Formal Approaches to Social Meaning, Variation and Identity Construction

Heather Burnett

Laboratoire de Linguistique Formelle
CNRS & Université Paris-Diderot

Limits of Variability in Language Summer School (Day 2)

Introduction

Topic of Class 2 and maybe Class 3

A ‘formal semantics/pragmatics’ for sociophonetic variation.

(1) (ING)

a. I’m working on my paper. [iŋ]
b. I’m workin’ on my paper. [in]

(2) /t/ release

a. We should mee[t^h]. released ‘t’
b. We should mee[t]. unreleased ‘t’

Game theory and sociolinguistics

Social Meaning Games (SMGs)

A new framework for the analysis of the social aspect of sociolinguistic variation.

- A formalization of the Third Wave approach to the meaning of variation (see Eckert, 2012) using Bayesian signalling games.

- An increasingly popular framework for analyzing pragmatic phenomena (Franke, 2009; Frank and Goodman, 2012; Lassiter and Goodman, 2015; Degen and Tanenhaus, 2015a; Franke and Jäger, 2016, among many many others).
- Such models have the potential for yielding a framework for unifying social meaning and with other kinds of linguistic meaning in context.

Plan

Bayesian game-theoretic pragmatics

Quantity implicatures

Social Meaning and sociolinguistic variation

'Third Wave' approach to variation

Social meaning games

Obama style shifting across 3 contexts (Labov, 2012)

Conclusion

Speaker agency and the social construction of identity

Game theory: A formalism for strategic interaction

The architecture

- There are (at least) two players.
- The players interact and the interaction results in a particular outcome.
- The outcome of the interaction depends on the choice of strategy of each player.
- Each player has a preference ordering over outcomes.

The solution concept

A rule/algorithm that determines how the game is played.

Signalling Games (Lewis, 1969)

1. There are two players: the speaker (S) and the listener (L).
2. S knows a piece of information that they want to communicate to L (their type).
3. L wants to learn the information that S is trying to communicate to them.
4. S has a set of messages (linguistic forms paired with semantic meanings) that they can choose to send to L in order to try to transmit their information.
    - S picks a message to send to L (i.e. says something), and L assigns an interpretation to the message (i.e. understands it in some way).

A Game of Cooperation/Coordination

Two outcomes:

1. L interprets the message in the way that S intended (so learns the information that S wants to tell them).
2. L doesn't interpret the message the way that S wanted (so doesn't learn the info).

Cooperation

- Outcome 1. is good for S and good for L.
- Outcome 2. is bad for S and bad for L.

Driving Example

(3) Question: Heather, do you drive?

- My type: '*I would prefer not to (and you would probably prefer me not to), but if you need me to move your car, or if it's an emergency, fine*.'
- I need to choose a message to communicate this.

(4) Messages

a. I would prefer not to (and you would probably prefer me not to), but if you need me to move your car, or if it's an emergency, or if it's just around the corner... Costly

b. I'm licensed by the province of Ontario. Less costly

Coordination $\Rightarrow$ Scalar Implicature

(5) I'm licensed by the province of Ontario.

a. Outcome 1: Heather knows how to drive but doesn't want to.

b. Outcome 2: Heather has a driver's license. She can be the driver on our road trip into the slippery foggy mountains!

- Outcome 1 is good for both of us.
- Outcome 2 is bad for both of us.

Iterated Best Response Models

A family of similar approaches which formalize Gricean reasoning (Grice, 1975) (particularly quantity and quality) using signalling games and a solution concept based on iterated reasoning (Franke, 2009; Frank and Goodman, 2012; Goodman and Stuhlmüller, 2013; Lassiter and Goodman, 2015; Degen et al., 2015; Bergen et al., 2016; Franke and Jäger, 2016, among many others).

Today's presentation

Following the Rational Speech Act model (Frank and Goodman, 2012).

Iterated best response signaling games

- When modelling communication, the solution concept that we use should make reference to reasoning process of the agents involved (Franke, 2009).

Hypothesis

Agents' reasoning is Bayesian (see Tenenbaum et al. (2011) for an overview).

- Wide applications across cognitive science: perception (Yuille and Kersten, 2006), memory (Shiffrin and Steyvers, 1997), sensorimotor systems (Steyvers et al., 2006), and language (Chater and Manning, 2006).

