
Organizing the People, Places, and Things in the Bible

Sean Boisen, <[\[myfirstname\]@logos.com](mailto:[myfirstname]@logos.com)>

BibleTech:2009, Seattle WA March 27 2009

Presentation URL: <http://semanticbible.org/other/talks/2009/bibletech/BK.html>



Advanced Technology for Eternal Truth

Outline

- Overview
- Modeling Entities in the Biblical Domain
- Agents
- Names
- Places
- Artifacts and Realia
- Data Representation
- Applications

Biblical Semantics

- The Bible is recorded language: meaning wrapped in linguistic signs ([de Saussure](#))
- The earliest recorded question is semantic in nature:

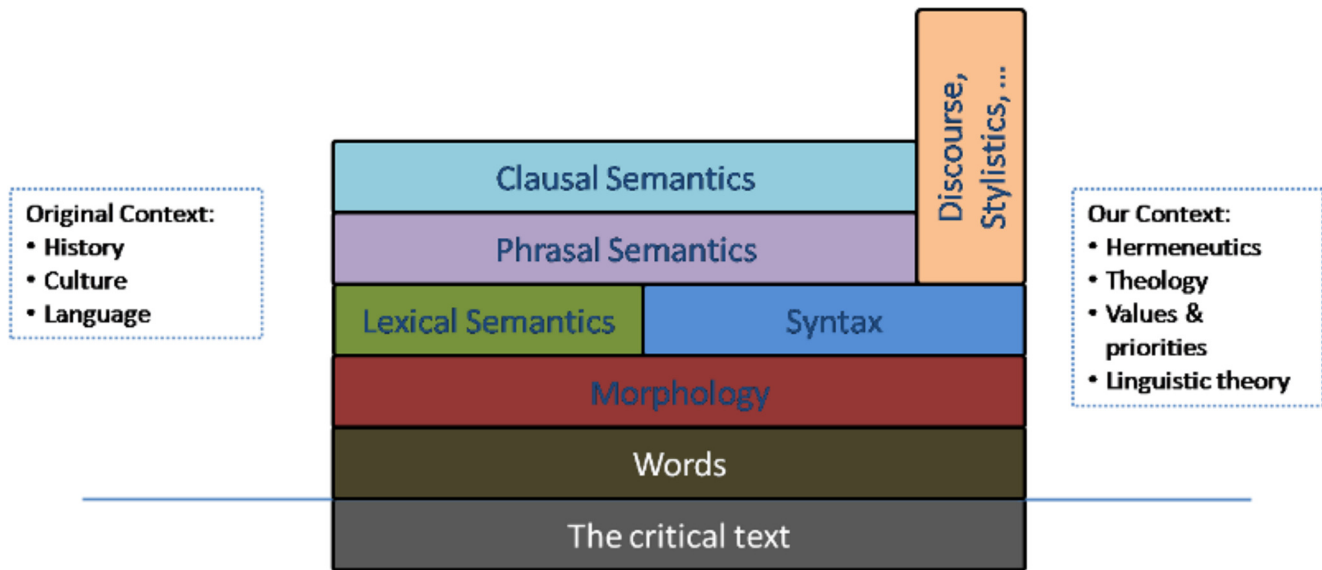
"Did God really say, 'You must not eat from any tree in the garden'?" ([Gen 3.1](#))

