Collation of Armenian manuscripts

A lone historian's approach

ց՝ իր արարանի այր գուներիան գովեստից մարդկան։ Որոճալով որոճեն զդիւական երբ Հայաստանիայի գուների հայերա բիներացի ասելով նե եր նորյոր և մերերի դեսեն և որո

Tara L Andrews | University of Oxford | 18 November 2009

արա « — արաթար» , արա արարագութագրութեն զկարգ ժամատեղացն։ լինին և բազումք ուսուն

Chronicle of Matthew of Edessa

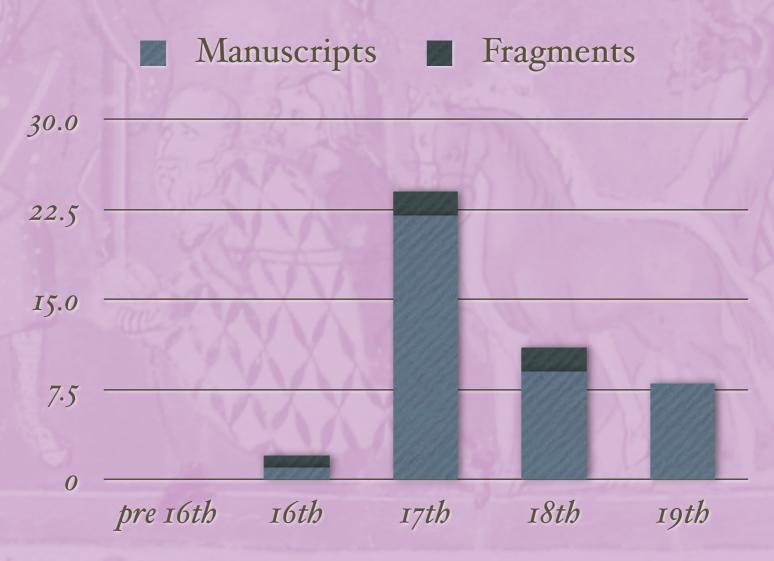
- Written around 1137
- 35 known manuscripts
- No manuscripts older than 1600

ց՝ իր արարանի արև հայի թարարան դովեստից մարդկան։ Որոճալով որոճեն զդիւական երա Հայաստանիայց՝ Հայի հերա բիկերութի ասելով նի եր կորորը և մերերի դեսեն և որո

Tara L Andrews | University of Oxford | 18 November 2009

ախա » — թերարար , մա ատապատաբարարան զկարգ ժամատեղացն։ լինին և բազումք ուսուն

Extant manuscripts of the Chronicle



ց՝ [երարարաները հայի շարիներին գովեստից մարդկան։ Որոճալով որոճեն զդիւական երբ Հայաստաները է է հերարան անհերան անհետ իր հեր հուսու և մերերի դեսեն և որո

Tara L Andrews | University of Oxford | 18 November 2009

արա » — արաթարը , արա առաարաբերակարորեն զկարգ ժամատեղացն։ լինին և բազումք ուսուն

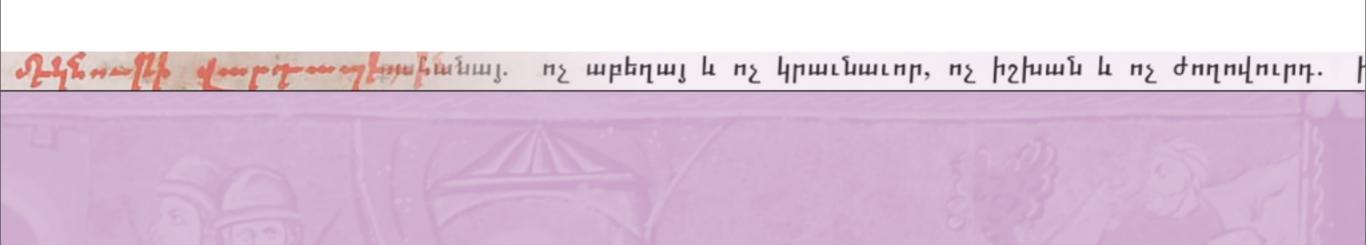
My collation problem

- No single manuscript can be the base
- No need to collate 'derivative' manuscripts
 - ...but difficult to identify these.
- Need solution: baseless collation and cladistic analysis.

