Subject:
Approaching a Theoretical Definition via Diachronic Typology

Spike Gildea
University of Oregon
The Problem: An Operational Definition without a Theoretical Definition

The Consequence: We can’t decide what Subject Properties really mean

- When they disagree
- When they are “partial”
- When they single out absolutive or obviative instead of nominative
This talk

- How definitions work
- How Keenan’s definition does and does not work
- Using diachrony to recognize valid properties
- Conclusions
Definitions in Philosophy of Science

- Theoretical Definitions
  - internally coherent perspectives
  - organize our understanding
  - Provide epistemological foundation for Operational Definitions
Definitions in Philosophy of Science

- **Operational Definitions**
  - The criteria for identifying categories in data

- **Logical Validity**
  - A criterion follows from theoretical definition
  - Community of experts agrees

- **Empirical Validity**
  - More than one criterion points to the same conclusion
Example of the PHONEME

- Theoretical Definition: A PHONEME is
  - a cognitive unit of sound
  - that *contrasts* with others
  - speakers combine phonemes into contrastive forms to associate with meaning

- Roughly, PHONEMES represent what we think we say, which is not identical to the phones we actually produce.
Example of the PHONEME

➢ Operational Definitions
  • Contrast in Distribution:
    ○ Contrastive Distribution
    ○ Complementary Distribution
  • Contrast in Meaning
    ○ Minimal Pairs
    ○ Free Variation
  • Lack of meaning contrast in Morphophonological alternation
Example of the PHONEME

- Logical Validity
  - Follows from presence or absence of contrast

- Empirical Validity
  - Follows from triangulation — the three tests identify the same phoneme

- Theoretical Challenge
  - When the three tests do not agree
Theories of Subject

- Morphological and syntactic category of most European languages
- Semantic category of AGENT
  - The one who does something
- Pragmatic category of TOPIC
  - What we are talking about
- Specified position in syntactic tree
  - NP external to VP
- Theoretical primitive
  - e.g., in Relational Grammar
Keenan’s 1976 Subject Properties

There is a large body of lore concerning the notion, and any proposed definition must largely agree with the traditional, and to some extent, pretheoretical usage of the term. Our approach then will be to collect a large and diverse set of cases from different Ls in which our pretheoretical judgments of subjecthood are clear. (Keenan 1976: 306)
Keenan’s 1976 Subject Properties

There is a large body of lore concerning the notion, and any proposed definition must largely agree with the traditional, and to some extent, pretheoretical usage of the term. Our approach then will be to collect a large and diverse set of cases from different Ls in which our pretheoretical judgments of subjecthood are clear. (Keenan 1976: 306)
Keenan’s 1976 Subject Properties

There is a large body of lore concerning the notion, and any proposed definition must largely agree with the traditional, and to some extent, pretheoretical usage of the term. Our approach then will be to collect a large and diverse set of cases from different Ls in which our pretheoretical judgments of subjecthood are clear. (Keenan 1976: 306)
Keenan’s approach

- Identify “Semantically Basic Sentences” (b-sentences) in each L
- Identify “basic subjects” (b-subjects)
- List properties of b-subjects in each L
- This list of properties can identify b-subjects in less clear cases
Keenan’s results

- **Pragmatic properties**
  - independent existence
  - indispensability
  - identifiability

- **Semantic Properties**
  - at least AGENT

- **Coding properties**
  - flagging
  - indexation
  - order

- **Behavior & Control Properties**
  - coreference patterns
  - “transformations”
The missing link: no theoretical definition

- Why these properties?
- Empirical Validity
  - When they point together to the same NP, they appear to validate each other
When they do not agree?

- There is no theoretical definition to guide us
- Some problematic consequences
  - “clear preponderance of properties” (p. 312)
  - Quasi-Subjects / Semi-Subjects
  - Languages without subjects
- We need theory to guide our choices!
An example of theoretical clarity

We suggest, that is, that the syntactic concept “subject” ought to be identified by syntactic means (in particular, the role of an NP in those transformational processes which seem most sensitive to grammatical relations); the more straightforward the correspondence between such syntactically defined categories and those of surface morphology, the better, of course, but this is definitely a secondary consideration.”

