

# Where do you mean?

Vector Semantics and Linguistic Processing

Cody Kingham

[github.com/codykingham/where\\_do\\_you\\_mean](https://github.com/codykingham/where_do_you_mean)

cak47@cam.ac.uk



UNIVERSITY OF  
CAMBRIDGE

Faculty of Asian and  
Middle Eastern Studies



ST JOHN'S COLLEGE  
UNIVERSITY OF CAMBRIDGE







# Vector Semantics

# Vector Semantics

- Context (location) and meaning are interdependent.

# Vector Semantics

- Context (location) and meaning are interdependent.
- Things that occur in similar places frequently are likely to be similar in meaning.

# Vector Semantics

- Context (location) and meaning are interdependent.
- Things that occur in similar places frequently are likely to be similar in meaning.
- This principle can be applied to lexical meaning.

# Vector Semantics

- Context (location) and meaning are interdependent.
- Things that occur in similar places frequently are likely to be similar in meaning.
- This principle can be applied to lexical meaning.
- distributional hypothesis – Wittgenstein (1953), Firth (1962), Sinclair (1991)

**KWIC:**  
**keyword in context**  
*War of the Worlds*

**“knew”**

/, few people in Woking even  
/for the Spotted Dog, for I  
/myself of what I already  
/man lit both lamps. Happily, I  
/. I do not know what they  
/been in active service and he  
/felt no anxiety about us, as he  
/a village. No one in London  
/, and began running; but I  
/us? (At that time no one  
/out of London, and all that he  
her in my absence. My cousin I  
/-ring. Then a heavy body I  
/going to Leatherhead, though I  
/the Surrey people had fled. I  
/not understand it, and then I  
/of despair, who so far as I  
/, the passing of something I  
/was incomprehensible. All I  
/frantic illuminations; they

**“thought”**

/do the same. No one gave a  
/as sources of human danger, or  
/fall of it, and, at most, have  
/the flash upon Mars. The  
/by Horsell Bridge. The fellow  
/any living creature. I  
/, I walked back, full of such  
/driving across the stars, and  
€ said my wife in terror. I  
but I was not too terrified for  
/cry of astonishment. I saw and  
of/ /communication. This was  
/Hampton Court station, but we  
/, on military details. He  
/fury of battle. No doubt the  
/we stared. Moved by a sudden

knew that the cylinder had opened/  
knew the landlord had a horse and/  
knew, that his wife had gone to/  
knew the road intimately. My wife/  
knew of the things happening/  
knew better than that. He made me/  
knew from the description in the/  
knew positively of the nature of/  
knew it was no good running from a/  
knew what food they needed.) A/  
knew of these Martians and their/  
knew was brave enough for any/ /to  
knew too well what was dragged/  
knew that there I had the poorest/  
knew I wanted to find my wife,/ /  
knew that it must be the red weed/  
knew, shared the city with myself/  
knew not what and then a/  
knew was that these things that/  
knew of it in Dublin, Edinburgh,/

thought to the older worlds of/  
thought of them only to dismiss/  
thought that another meteorite had/  
thought of the confined creature/  
thought he was a lunatic at large/  
thought the unscrewing might be/  
thought, to my home in Maybury./  
thought that nothing more serious/  
thought perplexed. Then I/ /go?  
thought. The terrible Heat-Ray/ /,  
thought nothing of the other four/  
thought to be due to the falling  
thought it was thunder. What the/  
thought of all those silent,/ /  
thought that was uppermost in a/  
thought, I looked northward, and/

**KWIC:**  
**keyword in context**  
*War of the Worlds*

**“knew”**

/, few people in Woking even  
/for the Spotted Dog, for **I**  
/myself of what I already  
/man lit both lamps. Happily, **I**  
/. I do not know what they  
/been in active service and he  
/felt no anxiety about us, as he  
/a village. No one in London  
, and began running; but **I**  
/us? (At that time no one  
/out of London, and all that he  
her in my absence. My cousin **I**  
/-ring. Then a heavy body **I**  
/going to Leatherhead, though **I**  
/the Surrey people had fled. **I**  
/not understand it, and then **I**  
/of despair, who so far as **I**  
, the passing of something **I**  
/was incomprehensible. All **I**  
/frantic illuminations; they

**“thought”**

/do the same. No one gave a  
/as sources of human danger, or  
/fall of it, and, at most, have  
/the flash upon Mars. The fellow  
/by Horsell Bridge. The fellow  
/any living creature. **I**  
/, I walked back, full of such  
/driving across the stars, and  
€ said my wife in terror. **I**  
but I was not too terrified for  
/cry of astonishment. I saw and  
of/ /communication. This was  
/Hampton Court station, but we  
/, on military details. He  
/fury of battle. No doubt the  
/we stared. Moved by a sudden

knew that the cylinder had opened/  
**knew** the landlord had a horse and/  
knew, that his wife had gone to/  
**knew** the road intimately. My wife/  
knew of the things happening/  
knew better than that. He made me/  
knew from the description in the/  
knew positively of the nature of/  
**knew** it was no good running from a/  
knew what food they needed.) A/  
knew of these Martians and their/  
**knew** was brave enough for any/ /to  
**knew** too well what was dragged/  
**knew** that there I had the poorest/  
**knew** I wanted to find my wife,/ /  
**knew** that it must be the red weed/  
**knew**, shared the city with myself/  
**knew** not what and then a/  
**knew** was that these things that/  
knew of it in Dublin, Edinburgh,/

thought to the older worlds of/  
thought of them only to dismiss/  
thought that another meteorite had/  
thought of the confined creature/  
thought he was a lunatic at large/  
**thought** the unscrewing might be/  
thought, to my home in Maybury./  
thought that nothing more serious/  
**thought** perplexed. Then I/ /go?  
thought. The terrible Heat-Ray/ /,  
thought nothing of the other four/  
thought to be due to the falling  
thought it was thunder. What the/  
thought of all those silent,/ /  
thought that was uppermost in a/  
thought, I looked northward, and/

**KWIC:**  
**keyword in context**  
*War of the Worlds*

**“knew”**

**“thought”**

/, few people in Woking even  
/for the Spotted Dog, for I  
/myself of what I already  
/man lit both lamps. Happily, I  
/. I do not know what they  
/been in active service and he  
/felt no anxiety about us, as he  
/a village. No one in London  
, and began running; but I  
/us? (At that time no one  
/out of London, and all that he  
her in my absence. My cousin I  
/-ring. Then a heavy body I  
/going to Leatherhead, though I  
/the Surrey people had fled. I  
/not understand it, and then I  
/of despair, who so far as I  
/, the passing of something I  
/was incomprehensible. All I  
/frantic illuminations; they

/do the same. No one gave a  
/as sources of human danger, or  
/fall of it, and, at most, have  
/the flash upon Mars. The  
/by Horsell Bridge. The fellow  
/any living creature. I  
/, I walked back, full of such  
/driving across the stars, and  
€ said my wife in terror. I  
but I was not too terrified for  
/cry of astonishment. I saw and  
of/ /communication. This was  
/Hampton Court station, but we  
/, on military details. He  
/fury of battle. No doubt the  
/we stared. Moved by a sudden

**knew that** the cylinder had opened/  
knew the landlord had a horse and/  
**knew, that** his wife had gone to/  
knew the road intimately. My wife/  
knew of the things happening/  
knew better than that. He made me/  
knew from the description in the/  
knew positively of the nature of/  
knew it was no good running from a/  
knew what food they needed.) A/  
knew of these Martians and their/  
knew was brave enough for any/ /to  
knew too well what was dragged/  
**knew that** there I had the poorest/  
knew I wanted to find my wife,/

**knew that** it must be the red weed/  
knew, shared the city with myself/  
knew not what and then a/

knew was that these things that/  
knew of it in Dublin, Edinburgh,/

thought to the older worlds of/  
thought of them only to dismiss/  
**thought that** another meteorite had/  
thought of the confined creature/  
thought he was a lunatic at large/  
thought the unscrewing might be/  
thought, to my home in Maybury./  
**thought that** nothing more serious/  
thought perplexed. Then I/ /go?  
thought. The terrible Heat-Ray/ /,  
thought nothing of the other four/  
thought to be due to the falling  
thought it was thunder. What the/  
thought of all those silent,/

**thought that** was uppermost in a/  
thought, I looked northward, and/

**KWIC:**  
**keyword in context**  
*War of the Worlds*

/, few people in Woking even  
/for the Spotted Dog, for I  
/myself of what I already  
/man lit both lamps. Happily, I  
/. I do not know what they  
/been in active service and he  
/felt no anxiety about us, as he  
/a village. No one in London  
/, and began running; but I  
/us? (At that time no one  
/out of London, and all that he  
her in my absence. My cousin I  
/-ring. Then a heavy body I  
/going to Leatherhead, though I  
/the Surrey people had fled. I  
/not understand it, and then I  
/of despair, who so far as I  
/, the passing of something I  
/was incomprehensible. All I  
/frantic illuminations; they

/do the same. No one gave a  
/as sources of human danger, or  
/fall of it, and, at most, have  
/the flash upon Mars. The  
/by Horsell Bridge. The fellow  
/any living creature. I  
/, I walked back, full of such  
/driving across the stars, and  
€ said my wife in terror. I  
but I was not too terrified for  
/cry of astonishment. I saw and  
of/ /communication. This was  
/Hampton Court station, but we  
/, on military details. He  
/fury of battle. No doubt the  
/we stared. Moved by a sudden

knew that the cylinder had opened/  
knew the landlord had a horse and/  
knew, that his wife had gone to/  
knew the road intimately. My wife/  
**knew of** the things happening/  
knew better than that. He made me/  
knew from the description in the/  
knew positively of the nature of/  
knew it was no good running from a/  
knew what food they needed.) A/  
**knew of** these Martians and their/  
knew was brave enough for any/ /to  
knew too well what was dragged/  
knew that there I had the poorest/  
knew I wanted to find my wife,/br/>knew that it must be the red weed/  
knew, shared the city with myself/  
knew not what and then a/  
knew was that these things that/  
**knew of** it in Dublin, Edinburgh,/

thought to the older worlds of/  
**thought of** them only to dismiss/  
thought that another meteorite had/  
**thought of** the confined creature/  
thought he was a lunatic at large/  
thought the unscrewing might be/  
thought, to my home in Maybury./  
thought that nothing more serious/  
thought perplexed. Then I/ /go?  
thought. The terrible Heat-Ray/ /,  
thought nothing of the other four/  
thought to be due to the falling  
thought it was thunder. What the/  
**thought of** all those silent,/br/>thought that was uppermost in a/  
thought, I looked northward, and/

**KWIC:**  
**keyword in context**  
*War of the Worlds*

**“knew”**

/, few people in Woking even  
/for the Spotted Dog, for I  
/myself of what I already  
/man lit both lamps. Happily, I  
/. I do not know what they  
/been in active service and he  
/felt no anxiety about us, as he  
/a village. No one in London  
, and began running; but I  
/us? (At that time no one  
/out of London, and all that he  
her in my absence. My cousin I  
/-ring. Then a heavy body I  
/going to Leatherhead, though I  
/the Surrey people had fled. I  
/not understand it, and then I  
/of despair, who so far as I  
/, the passing of something I  
/was incomprehensible. All I  
/frantic illuminations; they

**“thought”**

/do the same. No one gave a  
/as sources of human danger, or  
/fall of it, and, at most, have  
/the flash upon Mars. The  
/by Horsell Bridge. The fellow  
/any living creature. I  
/, I walked back, full of such  
/driving across the stars, and  
€ said my wife in terror. I  
but I was not too terrified for  
/cry of astonishment. I saw and  
of/ /communication. This was  
/Hampton Court station, but we  
/, on military details. He  
/fury of battle. No doubt the  
/we stared. Moved by a sudden

knew that the cylinder had opened/  
**knew the** landlord had a horse and/  
knew, that his wife had gone to/  
**knew the** road intimately. My wife/  
knew of the things happening/  
knew better than that. He made me/  
knew from the description in the/  
knew positively of the nature of/  
knew it was no good running from a/  
knew what food they needed.) A/  
knew of these Martians and their/  
knew was brave enough for any/ /to  
knew too well what was dragged/  
knew that there I had the poorest/  
knew I wanted to find my wife,/ /  
knew that it must be the red weed/  
knew, shared the city with myself/  
knew not what and then a/  
knew was that these things that/  
knew of it in Dublin, Edinburgh,/

thought to the older worlds of/  
thought of them only to dismiss/  
thought that another meteorite had/  
thought of the confined creature/  
thought he was a lunatic at large/  
**thought the** unscrewing might be/  
thought, to my home in Maybury./  
thought that nothing more serious/  
thought perplexed. Then I/ /go?  
thought. The terrible Heat-Ray/ /,  
thought nothing of the other four/  
thought to be due to the falling  
thought it was thunder. What the/  
thought of all those silent,/ /  
thought that was uppermost in a/  
thought, I looked northward, and/

**KWIC:**  
**keyword in context**  
*War of the Worlds*

**“knew”**

/, few people in Woking even  
/for the Spotted Dog, for I  
/myself of what I already  
/man lit both lamps. Happily, I  
/. I do not know what they  
/been in active service and he  
/felt no anxiety about us, as he  
/a village. No one in London  
/, and began running; but I  
/us? (At that time no one  
/out of London, and all that he  
her in my absence. My cousin I  
/-ring. Then a heavy body I  
/going to Leatherhead, though I  
/the Surrey people had fled. I  
/not understand it, and then I  
/of despair, who so far as I  
/, the passing of something I  
/was incomprehensible. All I  
/frantic illuminations; they

knew that the cylinder had opened/  
knew the landlord had a horse and/  
knew, that his wife had gone to/  
knew the road intimately. My wife/  
knew of the things happening/  
knew better than that. He made me/  
knew from the description in the/  
knew positively of the nature of/  
**knew it** was no good running from a/  
knew what food they needed.) A/  
knew of these Martians and their/  
knew was brave enough for any/ /to  
knew too well what was dragged/  
knew that there I had the poorest/  
knew I wanted to find my wife,/ /  
knew that it must be the red weed/  
knew, shared the city with myself/  
knew not what and then a/  
knew was that these things that/  
knew of it in Dublin, Edinburgh,/

**“thought”**

/do the same. No one gave a  
/as sources of human danger, or  
/fall of it, and, at most, have  
/the flash upon Mars. The  
/by Horsell Bridge. The fellow  
/any living creature. I  
/, I walked back, full of such  
/driving across the stars, and  
€ said my wife in terror. I  
but I was not too terrified for  
/cry of astonishment. I saw and  
of/ /communication. This was  
/Hampton Court station, but we  
/, on military details. He  
/fury of battle. No doubt the  
/we stared. Moved by a sudden

thought to the older worlds of/  
thought of them only to dismiss/  
thought that another meteorite had/  
thought of the confined creature/  
thought he was a lunatic at large/  
thought the unscrewing might be/  
thought, to my home in Maybury./ /  
thought that nothing more serious/  
thought perplexed. Then I/ /go?  
thought. The terrible Heat-Ray/ /,  
thought nothing of the other four/  
thought to be due to the falling  
**thought it** was thunder. What the/  
thought of all those silent,/ /  
thought that was uppermost in a/  
thought, I looked northward, and/

# Jupyter Notebook

**meat\_bread\_gold.ipynb**

# Vector Semantics

# Vector Semantics

- measures word behavior within a sample of text data
  - measurement limited by sample size

# Vector Semantics

- measures word behavior within a sample of text data
  - measurement limited by sample size
- Word behavior is a byproduct of human behavior:  
**category making**

# Vector Semantics

- measures word behavior within a sample of text data
  - measurement limited by sample size
- Word behavior is a byproduct of human behavior:  
**category making**
- cognitive semantics & construction grammar — language forms are patterns linked to meaning (Goldberg 1995)

# Vector Semantics

- measures word behavior within a sample of text data
  - measurement limited by sample size
- Word behavior is a byproduct of human behavior:  
**category making**
- cognitive semantics & construction grammar — language forms are patterns linked to meaning (Goldberg 1995)
- Vector semantics is a way to empirically study the interaction of categories in the mind.

