

Situated Dialogue

Nicholas Asher[†] and Julie Hunter[‡]

[†]IRIT, CNRS, Toulouse

[‡]GLiF, University Pompeu Fabra, Barcelona &
IRIT, University Paul Sabatier, Toulouse

Meeting 1
July 9, 2016

Who are we?

Nicholas Asher

- IRIT, CNRS, Toulouse
- linguistic context
- discourse structure and dependence

Julie Hunter

- GLiF, Universitat Pompeu Fabra, Barcelona & IRIT, Université Paul Sabatier, Toulouse
- extra-linguistic context
- discourse structure

What is situated dialogue?

Dialogue in which interlocutors exploit the world around them in semantically significant ways

- Events and objects perceived or inferred from the extra-linguistic context affect the interpretation of linguistically expressed contents or of the discourse as a whole
- semantic dependence encoded in arbitrary sign (deixis)
- co-verbal or iconic gestures
- but also events that are not picked out by deixis or gestures can sometimes do the job of linguistically expressed contents in discourse

A very simple example

Example

A: Are you thirsty?

B: [Nods very vigorously]

An extended example

(Seinfeld)

Overview of course

- ① Various ways in which the extra-linguistic context is semantically relevant and standard theories for modelling these dependencies
- ② Show some inadequacies in the standard picture; argue for a vastly more interactive approach to extra-linguistic/linguistic interaction
- ③ Exploit a corpus study to develop parts of a more holistic approach.

Outline for this class

- ① language and the world
- ② deixis
- ③ gestures and iconicity
- ④ examples without deixis or gesture

Language and the world

- Sellarsian distinction between three types of language use: language-entry transitions, intralinguistic transitions, and language-exit transitions.
- Contexts of perception and observation give rise to language-entry transitions; e.g., one gives a linguistic response to some form of sensory stimulation. (cf. Quine's occasion sentences)
- Inferential and what we will call *coherence* relations are instances of intralinguistic transitions, transitions between one instance of a unit of content and another.
- The announcement of a resolution followed by an action that attempts to carry it out is a case of exiting a language.

Sellars's idealized view

- Sellars's paradigm of language use was that of a scientist, observing the world.
- observations yield assertions about what was observed.
- the scientist then infers from these observations certain predictions
- which the scientist goes off to test.

Even so, conversational analysts have found the language entry and exit metaphors useful (*hello, goodbye*) (and probably even postulated them independently).

Our view of language and the world

The Sellarsian picture is very much an idealization. Philosophers and linguists have shown linguistic/extra-linguistic interactions that are not language entry or exit rules as we have portrayed them:

- deictic gestures have been postulated to be essential for the semantics of certain expressions
- other gestures, which one might take to be extra-linguistic events (sign languages being a case apart), can affect the interpretation of speech or even replace linguistic moves.
- finally, dialogue agents can exploit extra-linguistic events using the same mechanisms used for intra-linguistic moves

All of this points to a complex and perhaps seamless interaction between an extra-linguistic (EL) context, linguistic moves, and discourse contexts.

To display, to show

- expressions or constructions with a constant meaning, but whose denotation changes in context.
- in particular, with the utterance context
- person, place, time

see Fillmore 1971 and Lyons 1977 for more

Personal deixis

I am teaching in this room.

- you are sitting
- she is sitting

Spatial deixis

I am teaching in **this** room.

- this/that
- here/there

Temporal deixis

I **am teaching** in this room

- tense
- now
- yesterday, this week

- EL context brought in/signalled via arbitrary linguistic sign.
- direct contribution to truth conditions/semantic content.

Back to the beginning

(Of the 20th century, at least)

Frege's theory:

- the sense of an expression e is what one grasps when one understands e and so should remain constant for each use of e .
 - constant sense for each deictic expression e .
- the sense of an expression e determines its reference.
 - different referent on different occasions of use, so different senses of e for each such occasion?

Frege (1956), Perry (1977)

Egocentric particulars: deictic expressions reduced to descriptions containing *this*, e.g. *now* means 'the time of this'

- how do we fix the reference of *this*?

