



「君の名は」に対する中国人の感情分析
-- ディープラーニングに基づいた自然言語処理を通じて

九州大学統合新領域学府 / 2017 年度論文中間発表

発表者：陳喴康

指導教員：麻生先生

目次

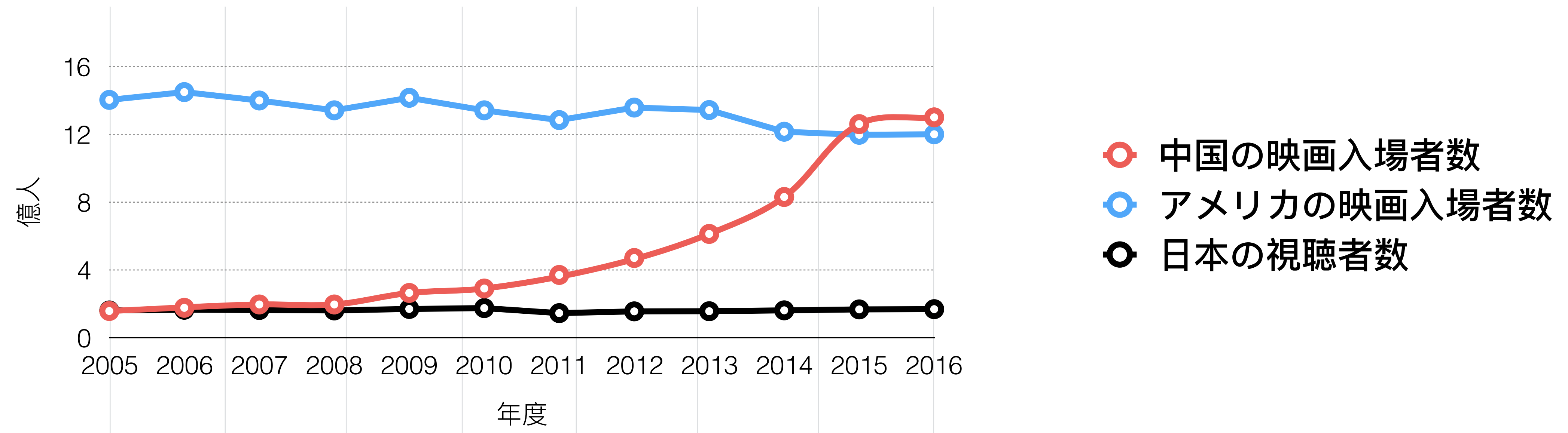
I.	研究背景	05
II.	研究目的と研究対象の選択	10
III.	研究方法の選択	14
IV.	研究の流れ	20