- Representable as a query in propositional logic: is it true that

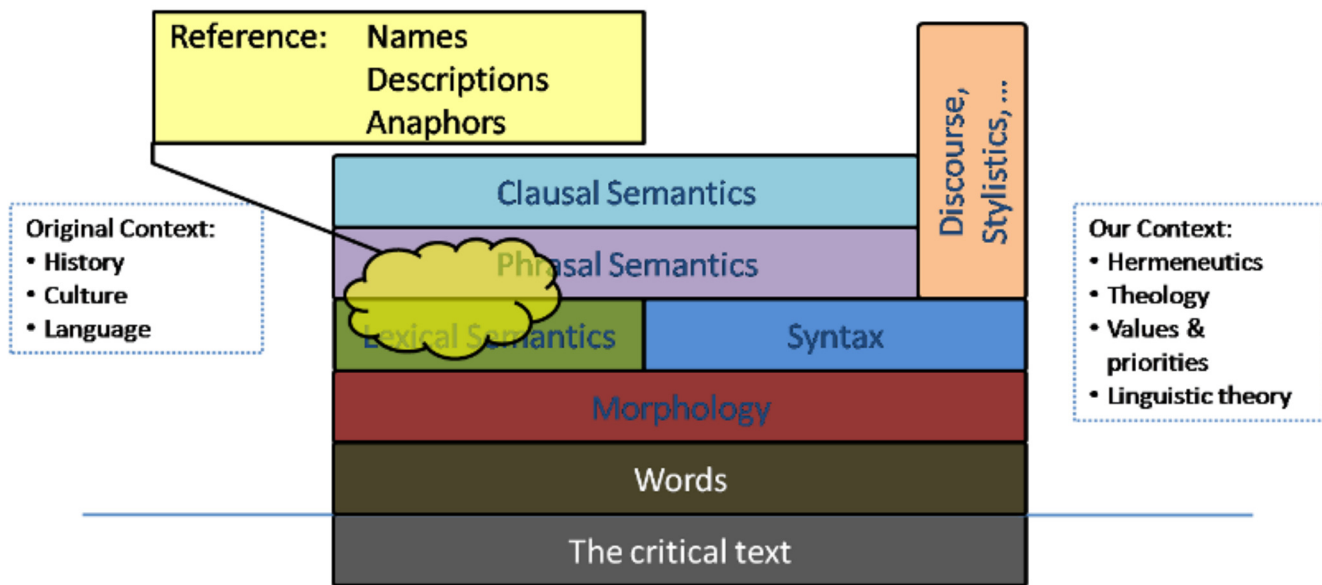
$\exists X: X = \text{"You must not ..."} \ \& \ \text{assert}(\text{God}, X)$

- The truth value of this query = `false`.
... which is irrelevant to the actual intent of God's initial utterance. So you could argue semantics has been both diabolical and irrelevant from the very beginning!

Linguistic Analysis and Biblical Interpretation



Linguistic Analysis and Biblical Interpretation (2)



The Referential Challenge

- Resolve linguistic references to the entities they refer to
 - Provide a consistent identifier (beyond their names)
- Disambiguate entities
 - 30 people named Zechariah
 - Jesus' disciple: Peter, Simon, Simeon, Cephas ...
 - Jebus, Ariel, Oholibah, Salem, Zion, "The Lord is There" ... = Jerusalem
 - Judah: person, tribe, territory

The Goals of the Bible Knowledgebase (BK)

 ... organizing the world's Biblical information to:

- Formalize referential semantics for important entities in the Biblical text
 - Current scope: named agents and places, concrete things
 - Consistent identification, including disambiguation
 - Tie entities back to their references in the text

Beyond Referential Semantics

- Once we've organized this data, we have a platform for going beyond linguistic reference to study aids
Provide a comprehensive set of Biblical references
- Link to source language terms
- Collect together other attributes and relationships
Now we've got a data structure to which other information can be associated
- Capture family/geographic/other context
Including the sources of these assertions in the text
- Provide a framework and metadata for information integration
Very comparable to the use of hyperlinks to Wikipedia pages
- Selective, machine-readable network of data

BK Use Cases

Some content and structure only makes sense in a digital world:

- Link references to entities for exploration and discovery
 - Facilitate Serendipitous Discovery™
- Report the aggregated information about entities
 - Provide convenient cultural and historical background information to readers
- Navigate by entity and data relationships

An important difference between search and navigation: you don't need to know how to identify the target, the link does that for you
- Index and search using semantics
- Provide reusable content

Prose isn't like this
- Visualize data

Modeling Entities in the Biblical Domain

- Ontology: "a specification of a conceptualization" ([Gruber](#))
- Describes concepts, their attributes, and relationships
- Captures just those aspects that are relevant to a particular purpose

Models are limited

Models are not pre-existent or ground truth: we create them for our own purposes

- Represents a compromise between:
 - what's useful (given our purposes)
 - what's knowable
 - what's feasible (given our resources)

What's Modeled by the BK?