Bayesian approach to cognitive science

An approach to answering questions concerning the nature of knowledge and cognition:

1. How does abstract knowledge guide learning and inference from sparse data?
2. What forms does abstract knowledge take, across different domains and tasks?
3. How is abstract knowledge itself acquired?

Structure + Statistics

1. How does abstract knowledge guide learning and inference from sparse data?
   - Powerful statistical inference engines.
2. What forms does abstract knowledge take, across different domains and tasks?
   - Richly structured, expressive knowledge representations.
3. How is abstract knowledge itself acquired?
   - Powerful statistical inference engines.

Bayesian Revolution in Cognitive Science

- *Until recently, cognitive modelers were forced to choose between two alternatives (Pinker, 1999): powerful statistical learning operating over the simplest, unstructured forms of knowledge [. . .] or richly structured symbolic knowledge equipped with only the simplest, non-statistical forms of learning [. . .].*
- *It appeared necessary to accept either that people's abstract knowledge is not learned or induced in a nontrivial sense from experience (hence essentially innate) or that human knowledge is not nearly as abstract or structured (as "knowledge-like") as it seems (hence simply associations).*

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science*, 331(6022), 1279-1285.

Fundamental Interpretation Rule: Bayesian inference

Humans draw a conclusion B after having observed event A ($P(B|A)$) through combining:

1. How likely they think A is to indicate B (P(A|B)).
2. How likely they thought B was to begin with (Pr(B)).

Bayes rule

$$P(B_i|A) = \frac{Pr(B_i) \times P(A|B_i)}{\sum_{j=1}^{|B|} Pr(B_j) \times P(A|B_j)} \tag{6}$$

$$P(B|A) \propto Pr(B) \times P(A|B) \tag{7}$$

Bayesian Game-Theoretic Pragmatics

(Some of the) Active Research Groups:

1. Language and Cognition, Computation and Cognition, Psychosemantics, and Interactive Language Processing Labs, Stanford University (Frank, Goodman, Lassiter, Degen)
2. Institute for Cognitive Science, Osnabrück (Franke)
   `http://www.home.uni-osnabrueck.de/michfranke/index.html`
3. Institute of Linguistics, Universität Tübingen (Jäger)
   `http://www.sfs.uni-tuebingen.de/~gjaeger/`
4. Institute for Logic Language and Computation, Amsterdam.
   `https://www.illc.uva.nl/`

Quantity implicatures

(8) a. Mary ate some of the cookies.
b. $\rightsquigarrow$ Mary did not eat all of the cookies.

Reasons to think that (8-b) is not encoded into the meaning of *some*

(9) a. If you eat some of the cookies, I'll be angry.
$\not\rightsquigarrow$ If you eat some but not all of the cookies, I'll be angry.
b. Did you eat some of the cookies?
$\not\rightsquigarrow$ Did you eat some but not all of the cookies?

Variable interpretation

Scalar enrichment is variable (Sperber and Wilson, 1986; Levinson, 2000; Degen, 2015; Degen and Tanenhaus, 2015b).

- Determiner strength, partitivity and contextual aspects determine participant judgements of *some* in corpus examples.

(10) (Degen, 2015, 17)
I wish my mother had had some of those opportunities, because I think she would have really, she rea-, would have succeeded in a lot of ways, that men, that women were not able to succeed in her generation.

Signalling game (RSA-style)

An RSA-style signalling game is a tuple $\langle \{S, L\}, W, M, [\![\cdot]\!], C, Pr \rangle$:

1. S, L are the players.
2. W is a set of possible worlds.
3. M is the set of messages.
4. $[\![\cdot]\!]$ is an interpretation function assigning a set of possible worlds to each message.
5. C is the set of message costs.
6. Pr is a probability distribution over worlds representing the listener's prior beliefs before hearing a message.

Models of what?

If we are modelling listener behaviour (Interpretation):

- *Pr* represents L's beliefs.

If we are modelling speaker behaviour (Production):

- *Pr* represents S's hypothesis concerning L's prior beliefs.

If we are modelling interaction (Evolution):

- *Pr* represents L's prior beliefs and is usually common knowledge.

The scenario

S and L baked three cookies, and then, while L was out, Mary stopped by and possibly ate some of them. Suppose that L calls the house and wants to know how many of the cookies Mary ate. What should S say and how should L understand what S says to them?

