诉	你	们	跟	罗	小	弟	半	个	月	不	见	,	今
rì		tā	jiǎng	le	hǎo	jǐ	gè	wěi	suǒ	xiǎo	huà	mǎ		bìng	qiě	hěn	yǒu	xiào	guǒ
日	,	她	讲	了	好	几	个	猥	琐	笑	话	吗	?	并	且	很	有	笑	果
				nǐ	zhǐ	tīng	le	yī	gè	duì	bù	duì		xiàn	zài	luō	xiǎo	dì	zhèng
!	!	!		你	只	听	了	一	个	对	不	对	?	现	在	罗	小	弟	正
zài	zhí	bō	xīn	duàn	zǐ														
在	直	播	新	段	子														



# Slangs & Puns

Pinyin

nǐ	yǐ	wéi	wǒ	huì	gào	su	nǐ	men	gēn	luō	xiǎo	dì	bàn	ge	yuè	bù	jiàn		jīn
你	以	为	我	会	告	诉	你	们	跟	罗	小	弟	半	个	月	不	见	,	今
rì		tā	jiǎng	le	hǎo	jǐ	gè	wěi	suǒ	xiǎo	huà	mǎ		bìng	qiě	hěn	yǒu	xiào	guǒ
日	,	她	讲	了	好	几	个	猥	琐	笑	话	吗	?	并	且	很	有	笑	果
				nǐ	zhǐ	tīng	le	yī	gè	duì	bù	duì		xiàn	zài	luō	xiǎo	dì	zhèng
!	!	!		你	只	听	了	一	个	对	不	对	?	现	在	罗	小		正
zài	zhí	bō	xīn	duàn	zǐ														
在	直	播	新	段	子														

Phonetic algorithm  
(Soundex?)



# Slangs & Puns

Pinyin

nǐ	yǐ	wéi	wǒ	huì	gào	su	nǐ	men	gēn	luō	xiǎo	dì	bàn	ge	yuè	bù	jiàn		jīn
你	以	为	我	会	告	诉	你	们	跟	罗	小	弟	半	个	月	不	见	,	今
rì		tā	jiǎng	le	hǎo	jǐ	gè	wěi	suǒ	xiǎo	huà	mǎ		bìng	qiě	hěn	yǒu	xiào	guǒ
日	,	她	讲	了	好	几	个	猥	琐	笑	话	吗	?	并	且	很	有	笑	果
				nǐ	zhǐ	tīng	le	yī	gè	duì	bù	duì		xiàn	zài	luō	xiǎo	dì	zhèng
!	!	!		你	只	听	了	一	个	对	不	对	?	现	在	罗	小		正
zài	zhí	bō	xīn	duàn	zǐ														
在	直	播	新	段	子														

笑 + 效果

Phonetic algorithm  
(Soundex?)





# Slangs & Puns

Pinyin

nǐ	yǐ	wéi	wǒ	huì	gào	su	nǐ	men	gēn	luō	xiǎo	dì	bàn	ge	yuè	bù	jiàn		jīn
你	以	为	我	会	告	诉	你	们	跟	罗	小	弟	半	个	月	不	见	,	今
rì		tā	jiǎng	le	hǎo	jǐ	gè	wěi	suǒ	xiǎo	huà	mǎ		bìng	qiě	hěn	yǒu	xiào	guǒ
日	,	她	讲	了	好	几	个	猥	琐	笑	话	吗	?	并	且	很	有	笑	果
				nǐ	zhǐ	tīng	le	yī	gè	duì	bù	duì		xiàn	zài	luō	xiǎo	dì	zhèng
!	!	!		你	只	听	了	一	个	对	不	对	?	现	在	罗	小		正
zài	zhí			àn	zǐ														
在	在			子															

Laughter Effects

笑 + 效果

Phonetic algorithm  
(Soundex?)



# Slangs & Puns

Positive

Pinyin

nǐ	yǐ	wéi	gēn	luō	xiǎo	dì	bàn	ge	yuè	bù	jiàn	jīn					
你	以	为	跟	罗	小	弟	半	个	月	不	见	，	今				
rì	tā	jiǎng	suǒ	xiào	huà	mǎ	bìng	qiě	hěn	yǒu	xiào	guǒ					
日	，	她	讲	猥	琐	笑	话	吗	？	并	且	很	有	笑	果		
！	！	！	zhǐ	tīng	le	yī	gè	duì	bù	duì	xiàn	zài	luō	xiǎo	dì	zhèng	
！	！	！	你	只	听	了	一	个	对	不	对	？	现	在	罗	小	正
zài	zhí	zài	zài	zài	zài	zài	zài	zài	zài	zài	zài	zài	zài	zài	zài	zài	zài
在	在	在	在	在	在	在	在	在	在	在	在	在	在	在	在	在	在

Laughter Effects

笑 + 效果

Phonetic algorithm  
(Soundex?)



# Take Two

Negative

你以为我会告诉你们跟罗小弟半个月不见，今日，她讲了好几个猥琐笑话吗？并且很有笑果!!! @天气这么好 你只听了一个对不对? 😁 现在罗小弟正在直播新段子《就算打扫卫生也会是一场爱情的开始》&《...》&《...》&《吃我们的...》

Positive

Positive

转发(1) | 收藏 | 评论



# Conclusions



Use every little details

Identify what variable are really necessary for the analysis

“Translate” it to sound





Chinese Social Media Monitoring

Ken Hu

ken@getsoshio.com

@whosbacon

@soshiolize

611 Broadway, Suite 704. New York, NY 10012