- Agents: individuals, groups of people, supernatural beings
- Places: cities, mountains, regions, rivers, roads ...
- Artifacts: specific objects that are relevant to the Biblical story
- Realia or "Things": *classes* of physical, depictable objects
- ... all as described within the Biblical text

Scope is limited: not the ancient near eastern world, not secondary literature on Biblical studies.

General Approach

- Be conservative in what we assert
- Provide references as evidence (provenance)
- Focus on end-user value rather than philosophy
 - Particularly hard to make hard distinctions among squishy values
- Give priority to relationships that increase connectedness ([FSD](#))

Data Sources

- Agents:
 - Logos' [Biblical People](#) data
 - [New Testament Names](#)
- Places:
 - Logos internal data
 - [OpenBible.info](#)
- Realia: UBS Realia volume, Louw-Nida, other specialized resources
- Scholarly resources
- ... and a whole lot of hard work

Core Attributes of BK Models: IDs

- Each BK entity belong to a single *application*, which functions as a namespace.
 - Applications are conventionally associated with a character prefix for conciseness
 - Agents: #
 - Places: @
 - Realia/Artifacts: \$
- Each entity has an ID which is unique within its application
 - Not meant to be human-readable
 - Readable, evocative IDs are more developer-friendly, but ultimately they're arbitrary

IDs, Namespaces, and Naming Conventions

Context	Composition	Function	Examples
App	ID	BK ID (≈XML NCName)	Adam Adam ArkOfNoah agate
BK	App prefix + ID	QID	#Adam @Adam \$ArkOfNoah \$agate
XML	namespace or prefix + ID	URI	http://www.libronix.com/ns/2007/bk-agent#Adam or agent:Adam place:Adam thing:ArkOfNoah thing:agate

Other Common Attributes

- Type (e.g. Man/Woman; Mountain/City/River)
 - Types are application-specific
 - Types can be arranged in an ontology
- Label: disambiguating name for entities whose names overlap or whose identity is otherwise unclear
 - "Eliel (priest)" vs "Eliel (the Mahavite)"
20% of names are used by more than 1 entity
- Description
Hard at both extremes of frequent and infrequent
- References
Half the agents have only a single reference!
- Rank: numeric score estimating importance, based on reference frequency

Agents (individuals, groups, supernatural beings)