Reichenbach on token-reflexives

Suppose I utter: 'I am teaching'

- yields a token, e.g. particular pattern of sound produced
- call this token t_1

⇒ The person who utters t_1 is teaching

Consequences:

- Two tokens of 'I am teaching' do not have the same meaning
- (11) If no one were to utter this token, I would not exist.
"Beliefs such as (11) could make one a compulsive talker." (Kaplan 1989, p. 520)

Reichenbach (1947)

Towards a solution

Introduce contexts, or contextual features, without reference to utterances.

The utterance time

Reichenbach: The E, R, S system (event point, reference point, speech point) included in principle a reference to an extra-linguistic point, the speech point. Add the reference point $R=S=E$ (present), $E=R < S$ (simple past), $E < R < S$ (pluperfect)

- Sam had passed the refuge at 9h. (the time of the passing occurred before 9), $E < 9h < S$
- Sam arrived at the refuge at 9h. (the time of the arrival occurs at 9), $E = 9h < S$
- Sam was at the party when Pat called. $R \subseteq_t E < S$
- I am at the party $S \subseteq E = R$.

Problems with Reichenbach's system

- not integrated into any formal semantics
- no semantics for adverbs.

Towards a formal model of the utterance time

Example

- (i) Last week I learned that there would be an earthquake.
- (ii) Last week I learned that there would *now* be an earthquake.

- *now* is semantically impervious to temporal operators that outscope it
- makes a truth-conditional contribution when in the scope of temporal operators

Kamp (1971)

A new operator: N

Example

(i) A child was born who would become ruler of the world.

$$P(\exists x)(x \text{ is born} \wedge F(x \text{ is ruler of the world}))$$

Example

(ii) A child was born who will become ruler of the world

??

- new operator N
- made explicit with *now*

The standard model for temporal operators

Standard model theory: interpretation of a formula ϕ in a language L .

Temporal logic: truth of a formula ϕ at a given time t .

- 'It rains' might be true at a time t but false at t' ($t' \neq t$)
- e.g., $\neg r$ but $F(r)$

Let $M = \langle Q, I \rangle$, where:

- Q is the set of atomic sentences q_1, q_2, \dots, q_n
- I is an interpretation function that assigns each q_n a function from $\mathcal{T} = \langle T, < \rangle$ to $\{0, 1\}$

$$[q_n]_{M,t}^{\mathcal{T}} = 1 \text{ iff } M(q_n)(t) = 1$$

Adding N to the standard model

For *now*, things are different.

- $\phi \rightarrow N\phi \wedge HN\phi \wedge GN\phi$

If it is the case that ϕ then **it is now the case that** ϕ , has always been the case that $N\phi$, and will always be the case $N\phi$.

Try to extend the standard model and assume $T = \{t_1, t_2\}$ and $t_1 < t_2$:

1. $q_n \leftrightarrow Nq_n$ axiom
2. $[q_n]_{M,t} = [Nq_n]_{M,t}$ id, 1
3. $[q_n]_{M,t_1} = 1, [q_n]_{M,t_2} = 0$ premise
4. $[Nq_n]_{M,t_1} = 1, [Nq_n]_{M,t_2} = 0$ 2,3
5. $[GNq_n]_{M,t_1} = 0$ G,4
6. $[q_n \rightarrow Nq_n \wedge HNq_n \wedge GNq_n]_{M,t} = 0$ 3,5

Time of utterance vs. time of evaluation

The problem: we are not interested in evaluating $N\phi$ relative to any point t ; we are interested in the truth at t of $N\phi$ at t_u , i.e. the utterance time.

Compare a token-reflexive treatment: we are interested in the truth at t of 'It rains at the time of t_1 ', where t_1 is the relevant token.

Double indexing: distinguish utterance time from evaluation time, but without reference to utterances.

A model for N : double indexing

The proposal in a nutshell: evaluate relative to two indices

- $[q_n]_{M,t,t'}^{\mathcal{F}} = M(q_n)(t)$
- $[N\phi]_{M,t,t'}^{\mathcal{F}} = 1$ iff $[\phi]_{M,t',t'}^{\mathcal{F}} = 1$

GNq_n can now quantify over times of evaluation $t'' > t$ as usual.