(Anderson 1976: 8)
An example of theoretical clarity

We suggest, that is, that the **syntactic concept** “subject” ought to be identified by **syntactic means** (in particular, the role of an NP in those transformational processes which seem most sensitive to grammatical relations); the more straightforward the correspondence between such syntactically defined categories and those of surface morphology, the better, of course, but this is definitely a secondary consideration.”

(Anderson 1976: 8)
An example of theoretical clarity

We suggest, that is, that the syntactic concept “subject” ought to be identified by syntactic means (in particular, the role of an NP in those transformational processes which seem most sensitive to grammatical relations); the more straightforward the correspondence between such syntactically defined categories and those of surface morphology, the better, of course, but this is definitely a secondary consideration.”

(Anderson 1976: 8)
How this guides practice

- Subjects can be identified in the absence of (even contradicting) nominative-accusative coding properties
  - “Deep” versus “Surface” Ergativity
  - Noncanonical subjects (and objects)
  - Differential argument marking/indexation
Remaining problems

- Syntactic subject properties can be inconsistent
  - Split between NOMINATIVE and DATIVE SUBJECTS
  - Split between NOMINATIVE and ABSOLUTIVE
  - Split between TOPIC/FOCUS and AGENTS
  - Split between PROXIMATE/OBVIATIVE and AGENTS

- It is not uncommon for “transformations” to contradict “control of coreference”

- How can we make progress towards a theoretical definition?
The approach in this talk

- Observe changes in subject properties over time
- Identify “motivated” versus “epiphenomenal” changes
- Examine the “motivated” changes
- Reverse-engineer a theoretical definition of subject
What might we gain?

- What insight follows from this theoretical definition?
  - Better understand why subjects are so prevalent?
  - Better organize our thinking about conflicting subject properties in individual languages?
  - Better understand why some languages might have no grammatical subject?
The sequence from here...

- Semantic and Pragmatic properties
- Mechanisms of syntactic change
- Three examples
  - Passive > Ergative
  - Nominalizations > Ergative
  - Nominalizations > Focus > Unmarked main clauses
- Conclusions
Semantic and Pragmatic Properties

- Appear to be fairly constant
  - Subject = Agent
  - Subject = Topic
- Claimed contradictions of these properties are interesting!
Semantic roles

- Intransitive subject (S) can be almost any semantic role!
- For bivalent predicates (A, P), several hierarchies have been proposed
  - Agent is (almost) always at the top
    - Absolutive P > Ergative A in a few languages?
  - Competition is found between
    - Experiencer and Stimulus
    - Possessor and Possessed
Pragmatic Roles

- Widespread agreement Subject is Topic
- No agreement on Operational Definition of Topic
  - Intuitive identification is not reliable
  - If grammar guides intuition, this is also circular
  - If grammar of translation guides intuition, this is even worse!
- Modern studies of Information Structure have not eliminated role of intuition
Non-intuitive tests of “Topic”? 

- Givón 1994/1997 text counts
  - Referential distance (Accessibility)
    - Not relevant to grammatical relations
  - Topic Persistence (Importance)
    - Highly correlated with grammatical relations
- Tomlin 1995/1997 fish film experiment
  - Stimulate Attentional Detection
  - In 19 languages, highly correlated (often at 100%!) with subject selection
Connection between Topic & Subject is suggested, but not definitional

- Intuitions are difficult to replicate
- Text counts apply to limited genres of text, and have not been widely used
- Fish Film is limited to one verb type and focuses on subject and voice.
Mechanisms of Change

- Construction Reanalysis
  - Is covert (can’t tell exactly when it happens)
  - Changes categories without changing surface form
  - Affects all elements of the construction
Mechanisms of Change

- Analogical Extension
  - Is overt (you see it immediately)
  - Changes surface form (not necessarily categories)
  - Can affect limited elements of the construction
A quick example

- English Progressive
- Locative predicate locates agent in an activity (12th, 13th century English)

  thær he wes an sloeting [hunting]

  ...thei weren at robbinge.

  thar he was in hontynge

1205 Layamon (Visser, p.1998, 2001)
One example of Syntax

- Various prepositions are attested marking the locative relation

  \textit{thær he wes an sloeting [hunting]}

  \textit{...thei weren at robbinge.}

  \textit{thar he was in hontynge}

1205 Layamon (Visser, p.1998, 2001)
Evidence for Reanalysis

- Semantics: Construction expands to other verb types, such as accomplishments and even states.
- Collocation restrictions: Construction becomes conventionalized with single preposition, *on*
An Analogical Extension

