A Computational & Behavioral Analysis of the Linguocultural Phenomenon of Food

Chelsea Madison McBride

Institut für Maschinelle Sprachverarbeitung
Universität Stuttgart

February 6, 2020

Universität
Stuttgart

What is a "*Linguocultural*" Phenomenon?

"*Linguocultural*" ...

What is a "*Linguocultural*" Phenomenon?

"*Linguocultural*" ...

- Coined by Russian linguist Vorobyov V.V (1997)

What is a "*Linguocultural*" Phenomenon?

"*Linguocultural*" ...

- Coined by Russian linguist Vorobyov V.V (1997)
- Refers to the relationship between language and culture

What is a "*Linguocultural*" Phenomenon?

"*Linguocultural*" ...

- Coined by Russian linguist Vorobyov V.V (1997)
- Refers to the relationship between language and culture

We define "*language*" and "*culture*" according to Encyclopedia Britannica (2019)

What is a "*Linguocultural*" Phenomenon?

"*Linguocultural*" ...

- Coined by Russian linguist Vorobyov V.V (1997)
- Refers to the relationship between language and culture

We define "*language*" and "*culture*" according to Encyclopedia Britannica (2019)

- "***Language***" → Only a single vessel of any given culture

What is a "*Linguocultural*" Phenomenon?

"*Linguocultural*" ...

- Coined by Russian linguist Vorobyov V.V (1997)
- Refers to the relationship between language and culture

We define "*language*" and "*culture*" according to Encyclopedia Britannica (2019)

- "***Language***" → Only a single vessel of any given culture
- "***Culture***" → All aspects of human life insofar as they are determined or conditioned by membership in a society

What is a "*Linguocultural*" Phenomenon?

"*Linguocultural*" ...

- Coined by Russian linguist Vorobyov V.V (1997)
- Refers to the relationship between language and culture

We define "*language*" and "*culture*" according to Encyclopedia Britannica (2019)

- "***Language***" → Only a single vessel of any given culture
- "***Culture***" → All aspects of human life insofar as they are determined or conditioned by membership in a society

"***Language as the manifestation of human thought and behavior in a society.***"

Food for Culture Prediction?

The relationship between language and culture is an important area of research...

Food for Culture Prediction?

The relationship between language and culture is an important area of research...

- **Food language** is one of the most culturally dependent phenomena able to be empirically tested (Langellier et al., 2015, Wiegand et al., 2014, Jurafsky, 2014)

Food for Culture Prediction?

The relationship between language and culture is an important area of research...

- **Food language** is one of the most culturally dependent phenomena able to be empirically tested (Langellier et al., 2015, Wiegand et al., 2014, Jurafsky, 2014)
- By creating **numerical human judgments in the form of food norms**, we can effectively measure and predict human behavioral patterns in regard to food language use

Food for Culture Prediction?

The relationship between language and culture is an important area of research...

- **Food language** is one of the most culturally dependent phenomena able to be empirically tested (Langellier et al., 2015, Wiegand et al., 2014, Jurafsky, 2014)
- By creating **numerical human judgments in the form of food norms**, we can effectively measure and predict human behavioral patterns in regard to food language use

Make educated predictions regarding human behavior based on historical, current trends

What Did We Do?

Two main aims of this thesis...

What Did We Do?

Two main aims of this thesis...

1. Perform a **behavioral analysis** of human thought and behavioral patterns pertaining to food items from two very different cultures

What Did We Do?

Two main aims of this thesis...

1. Perform a **behavioral analysis** of human thought and behavioral patterns pertaining to food items from two very different cultures
2. Perform a **computational analysis** of linguistic features encoded in food-word representations

What Did We Do?

Two main aims of this thesis...

1. Perform a **behavioral analysis** of human thought and behavioral patterns pertaining to food items from two very different cultures
2. Perform a **computational analysis** of linguistic features encoded in food-word representations

Agenda

We will present this work as follows...

Agenda

We will present this work as follows...

1. Motivation + Research questions

Agenda

We will present this work as follows...

1. Motivation + Research questions
2. Multi-domain Corpora & Dataset

Agenda

We will present this work as follows...

1. Motivation + Research questions
2. Multi-domain Corpora & Dataset
3. Behavioral & Computational Methods/ Models

Agenda

We will present this work as follows...

1. Motivation + Research questions
2. Multi-domain Corpora & Dataset
3. Behavioral & Computational Methods/ Models
4. Analysis: Behavioral & Computational

Agenda

We will present this work as follows...

1. Motivation + Research questions
2. Multi-domain Corpora & Dataset
3. Behavioral & Computational Methods/ Models
4. Analysis: Behavioral & Computational
5. Discussion

Agenda

We will present this work as follows...