● 研究背景

- 研究目的と研究対象の選択
- 研究方法の選択
- 研究の流れ

/ 研究背景：中国映画業界の発展状況

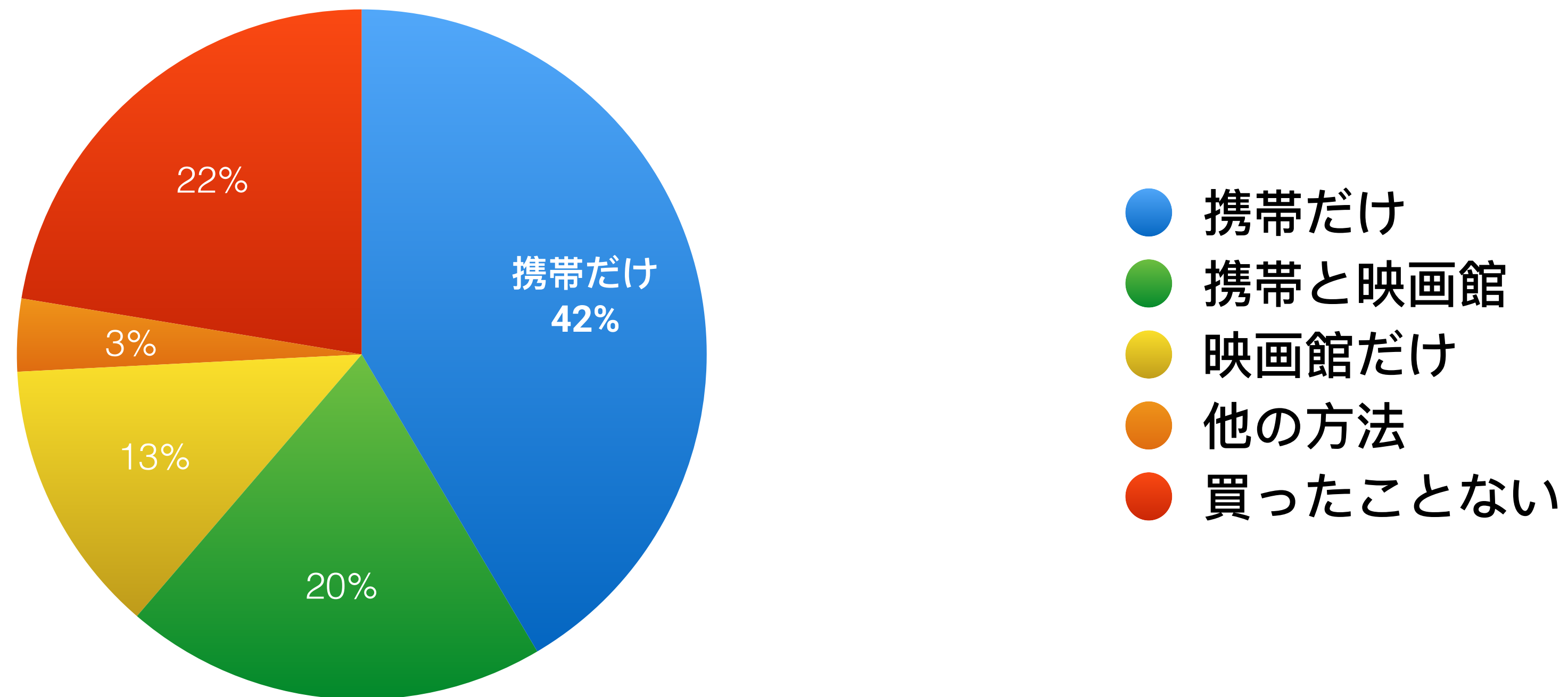
日・中・米の映画入場者数



出典：日本映画制作者連盟のデータより作成

/ 研究背景：中国映画業界の発展状況

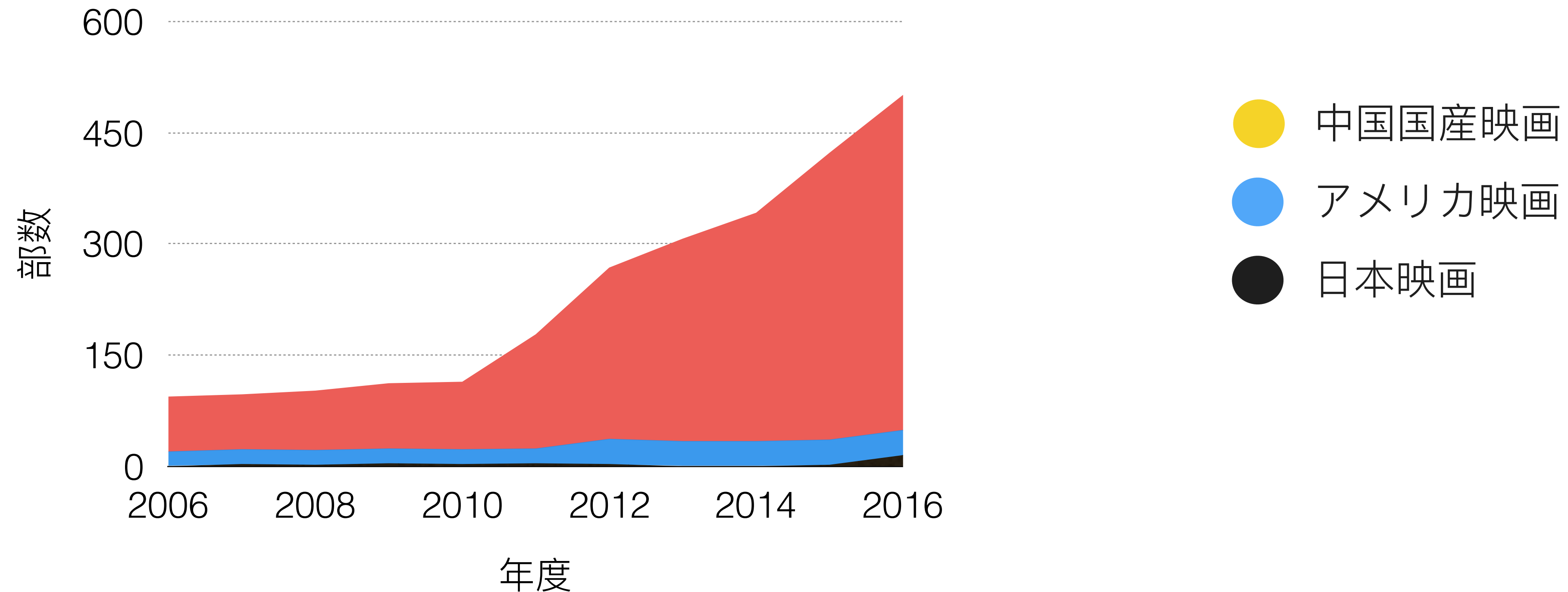
中国人チケットを購入する方法



出典：wechatチケット2016年のアンケートデータより作成（アンケート数：11261）

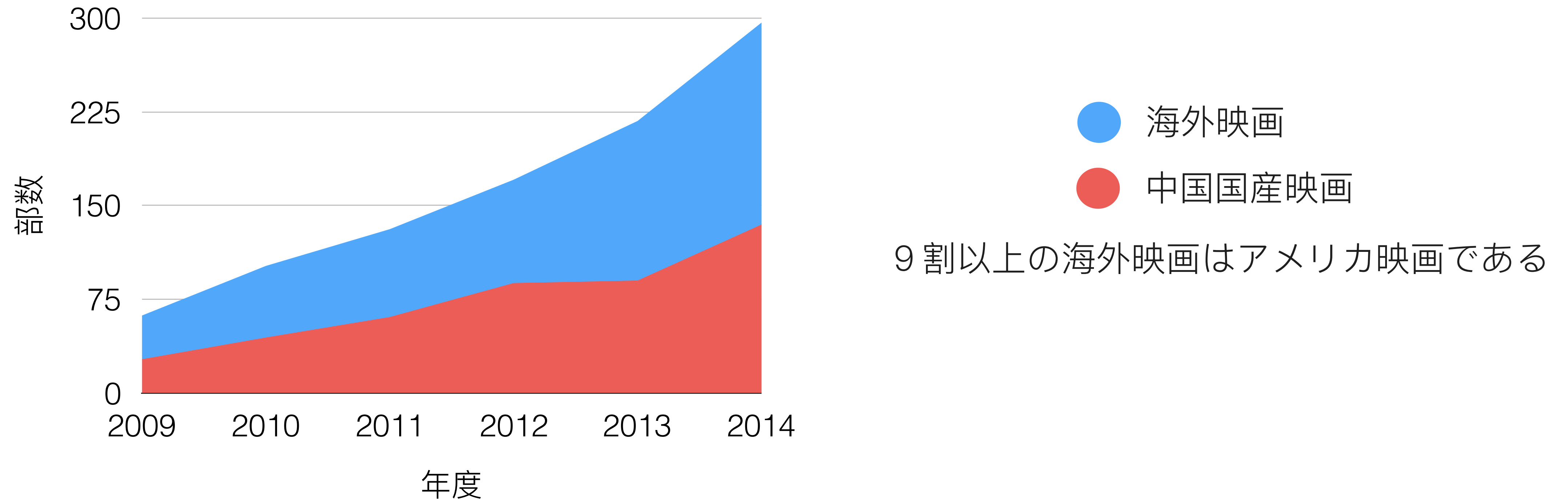
