Grammaticality Judgments as Linguistic Evidence
Course Outline and Introduction

Brian Murphy

Language, Interaction and Computation Lab
Centre for Mind/Brain Sciences (CIMeC)
University of Trento

ESSLLI 2009, Bordeaux
Course Outline

- Notions of Grammaticality, and Current Practice in Linguistics
- Alternative Sources of Linguistic Evidence
- Guidelines for Eliciting Judgements of Grammaticality
- Scales for Measuring Grammaticality
- Theoretical Implications
Today’s Outline

- Quick History of the Notion of Grammaticality
- Grammaticality of Syntax, Semantics and more
- Absolute/Relative and Binary/Continuous Grammaticality
- Expert and Non-Expert Grammaticality
- Bias in Judgement and Interpretation of Grammaticality
Prehistory: all speakers know you can’t say certain things, ...

▶ intuitively - because their internalised grammar tells them so
▶ prescriptively - because someone told them so (Académie française, Chicago Style Guide, Duden, school teachers, witch-doctors, people who write outraged letters to newspapers ...)
▶ and they have some sense of this difference (e.g. expressions you use day-to-day, but wouldn’t write in an essay, say in a job interview, ...)
All speakers can evaluate the well-formedness and overall appropriateness of their own and other people’s utterances, and often do so consciously or unconsciously, e.g.

- choosing the most suitable construction/phrase/work for a communicative task
- proof-reading a document
- helping a language learner express themselves correctly
Sībawayhi (8c Persian scholar of Arabic) gave this typology [Bohas et al., 1990] of well-formedness:

- **straight** (meaningful)
  - **good** (well-formed)
  - **bad** (ill-formed)
- **crooked** (unintelligible)
- also appealed to accepted usage; and felicity in context
Chomskyan grammars

- early days [Chomsky, 1964]: violations of semantics are less severe than violations of syntax
  - John plays gold [grammatical]
  - Golf plays John [ungrammatical]
  - John plans and [more ungrammatical]

- nowadays: well-formedness exclusively syntactic and unitary; interpretability and plausibility play no role
  - Colorless green ideas sleep furiously
  - Adj-Adj-N-V-Adv
  - Several large estates grew outwards [Wacky internet corpus]
  - Furiously sleep ideas green colorless
  - Adv-V-N-Adj-Adv
  - well done u succeeded good [Wacky internet corpus]

- grammaticality (competence) distinguished from acceptability (performance)
Mid-nineties:

- Schütze questions empirical basis of linguistics [Schütze, 1996]
- Cowart proposes methodological guidelines for acceptability experiments [Cowart, 1997]
- Bard et al introduce continuous marking scale, borrowed from psychophysics [Bard et al., 1996]

Mid-noughties:

- Methodological practices have continued much as before
- Edelman and Christiansen attack theory and practice of Minimalism head-on [Edelman and Christiansen, 2003]

★ putting forward a theory is like taking out a loan, to be repaid by gleaning an empirical basis for it; theories that fail to do so (or their successors that may have bought their debts) are declared bankrupt

Brian Murphy (CIMEC)
Grammaticality: Outline & Intro
ESSLLI 2009
Minimalism is not even mentioned in recent reviews of and opinions on various aspects of language research in this journal [Trends in Cog Sciences], ranging from sentence processing and production and syntactic acquisition to the brain mechanisms of syntactic comprehension.

Other linguists are more polite, but make similar points [Wasow and Arnold, 2005, Bornkessel-Schlesewsky and Schlesewsky, 2007].

- First comprehensive empirical studies of experimental methodology for judgement experiments (including Featherston, Fanselow, Murphy, Sprouse, Weskott, Bader)
- Interest in internet-based experimentation (Keller, Murphy, Myers)
In general, linguists and other cognitive scientists agree that some structures/sentences/utterances are permitted in a language, while others are not.

The rules/patterns/conventions that govern this can be collectively described as grammar.