- $[G\phi]_{M,t,t'} = 1$ iff for all $t'' \in T, t < t'', [\phi]_{M,t'',t'} = 1$

However, we do not need to ask if $N\phi$ is true at each such t'' , but only if it is true at each such t'' that $N\phi$ is true at t'

- $[GN\phi]_{M,t,t'} = 1$ iff for all $t'' \in T, t < t'', [N\phi]_{M,t'',t'} = 1$
- iff for all $t'' \in T, t < t'', [\phi]_{M,t',t'} = 1$

Kaplan (1989)

Like Kamp (1971):

- expression (rather than utterance) based; logic for *occurrences* of sentences, i.e. sentence + context pairs
- double index account
- operator for temporal indexicals (+ *actually*)

More general:

- Context: agent, time, location, world
- Circumstance of evaluation: world-time pair

Two kinds of meaning

Character: the role associated with an indexical expression; what all occurrences of a given indexical expression i have in common

- e.g., I : the speaker/agent of the context

Content: what an indexical occurrence contributes to the truth conditions of the sentence in which the indexical figures; can change from one occurrence to another.

- Julie says, 'I am a woman'; true just in case **Julie** is a woman
- Nicholas says, 'I am a woman'; true just in case **Nicholas** is a woman

A two-dimensional semantics for indexicals

Character: function from contexts to contents

Content: function from circumstances of evaluation to extensions

An expression will contribute its meaning to either character or content, but not both. In Kaplan's words:

Indexicals have a context-sensitive character. It is characteristic of an indexical that its content varies with context.

Non-indexicals have a fixed character. The same content is evoked in all contexts. This content will typically be sensitive to circumstances, that is, the non-indexicals are typically not-rigid designators but will vary in extension from circumstance to circumstance. (p. 506)

Rigidity (for Kaplan)

A **rigid designator**: designates the same thing in all possible worlds

A **directly referential** expression: semantical rules determine rigid designation of the *actual* referent. (p. 493)

I, *here*, temporal adverbials and *actually* have constant contents, and because of *how* this happens, they are directly referential (and so rigid)

- indexicals contribute individuals to structured propositions in Kaplan's philosophical theory (contrast with Frege)
- *I* and *here* are associated with constant functions from contexts to individuals in Kaplan's Logic of Demonstratives; *now* and *actually* treated as operators à la Kamp.
- in any case, indexical contents are fixed independently of a circumstance of evaluation and are not the sorts of things that can be manipulated by logical operators.

Contexts vs. circumstances

In Kamp's logic, operators do not affect the second time index, but why not introduce operators that do?

Kaplan argues for a deep division between contexts and circumstances:

- I am here now vs. \square I am here now.
(proper contexts vs. improper circumstances)
- I do not exist
(fixing content vs. evaluating content)
- In some contexts, it is true that I am not tired now
(outside of quotation, this should be about me and the current time – no monsters!)

Kaplan in sum

Kaplan captured many important aspects of indexical semantics:

- the role of indexicals
- the distinction between the role and the contribution to truth-conditions (compare Frege, Russell, Reichenbach)
- special kind of rigidity

A gap in the account

But the way that the captured them assumed relied on a deep division between linguistic context sensitivity and extra-linguistic context sensitivity.

Where does linguistic context sensitivity fit into this picture?

- Anaphoric uses of demonstratives and pronouns have a context-sensitive meaning in a very important sense, but Kaplan's contexts don't cover them
- and they certainly couldn't cover bound uses of pronouns!

Bound vs. free variables

Kaplan:

- Deictically/demonstratively used pronouns: free variables under an assignment
- Anaphorically used pronouns: syntactically bound

A division echoed in many theories, despite intuitive similarity of the uses:

- [Pointing at Nicholas] He is from Brooklyn.
- I'm going to New York with Nicholas. He is from Brooklyn.

Dependence on linguistic context

Independent work on and models for linguistic context dependency

- what do we mean by this? Expressions with a semantics dependent on prior discourse context are those whose meaning is determined by the interpretation of the prior context.
- anaphoric pronouns, definite descriptions, presupposed material, ellipsis, but also tense.

Dynamic semantics

A leading approach to pronominal anaphora.

Motivated by cross-propositional anaphora and data on the accessibility of anaphoric antecedents.

Example

- (i) A dog_i began barking. It_i kept barking for hours.
- (ii) # No dog_i was barking. It_i was shut indoors.

In dynamic accounts, an indefinite introduces an element of interpretation that remains available for the interpretation of subsequent discourse unless it is blocked by the presence of logical or modal operators.