Possible World	Description
w_0	Mary ate 0 cookies
w_1	Mary ate 1 cookie
w_2	Mary ate 2 cookies
w_3	Mary ate 3 cookies

Table: Universe (W) in cookie example

Messages

Short name	message	$[\![\text{message}]\!]$
NONE	Mary ate none of the cookies	$\{w_0\}$
SOME	Mary ate some of the cookies	$\{w_1, w_2, w_3\}$
ALL	Mary ate all of the cookies	$\{w_3\}$

Table: Messages in cookie example

Prior beliefs

Suppose L has no prior expectations about how many cookies Mary ate.

- *Pr* is uniform over the set of possible worlds.

w_0	w_1	w_2	w_3
0.25	0.25	0.25	0.25

Table: L has uniform prior beliefs (Pr(w)).

Formalization of Quality Maxim

When they hears a message m, L restricts their attention to the worlds in which m is true.

- L conditions on $[\![m]\!]$: intersection followed by renormalization of the measure.

Message	w_0	w_1	w_2	w_3
NONE	1	0	0	0
ALL	0	0	0	1
SOME	0	0.3	0.3	0.3

Table: L's beliefs immediately after hearing m (Pr(w|m)).

Formalization of Quantity Maxim

Coordination (i.e. communication) occurs because speakers try to say the most informative statement possible. And listeners know this.

- Informativity is encoded as part of speaker's utility function (u_S).
- (Frank and Goodman, 2012, et seq.): The informativity of m is its negative surprisal (positive natural log probability (Shannon, 1948)) of the prior conditioned on the truth of the message.
- Costs can encode grammatical/psychological constraints on utterances (length, markedness etc.).

(11) $$u_S(m, w) = ln(Pr(w|m)) - c(m)$$

Speaker Utility

Message	w_0	w_1	w_2	w_3
NONE	0	$-\infty$	$-\infty$	$-\infty$
ALL	$-\infty$	$-\infty$	$-\infty$	0
SOME	$-\infty$	-0.11	-0.11	-0.11

Table: S's utility for *m* for communicating *w* (u_S(w, m)).

If we are in w_2 (two of three cookies eaten):

(12) a. $u_S(w_2, \text{SOME}) = ln(0.\dot{3}) = -0.11$
b. $u_S(w_2, \text{NONE}) = ln(0) = -\infty$
c. $u_S(w_2, \text{ALL}) = ln(0) = -\infty$

Predicting linguistic production

To account for variability in action selection:

Soft-Max Choice (Luce, 1959; Sutton and Barto, 1998)

For a world w, a message m and a value λ (the temperature).

$$P_S(m|w) = \frac{exp(\lambda * u_S(w,m))}{\sum_{m' \in M} exp(\lambda * u_S(w,m'))}$$