- The erstwhile preposition *on* begins to reduce: 
  
  \[ \text{on} > \text{an} > \text{a-} > \emptyset \]

  *now whyll I am awhryttyng of thys letter...* 
  
  *I am \( \emptyset \) doynge of my nedynges* 
  
  1475 Cely Papers (Visser, p.2004, 2002)

- Evidence for reanalysis: this reduction only occurs in the Progressive Construction.
Relevance for Argument Structure

- Inherited subject property: auxiliary agreement with nominative (S+A)

\[
\begin{array}{c}
[ & S & ] & \text{AUX-S} & [ & \text{V}_{\text{INTR}} & ] \\
John \text{ Cheynye} & \textit{is} & \textit{owt a} & \textit{hawking} \\
I \textit{am} & \text{doynge} & \text{of my nedynges} \\
\end{array}
\]

1481/1475 (Visser, p.1998, 2002)
Relevance for Argument Structure

- Inherited (but innovative) object property: of as Accusative preposition

\[
[ \text{S} ] \text{AUX-S} [ \text{VINTR} ]
\]

*John Cheynye is owt a hawking*

*A AUX-A [ VTR [ [ P ]]]

*I am doynge of my nedynges*

1481/1475 (Visser, p.1998, 2002)
General Lesson

- Grammatical properties can be construction-specific
- Reanalysis can create “epiphenomenal” grammatical properties
  - The addition of object flagging is not “motivated” by the object relation
  - The object of progressive constructions is not “more” of an object because of flagging
- Analogical Extension can make incremental changes one feature at a time
Three types of examples from South America

- Passive > Ergative main clause
  - Parallel in Indic

- Action Nominalizations > Ergative main clauses

- Participant Nominalizations > Focus Clauses (> ordinary main clauses)
  - Parallel in Austronesian
Passive > Ergative main clause

- Carib, Apalaí, Wayana, and Tiriyó (Cariban) Gildea (1997)

- The evolutionary chain
  - Stative participle in nonverbal predicate
  - Gains eventive reading, becomes agentless passive
  - Adds option of oblique (demoted) A
  - Oblique A becomes frequent (more topical)
  - Construction becomes ergative past tense
Argument structure of passive

- The passive (Carib of Suriname)

  PAT   V       AUX    AGT
  S     PARTICIPLE  S-AUX  [ OBLIQUE ]
  seeri t-owa’ma-Ø  n-eei Baaku ’wa
  Silvia AD-embrace-PRTCP 3-AUX Baaku AGT

  ‘Sylvia was embraced by Baaku’
Argument structure of passive

- The passive (Carib of Suriname)

PAT V AUX AGT

S PARTICIPLE S-AUX [ OBLIQUE ]

Seeri t-owa’ma-Ø n-eei Baaku ’wa

Silvia AD-embrace-PRTCP 3-AUX Baaku AGT

‘Sylvia was embraced by Baaku’

- Patient is Nominative S, unmarked case, controls auxiliary agreement
Argument structure of passive

- The passive (Carib of Suriname)

  PAT          V         AUX       AGT
  S            PARTICIPLE S-AUX [OBLIQUE]

  seeri  t-owa’ma-Ø      n-eei  Baaku ’wa

Silvia AD-embrace-PRTCP 3-AUX Baaku AGT

‘Sylvia was embraced by Baaku’

- Agent is oblique, marked by postposition ’wa
Argument structure before reanalysis

- Patient is subject, has four subject properties
  - Unmarked case
  - Verb agreement
  - Controls coreference with reflexive possessive prefix
  - Controls coreference in conjunction reduction
- Agent is oblique, has no subject properties
Argument structure in Carib (Hoff 1995)

- Patient is subject, has **three** subject properties
  - Unmarked case
  - Verb agreement
  - Controls coreference in conjunction reduction

- Agent is oblique, but has **one** subject property
  - Controls coreference with reflexive possessive prefix
Argument structure in Tiriyó (Gildea 1997)

- Absolutive has **two** subject properties
  - Unmarked case
  - Aux agreement
- Ergative has **two** subject properties
  - Controls coreference with reflexive possessive prefix
  - Controller and target of “while” clause coreference
- Ergative properties are syntactic, so **Ergative is subject**
Another way to look at it...