The basic idea

- the meaning of an expression is given in terms of what it *does* to the incoming discourse context.
- and so meaning, e.g. of a sentence s is represented as a relation between contexts, or a function between a set of input contexts and a set of output contexts.
- The meaning of a text is the result of composing sentential meaning via relational composition.
- Various implementations of this idea exist.

Some dynamic semantics for tense and anaphoric expressions

- Kamp's discourse representation theory: a theory in which the construction of a logical form for a text is dynamic (the construction of a DRS) but the semantics for the logical form is not (embedding).
- Groenendijk and Stokhof (1991), Muskens (1996). Theories where the semantics is dynamic, and meanings are actions as in dynamic logic.
- Barker and Shan (2012), de Groote (2006). Theories where the types of lexical entries are made more complex to take account of an input context and an output context; can define both a classical and a dynamic semantics within the framework, classical truth conditions for sentences or texts.

Interpreting tense in discourse

Let's go back to our Reichenbachian examples

- Sam arrived at the refuge at 9h. Pat had prepared the meal and was waiting for him. (the preparation occurs before the arrival)
- How to extend and formalize Reichenbach's picture?
- Kamp 1979 married a Davidsonian view of verbs as event predicates with dynamic semantics (Kamp and Reyle 1993).

Example

A man entered. He smoked a cigarette. He left.

- we will illustrate with Kamp's DRT
- we will build a DRS for the discourse incrementally, sentence by sentence.
- a DRS is a pair of sets (U, Con) where U is a set of *discourse referents* or variables and Con is a set of formulas in U .
- Some of these formulas will exploit parameters whose values (discourse referents) can be updated as the DRS is constructed; their names are like Reichenbach's.

The DRS for the first sentence

x, e, n
$man(x)$ $enter(e, x)$ $e < n$ $Rpt := TPpt = Ept := e$

After the second sentence:

We build a DRS for the second sentence and then merge it with the DRS for the first, resolving anaphoric dependencies:

x, e, e_1, y, n
$man(x)$
$enter(e, x)$
$e < n$
$cigarette(y) \text{ smoke}(e_1, x, y)$
$e < e_1$
$e_1 < n$
$Rpt := TPpt = Ept := e_1$

After the third sentence

x, e, e_1, e_2, y, n

$man(x)$

$enter(e, x)$

$e < n$

$cigarette(y) \text{ smoke}(e_1, x, y)$

$e < e_1$

$e_1 < n$

$leave(e_2, x)$

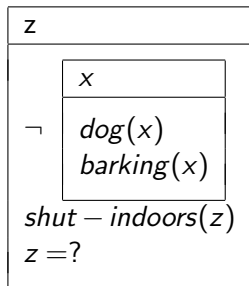
$e_1 < e_2$

$e_2 < n$

$Rpt := TPpt = Ept := e_2$

Constraints on accessibility

No dog was barking. It was shut indoors:



The variable introduced by the anaphor (z) cannot be identified with its antecedent (x) due to the semantics of the DRS.

Dynamic semantics: assessment

- though alternative ways of treating nominal anaphors exist that do not use dynamic methods, dynamic semantics has become an accepted way of modeling linguistic context dependence.
- a uniform treatment of many phenomena in which context dependence has been observed: presupposition, ellipsis, attitude and modal dependencies (modal subordination), plurals, as well as nominal and temporal anaphors.
- also used not only for spoken languages but sign languages (Schlenker's work)

The Received view of EL and L contexts

- Kaplanian semantics for indexicals and demonstratives
- dynamic semantics for elements dependent for their interpretation on the linguistic discourse context. (Sellars intra-linguistic moves developed)
- other EL influences ignored

Mechanism of deixis and mechanism of anaphora completely independent.

A complication in Kaplan's account

What is the character of a demonstrative?

- contexts are not so easy to set up
- what is the role of demonstrations?

The theory of 'Demonstratives'

A demonstration is used to *fix the reference* of a demonstrative.

- demonstrations are necessary
- context and character are not enough; demonstratives are incomplete

Frege's theory of demonstrations

Frege's theory of *demonstrations* was correct.

- a demonstration presents a demonstratum in a particular manner
- the demonstratum picked out by a demonstration can vary from one circumstance of evaluation to another; the same demonstration might pick out one entity in actual circumstances, but another entity in counterfactual circumstances.
- a demonstration without a demonstratum might have had a demonstratum; and one with a demonstratum might have failed to have one

Fregean theory of demonstratives

But the theory of *demonstratives*:

- The sense of the demonstration is the sense of the demonstrative.
- Thus, the contribution of the demonstrative to the proposition expressed is something that can vary from one circumstance to another.

was wrong.