# Recipe for a Semantic Vector Space



# Recipe for a Semantic Vector Space

- text corpus



# Recipe for a Semantic Vector Space

- text corpus
- sample of target words



# Recipe for a Semantic Vector Space

- text corpus
- sample of target words
- strategy for summarizing a target word's context
  - n-grams, “bag of words”
  - surface forms, lexemes, syntactic roles
  - e.g. **The dog ran** > context = “the” “run”



# Recipe for a Semantic Vector Space

- text corpus
- sample of target words
- strategy for summarizing a target word's context
  - n-grams, “bag of words”
  - surface forms, lexemes, syntactic roles
  - e.g. **The dog ran** > context = “the” “run”
- method of normalizing raw context frequencies



# Recipe for a Semantic Vector Space

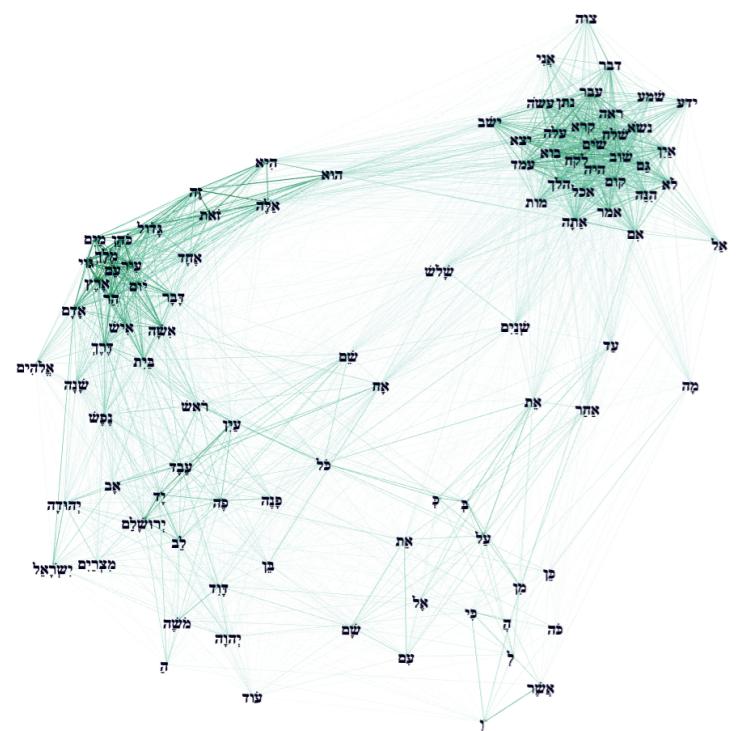
- text corpus
- sample of target words
- strategy for summarizing a target word's context
  - n-grams, “bag of words”
  - surface forms, lexemes, syntactic roles
  - e.g. **The dog ran** > context = “the” “run”
- method of normalizing raw context frequencies
- method of analysis: e.g. distance formulae, clustering methods





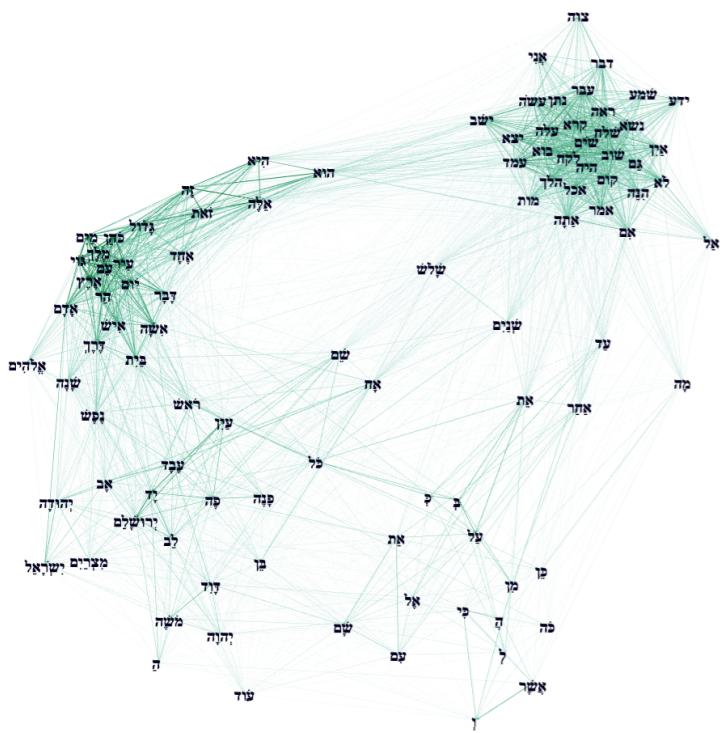


## Parts of Speech Model

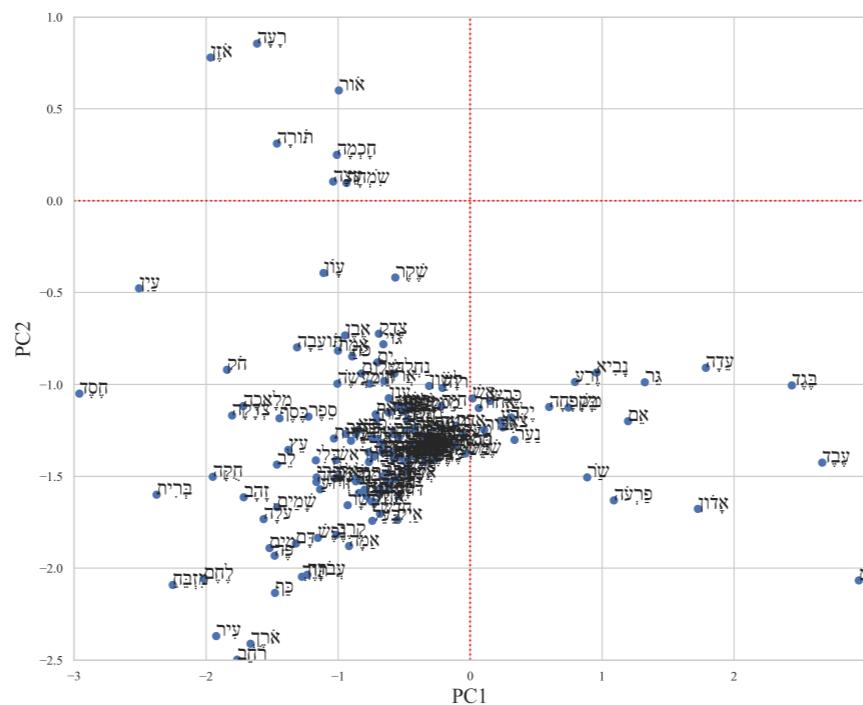




# Parts of Speech Model

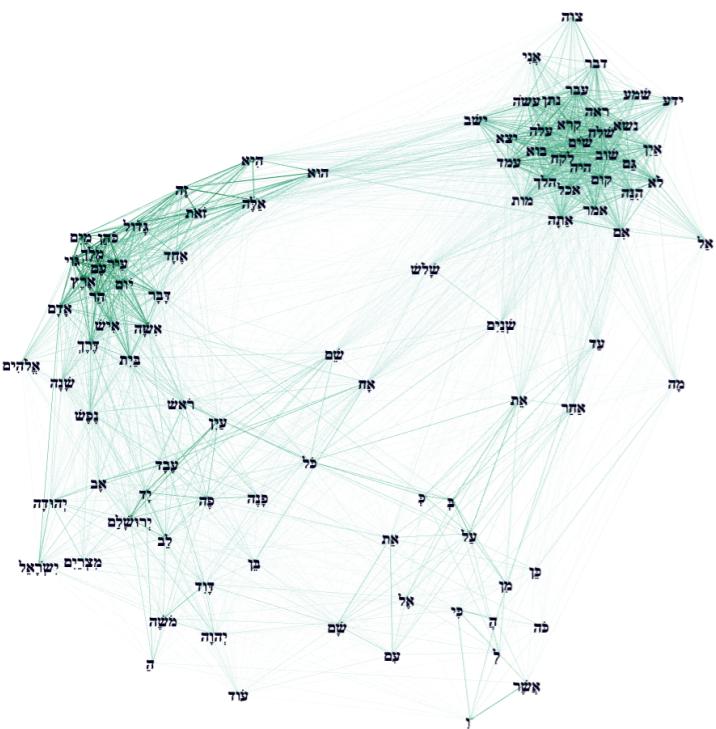


# Noun Classes Model

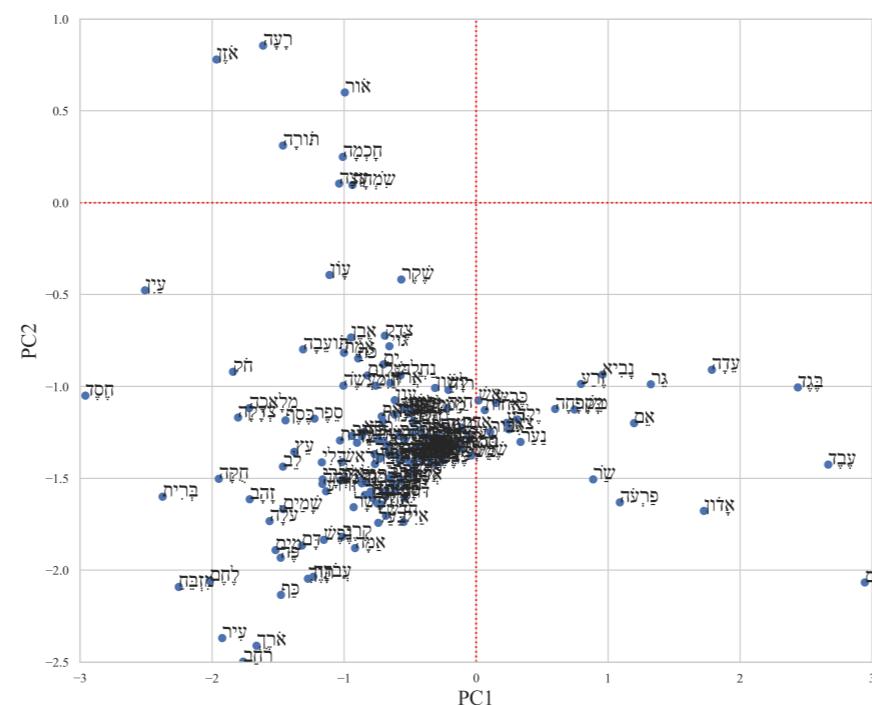




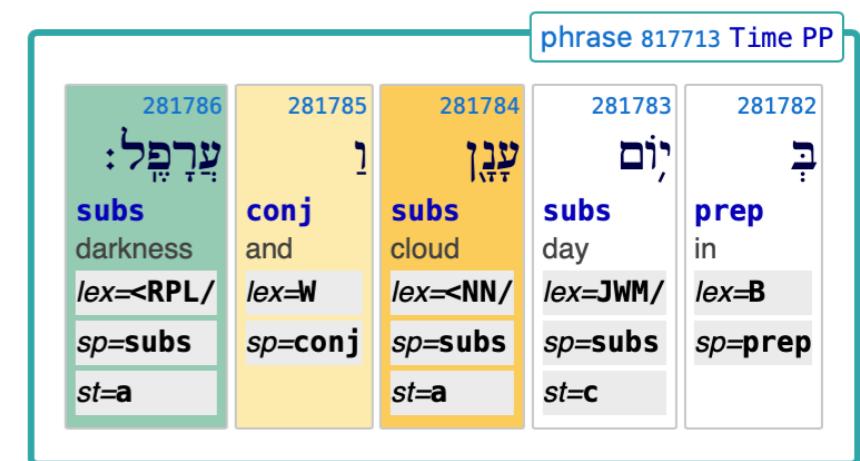
# Parts of Speech Model



# Noun Classes Model



# Coordinate Parsing Model



# Parts of Speech Model



# Parts of Speech Model



# Parts of Speech Model

- **corpus:** Hebrew Bible, ETCBC's BHSA



# Parts of Speech Model

- **corpus:** Hebrew Bible, ETCBC's BHSA
- **sample:** top 100 most distributed lexemes (DP,  
Gries 2008)



# Parts of Speech Model

- **corpus:** Hebrew Bible, ETCBC's BHSA
- **sample:** top 100 most distributed lexemes (DP, Gries 2008)



- **context:** n-gram, 3 word window in sentence



# Parts of Speech Model

- **corpus:** Hebrew Bible, ETCBC's BHSA
- **sample:** top 100 most distributed lexemes (DP, Gries 2008)
- **context:** n-gram, 3 word window in sentence
- **normalization:**  $\Delta P$  – unidirectional measure of statistical association (Ellis 2006)
- 