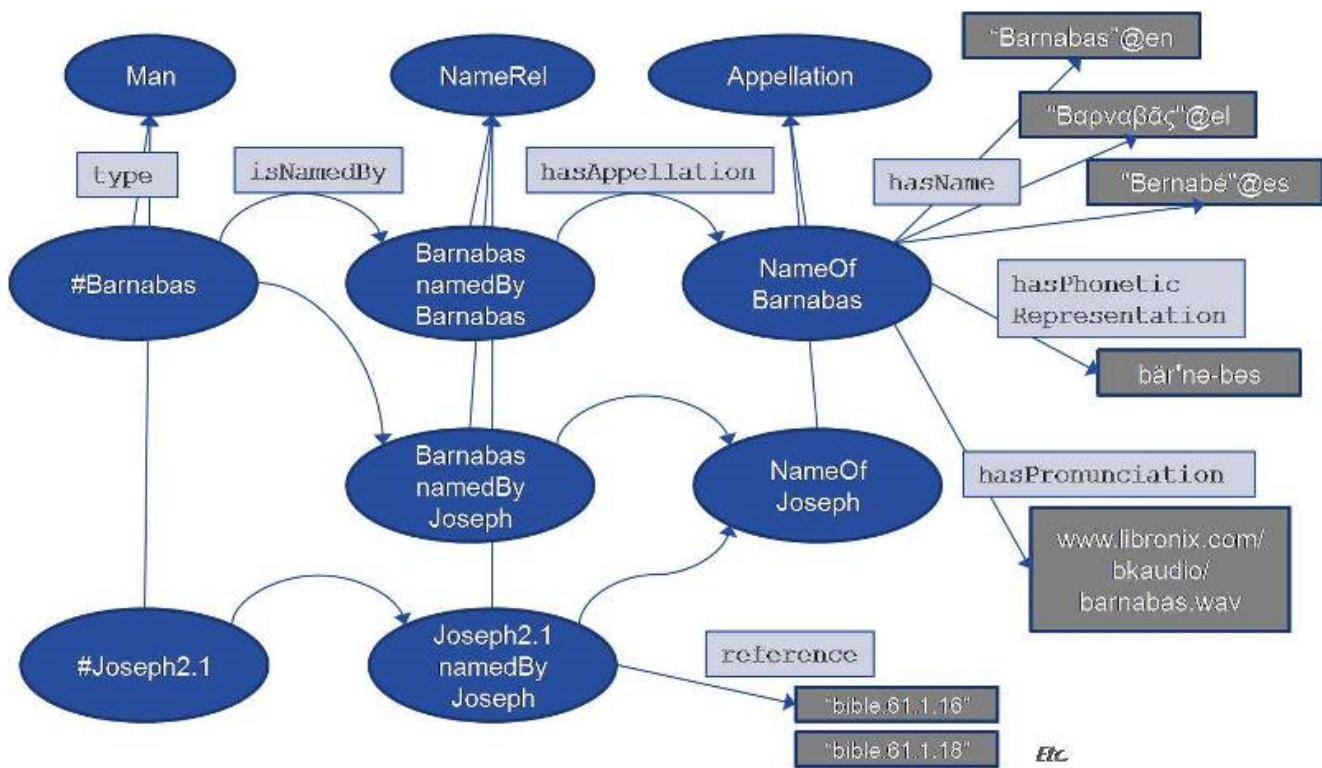
- Entities which have *agency*, that is, they can do things
- Individuals, but also groups of people, and supernatural beings (whether real or supposed)
- Other attributes:
 - Roles
 - Ethnicity
- Human individuals have family relations
- Groups are often related to individuals

Names and Naming Relations

We tend to merge names with what they name

- Names in BK are first-class objects with their own (linguistic) attributes
 - Defined relative to source language lemmas
about 150 cross-testamental Agents
 - Keeps linguistic properties in one place
e.g. pronunciation, literal meaning
- A naming relation pairs entities and names:
 - Separates names from their uses in naming
10% of agents, 25% of places have more than one name
 - References are attached to naming relations, not names themselves
 - If there are multiple naming relations, one is designated as the *preferred* name

Naming Relations Example



Places

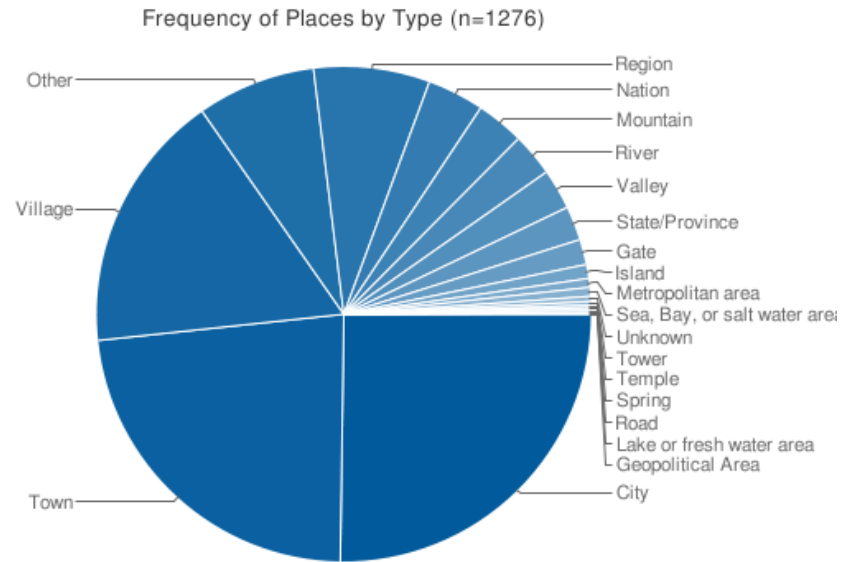
- Fixed, locatable entities in the natural realm with geo-coordinates
- Three primary sub-types, based on geometry:
 - Place (Point):
 - Metropolis, City, Town, Village
 - Mountain, Spring
 - Other, Unknown