Pat and Mike

'He now lives in Princeton, NJ.'

- → **Paul**, who lives in Princeton
 - **Pat**, the proposition expressed, is true.
- ① What if Paul had moved away?
 - Pat would be false
 - ② What if Paul and Charles had sneakily switched places?
 - Pat would be true! (Even if Charles doesn't live in Princeton!)

'He now lives in Princeton, NJ.'

- → **Charles**, who does not live in Princeton
- **Mike**, the proposition expressed, is false.

Dthat

Demonstrations are like descriptions (sort of), but demonstratives are not

- could there be a demonstrative that requires completion by a description?
- yes!
- dthat[α]

'Afterthoughts': directing intentions

Push the analogy with descriptions even further:

- Donnellan's martini case: 'Who is the guy drinking a martini?'

The description doesn't need to fit the man for the question to be clear

- question is really: 'Who is that guy?' and the description is only there to help the interlocutor figure out who the referent is.
- not a semantic requirement
- the obvious linguistic cues seem to fail us in this case.

Similarly, a demonstration is only there to help reveal a speaker's inner intentions; fitting the demonstratum is not a semantic requirement.

An exotic kind of ambiguity

I'll take **that**, **that**, and **that**, but not **that**.

- different context for each token?
- that would reintroduce problems that Kaplan was avoiding in adopting occurrences in the first place

I'll take **that**₁, **that**₂, and **that**₃, but not **that**₄.

- each syntactic occurrence coupled with a directing intention.
- intention is neither context nor character nor referent

A can of worms

'Pure' indexicality is only one way in which the extra-linguistic context is made relevant for semantic content.

- 1 What is the semantic role of demonstrations and other co-verbal gestures?
- 2 What other influences from the extra-linguistic context are left out of Kaplan's account?

A description of relevant phenomena

- 1 demonstrative gestures
- 2 co-verbal gestures more generally
- 3 iconic gestures and actions
- 4 exploitation of EL events without deixis

Demonstratives without demonstrations

Demonstrations are a well accepted violation of the insularity of intralinguistic moves.

But are they always necessary for demonstratives?

- No. (More on that later.)

When are they necessary?

Demonstrations without demonstratives

They can also accompany other expressions

Example

I can give you OTHER BOOKS [demonstration gesture]

- gesture could attach to *books*, the NP, or the VP

What is the relation between the gesture and the phrase?

Generalizing: demonstrations are a particular kind of co-verbal gesture

- coverbal gestures have been studied (McNeil 1992, Kendon 2004); some corpora with annotations of gestures by the Bielefeld group among others (Kranstedt et al. 2005) exist.
- demonstrations are a paradigm co-verbal gesture.
- demonstrations involve a motion of the hands and perhaps also of the body that accompanies speech.
- demonstrations have their own structural characteristics and two general functions: object pointing and region pointing
- Davidson (2016) proposes a semantics integrating pointing gestures with linguistic input.

Co-verbal gesture generally

Many sorts of gestures accompany linguistic production.

- described by features like hand (and finger) shape, location and movement, and palm orientation.
- deictic but also depictive or iconic co-verbal gestures.
- these gestures typically amplify or reinforce the content of the verbal message (Cassell & Prevost 1996).

Gestures and language

The linguistic/extra-linguistic divide:

- (Goffman, 1963, Ekman and Friesen, 1969, Kendon, 2004) all distinguish communicative actions from other behaviors that people do in conversation, such as practical actions and incidental 'nervous' movements.
- e.g., cooking an omelette while talking about one's day; a nervous tick while talking.

Gestures are **communicative actions**.

Co-verbal gesture tightly coupled to linguistic features

Prosodic and temporal overlap constraints on where coverbal gestures, the "stroke" may appear (McNeill 1992, Alahverdzhieva et al. 2012)

Example

- (i) Your mother CALLED [hand lifts to ear imitating a telephone]
- (ii) #Your MOTHER called [hand lifts to ear imitating a telephone]

Example

And um I thought not too edgy and like a box, more kind of hand-held more um . . . not as uh computery [typing gesture] and organic, yeah, more organic shape I think.