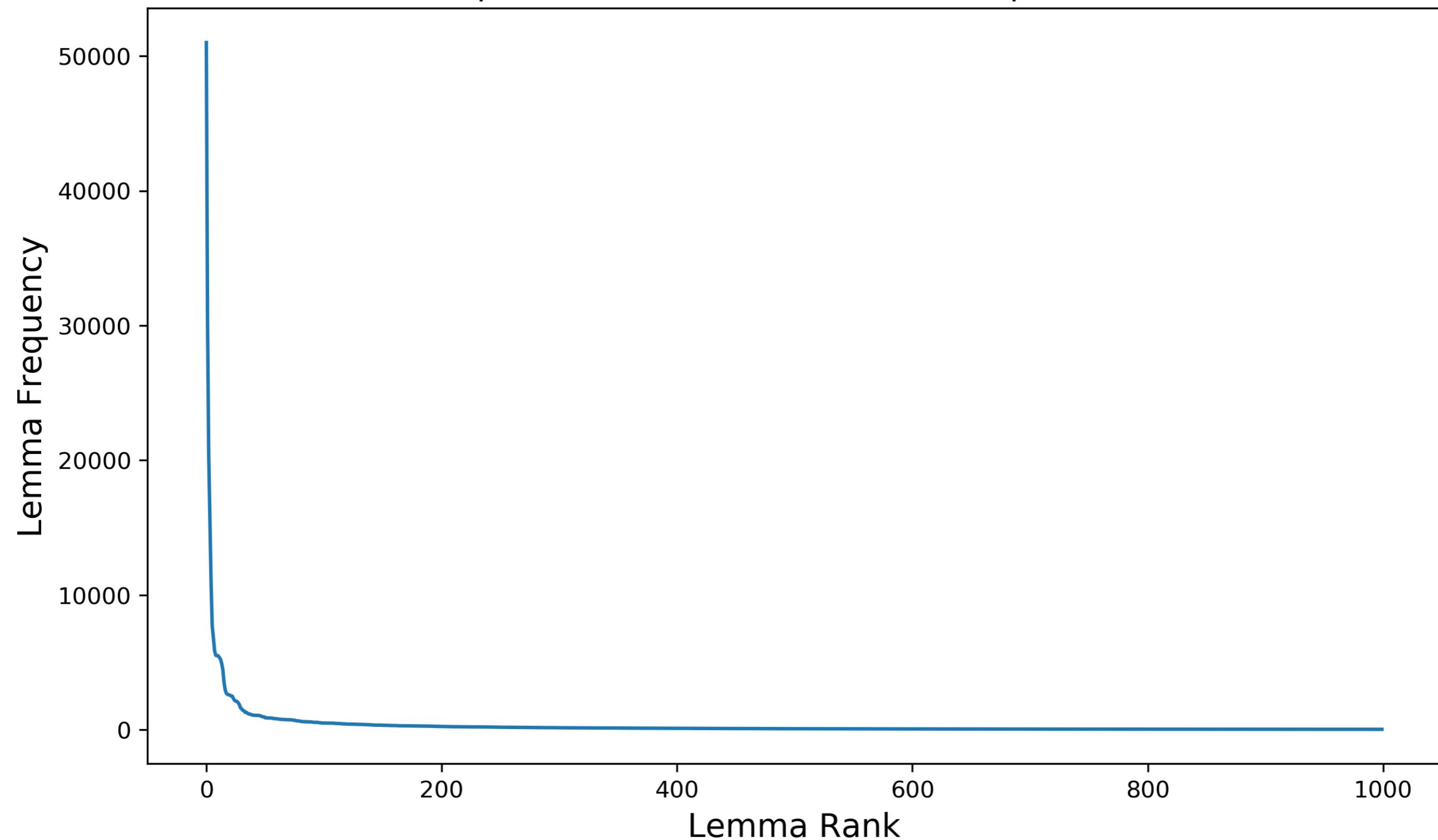


# Parts of Speech Model

- **corpus:** Hebrew Bible, ETCBC's BHSA
- **sample:** top 100 most distributed lexemes (DP, Gries 2008)
- **context:** n-gram, 3 word window in sentence
- **normalization:**  $\Delta P$  – unidirectional measure of statistical association (Ellis 2006)
- **analysis:** cosine similarity + Gephi & ForceAtlas2



## Zipfian Curve in Hebrew Word Frequencies



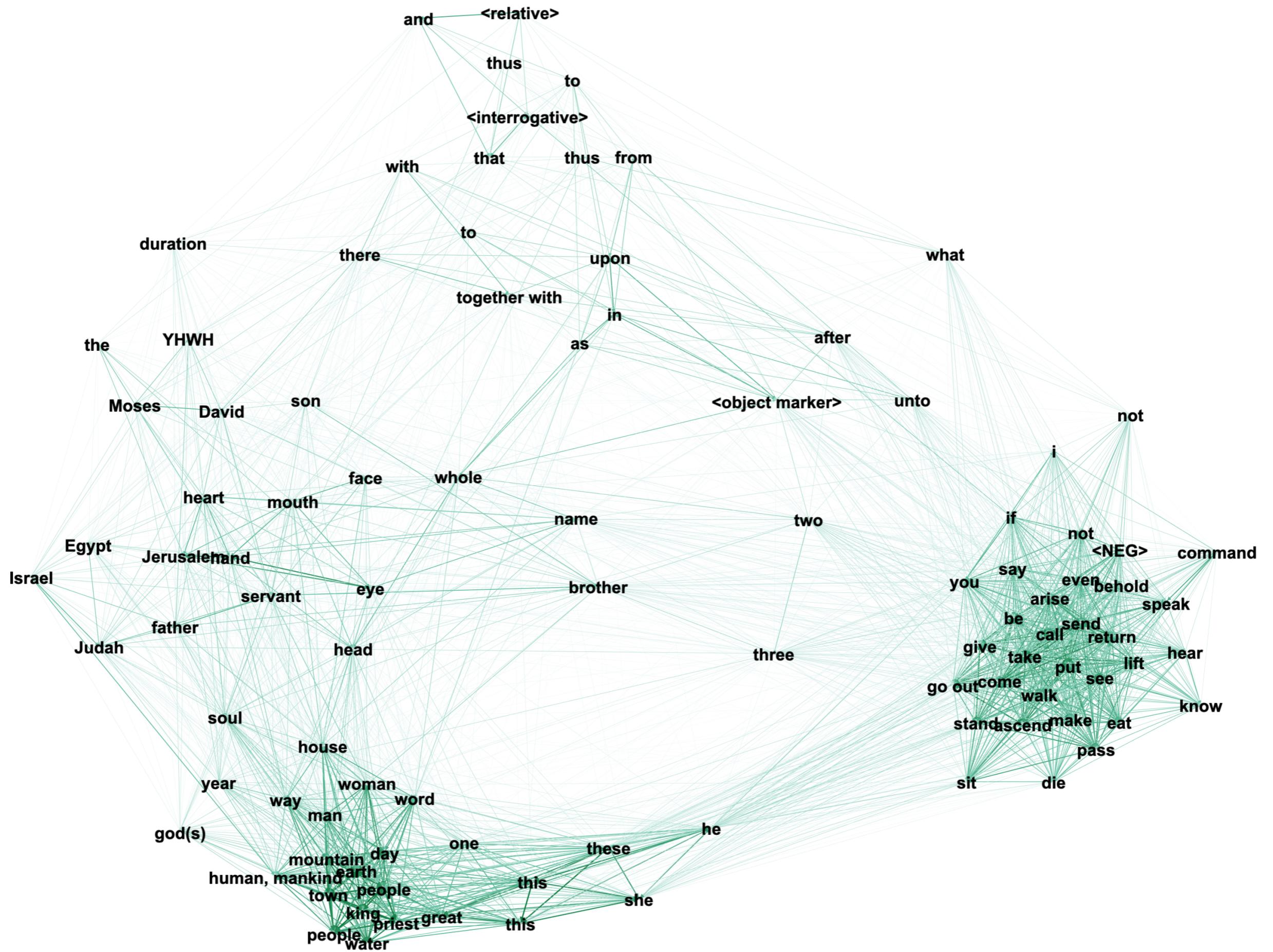
full dimensions: (100, 41204)

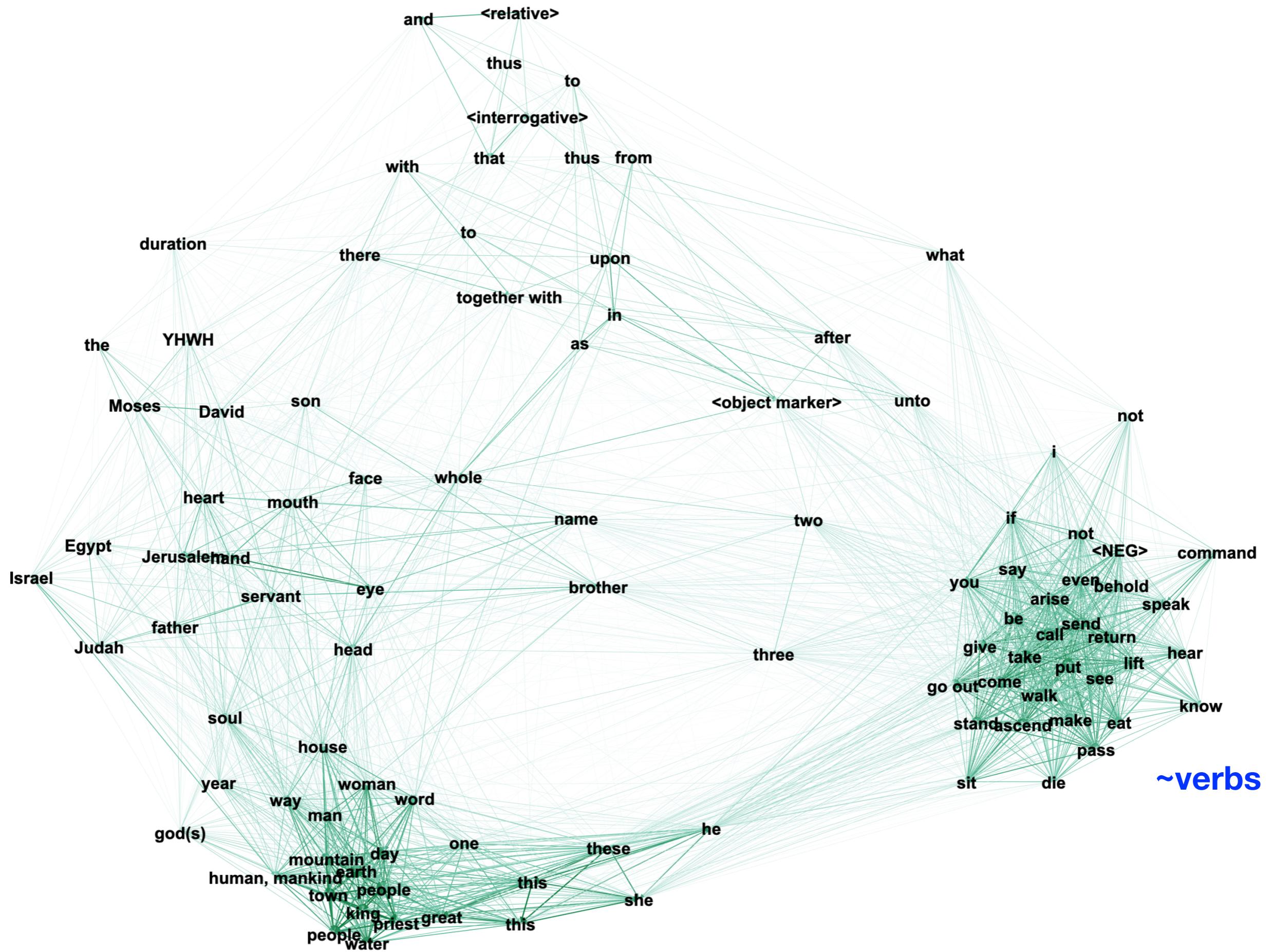
	<b>-3.&gt;T</b>	<b>-2.H</b>	<b>-1.&gt;RY/</b>	<b>1.R&gt;CJT/</b>	<b>2.BR&gt;[</b>	<b>3.&gt;LHJM/</b>	<b>-3.MJM/</b>	<b>-2.HJH[</b>
<b>B</b>	-0.001036	0.002353	-0.000477	0.000049	0.000029	0.000127	0.000076	-0.000116
<b>&gt;LHJM/</b>	0.001860	-0.016776	-0.001043	-0.000025	-0.000021	-0.000085	-0.000004	-0.000455
<b>&gt;T</b>	0.012185	-0.007396	-0.000730	0.000026	-0.000021	0.000444	0.000051	-0.001489
<b>H</b>	0.000371	0.004738	-0.000217	-0.000021	-0.000023	-0.000861	0.000142	0.000298
<b>W</b>	0.000662	0.017336	0.000734	-0.000004	-0.000008	-0.000143	-0.000088	-0.001180
<b>&gt;RY/</b>	0.003036	-0.016468	-0.000686	0.000062	0.000066	0.002131	0.000443	0.000134
<b>HJH[</b>	-0.003413	-0.009327	-0.000528	0.000080	-0.000021	0.001301	0.000011	0.011531
<b>&lt;L</b>	-0.000648	0.001638	-0.000588	-0.000025	-0.000021	-0.000452	0.000313	-0.000576
<b>PNH/</b>	-0.001612	-0.016103	-0.001107	-0.000025	-0.000021	-0.000213	0.000047	0.001090
<b>MJM/</b>	0.000211	-0.011837	-0.001200	0.000346	-0.000021	-0.000532	0.010140	0.000176