1. Motivation + Research questions
2. Multi-domain Corpora & Dataset
3. Behavioral & Computational Methods/ Models
4. Analysis: Behavioral & Computational
5. Discussion
6. Conclusions

Motivation

Motivation

Novel investigation how language interacts with culture using a data-driven, computational approach by capturing linguistic information using continuous space word representations

Motivation

Novel investigation how language interacts with culture using a data-driven, computational approach by capturing linguistic information using continuous space word representations

Main research questions:

Motivation

Novel investigation how language interacts with culture using a data-driven, computational approach by capturing linguistic information using continuous space word representations

Main research questions:

1. Can we predict cultural information in the form of norms from linguistic information in the form of word embeddings?

Motivation

Novel investigation how language interacts with culture using a data-driven, computational approach by capturing linguistic information using continuous space word representations

Main research questions:

1. Can we predict cultural information in the form of norms from linguistic information in the form of word embeddings?
2. How do corpora from different domains effect model performance?

Motivation

Novel investigation how language interacts with culture using a data-driven, computational approach by capturing linguistic information using continuous space word representations

Main research questions:

1. Can we predict cultural information in the form of norms from linguistic information in the form of word embeddings?
2. How do corpora from different domains effect model performance?

Contributions:

Motivation

Novel investigation how language interacts with culture using a data-driven, computational approach by capturing linguistic information using continuous space word representations

Main research questions:

1. Can we predict cultural information in the form of norms from linguistic information in the form of word embeddings?
2. How do corpora from different domains effect model performance?

Contributions:

Improve localization techniques, advance marketing and economic strategies, assist medical and public health sectors

Dataset & Corpora

Food Norms

While large norm sets are available in the field of psychology (Bradley et al., 1999, Westbury et al., 2015)... there are no food norm sets that we know of

Food Norms

While large norm sets are available in the field of psychology (Bradley et al., 1999, Westbury et al., 2015)... there are no food norm sets that we know of

- Evaluation data collected by Chen Xu from Computer Science

Food Norms

While large norm sets are available in the field of psychology (Bradley et al., 1999, Westbury et al., 2015)... there are no food norm sets that we know of

- Evaluation data collected by Chen Xu from Computer Science
    - Novel data set of 200 word food items, evaluated by 25 participants from Britain and China, according to 6 categories on a scale of 1 to 9 (ex: 1 = happy, 9 = sad)

Food Norms

While large norm sets are available in the field of psychology (Bradley et al., 1999, Westbury et al., 2015)... there are no food norm sets that we know of

- Evaluation data collected by Chen Xu from Computer Science
  - Novel data set of 200 word food items, evaluated by 25 participants from Britain and China, according to 6 categories on a scale of 1 to 9 (ex: 1 = happy, 9 = sad)

1. *Happy/Sad*
2. *Healthy/Unhealthy*
3. *Cheap/Expensive*
4. *Environmentally-Friendly/Environmentally-Unfriendly*
5. *Casual-Simple/Posh-Elegant*
6. *Local-Traditional/Foreign-Not Traditional*

Food Norms

While large norm sets are available in the field of psychology (Bradley et al., 1999, Westbury et al., 2015)... there are no food norm sets that we know of

- Evaluation data collected by Chen Xu from Computer Science
  - Novel data set of 200 word food items, evaluated by 25 participants from Britain and China, according to 6 categories on a scale of 1 to 9 (ex: 1 = happy, 9 = sad)

1. *Happy/Sad*
2. *Healthy/Unhealthy*
3. *Cheap/Expensive*
4. *Environmentally-Friendly/Environmentally-Unfriendly*
5. *Casual-Simple/Posh-Elegant*
6. *Local-Traditional/Foreign-Not Traditional*

Socially constructed human judgment for the food item based on ingrained cultural predispositions of participants → empirically measurable ***norm*** as a numerical value

Multi-Domain Corpora

Three domains: Wikipedia, news, and web

Multi-Domain Corpora

Three domains: Wikipedia, news, and web

- WaCkypedia:

Multi-Domain Corpora

Three domains: Wikipedia, news, and web

- WaCkypedia:
  - 2009 dump of the English Wikipedia entries consisting of about 800 million tokens

Multi-Domain Corpora

Three domains: Wikipedia, news, and web

- WaCkypedia:
  - 2009 dump of the English Wikipedia entries consisting of about 800 million tokens
  - We used the full corpus (800M, 14 GB)

Multi-Domain Corpora

Three domains: Wikipedia, news, and web

- WaCkypedia:
  - 2009 dump of the English Wikipedia entries consisting of about 800 million tokens
  - We used the full corpus (800M, 14 GB)
- Gigaword:

Multi-Domain Corpora

Three domains: Wikipedia, news, and web

- WaCkypedia:
  - 2009 dump of the English Wikipedia entries consisting of about 800 million tokens
  - We used the full corpus (800M, 14 GB)
- Gigaword:
  - Over 12 GB of text from seven distinct sources of international English news-wire outlets