Grammaticality is the measure of this.
What can make a sentence “bad” I

- Gobbledygook
  - Buses grave expectedly novice chute spoken?
- Number, gender, tense agreement, word order
  - He am here already
  - I saw two student
  - Jane wants some grape
  - The penguin lives in zoo the
- Ambiguous, implausible, uninterpretable sentences
  - Colorless green ideas sleep furiously
  - I’ll get him for three (thousand dollars)
- Discourse constraints
  - I met Mary yesterday. She bought a second hand car
  - I met Mary yesterday. A second hand car was sold to her
What can make a sentence “bad” II

- Processing load
  - The man is here again
  - The man that the security guard saw is here again
  - The man that the security that guard that the manager hired saw is here again
  - The man that the security guard that the manager that you told me about hired saw is here again

- Garden path sentences
  - The cotton clothing is usually made of grows in Mississippi
Range of Theoretical Varieties

- Absolute, binary, syntax based
  - Chomskyan grammars post 1965,

- Absolute, binary, broadly based
  - Lexical Functional Grammar, [Kaplan and Bresnan, 1982]
  - Role and Reference Grammar, [van Valin, 1993]
  - Phrase Structure Grammars,
    [Gazdar et al., 1985, Pollard and Sag, 1994]

- Relative, binary
  - Optimality Theory, [Prince and Smolensky, 1997]
  - Variants: Linear/Stochastic Optimality Theory,
    [Boersma, 2004, Keller, 2006]

- Relative, continuous
  - Harmonic Theory, [Legendre et al., 1990]
Theoretical: Admissibility

- The considered informed opinion of an expert whether a structure is licensed by the grammar
- Strives to generalise away from particular token (e.g. lexical choices, possible alternative interpretations)
- Pros: trained, finer sense of well-formedness (like a sommelier)
- Cons: risk of confirmation bias, and selection bias
- Paradox: expert judgements must be informed by overall theoretical framework, but should be neutral with respect to the theoretical question being explored
Empirical: Acceptability

- The spontaneous reaction of an average speaker on how well-formed a sentence or utterance is
- Pros: less subject to bias, more naturalistic, theory neutral
- Cons: cannot distinguish between different kinds of ill-formedness

Reality is somewhere in between

- Non-expert informants are likely to think there is a “right” answer, and will strive to give it, based partly on prescriptive grammars
- Experts are likely to be influenced by their spontaneous interpretation and gut reaction to the sentence, even if they try not to be (Featherston)
Gathering judgements, theorists often rely on:

- their own intuitions
- or those of a small number of colleagues, who may have similar theoretical stances
- of constructed sentences, in isolation
- annotated with ad-hoc marginal categories (*, ?, ??, *?, etc)

- borderline grammatical
- not sure if it is grammatical
- grammatical in some circumstances (e.g. interpretations, contexts, intonations, domains, registers)
- grammatical for some (groups of) speakers
Conventional Practice II

- **Interpreting the evidence** [Wasow and Arnold, 2005, Bornkessel-Schlesewsky and Schlesewsky, 2007]
  - linguists often make strong formal assumptions, that rule out otherwise plausible explanations
  - linguistics has a history of ignoring inconvenient evidence
  - this has led to false conclusions

- In fairness, all researchers are subject to theoretical preconceptions and other bias. But linguists are unusual in that they often interpret evidence that they themselves have generated.
Knowing the theoretical point being made, the theorist-informant may be biased in:

- selecting illustrative examples
- creating illustrative examples
- judging borderline examples
- wrongly attributing grammaticality to competence or performance
- selecting like-minded informants
- group-think
Judgement bias is a well established cognitive phenomenon

- priming effects of social attitudes - e.g. Implicit Association tests
- pharmacology - double-blind testing
- education - grades affected by teacher’s attitudes towards student
- personality tests, astrology - accept that which confirms beliefs about self, ignore others

Again, bias is a problem all disciplines struggle with - something to control rather than be eliminated
But Even Linguists Don’t Agree

- Grammaticality judgements from 22 native Dutch professional linguists (Gisbert Fanselow)