ց՝ իր արարանի այր գուներիան գովեստից մարդկան։ Որոճալով որոճեն զդիւական երբ Հայաստանիայի գուների հայեստեն ասելով նե ես հոս որ և մերերի դեսեն և որո

Tara L Andrews | University of Oxford | 18 November 2009

ախա » — թերարար, մա ատապատաբարարան զկարգ ժամատեղացն։ լինին և բազումք ուսուն



Is it worth it?

ց՝ [երթութուները է գրինան գովեստից մարդկան։ Որոճալով որոճեն զդիւական երա հայաստանի անտանական։ Որոճալով որոճեն զդիւական երա հայաստանի անտանական։ Որոճալով որոճեն զդիւական երա

Tara L Andrews | University of Oxford | 18 November 2009

արա » — արաթարը , արա առաարաբերակարորեն զկարգ ժամատեղացն։ լինին և բազումք ուսուն

Tools I need

- Way to collate transcribed manuscripts
- Way to review results for editorial decisions
 - Help keeping my decisions consistent
- Publication of finished edition

ց, իր արագրուն հայր թարգեր երնան գովեստից մարդկան։ Որոճալով որոճեն զդիւական երբ Հայաստանիայց, *5, իվերա հա*նիերացն ասելով ներ եր հրուսը և մերերի դեսեն և դու

Tara L Andrews | University of Oxford | 18 November 2009

արա » — արաթարը , արա առաարաբարարարին ըկարգ ժամատեղացն։ լինին և բազումք ուսուն

Specific collation requirements

- Unicode for Armenian script
- TEI files for manuscript transcriptions
- Output suitable for stemmatic analysis, direct publishing, or manipulation by other programs

ց՝ իր արարանի այր շարինան գովեստից մարդկան։ Որոճալով որոճեն զդիւական երբ Հայաստանիայց՝ Հ, իվերա հիմերացն ասելով նե ես կողմու և ներերի որոեն և որո

Tara L Andrews | University of Oxford | 18 November 2009

արա » — թերարաթ, այս արասարաբարարարեն զկարգ ժամատեղացն։ լինին և բազումք ուսուն

- COLLATE
- Juxta
- TUSTEP

ց՝ [երարարաները հայի թարիներին երինան գովեստից մարդկան։ Որոճալով որոճեն զդիւական երբ Հայաստաները 5 , իվերան հովեստին ասելով նե ես կողմու և ներերի դեսոեմ և դու

Tara L Andrews | University of Oxford | 18 November 2009

ւիս » — թերաթար, այն արասարագրարան ակարդ ժամատեղացն։ լինին և բազումք ուսուն

- Juxta
- TUSTEP

Tara L Andrews | University of Oxford | 18 November 2009

ախա » — թեթաթար» , ար առաարագութագրութեն զկարգ ժամատեղացն։ լինին և բազումք ուսուն

• TUSTEP

ց՝ իր արարանի այր գուներիան գովեստից մարդկան։ Որոճալով որոճեն զդիւական երբ Հայաստանիայց՝ Հայ է վերան անկերացն ասելով նե ես կորորը և մերերի դեսեն և որո

Tara L Andrews | University of Oxford | 18 November 2009

ախա » — թեթաթար» , ար առաարագրագրութեն ըկարգ ժամատեղացն։ լինին և բազումք ուսուն

ց՝ իր արարանի այր գուներիան գովեստից մարդկան։ Որոճալով որոճեն զդիւական երբ Հայաստանիայի գուների հայեստեն ասելով նե ես հոս որ և մերերի դեսեն և որո