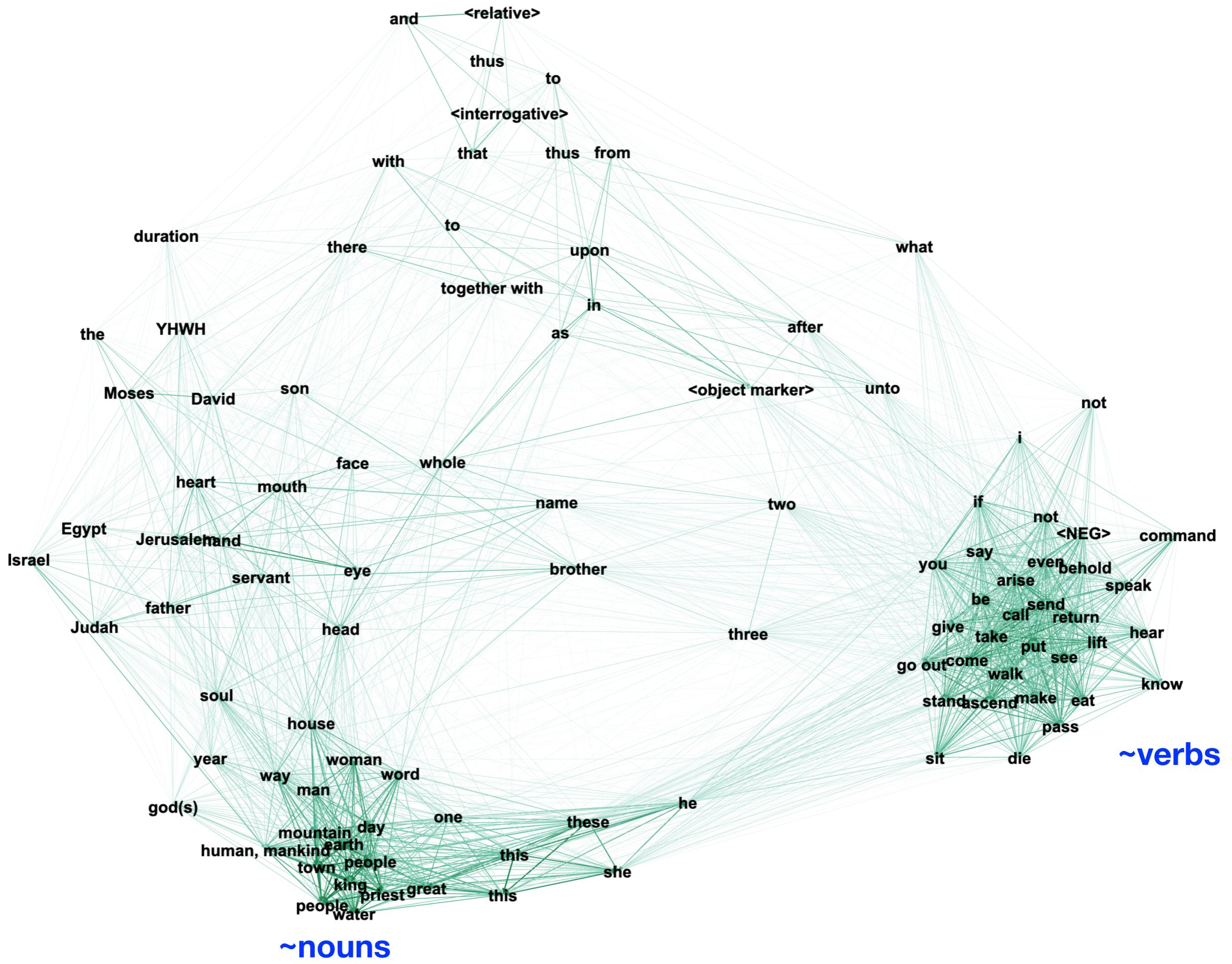
	B	>LHJM/	>T	H	W	>RY/	HJH[	<L	PNH/
B	1.000000	-0.136176	0.550530	-0.274051	-0.091999	-0.208063	-0.185976	0.626086	-0.029701
>LHJM/	-0.136176	1.000000	-0.171587	-0.129920	-0.002669	0.300991	-0.153233	-0.158390	0.047568
>T	0.550530	-0.171587	1.000000	-0.272447	-0.376742	-0.266840	0.165530	0.570215	-0.079837
H	-0.274051	-0.129920	-0.272447	1.000000	0.112907	-0.069290	-0.288102	-0.187305	0.091368
W	-0.091999	-0.002669	-0.376742	0.112907	1.000000	-0.189489	-0.548075	-0.095528	-0.198414
>RY/	-0.208063	0.300991	-0.266840	-0.069290	-0.189489	1.000000	-0.204252	-0.241469	-0.027057
HJH[	-0.185976	-0.153233	0.165530	-0.288102	-0.548075	-0.204252	1.000000	-0.125302	-0.105253
<L	0.626086	-0.158390	0.570215	-0.187305	-0.095528	-0.241469	-0.125302	1.000000	-0.047564
PNH/	-0.029701	0.047568	-0.079837	0.091368	-0.198414	-0.027057	-0.105253	-0.047564	1.000000
MJM/	-0.125372	0.252663	-0.159008	-0.166524	-0.190084	0.827957	-0.155141	-0.149116	-0.001856
>MR[	-0.220878	-0.104413	0.106645	-0.242685	-0.514283	-0.170039	0.664027	-0.218460	0.140998
R>H[	-0.128572	-0.145078	0.202005	-0.222007	-0.465541	-0.194255	0.654897	-0.061707	0.004395
KJ	-0.092597	0.177883	-0.385489	0.008535	0.456054	-0.099231	-0.247210	-0.149167	-0.054475
QR>[	-0.248269	-0.073924	0.047789	-0.279372	-0.471859	-0.152312	0.820903	-0.196548	-0.085214
L	0.324513	-0.086371	0.183148	-0.202406	0.275984	-0.302039	-0.214334	0.335124	-0.109457

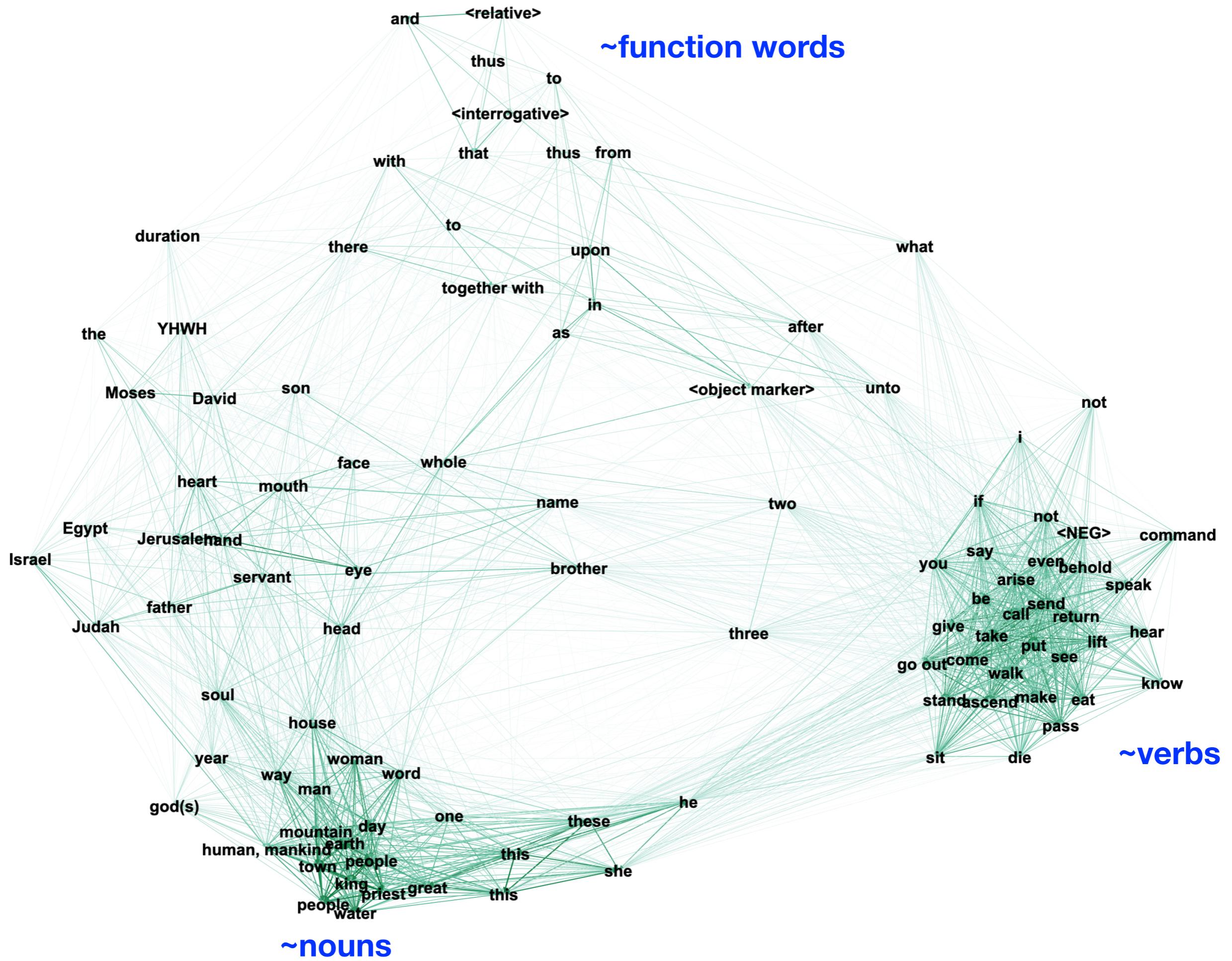
	B	>LHJM/	>T	H	W	>RY/	HJH[	<L	PNH/
B	1.000000	-0.136176	0.550530	-0.274051	-0.091999	-0.208063	-0.185976	0.626086	-0.029701
>LHJM/	-0.136176	1.000000	-0.171587	-0.129920	-0.002669	0.300991	-0.153233	-0.158390	0.047568
>T	0.550530	-0.171587	1.000000	-0.272447	-0.376742	-0.266840	0.165530	0.570215	-0.079837
H	-0.274051	-0.129920	-0.272447	1.000000	0.112907	-0.069290	-0.288102	-0.187305	0.091368
W	-0.091999	-0.002669	-0.376742	0.112907	1.000000	-0.189489	-0.548075	-0.095528	-0.198414
>RY/	-0.208063	0.300991	-0.266840	-0.069290	-0.189489	1.000000	-0.204252	-0.241469	-0.027057
HJH[	-0.185976	-0.153233	0.165530	-0.288102	-0.548075	-0.204252	1.000000	-0.125302	-0.105253
<L	0.626086	-0.158390	0.570215	-0.187305	-0.095528	-0.241469	-0.125302	1.000000	-0.047564
PNH/	-0.029701	0.047568	-0.079837	0.091368	-0.198414	-0.027057	-0.105253	-0.047564	1.000000
MJM/	-0.125372	0.252663	-0.159008	-0.166524	-0.190084	0.827957	-0.155141	-0.149116	-0.001856
>MR[	-0.220878	-0.104413	0.106645	-0.242685	-0.514283	-0.170039	0.664027	-0.218460	0.140998
R>H[	-0.128572	-0.145078	0.202005	-0.222007	-0.465541	-0.194255	0.654897	-0.061707	0.004395
KJ	-0.092597	0.177883	-0.385489	0.008535	0.456054	-0.099231	-0.247210	-0.149167	-0.054475
QR>[	-0.248269	-0.073924	0.047789	-0.279372	-0.471859	-0.152312	0.820903	-0.196548	-0.085214
L	0.324513	-0.086371	0.183148	-0.202406	0.275984	-0.302039	-0.214334	0.335124	-0.109457

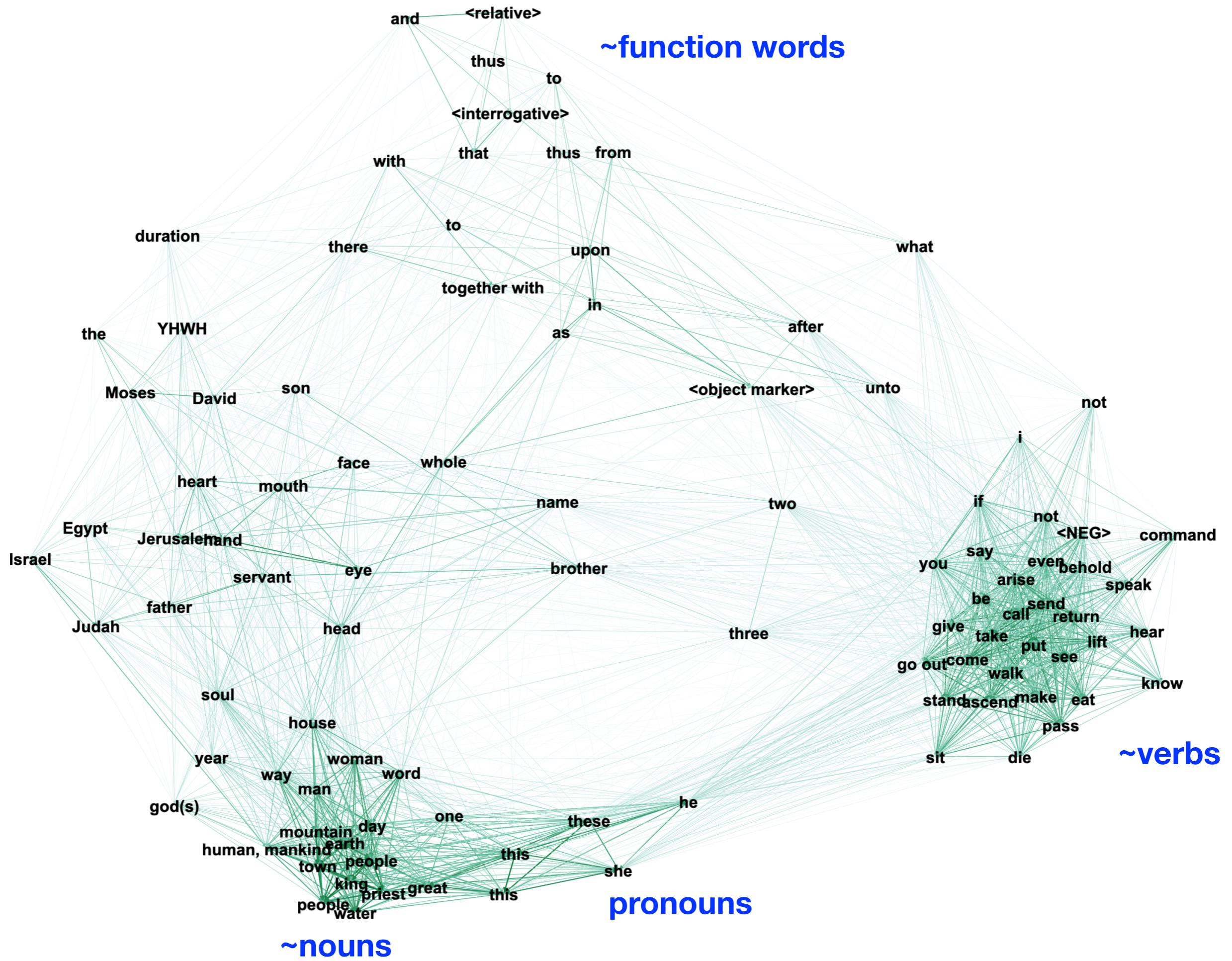
	B	>LHJM/	>T	H	W	>RY/	HJH[	<L	PNH/
B	1.000000	-0.136176	0.550530	-0.274051	-0.091999	-0.208063	-0.185976	0.626086	-0.029701
>LHJM/	-0.136176	1.000000	-0.171587	-0.129920	-0.002669	0.300991	-0.153233	-0.158390	0.047568
>T	0.550530	-0.171587	1.000000	-0.272447	-0.376742	-0.266840	0.165530	0.570215	-0.079837
H	-0.274051	-0.129920	-0.272447	1.000000	0.112907	-0.069290	-0.288102	-0.187305	0.091368
W	-0.091999	-0.002669	-0.376742	0.112907	1.000000	-0.189489	-0.548075	-0.095528	-0.198414
>RY/	-0.208063	0.300991	-0.266840	-0.069290	-0.189489	1.000000	-0.204252	-0.241469	-0.027057
HJH[	-0.185976	-0.153233	0.165530	-0.288102	-0.548075	-0.204252	1.000000	-0.125302	-0.105253
<L	0.626086	-0.158390	0.570215	-0.187305	-0.095528	-0.241469	-0.125302	1.000000	-0.047564
PNH/	-0.029701	0.047568	-0.079837	0.091368	-0.198414	-0.027057	-0.105253	-0.047564	1.000000
MJM/	-0.125372	0.252663	-0.159008	-0.166524	-0.190084	0.827957	-0.155141	-0.149116	-0.001856
>MR[	-0.220878	-0.104413	0.106645	-0.242685	-0.514283	-0.170039	0.664027	-0.218460	0.140998
R>H[	-0.128572	-0.145078	0.202005	-0.222007	-0.465541	-0.194255	0.654897	-0.061707	0.004395
KJ	-0.092597	0.177883	-0.385489	0.008535	0.456054	-0.099231	-0.247210	-0.149167	-0.054475
QR>[	-0.248269	-0.073924	0.047789	-0.279372	-0.471859	-0.152312	0.820903	-0.196548	-0.085214
L	0.324513	-0.086371	0.183148	-0.202406	0.275984	-0.302039	-0.214334	0.335124	-0.109457