Quantitative predictions for language use

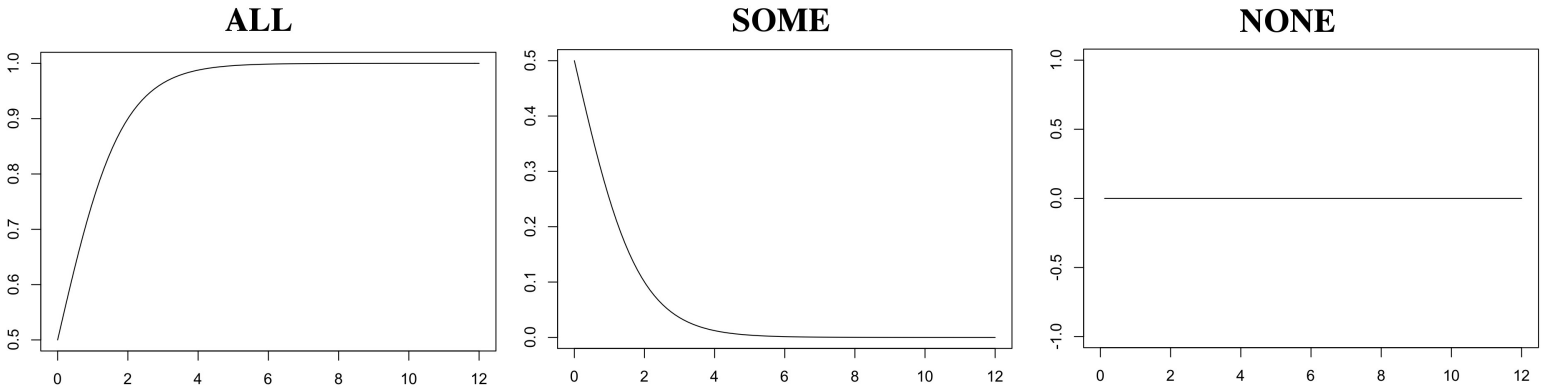

Figure: Predictions for ALL, SOME, NONE communicating w_3, by λ

Quantitative Predictions for Language Use

Message	w_0	w_1	w_2	w_3
NONE	1	0	0	0
ALL	0	0	0	0.99
SOME	0	1	1	0.01
Prediction	Cat. NONE	Cat. SOME	Cat. SOME	Favored ALL

Table: S's predicted use of m, given w with $\lambda = 10$ ($P_S(m|w)$).

Quantitative Predictions for Language Interpretation

Interpretation as Bayesian Inference

$$P_L(w|m) = \frac{Pr(w) \times P_S(m|w)}{\sum_{w'} Pr(m|w') \times P_S(m|w')}$$

Message	w_0	w_1	w_2	w_3	PREDICTION
NONE	1	0	0	0	Categorical w_0
ALL	0	0	0	1	Categorical w_3
SOME	0	0.498	0.498	0.005	Favoured w_1, w_2

Table: L's predicted interpretation of w, given m ($P_L(w|m)$).

Heavily Weighted Priors

Suppose that L knows that Mary usually likes to have two cookies for her dessert.

w_0	w_1	w_2	w_3
0.1	0.1	0.7	0.1

Table: L's priors heavily weighted on w_2.

Prior beliefs influence interpretation

L's interpretation probabilities change.

- L's probability of interpreting w_2 after SOME is now 0.87 from 0.498.

Computational resources for Bayesian pragmatics

To facilitate calculations and prediction testing, a number of computational implementations have been developed:

1. Chris Potts' implementation in python:
   `https://github.com/cgpotts/pypragmods`
2. Goodman and Tenenbaum's implementation in Church:
   `https://probmods.org/`.
   - Also comes with a textbook.
3. Goodman and Stuhlmüller's implementation in WebPPL:
   `http://dippl.org/examples/pragmatics.html`
   - Also comes with a textbook for Scontras & Tessler's 2016 ESSLLI course:
     `http://gscontras.github.io/ESSLLI-2016/`

Summary

Bayesian game-theoretic models provide a framework for:

1. Formalizing pragmatic theories (in this case Gricean pragmatics).
2. Making both qualitative and quantitative predictions about (possibly variable) language use and interpretation.
3. Capturing interactive co-construction of meaning (in this case truth-conditional).
   - The inference (13-b) arises as a product of coordination between the speaker and listener.

   (13) a. Mary ate some of the cookies.
   b. $\rightsquigarrow$ Mary didn't eat all of the cookies.

4. Capturing the contribution that speaker/listener prior beliefs make to pragmatic interpretation.

Recall from Class 1

Generalization from perception studies

Hearers make judgments about the properties that characterize speakers based on the linguistic forms that they use.

In MGT studies,

- -ing was associated with properties like *competence* and *articulateness*.
- -in' was associated with properties like *sincerity* and *friendliness*.

Generalization from production studies

Speakers strategically exploit hearer's interpretation process to communicate properties about themselves to their interlocutors.

Figure 3. President Obama's use of (ING) in three contextual styles.