Multi-Domain Corpora

Three domains: Wikipedia, news, and web

- WaCkypedia:
  - 2009 dump of the English Wikipedia entries consisting of about 800 million tokens
  - We used the full corpus (800M, 14 GB)
- Gigaword:
  - Over 12 GB of text from seven distinct sources of international English news-wire outlets
  - Corpus was limited to the *New York Times* US English (4B, 6.7 GB)

Multi-Domain Corpora

Three domains: Wikipedia, news, and web

- WaCkypedia:
  - 2009 dump of the English Wikipedia entries consisting of about 800 million tokens
  - We used the full corpus (800M, 14 GB)
- Gigaword:
  - Over 12 GB of text from seven distinct sources of international English news-wire outlets
  - Corpus was limited to the *New York Times* US English (4B, 6.7 GB)
- ukWaC:

Multi-Domain Corpora

Three domains: Wikipedia, news, and web

- WaCkypedia:
  - 2009 dump of the English Wikipedia entries consisting of about 800 million tokens
  - We used the full corpus (800M, 14 GB)
- Gigaword:
  - Over 12 GB of text from seven distinct sources of international English news-wire outlets
  - Corpus was limited to the *New York Times* US English (4B, 6.7 GB)
- ukWaC:
  - 2 billion English words collected from a web crawl limited to the .uk domain

Multi-Domain Corpora

Three domains: Wikipedia, news, and web

- WaCkypedia:
  - 2009 dump of the English Wikipedia entries consisting of about 800 million tokens
  - We used the full corpus (800M, 14 GB)
- Gigaword:
  - Over 12 GB of text from seven distinct sources of international English news-wire outlets
  - Corpus was limited to the *New York Times* US English (4B, 6.7 GB)
- ukWaC:
  - 2 billion English words collected from a web crawl limited to the .uk domain
  - We used the full corpus (1.9B, 12 GB)

Methods

Computational Methods: Continuous Space Word Representations

Based on co-occurrence as described in the distributional hypothesis:

Computational Methods: Continuous Space Word Representations

Based on co-occurrence as described in the distributional hypothesis:

"*Linguistic items with similar distributions have similar meanings.*"

Computational Methods: Continuous Space Word Representations

Based on co-occurrence as described in the distributional hypothesis:

"*Linguistic items with similar distributions have similar meanings.*"

Continuous space word representations:

Computational Methods: Continuous Space Word Representations

Based on co-occurrence as described in the distributional hypothesis:

"*Linguistic items with similar distributions have similar meanings.*"

Continuous space word representations:

- Compact, dense, low dimensional embeddings

Computational Methods: Continuous Space Word Representations

Based on co-occurrence as described in the distributional hypothesis:

"*Linguistic items with similar distributions have similar meanings.*"

Continuous space word representations:

- Compact, dense, low dimensional embeddings
- Each dimension represents a distinctly informative property of the word

Computational Methods: Continuous Space Word Representations

Based on co-occurrence as described in the distributional hypothesis:

"*Linguistic items with similar distributions have similar meanings.*"

Continuous space word representations:

- Compact, dense, low dimensional embeddings
- Each dimension represents a distinctly informative property of the word
- Jointly represent the space in an efficient manner

Computational Methods: Continuous Space Word Representations

Based on co-occurrence as described in the distributional hypothesis:

"*Linguistic items with similar distributions have similar meanings.*"

Continuous space word representations:

- Compact, dense, low dimensional embeddings
- Each dimension represents a distinctly informative property of the word
- Jointly represent the space in an efficient manner

Advantage → Reduce sparsity of traditional vector-space representations

Computational Methods: Continuous Space Word Representations

Based on co-occurrence as described in the distributional hypothesis:

"*Linguistic items with similar distributions have similar meanings.*"

Continuous space word representations:

- Compact, dense, low dimensional embeddings
- Each dimension represents a distinctly informative property of the word
- Jointly represent the space in an efficient manner

Advantage → Reduce sparsity of traditional vector-space representations

Disadvantage → Computationally expensive, complex algorithms i.e. neural networks

Computational Methods: Multi-Linear Regression

Multi-Linear Regression (MLR) is used to forecast a response, investigate how several variables may be interrelated

Computational Methods: Multi-Linear Regression

Multi-Linear Regression (MLR) is used to forecast a response, investigate how several variables may be interrelated

MLR model was trained on participant scores (norms) to deduce statistical relationships between x, the food word in the form of a vector and y, a norm evaluated according to one of six categories

Computational Methods: Multi-Linear Regression

Multi-Linear Regression (MLR) is used to forecast a response, investigate how several variables may be interrelated

MLR model was trained on participant scores (norms) to deduce statistical relationships between x, the food word in the form of a vector and y, a norm evaluated according to one of six categories

$$\boxed{y = m + b_1 x_1 + \cdots + b_{100} x_{100}} \tag{1}$$

Computational Methods: Multi-Linear Regression

Multi-Linear Regression (MLR) is used to forecast a response, investigate how several variables may be interrelated