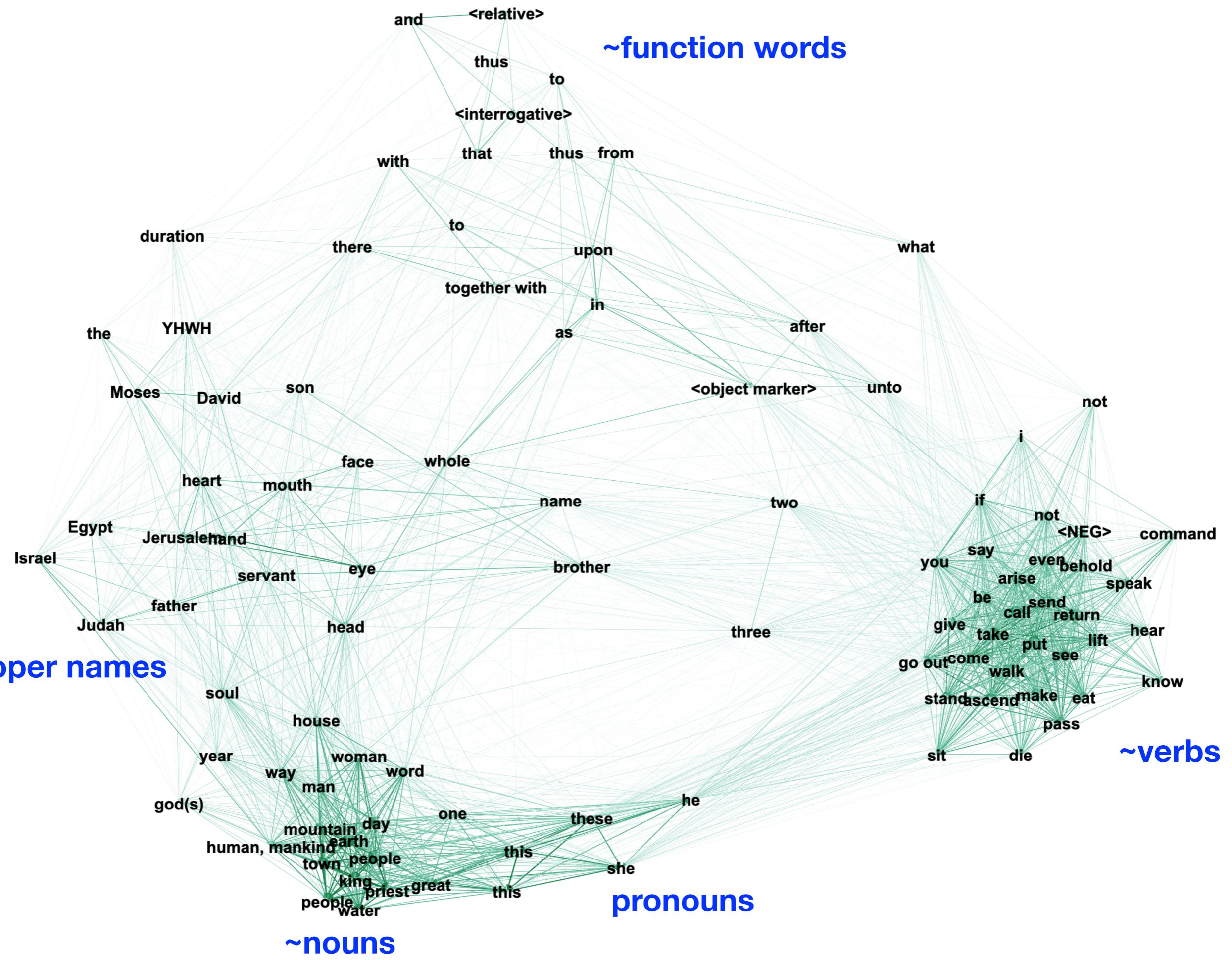


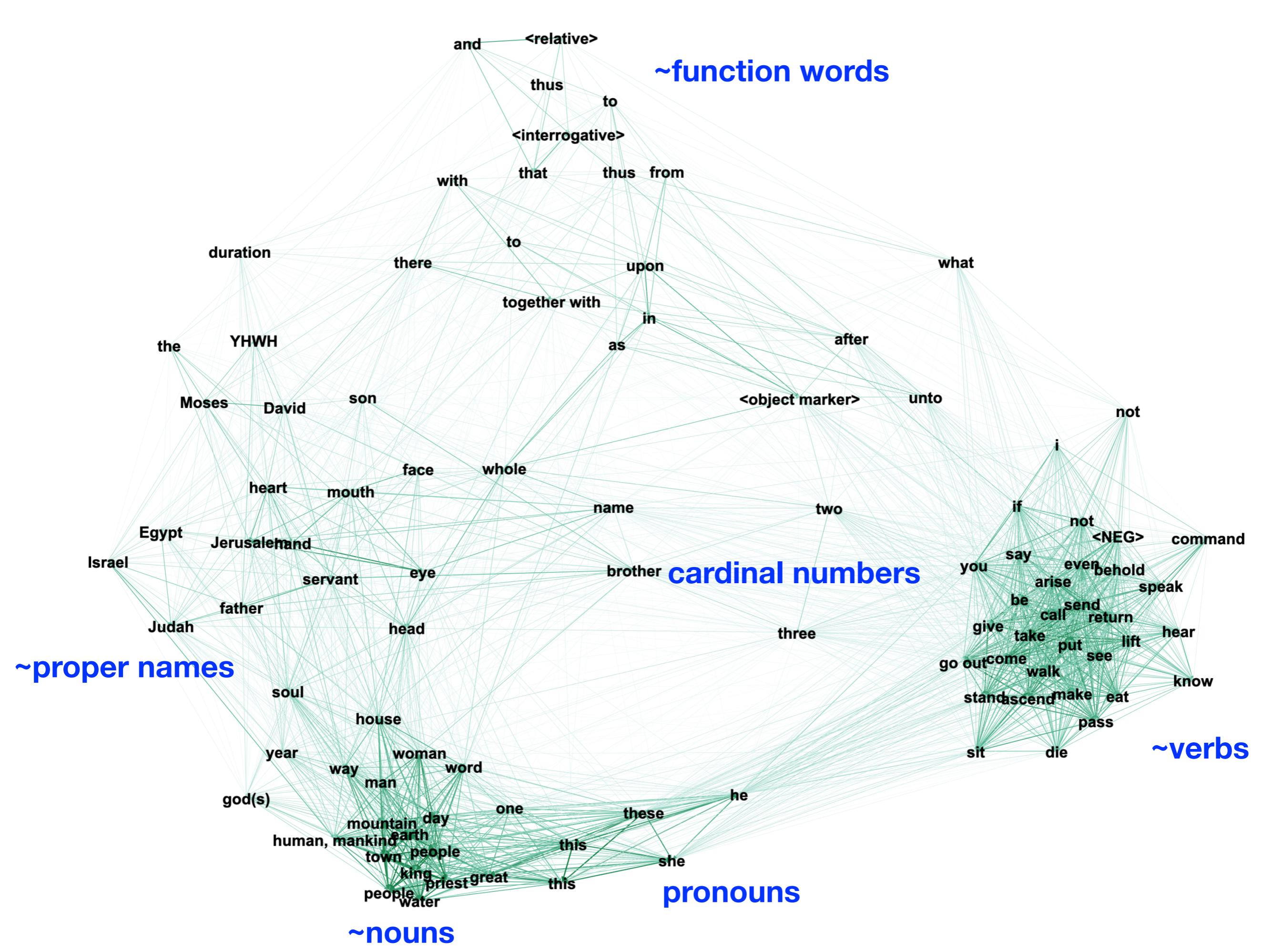












marker>

unto

not

~verbs

two

if

not

<NEG>

command

say

ever

behold

you

arise

be

send

call

return

speak

give

take

put

see

hear

go outcome

walk

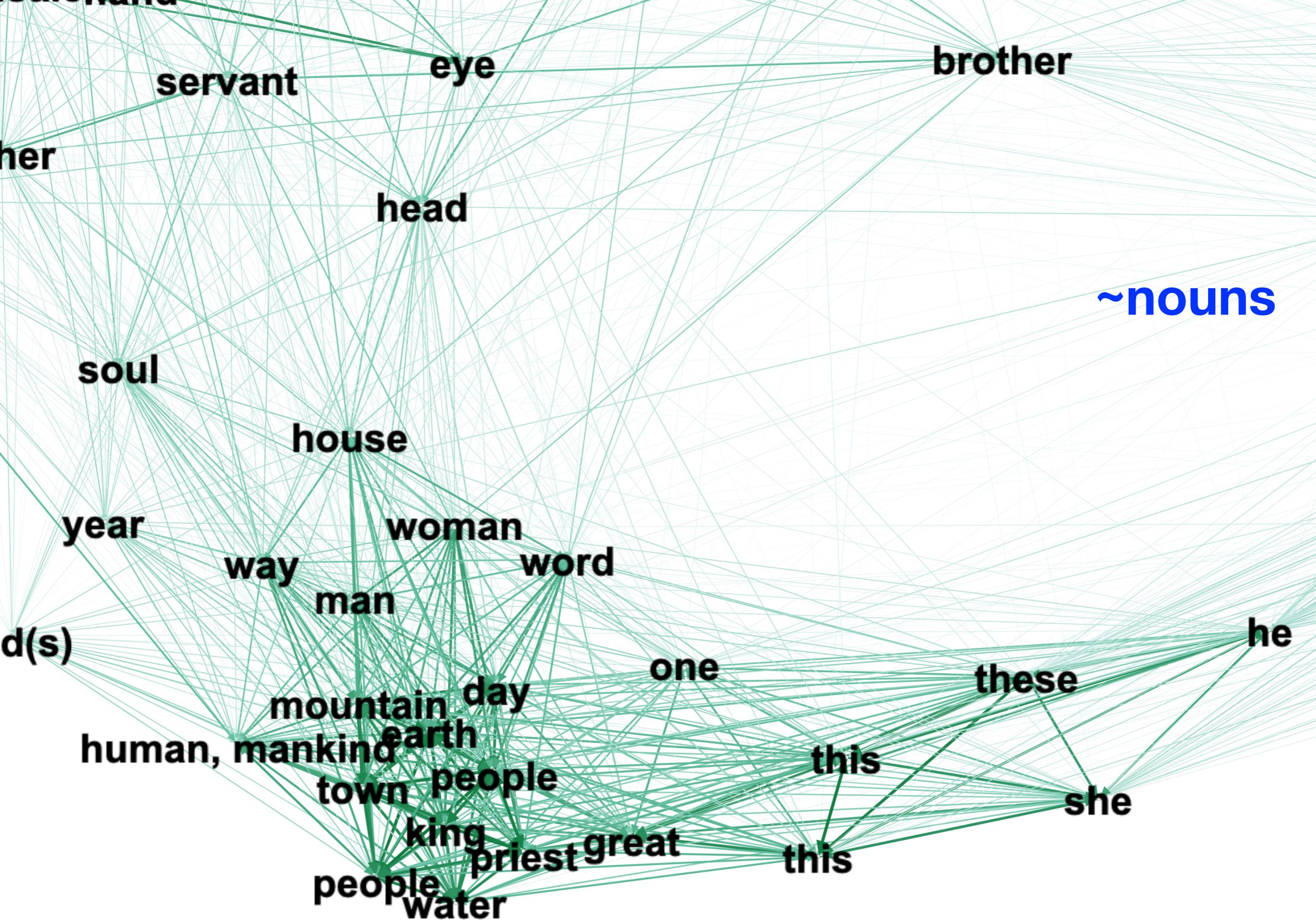
know

stand ascend make eat

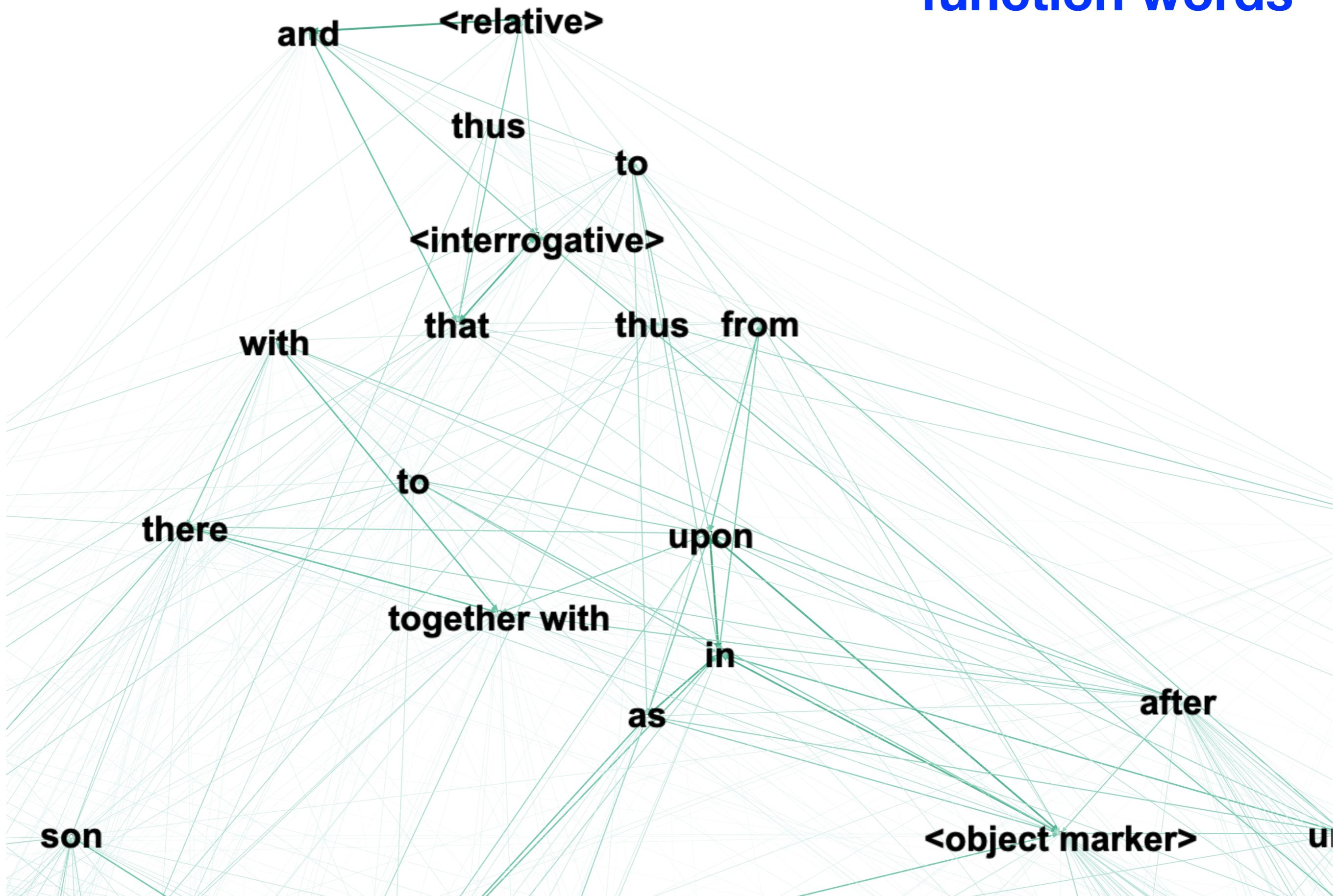
pass

sit

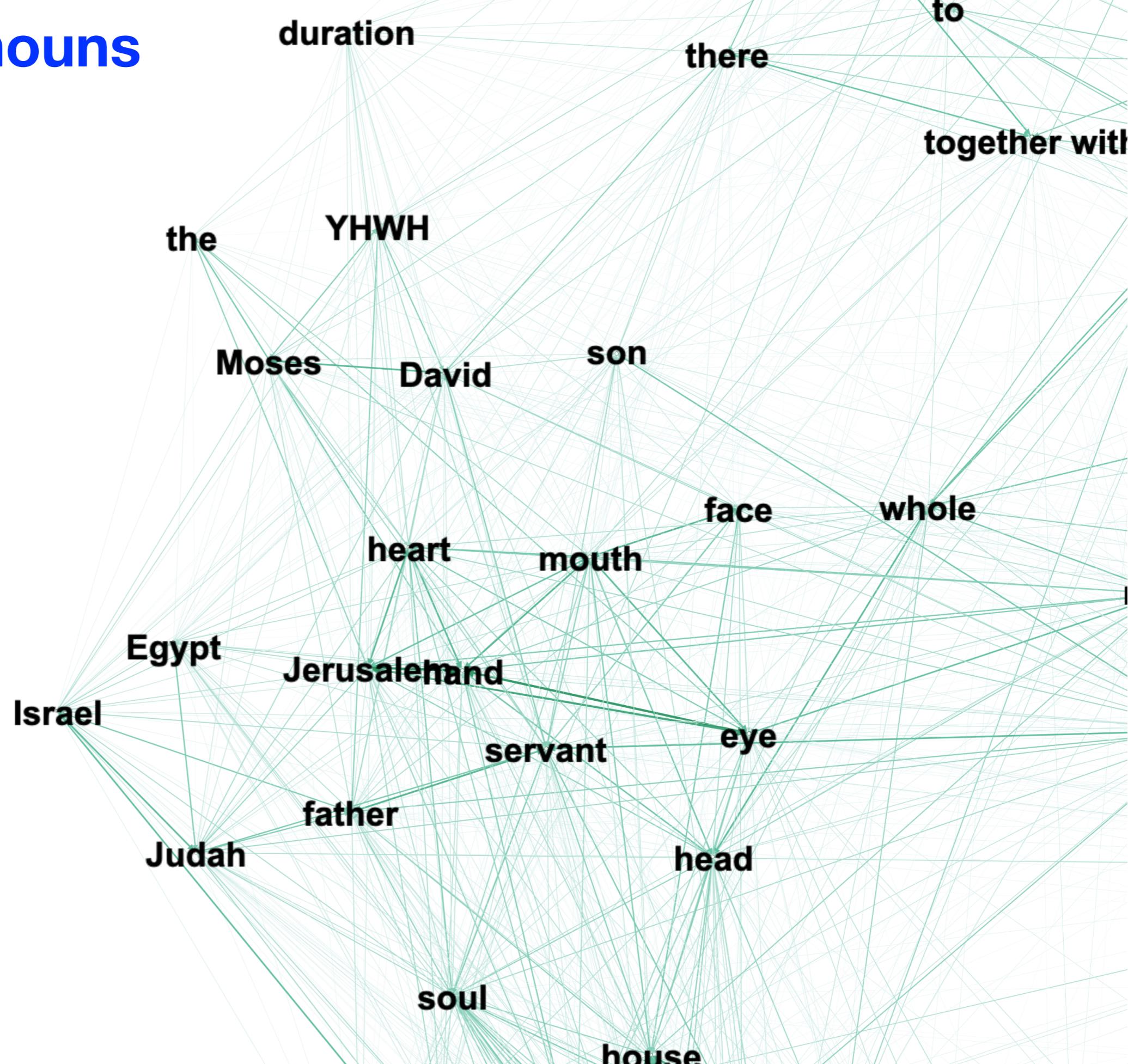
die



## ~function words



## ~proper nouns



# Noun Classes Model



# Noun Classes Model



# Noun Classes Model

- **corpus:** Hebrew Bible, ETCBC's BHSA



# Noun Classes Model

- **corpus:** Hebrew Bible, ETCBC's BHSA

- **sample:** top 199 most common nouns



# Noun Classes Model

- **corpus:** Hebrew Bible, ETCBC's BHSA



- **sample:** top 199 most common nouns

- **context:** syntactic context of target lexeme (BHSA)



# Noun Classes Model

- **corpus:** Hebrew Bible, ETCBC's BHSA



- **sample:** top 199 most common nouns

- **context:** syntactic context of target lexeme (BHSA)

- **normalization:** Fisher's Exact; bidirectional measure of statistical association (Stefanowitsch and Gries 2003)