- the overall message is not [computery with a keyboard].
- distinct but related content to the lexical item.
- this requires a theory in which the content of the gesture falls under the scope of a linguistically given element.

Coverbal gesture and meaning

- the meaning of the gesture seems heavily dependent upon the meaning of the linguistic element *computery* with which it co-occurs.
- the motion in another context might have picked out a keyboard instrument.
- in general the meaning of the gesture on its own is highly underspecified
- and is typically determined at least in part by the linguistic content it modifies.

Gestures and iconicity

Not all gestures or communicative EL actions are arbitrary signs (like demonstrations)

Example

I looked at Kate during the talk and she was like [models sleeping]

- The gesture, [models sleeping], resembles the activity of sleeping.
- Sign language expressions for *grow* and their adverbial modifiers—e.g., *grow a little*, *grow a lot*—are often iconic. (see Schlenker p. 28 for a picture).
- hands, but also body position, facial expression are used.

Iconicity

- Iconicity has to do with the semantics of an expression.
- the connection between a linguistic expression and its denotation is typically arbitrary.
- e.g., the meaning of the word *apple*, $\|apple\|$ is a set of fruits that do not in themselves and whose appearance or genetic make-up bear no similarity to the shape of the word 'apple'.
- However, quotation is a device whose meaning is a function that converts an expression α into an expression β such that $\|\beta\|$ has a structural similarity to the expression α . 'apple' has two 'p's in it.
- Quotation is an iconic device (Davidson).

Other iconic expressions

Example

- (i) The talk was loooooong.
- (i) The talk was shooooooort.

- The excessive duration of the vowel in (i) gives a vivid idea of the real or experienced duration of the talk and (ii) in contrast yields a rather odd effect (Schlenker).
- *loooooong* is an iconic expression.

Iconic gestures and grammar

Iconic gestures can complete linguistic context:

Example

I looked at Kate during the talk and she was like [models sleeping]

Need not even be signalled by an expression such as *like*

Example

The flute plays [insert music] and then the oboe enters.

The music provides a particular point at which the oboe should enter, and it need not be the end of the flute's role (Clark).

Beyond character

More than a question of supplementing character; no expression whose character can be analyzed.

- extra-linguistic context fills in gaps in LF.
- extra-linguistic content changes the interpretation: 'The flute plays and the oboe enters' would have a different interpretation.

More interesting EL/linguistic interactions

- Semantic effects: the phrase *the flute plays [insert music]* has an interpretation on which the sound is what the flute plays.
- a slight change in the example: *the flute plays the first three bars [flute playing] and then the oboe enters to pick up the melody.*
- and the interpretation changes: the flute playing elaborates or specifies what the first three bars are.
- yet another change *the flute plays this [flute playing]*, and we get something like the first interpretation but by a different linguistic means.
- how do you compose the sound together with different linguistic content to get the different meanings?

Actions in lieu of verbal discourse moves

- Sometimes an action (gesture) takes the place of a verbal response or fills in for a linguistic expression.

Example

A: Do you love me? (i) B: [shakes head from side to side], (ii) [nods], (iii)
B: [kisses A]

- the actions in (i) and (ii) have a conventional meaning, but the action in (iii) does not.
- the action in (iii) has implicatures that answer the question.
- there isn't a model of these kinds of interactions between EL events and L discourse moves, as far as we know.

Stepping back

We've seen various roles that EL events and objects can play in discourse.

- With iconic expressions, EL structure affects the interpretation of the iconic expression.
- With indexical and demonstrative expressions, the EL context directly determines interpretation.
- Sometimes EL events (gestures) accompany and affect the interpretation of discourse moves.
- Sometimes EL events simply replace L discourse moves.

(Seinfeld again)

Next class

A broad range of ways that the extra-linguistic context influences semantic content—far beyond Kaplan's well-behaved indexicals or Sellars' language entrance and exit moves

Next class:

- 1 further chip away at the linguistic/extra-linguistic divide
- 2 show that dynamic models of presupposition can supply a uniform mechanism for extra-linguistic and linguistic context sensitivity
- 3 argue that a full uniform account of this mechanism requires supplementation with rhetorical relations and structure

For slides from the class, visit: <http://www.juliejhunter.com/teaching.html>