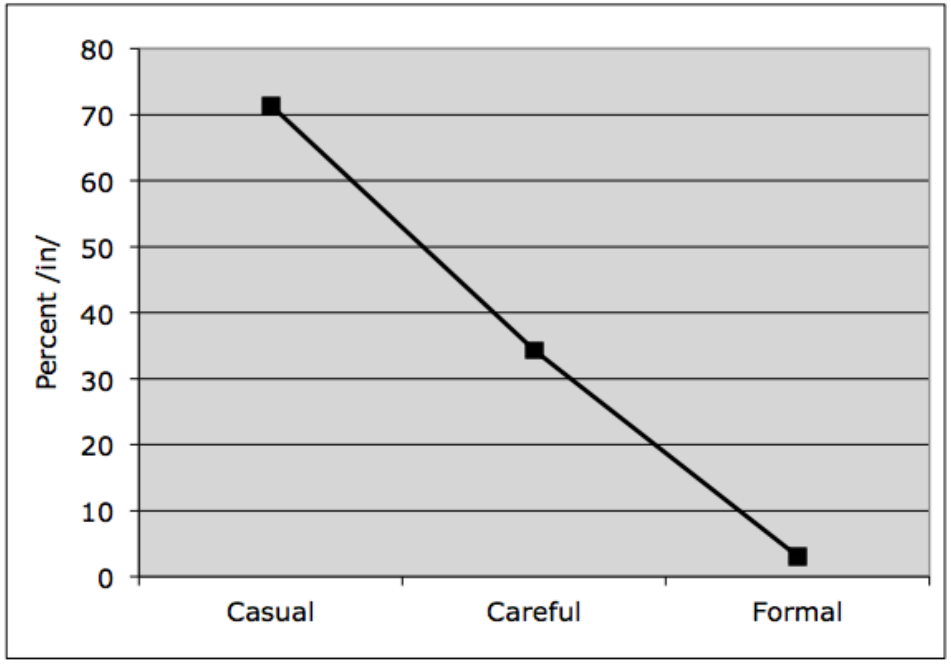

Figure: Obama's use of (ING) across contexts

What do we want in a formal model?

1. A framework that can capture the context dependent interplay between conversational participants (i.e. both the speaker and the listener). (Interactivity)
2. A model in which the speaker (tries to) choose the variant that has the best chance to construct their desired persona. (Approximate rationality)
3. A model that predicts quantitative patterns of variation/interpretation. (Variability)

Proposal

Game theoretic models have these properties.

Third Wave approach to variation

Empirical discovery from sociolinguistics

The linguistic resources speakers vary across situations **coincide** with those used to distinguish social groups (Labov, 1966, 1972; Bell, 1984; Eckert, 1989, among many others).

Variation as social practice (Eckert, 2000, 2008, 2012)

TW pursues unified treatment of social stratification and style shifting as interactive rational language use.

Two empirical phenomena or one?

Labov, W. (1966). *The Social Stratification of English in New York City*. CFAL.

TW in a nutshell. . .

- Variants are related to abstract mental representations (meanings) which mediate the relationship between language and personae (identity/social types) (Ochs, 1992, 1993; Silverstein, 1979, 2003; Eckert, 2008, among many others).
- These abstract mental representations are made up of sets of properties, stances or other concepts/ideas that are to be attributed to the speaker (indexical fields (Eckert, 2008)).
- Speakers use these linguistic resources to (attempt to) construct the persona that will be the most useful to them in their context-specific goals. interactivity/rationality
  - The properties indexed by -ing are more useful to Obama in a formal setting than in an informal setting.

Stratification in TW (Eckert, 2000, 2008, 2012)

Speakers of different social groups (gender, class, age etc.) have very different experiences and live very different lives.

- As such, speakers value different properties in objects and people (see also Bourdieu and Passeron, 1970; Bourdieu, 1979; Lamont, 1992, 2009, among many others).
- Personae/identities that may be useful/desirable to individuals of certain social groups may be less so to individuals of other social groups.
- Linguistic expressions with social meanings that can be used to construct these personae are predicted be more useful to individuals of certain social groups than to others.
    - The properties indexed by -ing are more useful to upper middle class speakers (being interviewed by a researcher) than to working class speakers.

The linguistic variable?

Sali's class: A linguistic variable is *two ways of saying the same thing*.

How does this relate to Third Wave's meaning-based approach?

- For sociophonetic variation: limit variables to truth-conditional equivalence.
- For morpho-syntactic variation: we may relax even this constraint.
- Assimilates the study of the linguistic variable to the study of alternatives in formal pragmatics (see Fox and Katzir, 2011, a.o.).

Social meaning games (Burnett, 2017)

1. The speaker (S) has a persona (an identity/social type) that they wish to communicate to the listener (L).
   (Where does the persona come from?)
2. S chooses a variant with which to signal their persona to L.
3. Variants are related to indexical fields (sets of properties).
4. L chooses a persona to attribute to S based on their prior beliefs about S and the variants' indexical field.