# Noun Classes Model

- **corpus:** Hebrew Bible, ETCBC's BHSA
- **sample:** top 199 most common nouns
- **context:** syntactic context of target lexeme (BHSA)
- **normalization:** Fisher's Exact; bidirectional measure of statistical association (Stefanowitsch and Gries 2003)
- **analysis:** cosine similarity; principle component analysis (PCA); Gephi ForceAtlas2



# Contexts Considered

# Contexts Considered

- primary verb lexeme & stem in a noun's clause
  - e.g. קול “voice” used with **qal stem קרא** “call out”

# Contexts Considered

- primary verb lexeme & stem in a noun's clause
  - e.g. קול “voice” used with **qal stem קרא “call out”**
- noun's syntactic role within the clause
  - e.g. לחם “bread” serves as a **direct object**
  - subject, object, complement, adjunct

# Contexts Considered

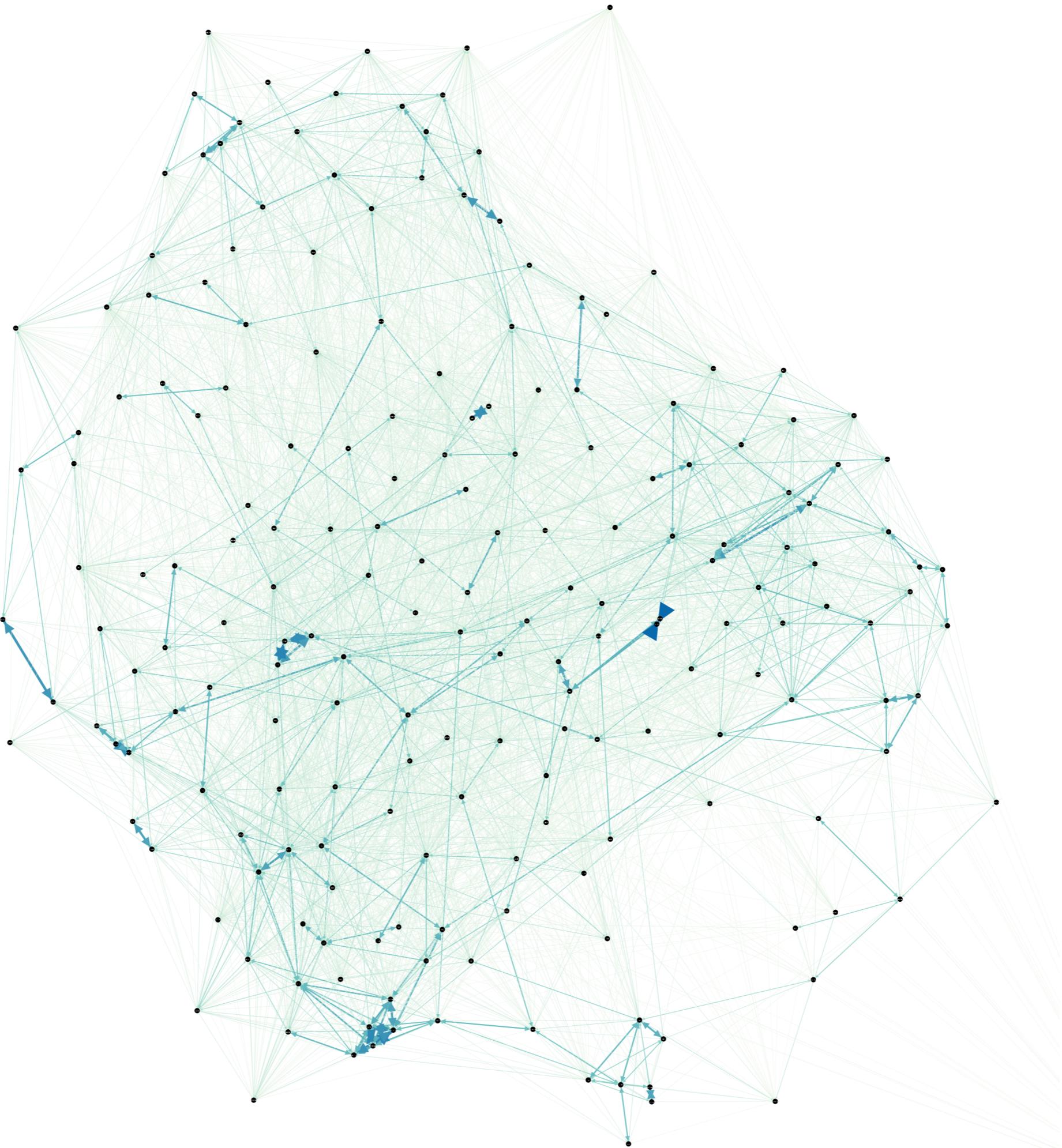
- primary verb lexeme & stem in a noun's clause
  - e.g. קול “voice” used with **qal stem קרא** “call out”
- noun's syntactic role within the clause
  - e.g. לחם “bread” serves as a **direct object**
  - subject, object, complement, adjunct
- lexeme of any attached preposition
  - e.g. דרך “road” occurs with **על** “on”