At sufficient resolution, all points are really regions, but not in our model. Most problematic for Jerusalem.

- Region (Box):
 - Region, GeopoliticalArea, Nation, State/Province
 - Lake, Sea
 - Island, Valley
- River/Road (Line)

Places (2)

- Geo-coordinates support map display and proximity calculation
- Coarse-grained era information for names
- Sub/super-region relationships



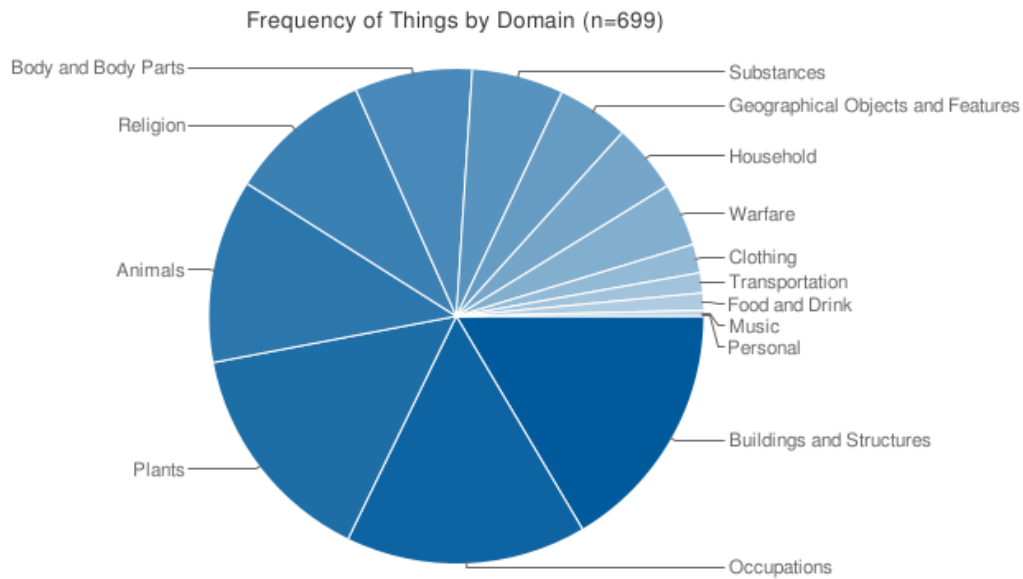
Artifacts

- Physical items with special significance to the Biblical narrative.
- A special subset of Realia
- Examples:
 - Noah's ark
 - the altar built by Abraham in Hebron
 - the crown of thorns
 - Goliath's sword

Realia/Things

- Physical objects capable of visual depiction
- Unlike Agents, Places, and Artifacts, these are not individual instances but *classes* of objects
- References are grouped by lemmas rather than names
- Organized coarsely and non-exclusively into domains
 - above the level of lemmas and Louw-Nida senses, but below the level of Louw-Nida subdomains
- Each Artifact is a type of Thing

Realia/Things (2)



Identity

- Same identifier = same entity: the real world isn't always so neat
- Principles for dealing with ambiguity:
 - sort out and resolve any ambiguity when you can
 - don't create any spurious ambiguity when you can avoid it
 - where's there's genuine ambiguity, *preserve it*

Joining is nearly always easier than splitting

- Possible identity:
 - Assign separate identifiers
 - Link as "possibly same"
- Use scholarly consensus and be conservative

Data Representation

- Original motivations were based on Semantic Web ideas:
 - URIs for Biblical entities
 - Standards-based representations (XML)
 - Formal ontological models (RDF, OWL)
 - See Semantic Technology talk ([PPT](#))
- Benefits:
 - Helpful, well-defined framework for thinking about classes, instances, properties, relationships, etc.
 - Linked-data ideas: a web-friendly URI, generality of RDF triples as a data structure
- Challenges:
 - Current tools better suited to publication-oriented, write-once/read-many model: not so well suited to development and maintenance.