Social Meaning Game

- S and L are the players.
- $\mathbb{P} = \{p_1, \ldots, p_n\}$ is a finite set of properties.
- $>$ is a relation on $\mathbb{P}$ that encodes antonymy.

Example: Obama across 3 contexts

$\mathbb{P}$ = {competent, incompetent, friendly, aloof}

(14) a. competent $>$ incompetent
b. friendly $>$ aloof

- The universe could be enriched with additional ideological structure...

Personae

Third Wave Variation Theory focuses on how variants combine together (styles), which construct particular social types (personae) (see Podesva, 2004; Eckert, 2008; Zhang, 2008, among many others).

- Possible personae are collections of properties that *go together*.

The personae are the set of largest consistent sets of properties.

Persona	Nickname
{competent, friendly}	'cool guy/gal'
{competent, aloof}	'stern leader'
{incompetent, friendly}	'doofus'
{incomptent, aloof}	'arrogant asshole'

Table: Universe in Obama example

Messages

- $M = \{m_1, \ldots, m_n\}$ is the set of messages (i.e. variants) that S can pick from.

Today's Example

M = {-ing, -in'}

Indexation and Indexical Fields

In Third Wave variation theory, individual variants have meaning that goes beyond their truth conditional meaning.

- Variants index sets of properties, called their indexical field (Eckert, 2008).

Variant	Eckert field
-ing	{competent, aloof}
-in'	{incompetent, friendly}

Eckert-Montague Fields

- In the spirit of Montague (1973), we can also look at indexical fields through the personae that they have the potential to construct.

Variant	Eckert field	Eckert-Montague field
-ing	{competent, aloof}	{comp., aloof}, {comp., friend.}, {incomp., aloof}
-in'	{incompetent, friendly}	{incomp., friend}, {comp., friend}, {incomp., aloof}

Table: Messages in Obama example

Obama at the BBQ

Suppose Obama wants to be perceived as the cool guy at the barbecue.

- He wants to construct the {competent, friendly} persona.

Listener prior beliefs

Obama at the BBQ

Obama is worried about coming off as too aloof (since he is the president).

stern leader	cool guy	asshole	doofus
{comp, aloof}	{comp, friend}	{incomp, aloof}	{incomp, friend}
0.30	0.20	0.30	0.20

Table: Obama worries about seeming aloof.

Contribution of Indexical Fields

When they hear a variant, L focuses their attention to the personae in the (Eckert-Montague) fields.

- L conditions on $[\![m]\!]$: intersection followed by renormalization of the measure.

	stern leader	cool guy	asshole	doofus
m	{comp, aloof}	{comp, frien}	{incomp, aloof}	{incomp, frien}
-ing	0.375	0.25	0.375	0
-in'	0	0.286	0.428	0.286

Table: L's beliefs immediately after hearing m (Pr(P|m)).

Speaker Utility as Informativity - Costs

(15) $u_S(m, P) = ln(Pr(P|m)) - c(m)$ RSA utility function

In a nutshell

1. The speaker tries to give the listener the most information possible about their persona.
2. The listener assumes that the speaker is (un)intentionally giving them the most information possible about S's persona.

Costs as linguistic conditioning

Message Costs

Costs can encode grammatical/psychological constraints on utterances.

- (ING) is conditioned by grammatical category and other abstract properties of morphological structure (Labov, 1966; Houston, 1985; Tamminga, 2014).
- Mathematical connections between game-theoretic syntax/semantics and OT syntax/semantics (also Linear OT, Harmonic Grammar).
- Since this requires more complicated message representations, we ignore costs here.

Obama at the BBQ (predictions)

- We obtain the speaker's probability distribution over variants through the Soft-max choice rule (based on u_S and λ).

Suppose $\lambda = 6$.

- P_{Obama}(-ing| {competent, friendly}) $\approx$ 0.31.
- P_{Obama}(-in'| {competent, friendly}) $\approx$ 0.69.

Obama after the BBQ

Suppose Obama is worried about coming off as incompetent when answering questions after the BBQ.

stern leader {comp, aloof}	cool guy/gal {comp, friend}	asshole {incomp, aloof}	doofus {incomp, friend}
0.20	0.20	0.30	0.30

Table: Obama worries about seeming incompetent.

Obama after the BBQ (predictions)

Suppose $\lambda = 6$.