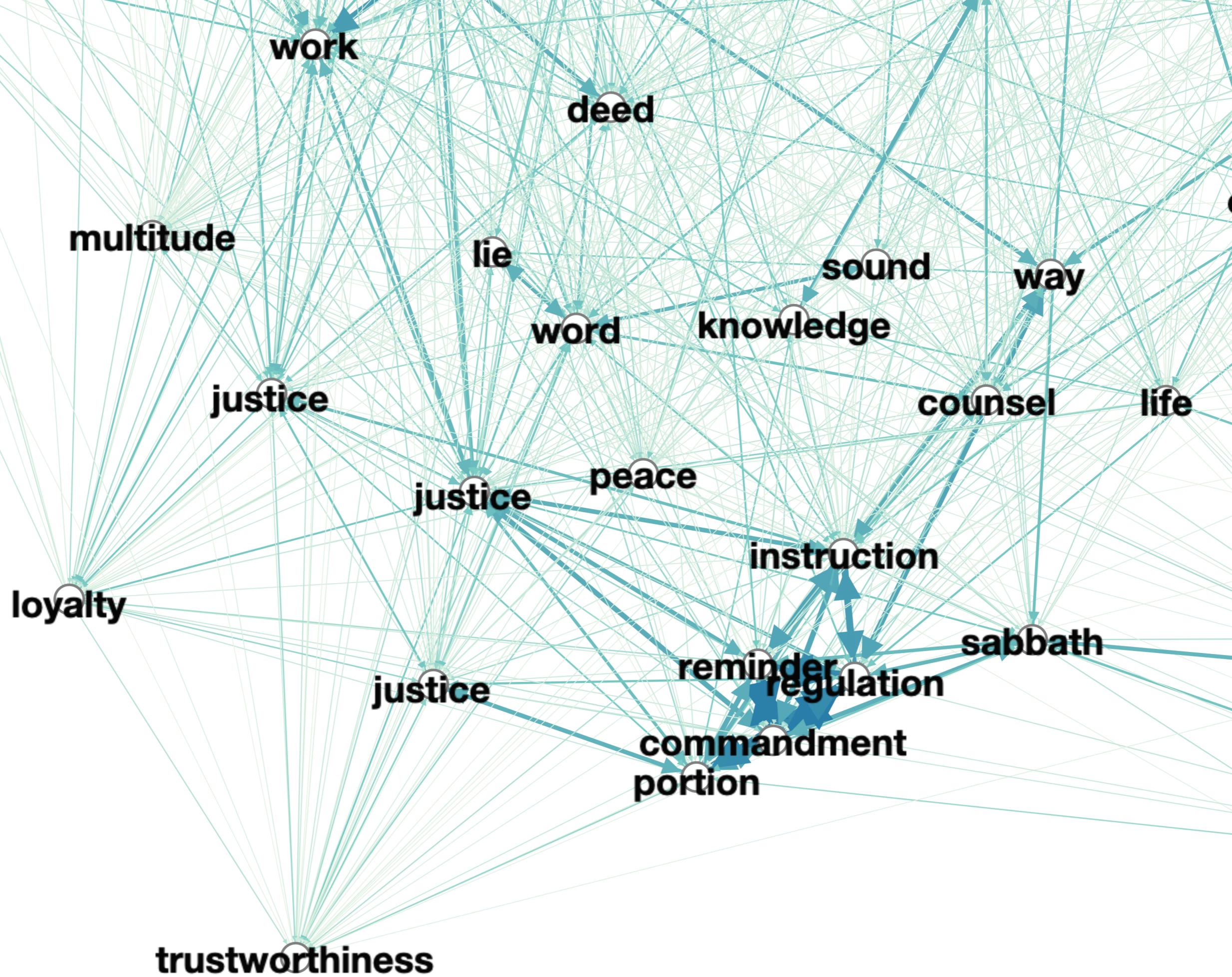
# Contexts Considered

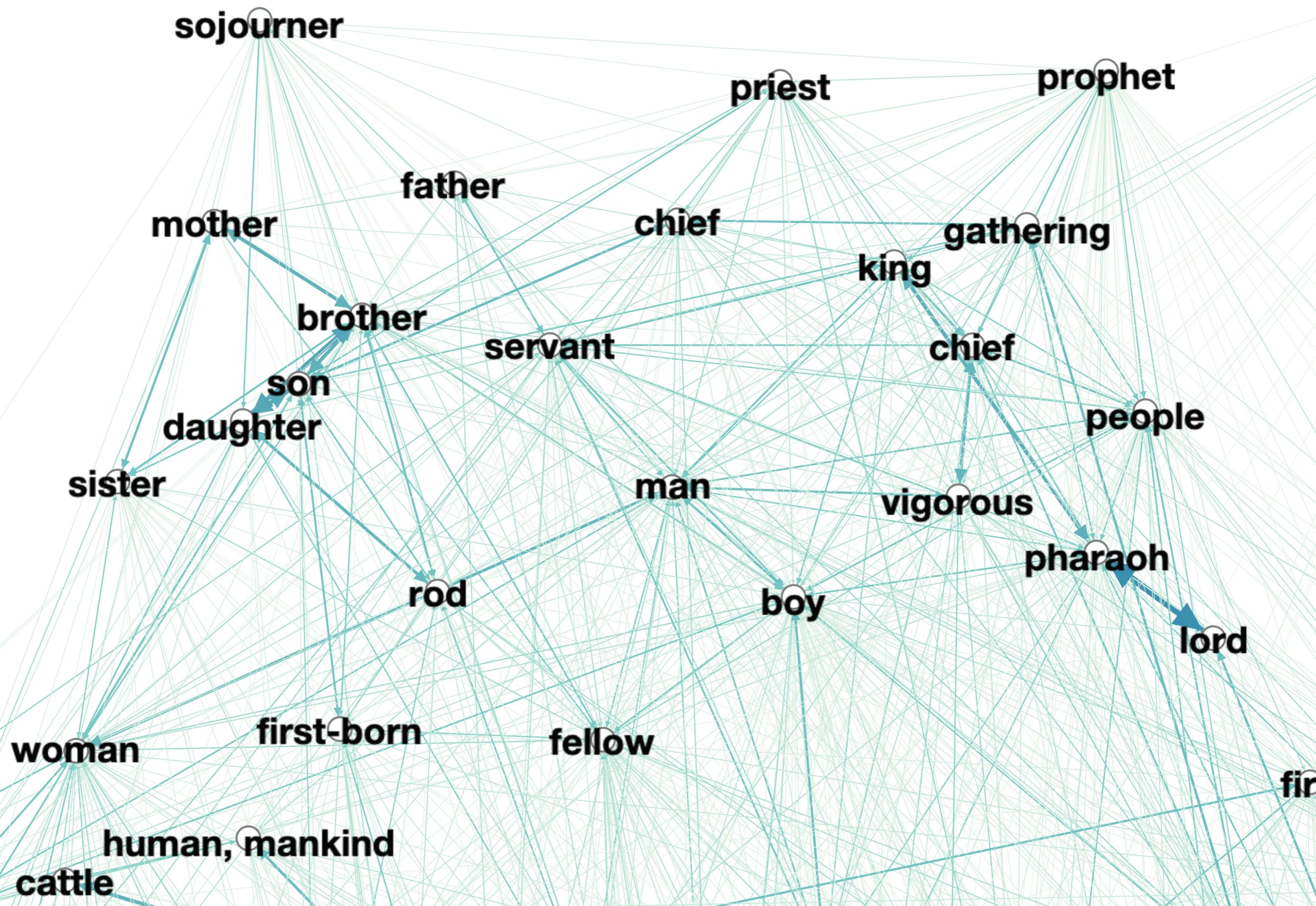
- primary verb lexeme & stem in a noun's clause
  - e.g. קול “voice” used with **qal stem קרא “call out”**
- noun's syntactic role within the clause
  - e.g. לחם “bread” serves as a **direct object**
  - subject, object, complement, adjunct
- lexeme of any attached preposition
  - e.g. דרך “road” occurs with **על “on”**
- phrase relations: adjunct, coordinate, adjectival
  - e.g. זהב “gold” is coordinate with **כסף “silver”**

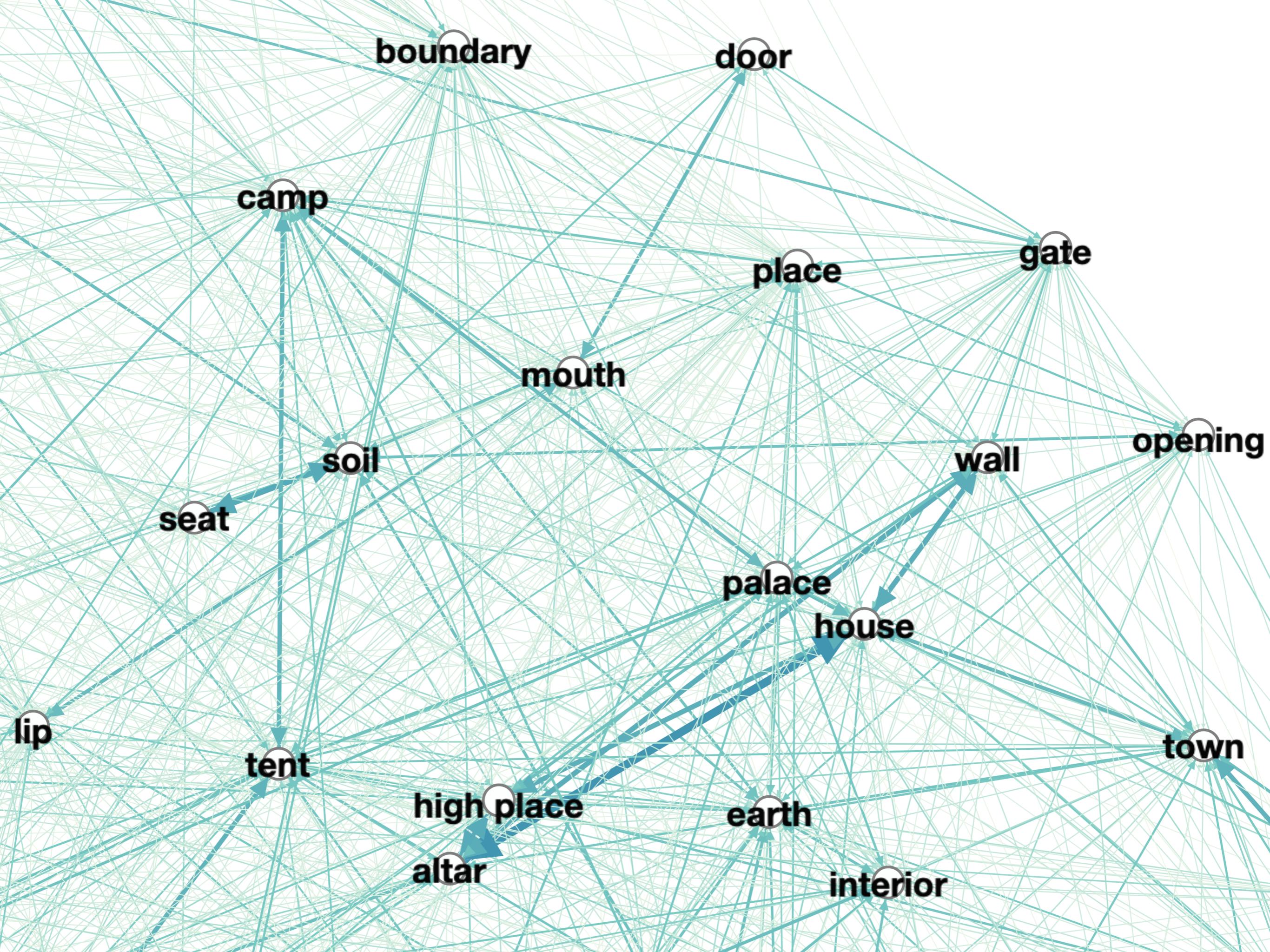
noun		context tag		frequency
לב	Objc ->	piel.	זכה.	1
פתח	Cmpl ->	qal.	עמד.	10
איש	Subj ->	hit.	פקד.	2
ארץ	Loca ->	qal.	גפל.	1

noun		context tag		frequency
גוי	Adju	- בָּ	amar.	2
יד	Cmpl	- מַנְ	ישע.	17
עת	Time	- בָּ	נתן.	1
דרך	Cmpl	- לְ	פונה.	1