<table>
<thead>
<tr>
<th>0vp</th>
<th>ik weet niet wie wat gekocht heeft</th>
<th>ik weet niet, wie wat aan gegeven heeft</th>
<th>wie weet wat gekocht heeft voor zijn zusje</th>
<th>wie weet wat wie gekocht heeft voor zijn zusje</th>
<th>wie weet wat wie gekocht heeft voor zijn zusje</th>
<th>wie weet wat wie gekocht heeft voor zijn zusje</th>
<th>wie weet wat wie gekocht heeft voor zijn zusje</th>
</tr>
</thead>
<tbody>
<tr>
<td>1c</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>1d</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>1p</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>1s</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>1u</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>2e</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>2f</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>2g</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>30l</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>30n</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>39r</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>39s</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>400u</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>401q</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>41q</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>41r</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>41t</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>41u</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>41v</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>41w</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>41x</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>41y</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>41z</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>5b</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>6h</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>6t</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>
Summary

- Grammaticality is an elastic notion, ranging from theoretically informed and idealised judgements by experts, to spontaneous gut-reactions from ordinary speakers.
- Conventional methodology in linguistics has led many to question the general validity of its theories.
- But as we will see later, these methodological issues are relatively easy to rectify.
- Discussion ...

  ▶ In your own work, do you find it difficult to judge the difference between grammatical/ungrammatical and performance/competence?
  ▶ Do you often disagree with judgments you see in the literature? (in your native or near-native languages)
  ▶ Are any of you already working on such methodological questions, or interested in doing so?
  ▶ Do you think grammaticality is/should be exclusively syntactic, or more broadly based
  ▶ Should we care if other cognitive sciences ignore us or not?

My Personal Opinion:

- Expert judgements of grammatical admissibility are different and complementary to non-expert judgements of acceptability.
- But expert judgements are dangerously vulnerable to subjective biases.
- Experts should validate their intuitions against the intuitions of other experts, and of non-experts.
- The onus is then on the expert to interpret and justify discrepancies.
- If performance is invoked, its precise effects should be demonstrated.
What comes next

- Notions of Grammaticality, and Current Practice in Linguistics
- **Others Sources of Linguistic Evidence**
  - Usage and frequency
  - Behavioural experiments: timing, eye-tracking
  - Neuroscience: EEG, MEG, fMRI
- Methodology for Eliciting Judgements of Grammaticality
- Scales for Measuring Grammaticality
- Theoretical Implications
Audience Participation

- What are your interests?
- Any topics missing?
- Would you like to vary the format? E.g. readings, seminars.
Acknowledgments

- Doctoral Supervisor at Trinity College Dublin, Carl Vogel
- Fellow group members Roberto Zamparelli, Marco Baroni, Massimo Poesio, and all participants in Linguistic Evidence seminar
- Current cabal of Linguistic Methodologists: Sam Featherston, Thomas Weskott, Gisbert Fanselow, Markus Bader, Jana Häussler, Jon Sprouse, James Myers, and probably some other people I’ve forgotten
- All the anonymous informants who have taken part in our experiments
References I


Boersma, P. (2004). A stochastic OT account of paralinguistic tasks such as grammaticality and prototypicality judgments. Unpublished manuscript.


References II

Knowledge of language: its nature, origin, and use. 
Praeger, New York.

The minimalist program. 

Experimental syntax. 

How seriously should we take minimalist syntax? a comment on Lasnik. 

Generalized Phrase Structure Grammar. 
Blackwell, Oxford.

Kaplan, R. M. and Bresnan, J. (1982). 
Lexical-functional grammar: A formal system for grammatical representation. 
Linear optimality theory as a model of gradience in grammar.

Harmonic grammar – a formal multi-level connectionist theory of linguistic well-formedness.

*Head-Driven Phrase Structure Grammar.*
CSLI, Stanford.

Optimality: From neural networks to universal grammar.

*The Empirical Basis of Linguistics.*
University of Chicago Press, Chicago.