- Absolutive *inherited* its subject properties *via construction reanalysis*
  - (Lack of) Case-marking
  - Control of verb agreement

- Ergative *gained* its subject properties *via analogical extension*
  - Control of coreference with reflexive prefix
  - Controller & target of ‘while’ clause coreference
Another way to look at it...

- Inherited subject properties reflect the source
- Gained subject properties reflect subconscious changes made by speakers
- If there is a conflict, the gained properties point to the modern system
Action Nominalizations > Ergative main clauses

- Cariban
  - Kapóng, Pemón, Makushi, Kuikúro, Cariña (Gildea 1998)
  - Yawarana, Mapoyo (Mattéi-Muller 2003)
  - Kari’nja (Yamada 2010)
  - Ye’kwana (Cáceres 2011)

- Trumai (isolate; Guirardello 2010; Gildea & Guirardello 2011)
Action Nominalizations >
Ergative main clauses

- Deverbal noun is possessed by its notional absolutive
- Agent is optionally added as an oblique
- This NP is used as a nonverbal predicate (subject is pleonastic pronoun)
- The construction becomes new main clause tense-aspect clause
Ergative Subordinate Clauses

- Action Nzns obligatorily possessed by S/P

\[ \text{P-V-NZR} \quad (A-OBL) \quad \text{S-V-NZR} \]
\[ t-üdü-dü-:ne \quad (kü-wya) \quad i-w-eetümü-dü \]
\[ 3\text{-do-NZR-EMPH} \quad 1\text{PL-AGT} \quad 3\text{-INTR-sing-NZR} \]

‘its doing (by us)’

Ye’kwana data from Cáceres (2011)

\[ [ P ] \quad [ S ] \]

cf. *The destruction of the city*  \quad *The glowing of the city*

Spike Gildea, University of Oregon
Ergative Subordinate Clauses

- Notional A is optional oblique

\[ \text{P-V-NZR} \quad (A-OBL) \quad \text{S-V-NZR} \]
\[ t-üdü-dü-:ne \quad (kü-wya) \quad i-w-ee tümü-dü \]
\[ 3\text{-do-NZR-EMPH} \quad 1\text{PL-AGT} \quad 3\text{-INTR-sing-NZR} \]

‘its doing (by us)’

Ye’kwana data from Cáceres (2011)

Ye’kwana data from Cáceres (2011)

\[ [ \text{P} ] \quad [ \text{S} ] \]

cf. The destruction of the city

The glowing of the city
Nominalizations enter main clauses

- Biclausal constructions > monoclausal
  - With phasal verbs
    \[
    \text{begin, finish} \quad \text{> INCHOATIVE, COMPLETIVE}
    \]
  - With locative predicates
    \[
    \text{be on [going]} \quad \text{> PROGRESSIVE}
    \]
  - As simililatives
    \[
    \text{That is [like [his going]]} \quad \text{> FUTURE}
    \]
  - As factives
    \[
    \text{That is [his going]} \quad \text{> IMPERFECTIVE/PRESENT, PAST}
    \]
Nominalizations enter main clauses

- Biclausal constructions > monoclausal
  - With phasal verbs
    - *begin, finish* > INCHOATIVE, COMPLETIVE
  - With locative predicates
    - *be on [going]* > PROGRESSIVE
  - As similatives
    - *That is [like [his going]]* > FUTURE
  - As factives
    - *That is [his going]* > IMPERFECTIVE/PRESENT, PAST
Nominalizations enter main clauses

- Biclausal constructions > monoclausal
  - With phasal verbs
    - *begin, finish* > INCHOATIVE, COMPLETIVE
  - With locative predicates
    - *be on [going]* > PROGRESSIVE
  - As similatives
    - *That is [like [his going]]* > FUTURE
  - As factives
    - *That is [his going]* > IMPERFECTIVE/PRESENT, PAST
Nominalizations enter main clauses