- Python-based web application framework



- MySQL backend with object-relational mapping



- HTML Template system
- Elegant RESTian URL design
- Automatic administrative interface

Applications: Retrieving Data by ... (Live Demos)

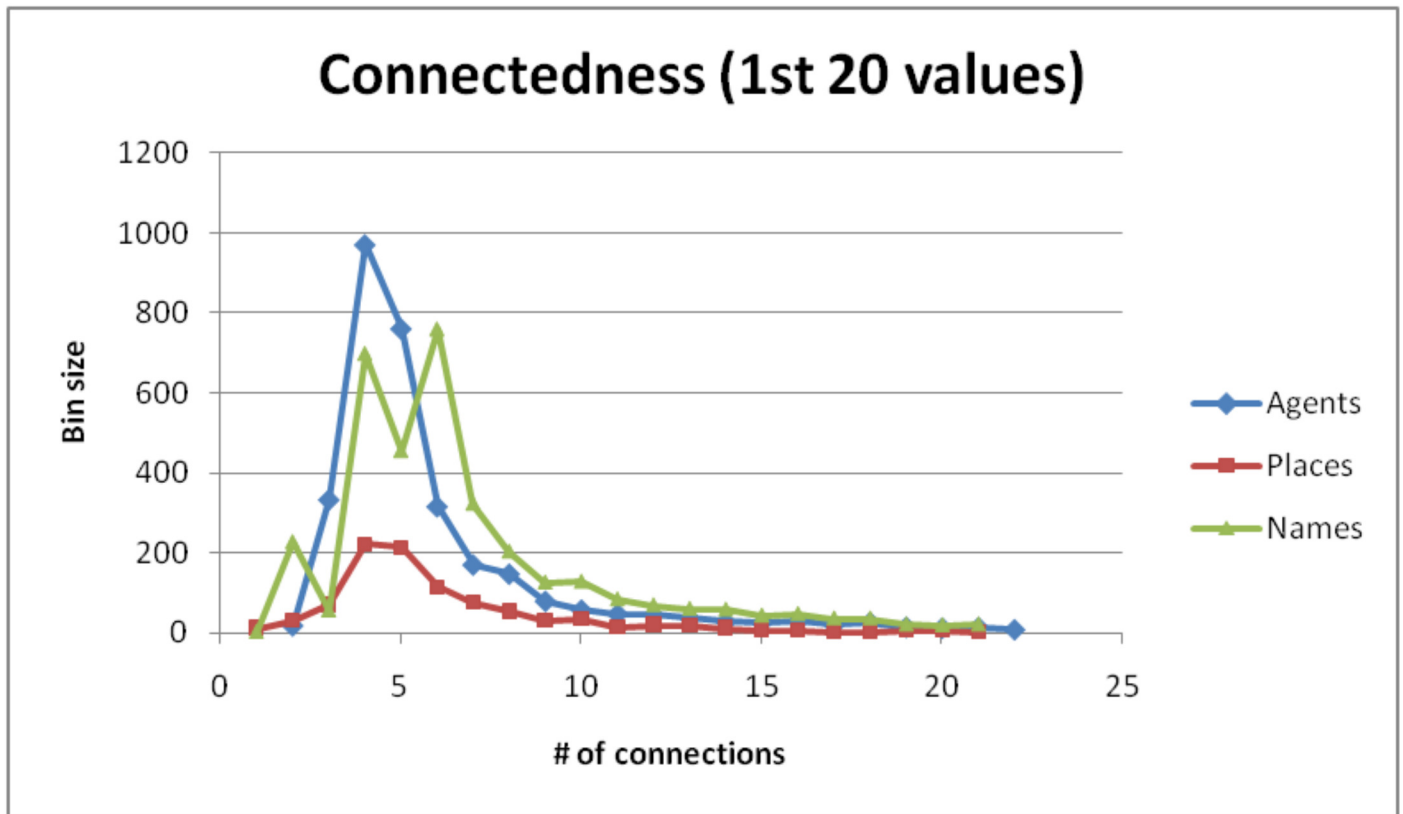
- Given the multi-dimensional representation, you can organize and retrieve entities in a variety of novel ways:
 - Entity by ID ([#Solomon](#), [@Hebron](#), [\\$crown](#))
 - Entities by reference ([Ge 44:12](#))
 - Entities sharing
 - a name ([NameOf_Adonijah](#))
 - a role ([king](#))
 - a type ([EthnicGroup](#), [FreshWaterAreas](#))
 - a lemma ([דָבִלָה](#))
 - Places in proximity ([within 1.5 degrees of Pisidian Antioch](#))
 - Things and Artifacts within a domain ([Warfare](#))