/ 研究背景：中国映画市場における海外映画

中国映画市場における日・米映画数量



広電総局のデータより作成

中国映画市場における海外映画の興行収入占比



広電総局のデータより作成

／ 研究背景：まとめ

中国映画市場の特徴

1

映画入場者数多い

未来の潜在力がまたある

3

外国映画の大ヒット

文化の包容性がある。国産映画の口コミが良くない

2

映画の輸入審査制度

国家広電総局等が映画に関して発布した管理法の規範化

4

携帯でチケットを購入する

映画の口コミが映画観賞の動機になる

/ 研究背景：まとめ

中国映画市場における日本映画の潜力

1

海外市場開発の必要性

少子化問題に伴って、日本国内映画
入場数低下する可能性がある

3

ファンズ基礎がある

中国映画市場7割以上の消費者は1985年
後に生まれた人、この人たちが日本文化
に対する好感を持ち

2

日本経済産業省の政策サポート

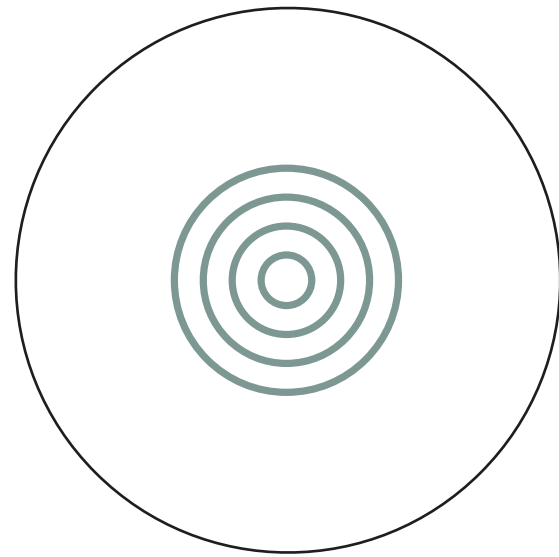
コンテンツ産業は日本のクールジャパ
ン戦略の一部になった

4

口コミの良さ

douban.comのデータによると、日本映画
の口コミは中国国産映画の口コミより高い

- 研究背景
- **研究目的と研究対象の選択**
- 研究方法の選択
- 研究の流れ



研究目的

本研究は歴代最も高い興行収入となった『君の名は』への感性印象形成の実態を明らかにすることを目的とする。