- Biclausal constructions > monoclausal
  - With phasal verbs
    - \textit{begin, finish} > INCHOATIVE, COMPLETIVE
  - With locative predicates
    - \textit{be on [going]} > PROGRESSIVE
  - As simililatives
    - \textit{That is [like [his going]} > FUTURE
  - As factives
    - \textit{That is [his going]} > IMPERFECTIVE/PRESENT, PAST
Nominalizations enter main clauses

- Biclausal constructions > monoclausal
  - With phasal verbs
    - *begin, finish* > INCHOATIVE, COMPLETIVE
  - With locative predicates
    - *be on [going]* > PROGRESSIVE
  - As simililatives
    - *That is [like [his going]]* > FUTURE
  - As factives
    - *That is [his going]* > IMPERFECTIVE/PRESENT, PAST
Argument structure before reanalysis

- Main clause subject has all subject properties
- Subordinate clause argument structure
  - Absolutive (S/P) possessor of deverbal N
    - Unmarked case
    - Obligatory preverbal position
    - Possessor-Possessed Constituent (NP)
  - Oblique A has no subject properties
Attested changes

- Expression of A becomes obligatory (3 cases)
  - A pronouns cliticize to the verb, creating ergative indexation
- A/S gains control of reflexive possessive prefix (all attested cases)
- Ergative flagging becomes optional (2 cases) or is lost (2 cases)
- \([S/P \ V]vp \ A{-erg} > [abs=V] \ S/A \ P \) (1 case)
Example 1: Akawaio (Fox 2003)

- Former main clause subject disappears
- Nzn becomes new main clause
  - Absolutive (S/P)
    - Unmarked NP (pronominal proclitic)
    - Absolutive-Verb Constituent (VP)
  - Ergative (A)
    - Marked NP (pronominal enclitic)
  - New subject (S/A) properties
    - Control of coreference with reflexive possessive prefix
    - Control of coreference with subject of -'ima ‘while’ clause
Example 2: Makushi (Abott 1990)

- Main clause subject disappears
- New main clause argument structure
  - Absolutive (S/P)
    - Unmarked NP (pronominal proclitic)
    - Absolutive-Verb Constituent (VP)
  - Ergative (A)
    - Marked NP (pronominal enclitic)
  - New subject (S/A) properties
    - Control of coreference with reflexive possessive prefix
    - Innovative Auxiliary agrees with S/A
Example 3: Trumai (Guirardello 2010)

- Main clause subject disappears
- Copula becomes 2\textsuperscript{nd} position focus/tense marker (no agreement)
- New main clause argument structure
  - Absolutive (S/P)
    - Unmarked NP (pronominal proclitic)
    - Absolutive-Verb Constituent (VP)
    - Absolutive to Absolutive raising
  - Ergative (A)
    - Marked NP (pronominal enclitic)
    - Free order
  - No new subject (S/A) properties
    - Maybe postural auxiliaries?
Another way to look at it...

- Absolutive and Ergative *inherited* coding properties (but no subject properties) *via* construction reanalysis
  - FLagging
  - Syntactic relation to verb
  - Absolutive to Absolutive raising (Trumai only)
Another way to look at it...

- In Cariban, S/A gained syntactic subject properties via analogical extension
  - Control of coreference with reflexive prefix
  - Controller of ‘while’ clause coreference
  - Auxiliary agreement
- In Trumai, possibly auxiliary agreement
- Subject can be more and less grammaticalized
- Is it therefore more and less real?
Participant Nominalizations > Focus clauses

- Cariban
  - Panare (Gildea 1998, 2011)
- Trumai (Guirardello 2010; Gildea & Guirardello 2011)
- Movima (Haude 2010; Gildea & Haude 2011; Gildea & Zúñiga 2016)
- Also a parallel in Austronesian languages (Starosta et al 1982; Kaufman 2009, 2011)
Participant Nominalizations > Focus clauses

- Deverbal participant noun is possessed by the other core argument
  - *He is my employer*
    
    SUBJ COP [GEN V-nzr ]
    AGT PAT
  - *He is my employee*
    
    SUBJ COP [GEN V-nzr ]
    AGT PAT
Participant Nominalizations > Focus clauses

- The subject of the nonverbal predicate is the item in FOCUS.
- The semantic role of the subject depends on the form of the nominalization:
  - Object Nominalization: $\text{SUBJ} = \text{AGT}$
  - Subject nominalization: $\text{SUBJ} = \text{PAT}$
- The construction becomes monoclausal:
  - Copula $\rightarrow$ focus marker
  - Symmetrical voice: both constructions are transitive, apparent reversal of GRs
Argument structure before reanalysis