Tara L Andrews | University of Oxford | 18 November 2009

ախա » — թեթաթար» , ար առաարագրագրութեն ըկարգ ժամատեղացն։ լինին և բազումք ուսուն

Virtues of a programmer

- Laziness
- Impatience
- Hubris

ց՝ իր արարանի այր գործեր երնան գովեստից մարդկան։ Որոճալով որոճեն զդիւական երբ Հայաստանիայց՝ Հայիներա հանրերացն ասելով նե ես կորորը և մերերի դեսեն և որո

Tara L Andrews | University of Oxford | 18 November 2009

ախա » — թեթաթար» , ար առաարագրագրութեն ըկարգ ժամատեղացն։ լինին և բազումք ուսուն

Specific collation requirements

- Unicode for Armenian script
- TEI files for manuscript transcriptions
- Output suitable for stemmatic analysis, direct publishing, or manipulation by other programs

ց՝ իր արարանի այր շարինան գովեստից մարդկան։ Որոճալով որոճեն զդիւական երբ Հայաստանիայց՝ Հ, իվերա հիմերացն ասելով նե ես կողմու և ներերի որոեն և որո

Tara L Andrews | University of Oxford | 18 November 2009

Specific collation requirements

- Unicode for Armenian script
- TEI files for manuscript transcriptions
- Output suitable for stemmatic analysis, direct publishing, or manipulation by other programs
- As little manual work as possible

ց, իր առագուները, գ, իների ան գովեստից մարդկան։ Որոճալով որոճեն զդիւական երբ Հայաստաները, գ, իների հիմերարի ասելով նե ես հրուսու և մերերի ուսրեն և ուս

Tara L Andrews | University of Oxford | 18 November 2009

ախա » — թերարար, այս արասարագուրդը արդարեն զկարգ ժամատեղացն։ լինին և բազումք ուսուն

Re-inventing the wheel

- 'Best-effort' collation algorithm
- Based on diff
- Was spared the 'transposition problem'

ց՝ իր արարանի այր գործեր երնան գովեստից մարդկան։ Որոճալով որոճեն զդիւական երբ Հայաստանիայց՝ Հայիներա հանրերացն ասելով նե ես կողմոր և մերերի դեսեն և որո

Tara L Andrews | University of Oxford | 18 November 2009

ախա » — թերարար, մա ատապատաբարարան զկարգ ժամատեղացն։ լինին և բազումք ուսուն

How diff works

The white cat was gingerly sitting atop a table. The white cat was sitting on top of a dining table.

SAME:	ADD:	SAME:	CHANGE:	SAME:	DELETE:	SAME:
The	+ gingerly	sitting	- atop	a	- dining	table
white		1301				A Warmer
cat		110115	+ on			1380
was	- man		+ top		1	1 833
100			+ of	31 111111111111111111111111111111111111		11/2011
		THE NAME OF THE				

ց՝ իր արարանի այր գուներիան գովեստից մարդկան։ Որոճալով որոճեն զդիւական երա Հայաստանիայց՝ Հայ է վերա հայիներացն ասելով նե ես կորորը և մերերի որոեմ և որո

Tara L Andrews | University of Oxford | 18 November 2009

ախա » — թերարար, այն արտարական արդարորեն զկարգ ժամատեղացն։ լինին և բազումք ուսուն

Transposition

The quick brown fox jumped over the lazy dogs. The lazy brown fox jumped over the quick dogs.

CHANGE:
- quick
--+ lazy
+ quick

Tara L Andrews | University of Oxford | 18 November 2009

ախա » — թերարար , մա ատապարաբարարան զկարգ ժամատեղացն։ լինին և բազումք ուսուն

Word movement

The quick brown fox jumped over the lazy dogs. The quick fox jumped over the lazy brown dogs.



ց՝ իր արարանի այր արդեն երիան գովեստից մարդկան։ Որոճալով որոճեն զդիւական երա Հայաստանիայի գուների հայաստան ասելով նե ես կողմու և ներերի ուսրեն և որո

Tara L Andrews | University of Oxford | 18 November 2009

ախա » — թերարար , մա ատապարաբարարան զկարգ ժամատեղացն։ լինին և բազումք ուսուն

401 of the Armenian era, there was a great famine in many places.