[In the actual presentation, these hyperlinks went to live demonstrations: there's no way to provide that in this web version.]

Applications: Visualizing Connectedness

- Relationships between entities creating a large, interconnected conceptual graph in Bible space
 - Entities as nodes
 - Relationships types as arcs
- You can use these to compute a "connectedness" metric to measure
 - for an individual entity, how connected is it?
 - across the data set (or subsets), what's the aggregate or mean connectedness?
 - the distribution of connectedness across passages
- Relative to our model and the completeness of the data

Connectedness Graph



Hyperlink Annotation

BARNABAS (PERSON) [*Gk Barnabas*

(*Βαρναβας*)]. An **apostle**, an associate of **Paul**, prominent in the church of **Antioch-on-the-Orontes** in **Syria**, and an early leader in the mission to gentiles. According to Acts, his **name was Joseph**, but he was **called Barnabas** by the apostles. **Luke**, the author of Acts, translates **Barnabas** to mean “son of encouragement, ” (from *Aram bar nēbû-â*) but it may simply mean “son of (the god) **Nebo**, ” or something similar. Acts reports that **Barnabas** was a **Levite** whose family came from **Cyprus**; hence he was a Diaspora **Jew**. He is first mentioned as a man who sold some land and donated the proceeds to the apostles in **Jerusalem** (4:36–37). Thus in vivid contrast to **Ananias** and **Sapphira**, who withheld a portion of their property (5:1–11), **Barnabas** is shown

Future Directions

- Other attributes (ethnicity, religious affiliation)
- All mentions (descriptive and pronominal), not just names
- Additional person-person relationships:
 - collaboration/antagonism,
 - service, other kinds of interaction
- Person-place relationships: native, resident, visitor
- Reported speech (with speaker/hearer)
- Narrative Events
- Chronology

Conclusions

- The beginnings of machine-readable semantic metadata for Biblical studies
- Enables users to more easily go from text to knowledge
- A powerful mechanism for information integration

Resources and References

- Boisen, S (2006). [Weaving the New Testament into the Semantic Web](#). Presented Nov 20, 2006 at the Society for Biblical Literature annual meeting.
- Boisen, S (2007). [Deploying Semantic Technologies for Digital Publishing: A Case Study from Logos Bible Software](#). Presented May 24, 2007 at the Semantic Technology Conference.
- Gruber, T.R. (1993). [A translation approach to portable ontologies](#). *Knowledge Acquisition*, 5(2):199-220.

About This Presentation

- This is an XHTML document, structured and presented with [HTML Slidy](#). You can use your browser to view source and see how it works.
- Hyperlinks courtesy of [Logos Bible Software's RefTagger](#)
- The Logos editorial staff who have contributed to the Bible Knowledgebase include:
 - Donna Boisen
 - Elizabeth Borries
 - Molly Crocker
 - SaraAnn Evans
 - Joshua Hou