- Main clause subject has all subject properties
- Subordinate clause argument structure
  - The other argument is the possessor of the deverbal N
    - Genitive or unmarked case
    - Obligatory preverbal position
    - Possessor-Possessed Constituent (NP)
  - Internal argument has no subject properties
Attested changes

- Very minimal in the South American cases
- A.FOCUS prefix extends to S (Panare)
- S/A gains agreement with postural auxiliaries (Trumai)
- No A/S properties identified in Movima
Example 1: Panare

- Object nominalizer only occurs with transitive verbs \( \rightarrow \) A focus

\( nəj \quad n\text{-}utu\text{-}jpə \quad mən \quad anə \)

who A\_FOCUS\_-give\_-PERFECT 3\_INAN\_be mother

Who gave (it)(to you), Mother?

(source: ‘Whose given thing is (it), mother?’)
Example 1: Panare

- Focus questions can now occur with intransitive verbs, uniting A & S

\( \text{nej} \quad n-\text{epi}-i \)

who \ S.FOC-come-PAST.INTER \n
‘Who came?’

(no source — *whose coming is it?)
Example 2: Movima

- Order restrictions are relaxed for external argument
  - When not focused, occurs postverbally
- Hierarchy introduced for selection of internal argument
  - $1 > 2 > 3$
  - HUMAN/ANIMATE $>$ INANIMATE
  - HUMAN $A >$ ANIMATE $P$
  - ANIMATE $A \geq$ HUMAN $P$
Example 2: Movima

- **1A → 2P requires etymological object nominalization**
  - *You are my employee → I (PROX) employ (DIRECT) you (OBV)*
  - *I am your employer → *I (OBV) employ (INVERSE) you (PROX)*

- **3A → 2P requires etymological subject nominalization**
  - *He is your employer → He (OBV) employs (INVERSE) you (PROX)*
  - *You are his employee → *He (OBV) employs (DIRECT) you (PROX)*
Example 2: Movima

- There is no grammar that identifies “subject”
- Most grammar privileges the OBVIATIVE
  - Extraction
    - Questions
    - Clefts
    - Relative clauses
  - External to the VP
- All of these properties are inherited
- This is a language with no evidence for a grammaticalized subject
Empirical Conclusions

- Not all subject properties are created equal
- Inherited properties tell us nothing about subject of the new clause type
  - Coding Properties: Flagging, Indexation, Constituency/Order
  - “Transformations” — “Extraction”, Relative clause formation, Passive, etc.
  - Control of coreference (e.g. subject of passive > absolutive with control)
Empirical Conclusions

- Innovative properties reflect the subconscious categories of the speakers:

  - Strong version: Any one innovative subject property is enough.
    - Lacks empirical validity

  - Weaker version: When two innovative subject properties agree, we can identify an innovative subject.
Theoretical Conclusions

- Which subject properties change?
  - Control of coreference
    - With the reflexive possessive prefix
    - With subjectless adverbial clauses (PRO)
  - Auxiliary agreement targets A+S
  - Word order shift (S joins A)
  - Expansion of morphological categories from A > A+S
  - Loss of Ergative flagging (A=S=P)
Theoretical Conclusions

- Why these subject properties?
- Control of coreference = TOPIC
  - AGENT > CONVENTIONALIZED TOPIC
- Word order shift (S joins A)
- Auxiliary agreement targets A+S
  - Innovative AUX agrees with CONVENTIONALIZED TOPIC
  - Some auxiliaries begin as matrix verbs that have an agentive relationship with subordinate verbs
    - Postural: ‘sit working’, ‘stand drinking’, etc.
    - Other: ‘try to V’, ‘begin V-ing’, ‘want to V’, etc.
Theoretical Conclusions

- What do these subject properties tell us?

Frequent TOPIC > SUBJECT

- cf. Greek & Roman tradition
  - What we talk about > grammatical subject
Theoretical Conclusions

- Why the AGENT?
  - As Jóhanna said, agent is instigator of event

AGENT = Frequent TOPIC

cf. Pāṇini’s KARTR

AGENT > SUBJECT
Thank you!


References


References