研究意義

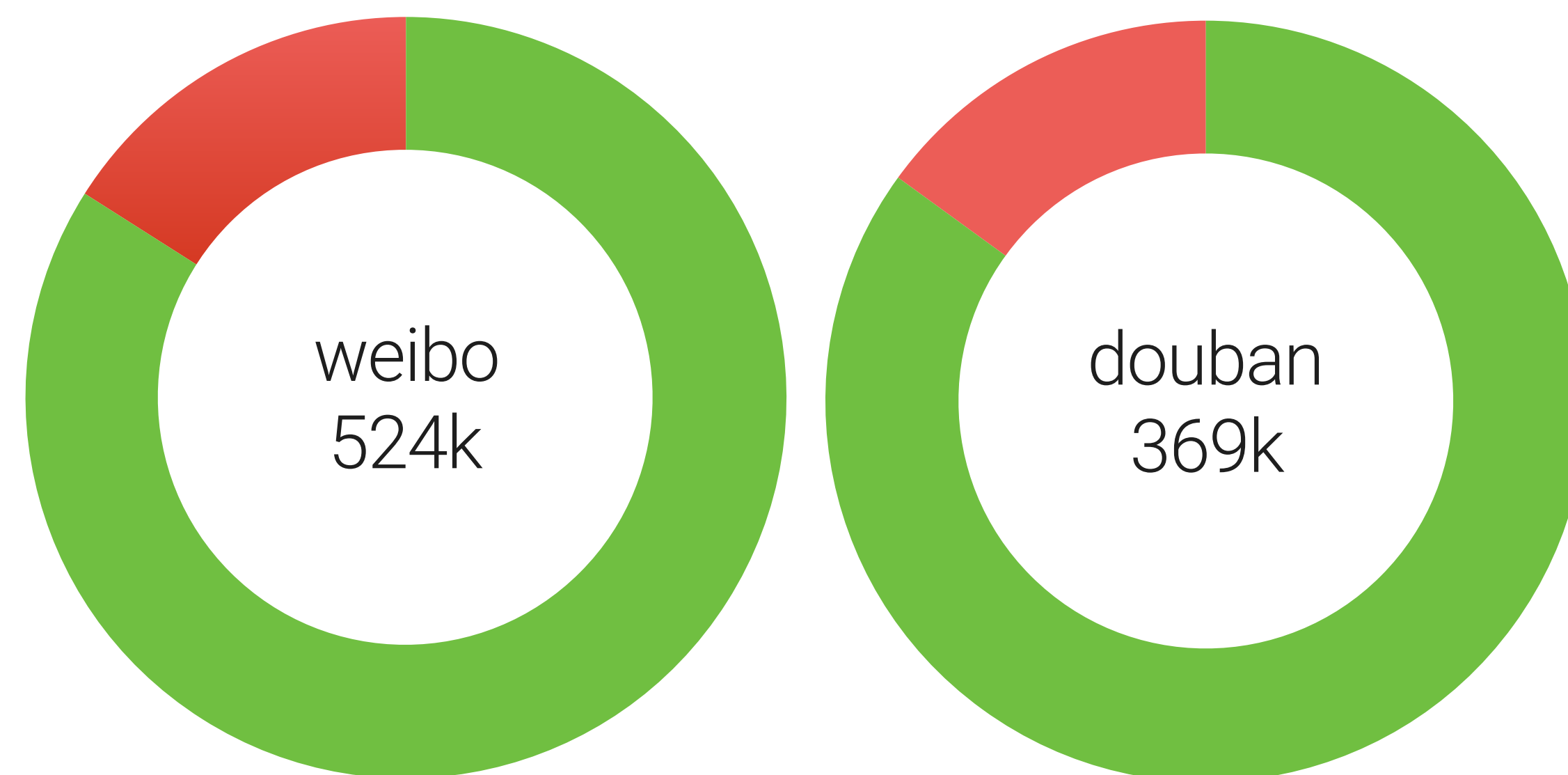
『君の名は』の映画自体が多くのファンを生み出したではなく、インターネットの口コミの原因で流行ってるので、この映画に対する中国人の感情分析ができれば、新しい視点による日本映画が海外市場における特別な魅力へ新たな分析枠組の提供もできる。

/ 研究対象の選択

中日映画歴史に前例がない話題性や高い評価

研究対象:

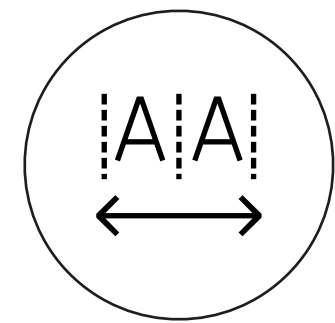
『君の名は』のdouban（映画評価サイト）の映画評価内容(140 文字制限の短評、約36.9万条)である。