- P_{Obama}(-ing| {competent, friendly}) ≈ 0.69.
- P_{Obama}(-in'| {competent, friendly}) ≈ 0.31.

Obama in front of Congress

Suppose Obama wants to be perceived as the stern leader in front of Congress.

- He wants to construct the {competent, aloof} persona.

Predictions

- P_{Obama}(-ing| {competent, aloof}) = 1.
- P_{Obama}(-in'| {competent, aloof}) = 0.

Summary

Bayesian game-theoretic models provide a framework for:

1. Formalizing sociolinguistic theories (in this case Third Wave variation theory (Eckert, 2000, 2008, 2012)).
2. Making both qualitative and quantitative predictions about (possibly variable) language use and interpretation.
3. Capturing interactive co-construction of meaning (in this case social).
   - The inference (16-b) arises as a product of coordination between the speaker and listener.

(16) a. I have been work[in] on my paper.
b. $\rightsquigarrow$ The speaker is friendly.

4. Capturing the contribution that speaker/listener prior beliefs make to social interpretation.

Where do personae come from?

In classic signalling games, S's type is determined by 'Nature'.

> *Game theorists like to think of the states of a signaling game as initial chance moves by a third player, called Nature, who selects any state $t \in T$ with probability $Pr(t)$, without any strategic concern of her own (cf. Harsanyi 1967, 1968a,b). In a signaling game, Nature reveals her choice to only the sender, but not the receiver. (Franke, 2009, 129)*

This doesn't seem quite right for identity construction...

Truth conditional meaning vs social meaning

We have reason to believe that propositional communication and persona/identity construction are different.

Propositional communication is reportative

S observes a fact about the external world and then tries to report it to L.

- S's type exists independently of both S's preferences and S's linguistic action reporting it.
- Nature metaphor is appropriate.

Speaker agency

Identity construction is performative

Aspects of S's identity are constructed (in part) through S's linguistic action.

- (Butler, 1990, 34): "There is no gender identity behind the expressions of gender; that identity is performatively constituted by the very "expressions" that are said to be its results."

Proposal (Burnett, 2016)

S's type should be should be chosen by human nature.

Next class. . .

Incorporating speaker agency into the model

- What determines the persona that S will choose in a given context? (Social theory)
- Extending SMGs with speaker agency.
- A speaker-agency model for social stratification.

References I

Bell, A. (1984). Language style as audience design. *Language in society*, 13(02):145–204.

Bergen, L., Levy, R., and Goodman, N. D. (2016). Pragmatic reasoning through semantic inference. *Semantics & Pragmatics*.

Bourdieu, P. (1979). *La distinction: critique sociale du jugement*. Les éditions de minuit, Paris.

Bourdieu, P. and Passeron, J.-C. (1970). *La reproduction: Éléments pour une théorie du système d'enseignement*. Les éditions de minuit, Paris.

Burnett, H. (2016). Signalling games, sociolinguistic variation and the construction of style. *accepted in Linguistics & Philosophy*.

Burnett, H. (2017). Sociolinguistic interaction and identity construction: The view from game-theoretic pragmatics. *Journal of Sociolinguistics*, 21(2):238–271.

Butler, J. (1990). *Gender trouble: Feminism and the subversion of identity*. Routledge.

Chater, N. and Manning, C. (2006). Probabilistic models of language processing and acquisition. *Trends in cognitive science*, pages 335–344.

Clark, R. L. (2014). *Meaningful games: Exploring language with game theory*. MIT Press.

Degen, J. (2015). Investigating the distribution of some (but not all) implicatures using corpora and web-based methods. *Semantics and Pragmatics*, 8(11):1–55.

Degen, J. and Tanenhaus, M. K. (2015a). Processing scalar implicature: A constraint-based approach. *Cognitive science*, 39(4):667–710.

Degen, J. and Tanenhaus, M. K. (2015b). Processing scalar implicature: A constraint-based approach. *Cognitive science*, 39(4):667–710.

Degen, J., Tessler, M. H., and Goodman, N. D. (2015). Wonky worlds: Listeners revise world knowledge when utterances are odd. *Proceedings of CogSci37*.

Dror, M., Granot, D., and Yaeger-Dror, M. (2013). Speech variation, utility, and game theory. *Language and Linguistics Compass*, 7(11):561–579.