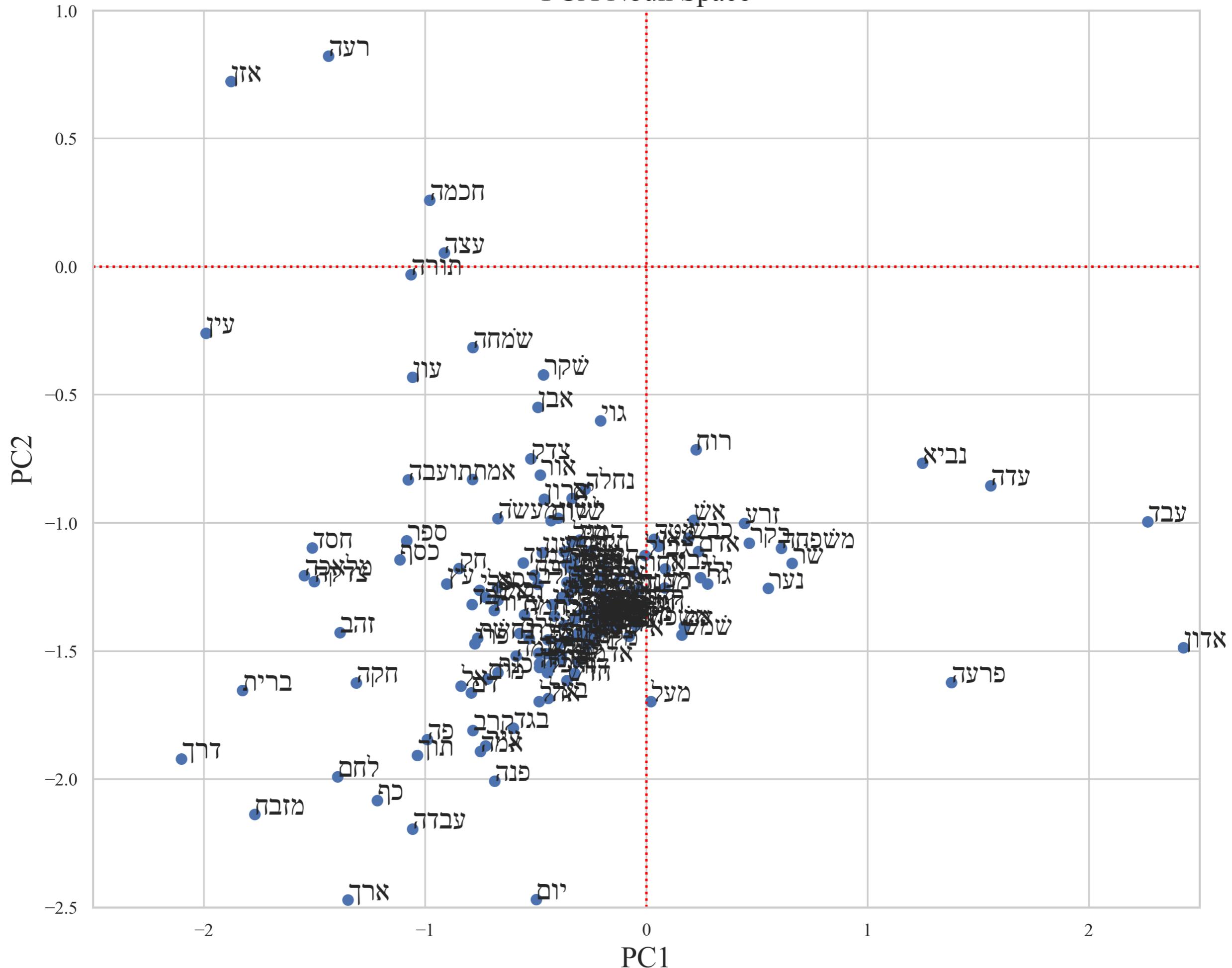




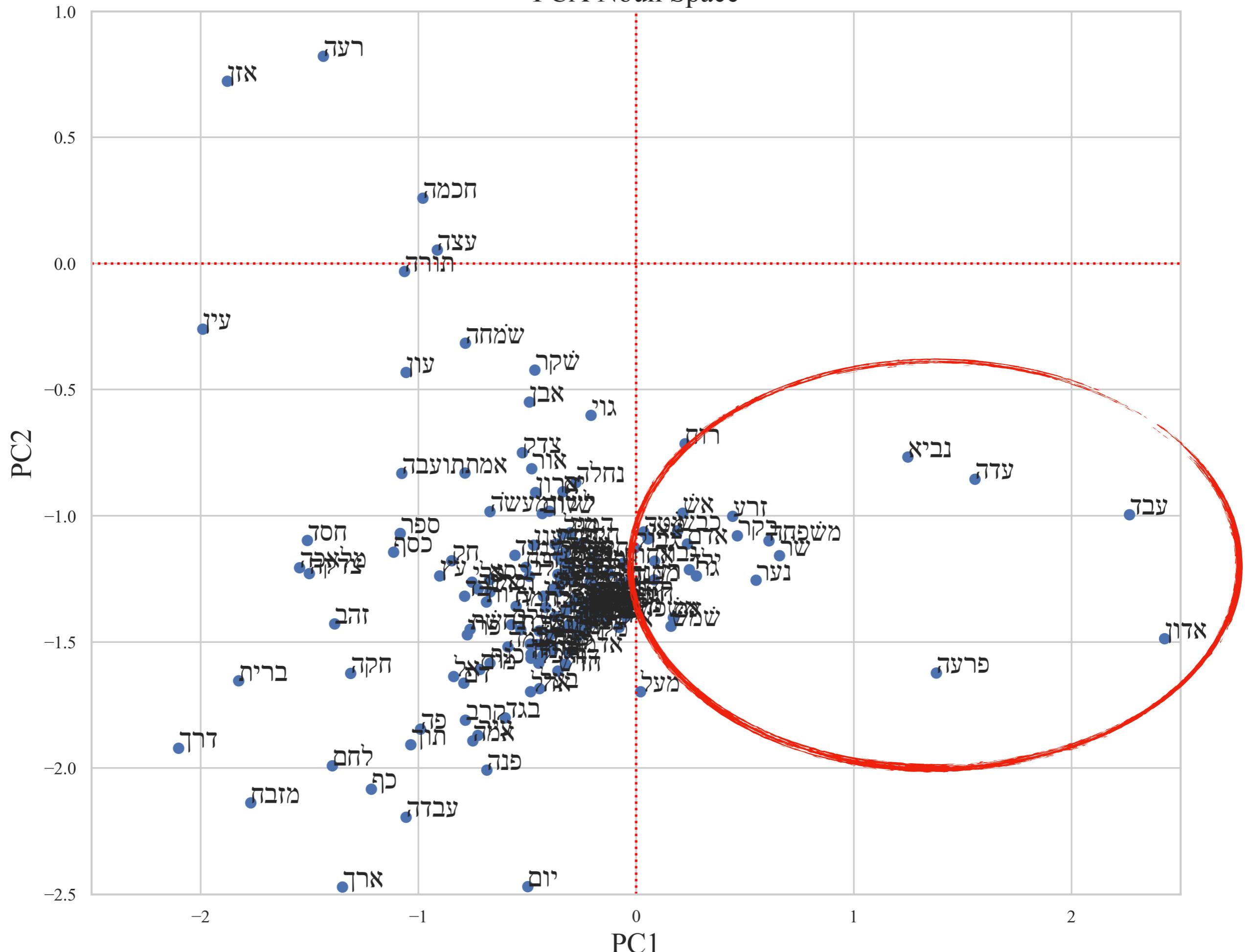




## PCA Noun Space



PCA Noun Space



# ~Animate Group

בן "son" (205.55)	שבט "tribe; rod" (2.94)	גָּרֶר "sojourner" (0.28)
בת "daughter" (26.0)	אָב "father" (2.9)	יַלְדֵּן "boy" (0.24)
אח "brother" (12.1)	אָדוֹן "lord" (2.43)	אָדָם "human, mankind" (0.24)
כהן "priest" (11.03)	עָבֵד "servant" (2.27)	רוֹחַ "wind" (0.22)
פר "young bull" (10.06)	עֲדָה "gathering" (1.56)	אֵשׁ "fire" (0.21)
מלך "king" (9.24)	פְּרָעהּ "pharaoh" (1.38)	כָּבֵשׂ "young ram" (0.19)
אִשָּׁה "woman" (4.62)	נָבִיא "prophet" (1.25)	אֶם "mother" (0.17)
אִישׁ "man" (4.24)	שָׁרֵךְ "chief" (0.66)	שֶׁמֶשׁ "sun" (0.16)
עם "people" (3.86)	מַשְׂפְּחָה "clan" (0.61)	גָּבּוֹר "vigorous" (0.08)
בָּכָר "first-born" (3.81)	גַּעַר "boy" (0.55)	רַע "fellow" (0.08)
נְשִׁיאָה "chief" (3.01)	בָּקָר "cattle" (0.46)	
	זַרְעָה "seed" (0.44)	

Noun	Context	Hits	Passages	Example
יד.n1	T.Subj→עֲשָׂה.v1.qal	7	Isa 31:7, 41:20, 66:2; Ps 119:73; Job 5:12, 12:9;	אֲשֶׁר עָשָׂה לְכֶם יְדִיכֶם חַטָּאת: (Isa 31:7)
שנה.n1	T.Subj→יִסְף.v1.hif	1	Prov 9:11	וַיֹּסֵפוּ לְךָ שְׁנָות חַיִים: (Prov 9:11)
יום.n1	T.Subj→חַנָּה.v1.qal	1	Judg 19:9	הַגָּה חַנּוֹת הַיּוֹם (Judg 19:9)
יום.n1	T.Subj→בָּוּא.v1.qal	14	1 Sam 26:10; Isa 13:9; Jer 46:21, 50:27, 50:31; Ezek	אוֹיָם בָּזָא (1 Sam 26:10)
בָּית.n1	T.Subj→מוֹת.v1.qal	2	1 Sam 22:16; 1 Chr 10:6	מוֹת תְּמֻמּוֹת אַתָּה וּכְלַבָּית אָבִיךָ: (1 Sam 22:16)
בָּית.n1	T.Subj→אָסְף.v1.nif	1	Judg 9:6	וַיַּאֲסַפֵּוּ בְּלִבְנֵי שָׁבֵט וּכְלַבָּית מְלוֹא (Judg 9:6)
בָּית.n1	T.Subj→אָכַל.v1.qal	3	Num 18:31; Deut 15:20; 1 Kgs 17:15	וְאָכַלְתֶּם אֹתָן בְּכָל-מִקְומָם אַתָּם וּבִתְּכֶם (Num 18:31)
ען.n1	T.Subj→שָׁמַר.v1.qal	1	Job 24:15	וְעַזְנָאָרְךָ שָׁמְרָה גַּשְׁךָ (Job 24:15)
ען.n1	T.Subj→שָׁאַל.v1.qal	1	Eccl 2:10	אֲשֶׁר שָׁאַלְוּ עִינֵּי (Eccl 2:10)
עיר.n1	T.Subj→יִשְׁבַּב.v1.qal	4	Jer 17:25, 31:24; Ezek 36:35; Lam 1:1	וַיִּשְׁבַּבְתָּה הָעִיר-הַזֹּאת לְעוֹלָם: (Jer 17:25)
עיר.n1	T.Subj→זַעַק.v1.qal	1	1 Sam 4:13	וַתַּזְעַק בְּלַהֲעֵיר: (1 Sam 4:13)
צדקה.n1	T.Subj→יִשְׁבַּב.v1.qal	1	Isa 32:16	וְצִדְקָה בְּפַרְמָל תִּשְׁבַּב: (Isa 32:16)
מים.n1	T.Subj→בָּוּא.v1.qal	6	Lev 11:34; Num 5:22, 5:24, 5:27; Ezek 47:9; Ps	אֲשֶׁר יָבֹא עַלְיוֹ מִים (Lev 11:34)
מים.n1	T.Subj→הַלְךָ.v1.qal	1	1 Kgs 18:35	וַיָּלְכוּ הַמִּים סְבִיב לְמִזְבֵּחַ (1 Kgs 18:35)
לב.n1	T.Subj→מוֹת.v1.qal	1	1 Sam 25:37	וַיִּמְתַּתְּ לְבָזְבָּז בְּקָרְבָּנוֹ (1 Sam 25:37)
לב.n1	T.Subj→זַעַק.v1.qal	1	Isa 15:5	לְבִי לְמוֹאָב זַעַק (Isa 15:5)

Noun	Context	Hits	Passages	Example
לֵב.n1	T.Subj→ הַלְךָ.v1.qal	2	2 Kgs 5:26; Job 31:7	לֹא־לֶבִי הַלְךָ (2 Kgs 5:26)
לֵב.n1	T.Subj→ נִכָּה.v1.hif	2	1 Sam 24:6; 2 Sam 24:10	וַיַּד לִב־דָּן אֲתָּה (1 Sam 24:6)
עַז.n1	T.Subj→ הַלְךָ.v1.qal	1	Judg 9:8	הַלְׁזֵד הַלְּבָבָ הַעֲצִים (Judg 9:8)
תוֹעַבָּה.n1	T.Subj→ עָשָׂה.v1.qal	1	Jer 32:35	לְעַשְׂוֹת הַתוֹעַבָּה הַזֹּאת (Jer 32:35)
סִפְר.n1	T.Subj→ בָּוֹא.v1.qal	2	2 Kgs 5:6, 10:7	בָּבָא הַסִּפְר הַזֶּה אָלִיךְ (2 Kgs 5:6)
נֶפֶש.n1	T.Subj→ מוֹת.v1.qal	3	Num 23:10; Judg 16:30; Job 36:14	תִּמְתַּחַנְתִּי מְתִתְנֵנִים יִשְׁרָאֵל (Num 23:10)
נֶפֶש.n1	T.Subj→ שָׁמַר.v1.qal	1	Ps 119:167	שָׁמַרְתִּי נֶפֶשִׁי עַד־תְּיִקְהָ (Ps 119:167)
נֶפֶש.n1	T.Subj→ אָכְל.v1.qal	1	Lev 17:12	כָּל־נֶפֶשׁ מִכֶּם לֹא־תִּאְכַּל דָּם (Lev 17:12)
נֶפֶש.n1	T.Subj→ שָׁאֵל.v1.qal	1	Deut 14:26	אֲשֶׁר תִּשְׁאַלְתָּךְ נֶפֶשׁ (Deut 14:26)
נֶפֶש.n1	T.Subj→ הַלְךָ.v1.qal	2	Isa 46:2; Eccl 6:9	וְנֶפֶשׁ בְּשֻׁבִּי הַלְּבָבָ: ס (Isa 46:2)

# Coordinate Parsing Model



# Coordinate Parsing Model



# Coordinate Parsing Model

- **corpus:** Hebrew Bible, PhD project corpus (derived from BHSA)



# Coordinate Parsing Model

- **corpus:** Hebrew Bible, PhD project corpus (derived from BHSA)
- **sample:** all lexemes in adverbial time phrases



# Coordinate Parsing Model

- **corpus:** Hebrew Bible, PhD project corpus (derived from BHSA)



- **sample:** all lexemes in adverbial time phrases

- **context:** bag of words, window = 5



# Coordinate Parsing Model

- **corpus:** Hebrew Bible, PhD project corpus (derived from BHSA)
- **sample:** all lexemes in adverbial time phrases
- **context:** bag of words, window = 5
- **normalization:**  $\Delta P$  – unidirectional measure of statistical association (Ellis 2006)
- 



# Coordinate Parsing Model

- **corpus:** Hebrew Bible, PhD project corpus (derived from BHSA)
- **sample:** all lexemes in adverbial time phrases
- **context:** bag of words, window = 5
- **normalization:**  $\Delta P$  – unidirectional measure of statistical association (Ellis 2006)
- **analysis:** cosine distance; parsing task



**Isaiah 17:11**

**בַּיּוֹם נִחְלָה וּכְאָב אֶנוֹשׁ**

**in a day of sickness and pain incurable**

## Cosine Distance

	day	sickness
pain	0.85	0.13

**Isaiah 17:11**

בַּיּוֹם נְחַלָּה וּכְאָב אֲנוֹשׁ

in a day of sickness and pain incurable

## **Proverbs 7:9**

**בְּאַיִשּׁוֹן לִילָה וּאֲפָלָה**

**in pupil of night and darkness**

## Cosine Distance

	night	pupil
darkness	0.47	0.66

## Proverbs 7:9

בְּאַיִשּׁוֹן לִילָה וְאַפְלָה

in pupil of night and darkness

## **Proverbs 7:9**

**ימֵי עֲנִיה וּמְרוֹדִיחַ**

**days of humiliation and roving**

## Cosine Distance

day      humiliation

---

roving    1.41    0.70

## Proverbs 7:9

ימֵי עֲנִיה וּמְרוֹדִיה :

days of humiliation and roving

## **Proverbs 7:9**

**בָּמוֹת אָבִיהָ וְאֶמְהָ**

**at death of her father and her mother**

## Cosine Distance

	death	father
mother	1.22	0.51

## Proverbs 7:9

בָּמוֹת אֲבִיהָ וְאֶמְהָ

at death of her father and her mother

## **Proverbs 7:9**

**בַּיּוֹם עֲנָנוֹ וְעַרְפָּלָן**

**in a day of cloud and darkness**

## Cosine Distance

	day	cloud
darkness	1.11	0.73

## Proverbs 7:9

בַּיּוֹם עָבֵן וְעַרְפֵּל

in a day of cloud and darkness

# Looking Ahead

~lemmatization

~word senses

**~part of speech**

~construction induction (“syntax”)

**~semantic classes**

# Looking Ahead

# Looking Ahead

- Semantic vector spaces capture various aspects of word meaning; but is it possible to consider aspects at once?

# Looking Ahead

- Semantic vector spaces capture various aspects of word meaning; but is it possible to consider aspects at once?
- future vision: a high dimensional map of Hebrew terms
  - word senses, registers, pragmatics?

# Looking Ahead

- Semantic vector spaces capture various aspects of word meaning; but is it possible to consider aspects at once?
- future vision: a high dimensional map of Hebrew terms
  - word senses, registers, pragmatics?
- Address data sparsity with smart strategies

# Looking Ahead

- Semantic vector spaces capture various aspects of word meaning; but is it possible to consider aspects at once?
- future vision: a high dimensional map of Hebrew terms
  - word senses, registers, pragmatics?
- Address data sparsity with smart strategies
- e.g. can Zipfian distributions be used as reference points for a machine learning model to classify rare terms?

# Works Cited

- Ellis, Nick. “Language Acquisition as Rational Contingency Learning.” *Applied Linguistics* 27/1 (2006): 1–24.
- Goldberg, Adele. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago.
- Gries, Stefan Th. “Dispersions and adjusted frequencies in corpora.” *International Journal of Corpus Linguistics* 13:4 (2008): 403–437.
- Stefanowitsch, Anatol, and Stefan Th. Gries. “Collostructions: Investigating the interaction of words and constructions.” *International Journal of Corpus Linguistics* 8:2 (2003): 209–243.