MLR model was trained on participant scores (norms) to deduce statistical relationships between x, the food word in the form of a vector and y, a norm evaluated according to one of six categories

$$\boxed{y = m + b_1 x_1 + \cdots + b_{100} x_{100}} \tag{1}$$

Goal: Infer what amount of variance in y (human score) can be explained by the dependence on x (vector)

Models

Model 1: Google Word2Vec

Google's Word2Vec model is a word-based, neural word embedding methodology for continuous space word representations

Model 1: Google Word2Vec

Google's Word2Vec model is a word-based, neural word embedding methodology for continuous space word representations

- Continuous Skip-gram model

Model 1: Google Word2Vec

Google's Word2Vec model is a word-based, neural word embedding methodology for continuous space word representations

- Continuous Skip-gram model
  - Instead of predicting the current word based on the context

Model 1: Google Word2Vec

Google's Word2Vec model is a word-based, neural word embedding methodology for continuous space word representations

- Continuous Skip-gram model
  - Instead of predicting the current word based on the context
  - Predicts words within a certain window before and after the current word

Model 1: Google Word2Vec

Google's Word2Vec model is a word-based, neural word embedding methodology for continuous space word representations

- Continuous Skip-gram model
  - Instead of predicting the current word based on the context
  - Predicts words within a certain window before and after the current word

Word2Vec was trained on all corpora with 100 dimensions, negative sampling of 5, sub-sampling of $1e-5$, a window size of 10

Model 2: Facebook fastText

Character-based neural network model from Facebook

Model 2: Facebook fastText

Character-based neural network model from Facebook

- Calculates word vectors based on the word's characters

Model 2: Facebook fastText

Character-based neural network model from Facebook

- Calculates word vectors based on the word's characters
  - Word vectors methodology enriched with subword information

Model 2: Facebook fastText

Character-based neural network model from Facebook

- Calculates word vectors based on the word's characters
  - Word vectors methodology enriched with subword information
  - Each character *n*-gram is assigned a vector and words are then represented as the sum of their character *n*-gram vectors

Model 2: Facebook fastText

Character-based neural network model from Facebook

- Calculates word vectors based on the word's characters
  - Word vectors methodology enriched with subword information
  - Each character *n*-gram is assigned a vector and words are then represented as the sum of their character *n*-gram vectors

fastText was trained on all corpora with 100 dimensions, subwords of 3 to 6, 5 epochs, learning rate of 0.05

Behavioral Analysis

Behavioral Analysis

What we did:

Behavioral Analysis

What we did:

- We investigate human cultural patterns of food items from two very different cultures with very different food staples, Britain and China

Behavioral Analysis

What we did:

- We investigate human cultural patterns of food items from two very different cultures with very different food staples, Britain and China

What we find:

Behavioral Analysis

What we did:

- We investigate human cultural patterns of food items from two very different cultures with very different food staples, Britain and China

What we find:

- Score distribution: British think more *happily*, *healthily*, *environmentally-friendly* about food items than Chinese

Behavioral Analysis

What we did:

- We investigate human cultural patterns of food items from two very different cultures with very different food staples, Britain and China

What we find:

- Score distribution: British think more *happily*, *healthily*, *environmentally-friendly* about food items than Chinese
- Categorical correlations: average correlation weaker for British participants than Chinese → higher value on *locality* for British than Chinese

Behavioral Analysis

What we did:

- We investigate human cultural patterns of food items from two very different cultures with very different food staples, Britain and China

What we find:

- Score distribution: British think more *happily*, *healthily*, *environmentally-friendly* about food items than Chinese
- Categorical correlations: average correlation weaker for British participants than Chinese $\rightarrow$ higher value on *locality* for British than Chinese
- Cultural food values: British participants love *potato*, hate *horsemeat*; Chinese love *artichoke*, hate *game meat*

Distribution of Participant Scores

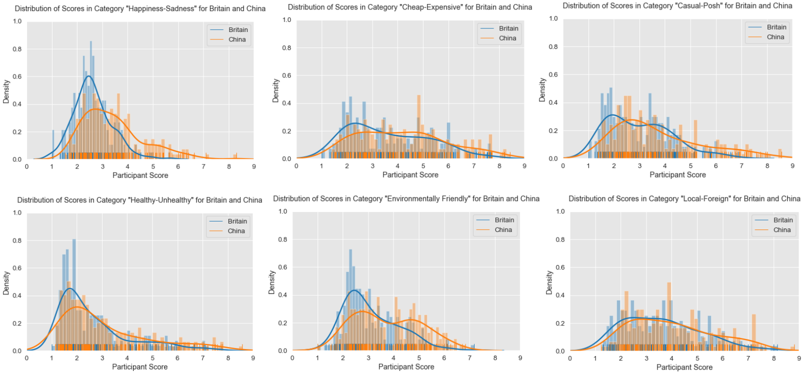


Figure: Distribution and density of British and Chinese participant scores for six evaluation categories (Example scale: 1 = happy and 9 = sad).