点数：両方とも8.5点

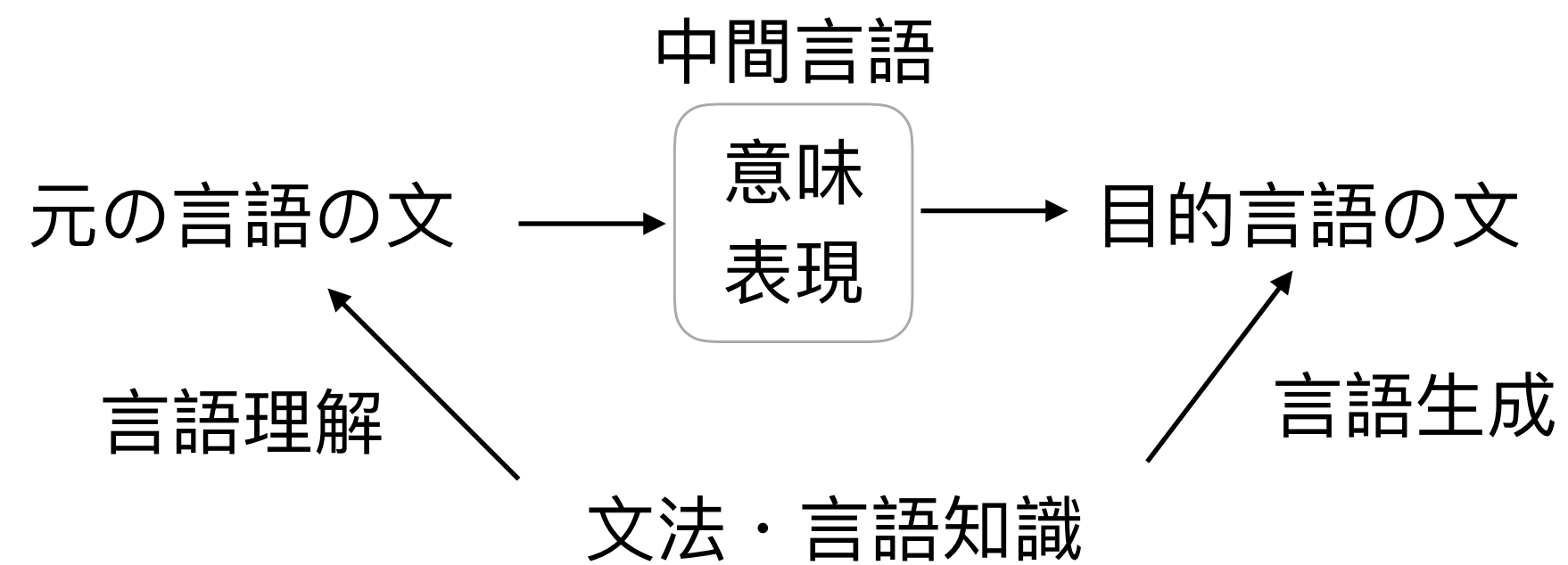
- 研究背景
- 研究目的と研究対象の選択
- **研究方法の選択**
- 研究の流れ