ց՝ իր արարանի այր գուներիան գովեստից մարդկան։ Որոճալով որոճեն զդիւական երբ Հայաստանիայց․ Հուների հայիսերացն ասելով նե ես կորուրը և մերերի որոեմ և որո

Tara L Andrews | University of Oxford | 18 November 2009

ախա » — թերարար , մա ատապարաբարարան զկարգ ժամատեղացն։ լինին և բազումք ուսուն

Ուրանի վարդագրանանալ. ոչ աբեղայ և ոչ կրաւնաւոր, ոչ իշխան և ոչ ժողովուրդ.

In 401 of the Armenian era, there was a great famine in many places.

In the Armenian year 401, there was a plague in many places.

ց՝ իր արարանիացի շարինան գովեստից մարդկան։ Որոճալով որոճեն զդիւական երբ Հայաստանիայց՝ Հայիների հայիներացն ասելով նե ես կողմու և ներերի ուսելու և ուս

Tara L Andrews | University of Oxford | 18 November 2009

ախա » — թերարար, այն արտարարարարարին զկարգ ժամատեղացն։ լինին և բազումք ուսուն

Constructing the 'superbase'

In	the	Armenian	year	401	of	the	Armenian	era,	there	was	а	great	famine plague	in	many	places.
		Jala .			11/53	31	Ka Killia						1 32	ΛE		
In				401	of	the	Armenian	era,	there	was	a	great	famine	in	many	places.

Tara L Andrews | University of Oxford | 18 November 2009

ախա » — թերաթարը , այս առաարարագուրակարորեն զկարգ ժամատեղացն։ լինին և բազումք ուսուն

<u> Քենսարի վարաագրանանալ</u>. ոչ աբեղայ և ոչ կրաւնաւոր, ոչ իշխան և ոչ ժողովուրդ.

Using the superbase

In	the	Armenian	year era	401	of	the	Armenian	era,	there	was	once	more	а	great	famine plague	in	many	places.
In				401	of	the	Armenian	era,	there	was	11/4	V	a	great	famine	in	many	places.
In	the	Armenian	year	401,		1			there	was			a	1	plague	in	many	places.
		Armenian	era	401,					there	was	once	more	a		plague.	- Y		