Eckert, P. (1989). *Jocks and burnouts: Social categories and identity in the high school*. Teachers College Press.

Eckert, P. (2000). *Language variation as social practice: The linguistic construction of identity in Belten High*. Wiley-Blackwell.

References II

Eckert, P. (2008). Variation and the indexical field. *Journal of sociolinguistics*, 12(4):453–476.

Eckert, P. (2012). Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual review of Anthropology*, 41:87–100.

Fox, D. and Katzir, R. (2011). On the characterization of alternatives. *Natural language semantics*, 19(1):87–107.

Frank, M. C. and Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, 336(6084):998–998.

Franke, M. (2009). *Signal to act: Game theory in pragmatics*. PhD thesis, Institute for Logic, Language and Computation.

Franke, M. and Jäger, G. (2016). Probabilistic pragmatics, or why bayes' rule is probably important for pragmatics. *Zeitschrift fr Sprachwissenschaft*, 35:3–44.

Goffman, E. (1967). *Interaction ritual: essays on face-to-face interaction.* Aldine.

Goodman, N. D. and Stuhlmüller, A. (2013). Knowledge and implicature: Modeling language understanding as social cognition. *Topics in cognitive science*, 5(1):173–184.

Grice, P. (1975). Logic and conversation. *Syntax and Semantics*, 3:41–58.

Houston, A. (1985). *Continuity and change in English morphology: The variable (ING)*. PhD thesis, University of Pennsylvania.

Labov, W. (1966). *The social stratification of English in New York city*. Center for Applied Linguistics.

Labov, W. (1972). *Sociolinguistic patterns*. University of Pennsylvania Press.

Labov, W. (2012). *Dialect diversity in America: The politics of language change*. University of Virginia Press.

Lamont, M. (1992). *Money, morals, and manners: The culture of the French and the American upper-middle class*. University of Chicago Press.

Lamont, M. (2009). *The dignity of working men: Morality and the boundaries of race, class, and immigration*. Harvard University Press.

Lassiter, D. and Goodman, N. D. (2015). Adjectival vagueness in a bayesian model of interpretation. *Synthese*, pages 1–36.

Levinson, S. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. MIT Press.

Lewis, D. (1969). *Convention*. Harvard UP, Cambridge.

References III

Luce, R. D. (1959). On the possible psychophysical laws. *Psychological review*, 66(2):81.

Montague, R. (1973). The proper treatment of quantification in ordinary english. In *Approaches to natural language*, pages 221–242. Springer.

Ochs, E. (1992). Indexing gender. *Rethinking context: Language as an interactive phenomenon*, 11:335.

Ochs, E. (1993). Constructing social identity: a language socialization perspective. *Research on Language and Social Interaction*, 26:287–306.

Pinker, S. (1999). How the mind works. *Annals of the New York Academy of Sciences*, 882(1):119–127.

Podesva, R. (2004). On constructing social meaning with stop release bursts. In *Sociolinguistics Symposium*, volume 15.

Shannon, C. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423.

Shiffrin, R. M. and Steyvers, M. (1997). A model for recognition memory: Rem-retrieving effectively from memory. *Psychonomic Bulletin & Review*, 4(2):145–166.

Silverstein, M. (1979). Language structure and linguistic ideology. *The elements: A parasession on linguistic units and levels*, pages 193–247.

Silverstein, M. (2003). Indexical order and the dialectics of sociolinguistic life. *Language & Communication*, 23(3):193–229.

Sperber, D. and Wilson, D. (1986). *Relevance: Communication and cognition*. Blackwell, Oxford.

Steyvers, M., Griffiths, T., and Dennis, S. (2006). Probabilistic inference in human semantic memory. *Trends in cognitive science*, page 327?334.

Sutton, R. S. and Barto, A. G. (1998). *Reinforcement learning: An introduction*. MIT press.

Tamminga, M. (2014). *Persistence in the production of linguistic variation*. PhD thesis, University of Pennsylvania.

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., and Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science*, 331(6022):1279–1285.

Yuille, A. and Kersten, D. (2006). Vision as bayesian inference: analysis by synthesis? *Trends in cognitive science*, page 301?308.

Zhang, Q. (2008). Rhotacization and the 'beijing smooth operator': the social meaning of a linguistic variable. *Journal of Sociolinguistics*, 12(2):201–222.