Correlation of Evaluation Categories

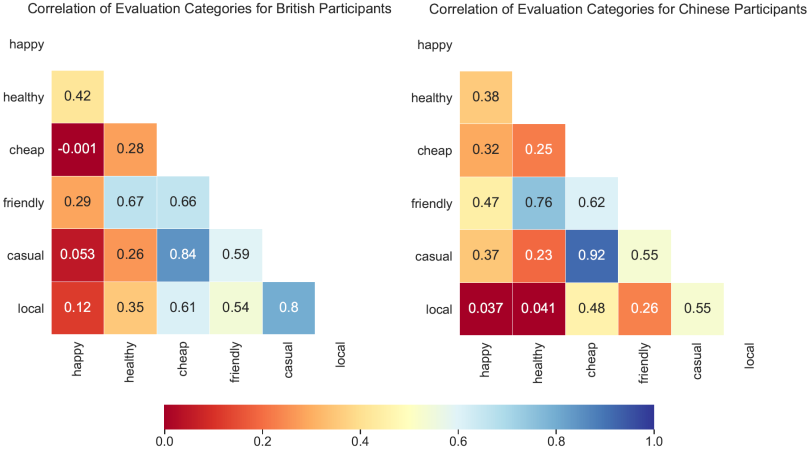


Figure: Spearman's correlation of aggregate evaluation scores for British and Chinese participants. Significant values < 0.5.

Food Scores by Country

		happy	*healthy*	*cheap*	*friendly*	*casual*	*local*
Britain	highest	strawberry	mandarin	**potato**	potherb	toast	**potato**
		1.52	1.15	1.38	1.00	1.38	1.24
	lowest	horsemeat	doughnut	caviar	veal	caviar	durian
		8.31	8.57	8.43	6.83	8.57	8.80
China	highest	quark cheese	**artichoke**	**artichoke**	**artichoke**	**artichoke**	puffed rice
		1.00	1.00	1.00	1.00	1.00	1.29
	lowest	game meat	doughnut	truffle	game meat	caviar	escargot
		5.67	7.75	7.8	7.17	6.63	8.17

Table: Highest and lowest scored food item per evaluation category for participants from Britain and China. (Ex. scale: 1 = happy, 9 = sad)

Computational Analysis

Computational Analysis

What we did:

Computational Analysis

What we did:

- We investigate the predictive power of linguistic information from word embeddings in relationship with cultural information in the form of food-norms

Computational Analysis

What we did:

- We investigate the predictive power of linguistic information from word embeddings in relationship with cultural information in the form of food-norms
- Train two state-of-the-art neural network word embedding models on English corpora from three different domains

Computational Analysis

What we did:

- We investigate the predictive power of linguistic information from word embeddings in relationship with cultural information in the form of food-norms
- Train two state-of-the-art neural network word embedding models on English corpora from three different domains
- Resulting vectors are used as independent variable in MLR, food-norms (scores) as dependent variable

Computational Analysis

What we did:

- We investigate the predictive power of linguistic information from word embeddings in relationship with cultural information in the form of food-norms
- Train two state-of-the-art neural network word embedding models on English corpora from three different domains
- Resulting vectors are used as independent variable in MLR, food-norms (scores) as dependent variable

What we find:

Computational Analysis

What we did:

- We investigate the predictive power of linguistic information from word embeddings in relationship with cultural information in the form of food-norms
- Train two state-of-the-art neural network word embedding models on English corpora from three different domains
- Resulting vectors are used as independent variable in MLR, food-norms (scores) as dependent variable

What we find:

- Multi-linear regression model trained on food-item vectors can predict cultural norms in the form of a participant score at $r = 0.65$ and $R^2 = 0.73$ on Word2Vec vectors trained on web domain corpus

Embedding Visualization

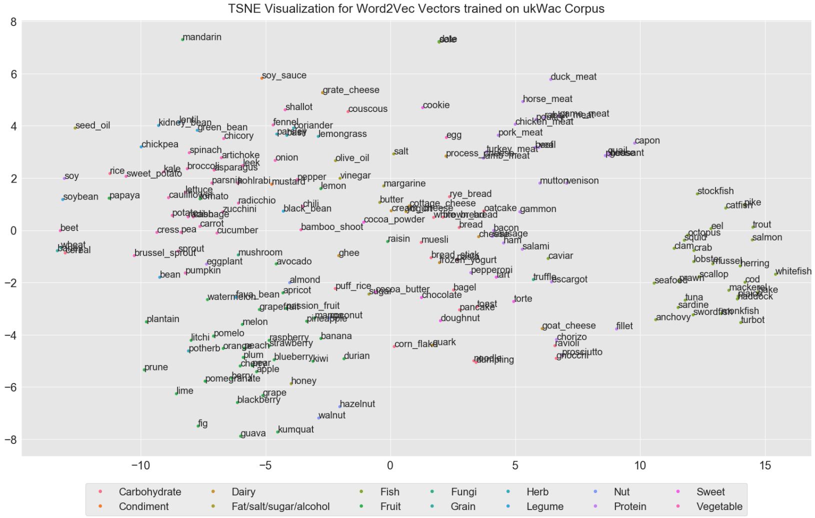


Figure: Word2Vec food norm vectors trained on UkWac web corpus.