コンピュータで言語の感情を分析する方法

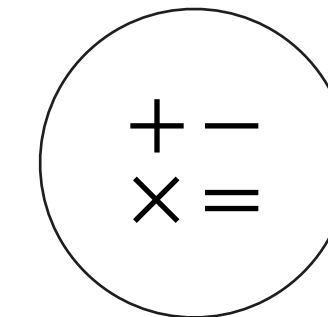


辞書・文法的アプローチ

文法，言語知識による「浅い」言語解析

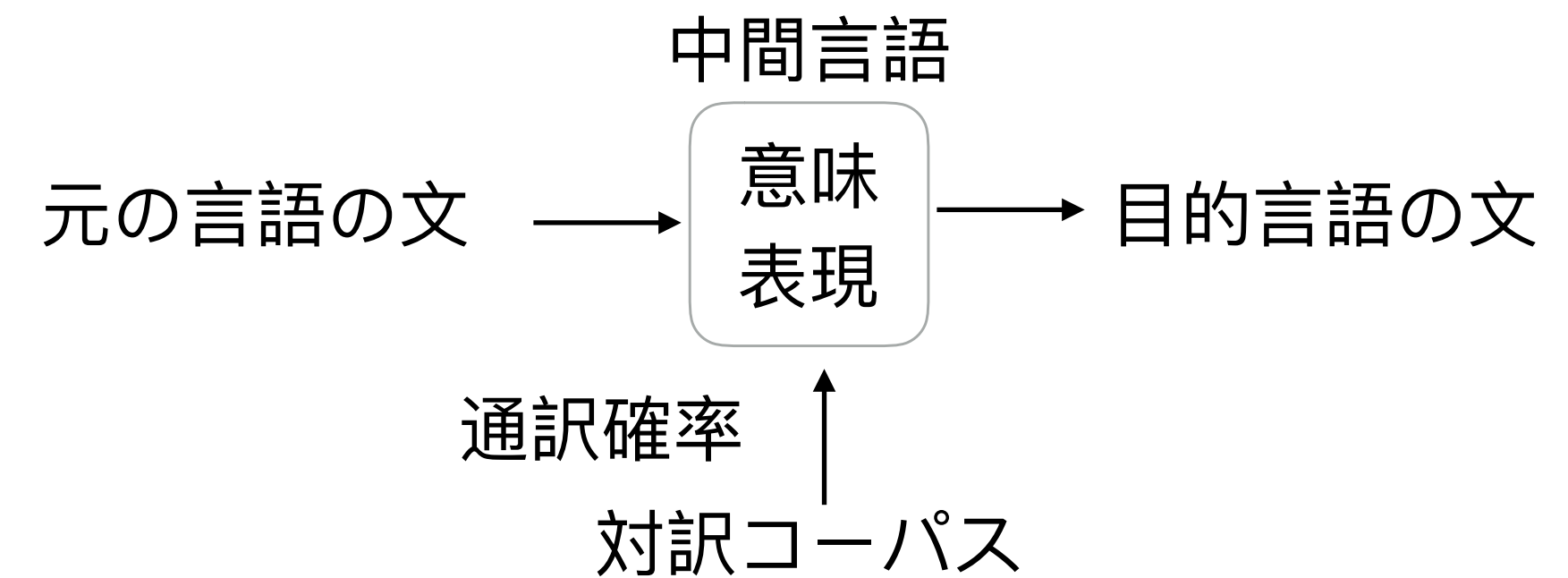


例：言語理解に基づく自動翻訳の正解率最大90%



数学的アプローチ

統計的手法による「浅い」言語解析



例：コーパス（用例）に基づく自動翻訳の正解率最大97%



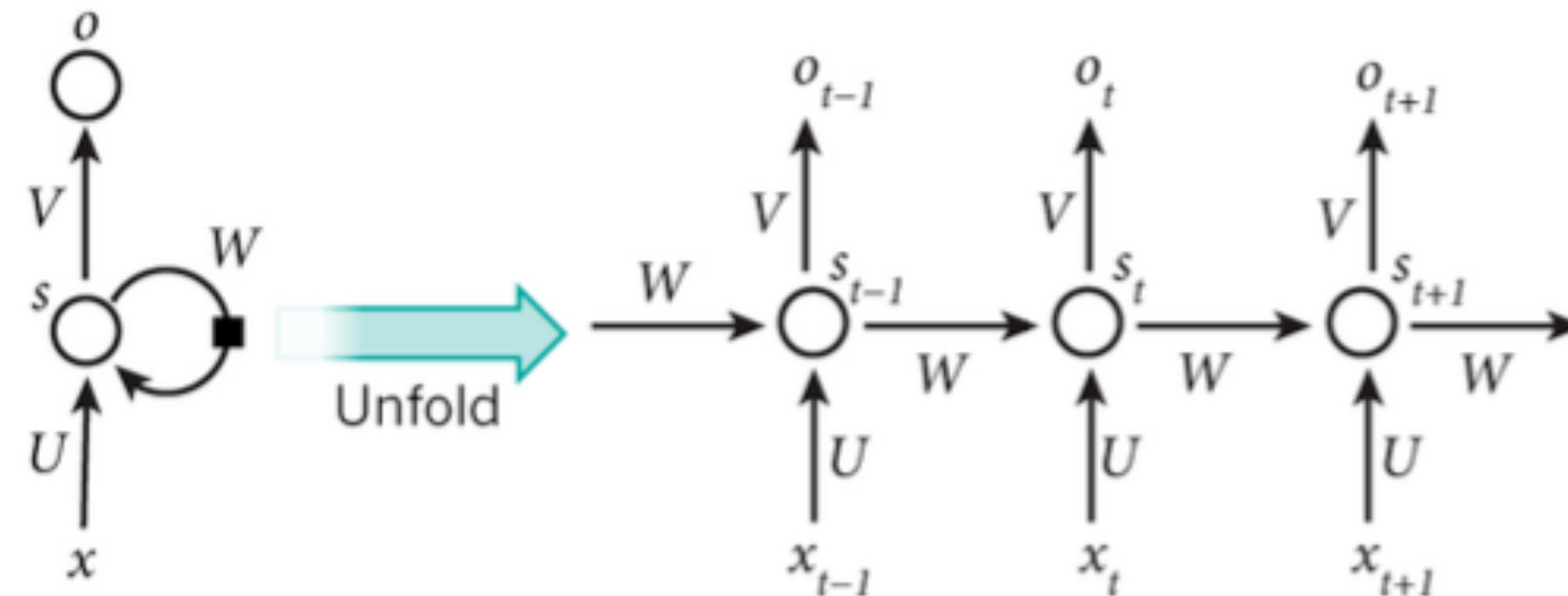
辞書・文法的アプローチの問題点

「深い」言語解析をどうするかという問題は依然として残る、辞書・文法的アプローチでは、多義語の意味、単語間の関係を理解すること可能ですが、言語の使用場合が変わる時に、誤解する可能性がある。例えば、「彼のことが大嫌いだ、かっこよすぎ」という言葉が辞書・文法的アプローチで理解すると、「大」は副詞である、「嫌い」はマナイス面の情感である、「かっこよすぎ」は形容詞である。全体的に見ると、間違いなくマナイス面の情感だった。でも実際に、この文書の情感は「大好き」なのだ。