Tara L Andrews | University of Oxford | 18 November 2009

ախա » — թերաթար, այն արասարարագրարեն ըկարգ ժամատեղացն։ լինին և բազումք ուսուն

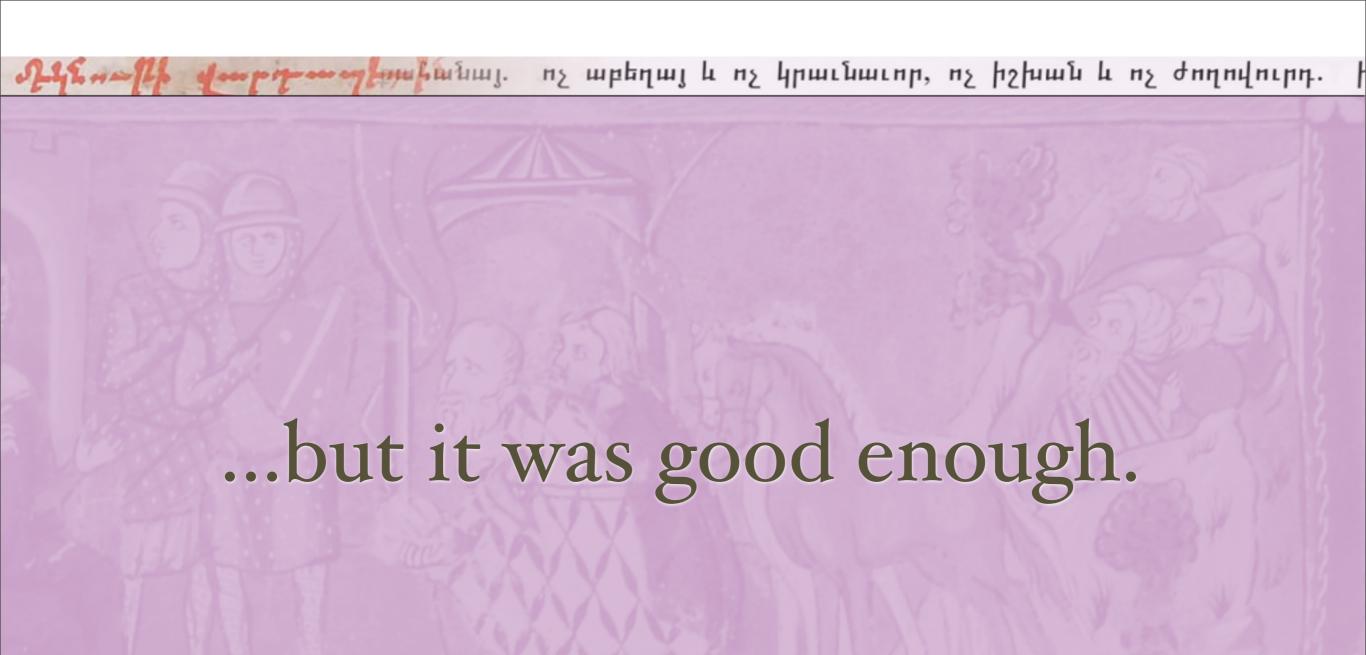
Straight to TEI

```
<app>
    <rdg wit="#A #B"><w>In</w><rdg>
    < rdg wit = "#C"/>
</app>
<app>
    <rdg wit="#B"><w>the</w></rdg>
    <rdg wit="#A #C"/>
</app>
<app>
    <rdg wit="#B #C"><w>Armenian</w></rdg>
    <rdg wit="#A"/>
</app>
<app>
    <rdg wit="#B"><w>year</w></rdg>
    <rdg wit="#C"><w>era</w></rdg>
    <rdg wit="#A"/>
</app>
<w>401</w>
<app>
    <rdg wit="#A"><w>of</w></rdg>
    <rdg wit="#B #C"/>
</app>
```

Tara L Andrews | University of Oxford | 18 November 2009

ւիս » — թերաթար, այս արասարագրակարարեն զկարգ ժամատեղացն։ լինին և բազումք ուսուն





ց՝ իր արարանի արև գրինան գրվեստից մարդկան։ Որոճալով որոճեն զդիւական երա հարարական հետ հուսու և հետերե դեսեն և դու

Tara L Andrews | University of Oxford | 18 November 2009

ախա » — թեթաթար» , ար առաարագրագրութեն ըկարգ ժամատեղացն։ լինին և բազումք ուսուն

What now?

- Use the collation results for stemmatic analysis
- Use the collation results to edit the text
- Send the collation results to another program for visualization

ց՝ իր արարանի այր գուներիան գովեստից մարդկան։ Որոճալով որոճեն զդիւական երբ Հայաստանիայց՝ Հայ հերա բաներան ասելով նե ես կողմու և ներերի դեսեն և որո