Computational Results: Word2Vec

Word2Vec

		happy	*healthy*	*cheap*	*friendly*	*casual*	*local*	Corp. Avg.
WaCky	r	0.39	0.56	0.46	0.61†	0.45	0.55	0.50
	R^2	0.48	0.57	0.50	0.64†	0.56	0.64	0.57
Giga	r	0.36	0.58	0.56	0.63†	0.51	0.48	0.52
	R^2	0.53	0.60	0.52	0.63†	0.59	0.54	0.57
ukWaC	r	0.49	0.57	0.63	0.63	0.65†	0.60	0.59
	R^2	0.56	0.64	0.65	0.54	0.72	0.73†	0.66
Cat. Avg.	r	0.41	0.57	0.51	**0.63**	0.53	0.54	**0.53**
Cat. Avg.	R^2	0.52	0.60	0.56	**0.64**	0.62	0.64	**0.60**

Table: Multi-linear regression results for Word2Vec over six evaluation categories per corpus. Bold values indicate strongest average correlation and highest variance overall with significance above a threshold of .5. Daggers indicate strongest r or highest R^2 of individual category or domain.

Computational Results: fastText

fastText

		happy	*healthy*	*cheap*	*friendly*	*casual*	*local*	Corp. Avg.
WaCky	r	0.32	0.49	0.42	0.56†	0.37	0.51	0.44
	R^2	0.40	0.54	0.51	0.60†	0.49	0.60	0.52
Giga	r	0.44	0.65†	0.56	0.61	0.50	0.40	0.53
	R^2	0.61	0.64	0.57	0.65†	0.60	0.50	0.60
ukWaC	r	0.33	0.50	0.62†	0.61	0.53	0.53	0.52
	R^2	0.41	0.54	0.70	0.66	0.70†	0.60	0.60
Cat. Avg.	r	0.48	0.59	0.51	**0.60**	0.46	0.43	**0.49**
Cat. Avg.	R^2	0.54	0.61	0.55	**0.63**	0.56	0.54	**0.57**

Table: Multi-linear regression results for fastText over six evaluation categories per corpus. Bold values indicate strongest average correlation and highest variance overall with significance above a threshold of .5. Daggers indicate strongest r or highest R^2 of individual category or domain.

Predicted Norms: Word2Vec

Word2Vec

	happy	*healthy*	*cheap*	*friendly*	*casual*	*local*
WaCky	kidney_bean	olive_oil	kohlrabi	green_bean	wheat	bamboo_shoot
	-0.11	-4.53	-0.26	-0.03	-0.99	0.24
	catfish	brown_bread	muesli	bagel	artichoke	strawberry
	7.70	7.82	9.53	6.78	8.42	9.23
Giga	walnut	**cress**	barley	cress	spinach	cress
	0.85	-1.64	-1.99	-0.27	-2.32	-0.38
	horse_meat	cocoa_butter	gammon	turbot	**caviar**	guava
	7.65	8.57	11.35	7.30	10.10	10.55
ukWaC	orange	**pomegranate**	soy_sauce	prune	soy_sauce	oatcake
	0.50	-2.51	-0.45	-0.25	-0.66	-0.61
	quark	**cookie**	scallop	cookie	quark	**durian**
	7.17	10.37	10.36	10.63	10.92	8.76

Table: Model prediction of highest and lowest norms for Word2Vec models per corpus per category. Food items in bold indicate a correct prediction: food item is among the five highest and lowest food items evaluated by British participants (Ex. scale: 1 = happy, 9 = sad).

Predicted Results: fastText

fastText

	happy	*healthy*	*cheap*	*friendly*	*casual*	*local*
WaCky	**chocolate**	lettuce	ghee	ghee	ghee	cod
	0.15	-1.67	-1.35	0.33	-0.87	-0.75
	trout	gnocchi	fillet	**fillet**	fillet	pomelo
	6.74	7.58	11.43	8.21	10.06	9.96
Giga	apple	**apple**	cereal	**barley**	white_bread	pepper
	-1.04	-1.77	-0.45	0.48	-0.93	0.14
	goose_meat	cocoa_butter	fillet	**game_meat**	**game_meat**	eel
	7.00	8.76	9.80	7.92	9.28	8.89
ukWaC	hazelnut	kiwi	soy_sauce	mushroom	soy_sauce	mustard
	-0.79	-3.24	-0.89	1.42	-1.52	-0.08
	stockfish	**cookie**	mussel	duck_meat	mussel	stockfish
	8.13	9.02	9.91	6.44	9.46	9.54

Table: Model prediction of highest and lowest norms for fastText models per corpus per category. Food items in bold indicate a correct prediction: food item is among the five highest and lowest food items evaluated by British participants (Ex. scale: 1 = happy, 9 = sad).