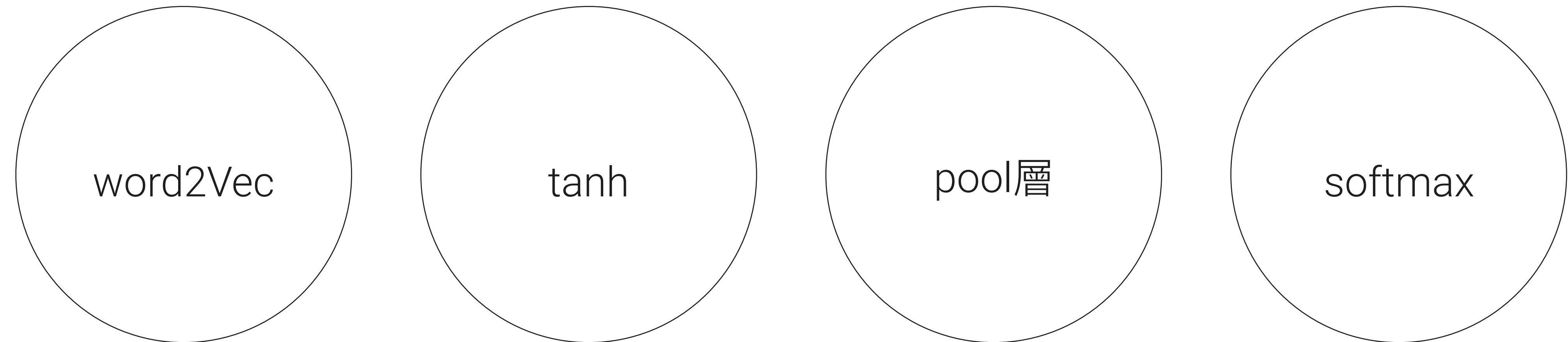
本研究では、数学的アプローチを使うので、文書に対する最初の処理はWord2Vecである。つまり、テキストコーパスを入力すると、出力結果は空間の中のベクトルのセットである。このベクトルのセットは空間の中にいるので、より精確の結果が出される。

数学的アプローチの解決方法

本研究では、数学的アプローチを使うので、文書に対する最初の処理はWord2Vecである。Word2Vecは、テキストをディープネットワーク解釈できる数値形式に変えた。つまり、テキストコーパスを入力すると、出力結果は空間の中のベクトルのセットである。このベクトルのセットは空間の中にあるので、より精確の結果が出される。



本研究のニューラルネットワークモデル（4層）



- 研究背景
- 研究目的と研究対象の選択
- 研究方法の選択
- **研究の流れ**

ご清聴ありがとうございます