Tara L Andrews | University of Oxford | 18 November 2009

ախա » — թերարար, այն արասարարարարարին զկարգ ժամատեղացն։ լինին և բազումք ուսուն

How not to do stemmatic analysis

Tara L Andrews | University of Oxford | 18 November 2009

արա « — արաթար» , արա առաարաբարարարարը ըկարգ ժամատեղացն։ լինին և բազումք ուսուն

```
Looking at app App452
Context: և մոլորեպ, Թողուն զնոգս ջինութեան տանց և զյաջողուածս գործոց, և միտ և նանապար ի գինարբուս
Following context:
#D
                             : դե գերին, վասն սիրո և բանկութեան չար
#W
                             : դե գերերեն վասն սիրո և բանկունեան չար
#X
                             : դէ դերերեն։ վասն սիրոյ և բանկունեան չար
#Y
                             : դեգերին, վասն սիրոց և ցանկունեանց չար՝
                             : դէ գերին, վասն սիրոց և ցանկութեան չար
#I #Jer #J #K #Z #V #A #B #0: դեպերին, վասն սիրոյ և բանկունեան չար
Available readings:
1 : դեդերին (#Jer #K #B #I #A #D #J #0 #V #Y #Z )
2 : դէ բերեն (#X )
4 : դէ գերին (#F)
6 : դե գերերերն (#W )
Manuscript details:
3: Witness(es) #X contain a : after #t#####
5: Witness(es) #F contain a , after դէդերին
Available commands:
accept (#)
accept detail (#)
subst detail (#) (value)
spelling (# alternate) (# canonical)
orth (# alternate) (# canonical)
emend (comment text)
note (note text)
next
help
save
quit
 -->
```

Tara L Andrews | University of Oxford | 18 November 2009

ախա » — թերաթարը , այս առաարարագուրակարորեն զկարգ ժամատեղացն։ լինին և բազումք ուսուն

```
Looking at app App452
Context: և մոլորեպ , Թողուն զնոգա ջինութեան տանց և զյաջողուածա գործոց, և փիտ և նանապատ ի գինաթբուս
Following context:
                             : դե գերին, վասն սիրո և բանկունեան չար
#D
#W
                             : դե գերերեն վասն սիրո և բանկունեան չար
#X
                             : դէ բերեն։ վասն սիրոյ և բանկուβեան չար
#Y
                             : դե գերին, վասն սիրոց և ցանկունեանց չար՝
                             : դէ գերին, վասն սիրոց և ցանկութեան չար
#I #Jer #J #K #Z #V #A #B #0: դեպերին, վասն սիրոյ և բանկունեան չար
Available readings:
1 : դեդերին (#Jer #K #B #I #A #D #J #0 #V #Y #Z )
2 : դէ բերեն (#X )
4 : դէ գերին (#F)
6 : դե գերերերն (#W )
Manuscript details:
3: Witness(es) #X contain a : after #t#####
5: Witness(es) #F contain a , after դէդերին
Available commands:
accept (#)
accept detail (#)
subst detail (#) (value)
spelling (# alternate) (# canonical)
orth (# alternate) (# canonical)
emend (comment text)
note (note text)
next
help
save
quit
 -->
```

ց՝ լե ա-ա-նե-այր՝ այր գորեն երիան գովեստից մարդկան։ Որոճալով որոճեն զդիւական երա Հ-ա-ա-նե-այց՝ այր է կերա անկերարի ասելով նել եր կրութը և մերերի ը հորեն և որո

Tara L Andrews | University of Oxford | 18 November 2009

ախա » — թերաթարը , այս առաարարագուրակարորեն զկարգ ժամատեղացն։ լինին և բազումք ուսուն

```
Looking at app App452
Context: և մոլորեպ , Թողուն զնոգա ջինութեան տանց և զյաջողուածա գործոց, և փիտ և նանապատ ի գինաթբուս
Following context:
                             : դե գերին, վասն սիրո և բանկունեան չար
#D
#W
                            : դե գերերեն վասն սիրո և բանկունեան չար
#X
                             : դէ բերեն։ վասն սիրոյ և բանկունեան չար
#Y
                            : դե գերին, վասն սիրոց և ցանկունեանց չար՝
                             : դէ գերին, վասն սիրոց և ցանկութեան չար
#I #Jer #J #K #Z #V #A #B #0: դե բերին, վասն սիրոյ և բանկութեան չար
Available readings:
1 ։ դեպերին (#Jer #K #B #I #A #D #J #0 #V #Y #Z )
2 : դէ բերեն (#X)
4 : դէ գերին (#F)
6 : դե գերէին (#W)
Manuscript details:
3: Witness(es) #X contain a : after #t#####
5: Witness(es) #F contain a , after դէդերին
Available commands:
accept (#)
accept detail (#)
subst detail (#) (value)
spelling (# alternate) (# canonical)
orth (# alternate) (# canonical)
emend (comment text)
note (note text)
next
help
save
quit
 -->
```