General Discussion

Research Answers

Answers to main research questions:

Research Answers

Answers to main research questions:

1. Can we predict cultural information in the form of norms from linguistic information in the form of word embeddings?

Research Answers

Answers to main research questions:

1. Can we predict cultural information in the form of norms from linguistic information in the form of word embeddings?
    - A multi-linear regression model trained on word representations can predict norms in the form of aggregate scores for food items at $r = 0.65$ and $R^2 = 0.73$

Research Answers

Answers to main research questions:

1. Can we predict cultural information in the form of norms from linguistic information in the form of word embeddings?
   - A multi-linear regression model trained on word representations can predict norms in the form of aggregate scores for food items at $r = 0.65$ and $R^2 = 0.73$
   - Best predicted category *Environmentally-Friendly*

Research Answers

Answers to main research questions:

1. Can we predict cultural information in the form of norms from linguistic information in the form of word embeddings?
   - A multi-linear regression model trained on word representations can predict norms in the form of aggregate scores for food items at $r = 0.65$ and $R^2 = 0.73$
   - Best predicted category *Environmentally-Friendly* $\rightarrow$ objective general knowledge

Research Answers

Answers to main research questions:

1. Can we predict cultural information in the form of norms from linguistic information in the form of word embeddings?
   - A multi-linear regression model trained on word representations can predict norms in the form of aggregate scores for food items at $r = 0.65$ and $R^2 = 0.73$
   - Best predicted category *Environmentally-Friendly* $\rightarrow$ objective general knowledge
   - Worst predicted category *Happiness-Sadness*

Research Answers

Answers to main research questions:

1. Can we predict cultural information in the form of norms from linguistic information in the form of word embeddings?
   - A multi-linear regression model trained on word representations can predict norms in the form of aggregate scores for food items at $r = 0.65$ and $R^2 = 0.73$
   - Best predicted category *Environmentally-Friendly* $\rightarrow$ objective general knowledge
   - Worst predicted category *Happiness-Sadness* $\rightarrow$ subjective, personal

Research Answers

Answers to main research questions:

1. Can we predict cultural information in the form of norms from linguistic information in the form of word embeddings?
   - A multi-linear regression model trained on word representations can predict norms in the form of aggregate scores for food items at $r = 0.65$ and $R^2 = 0.73$
   - Best predicted category *Environmentally-Friendly* $\rightarrow$ objective general knowledge
   - Worst predicted category *Happiness-Sadness* $\rightarrow$ subjective, personal
2. How do corpora from different domains effect model performance?

Research Answers

Answers to main research questions:

1. Can we predict cultural information in the form of norms from linguistic information in the form of word embeddings?
    - A multi-linear regression model trained on word representations can predict norms in the form of aggregate scores for food items at $r = 0.65$ and $R^2 = 0.73$
    - Best predicted category *Environmentally-Friendly* $\rightarrow$ objective general knowledge
    - Worst predicted category *Happiness-Sadness* $\rightarrow$ subjective, personal
2. How do corpora from different domains effect model performance?
    - Word2Vec vectors trained on web-based domain (ukWaC) elicit higher correlations than either news- or Wikipedia-based domains

Research Answers

Answers to main research questions:

1. Can we predict cultural information in the form of norms from linguistic information in the form of word embeddings?
    - A multi-linear regression model trained on word representations can predict norms in the form of aggregate scores for food items at $r = 0.65$ and $R^2 = 0.73$
    - Best predicted category *Environmentally-Friendly* $\rightarrow$ objective general knowledge
    - Worst predicted category *Happiness-Sadness* $\rightarrow$ subjective, personal
2. How do corpora from different domains effect model performance?
    - Word2Vec vectors trained on web-based domain (ukWaC) elicit higher correlations than either news- or Wikipedia-based domains
    - Largest corpus, most general domain, most food-related information

Potential Drawbacks: Sense Ambiguity

A limitation of word vectors: conflate all senses of a word into a single vector

Potential Drawbacks: Sense Ambiguity

A limitation of word vectors: conflate all senses of a word into a single vector

We did not consider sense variability in the scope of this thesis

Potential Drawbacks: Sense Ambiguity

A limitation of word vectors: conflate all senses of a word into a single vector
We did not consider sense variability in the scope of this thesis

- *date*:

Potential Drawbacks: Sense Ambiguity

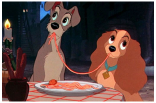
A limitation of word vectors: conflate all senses of a word into a single vector
We did not consider sense variability in the scope of this thesis

- *date*:

Potential Drawbacks: Sense Ambiguity

A limitation of word vectors: conflate all senses of a word into a single vector
We did not consider sense variability in the scope of this thesis

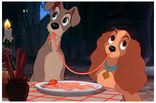
- *date*:



Potential Drawbacks: Sense Ambiguity

A limitation of word vectors: conflate all senses of a word into a single vector
We did not consider sense variability in the scope of this thesis

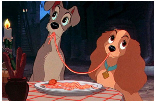
- *date*:



Potential Drawbacks: Sense Ambiguity

A limitation of word vectors: conflate all senses of a word into a single vector
We did not consider sense variability in the scope of this thesis

- *date*:

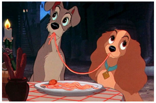


- *sole*:

Potential Drawbacks: Sense Ambiguity

A limitation of word vectors: conflate all senses of a word into a single vector
We did not consider sense variability in the scope of this thesis

- *date*:

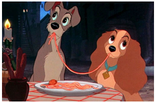

- *sole*:

Potential Drawbacks: Sense Ambiguity

A limitation of word vectors: conflate all senses of a word into a single vector

We did not consider sense variability in the scope of this thesis

- *date*:

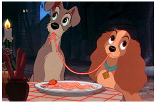


- *sole*:

Potential Drawbacks: Sense Ambiguity

A limitation of word vectors: conflate all senses of a word into a single vector
We did not consider sense variability in the scope of this thesis

- *date*:

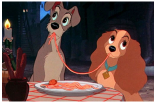


- *sole*:

Potential Drawbacks: Sense Ambiguity

A limitation of word vectors: conflate all senses of a word into a single vector
We did not consider sense variability in the scope of this thesis

- *date*:



- *sole*:

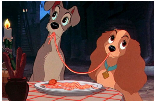
- *kiwi*:

Potential Drawbacks: Sense Ambiguity

A limitation of word vectors: conflate all senses of a word into a single vector

We did not consider sense variability in the scope of this thesis

- *date*:


- *sole*:



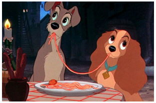
- *kiwi*:

Potential Drawbacks: Sense Ambiguity

A limitation of word vectors: conflate all senses of a word into a single vector

We did not consider sense variability in the scope of this thesis

- *date*:


- *sole*:

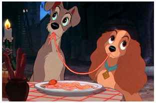


- *kiwi*:


Potential Drawbacks: Sense Ambiguity

A limitation of word vectors: conflate all senses of a word into a single vector
We did not consider sense variability in the scope of this thesis

- *date*:


- *sole*:



- *kiwi*:



Take-Home

Take-Home

In this thesis, we presented a ...

Take-Home

In this thesis, we presented a ...

Behavioral analysis of human food culture based on British, Chinese norms

Take-Home

In this thesis, we presented a ...

Behavioral analysis of human food culture based on British, Chinese norms

- British participants think food items to be *happier*, *healthier*, and more *environmentally friendly* than Chinese participants

Take-Home

In this thesis, we presented a ...

Behavioral analysis of human food culture based on British, Chinese norms

- British participants think food items to be *happier*, *healthier*, and more *environmentally friendly* than Chinese participants
- British participants place more value on food item's *locality*, Chinese on food item's *health* properties

Take-Home

In this thesis, we presented a ...

Behavioral analysis of human food culture based on British, Chinese norms

- British participants think food items to be *happier*, *healthier*, and more *environmentally friendly* than Chinese participants
- British participants place more value on food item's *locality*, Chinese on food item's *health* properties

Computational analysis of linguistic features encoded in food-word representations

Take-Home

In this thesis, we presented a ...

Behavioral analysis of human food culture based on British, Chinese norms

- British participants think food items to be *happier*, *healthier*, and more *environmentally friendly* than Chinese participants
- British participants place more value on food item's *locality*, Chinese on food item's *health* properties

Computational analysis of linguistic features encoded in food-word representations

- British participant norms in the form of aggregate scores for food items can be predicted from food item word representations using Multi Linear Regression

Take-Home

In this thesis, we presented a ...

Behavioral analysis of human food culture based on British, Chinese norms

- British participants think food items to be *happier*, *healthier*, and more *environmentally friendly* than Chinese participants
- British participants place more value on food item's *locality*, Chinese on food item's *health* properties

Computational analysis of linguistic features encoded in food-word representations

- British participant norms in the form of aggregate scores for food items can be predicted from food item word representations using Multi Linear Regression
- Norm easiest to predict for participant-objective category *Environmentally-Friendly* and least predictable for participant-subjective category *Happiness-Sadness*