յ՝ [երարագրել երինան գովեստից մարդկան։ Որոճալով որոճեն զդիւական երա Հայաստաները է է է է է երև հայան ասելով հե ես կողմող և հեղերի դեղեն և որո

Tara L Andrews | University of Oxford | 18 November 2009

ւրա » — թերթորդ» , մա արապատարարարին զկարգ ժամատեղացն։ լինին և բազումք ուսուն

```
Looking at app App452
Context: և մոլորեպ , Թողուն զնոգա ջինութեան տանց և զյաջողուածա գործոց, և փիտ և նանապազ ի գինաթբուս
Following context:
                             : դե գերին, վասն սիրո և բանկունեան չար
₩W
                            : դե գերերեն վասն սիրո և ցանկունեան չար
μX
                             : դէ բերեն։ վասն սիրոյ և բանկունեան չար
                            : դե գերին, վասն սիրոց և ցանկունեանց չար՝
                             : դէ գերին, վասն սիրոց և ցանկունեան չար
#I #Jer #J #K #Z #V #A #B #0: դե երիրն, վասն սիրոյ և բանկութեան չար
Available readings:
1 ։ դեպերին (#Jer #K #B #I #A #D #J #0 #V #Y #Z )
2 : դէ դերերեն (#X)
4 : դէ գերին (#F)
6 : դե գերէին (#W)
Manuscript details:
3: Witness(es) #X contain a : after #t#####
5: Witness(es) #F contain a , after դէդերին
Available commands:
accept (#)
accept detail (#)
subst detail (#) (value)
spelling (# alternate) (# canonical)
orth (# alternate) (# canonical)
emend (comment text)
note (note text)
next
help
save
quit
-->
```

յ՝ [երարագրել երինան գովեստից մարդկան։ Որոճալով որոճեն զդիւական երա Հայաստաները է է է է է երև հայան ասելով հե ես կողմող և հեղերի դեղեն և որո

Tara L Andrews | University of Oxford | 18 November 2009

իս » — թեթորթը, մա արապարաբարարորեն զկարգ ժամատեղացն։ լինին և բազումք ուսուն

From collation results...

```
<app>
    <rdg wit="#A #B"><w>In</w><rdg>
    < rdg wit = "#C"/>
</app>
<app>
    <rdg wit="#B"><w>the</w></rdg>
    <rdg wit="#A #C"/>
</app>
<app>
    <rdg wit="#B #C"><w>Armenian</w></rdg>
    <rdg wit="#A"/>
</app>
<app>
    <rdg wit="#B"><w>year</w></rdg>
    <rdg wit="#C"><w>era</w></rdg>
    <rdg wit="#A"/>
</app>
<w>401</w>
<app>
    <rdg wit="#A"><w>of</w></rdg>
    <rdg wit="#B #C"/>
</app>
```

Tara L Andrews | University of Oxford | 18 November 2009

ւիս » — թերթթթթ, մա արապարաբարարորեն զկարգ ժամատեղացն։ լինին և բազումք ուսուն

...to edited text

```
<app>
    <lem wit="#A #B"><w>In</w><rdg>
    < rdg wit = "#C"/>
</app>
<app>
    <le>wit="#B"><w>the</w></rdg>
    <rdg wit="#A #C"/>
</app>
<app>
    <rdg wit="#B #C"><w>Armenian</w></rdg>
    <lem wit="#A"/>
</app>
<app>
    <le>wit="#B"><w>year</w></rdg>
    <rdg wit="#C"><w>era</w></rdg>
    <rdg wit="#A"/>
</app>
<w>401</w>
<app>
    <lem wit="#A"><w>of</w></rdg>
    <rdg wit="#B #C"/>
</app>
```

Tara L Andrews | University of Oxford | 18 November 2009

ւիս » — թերաթար, այս արասարագրակարարեն զկարգ ժամատեղացն։ լինին և բազումք ուսուն

...to formatted text and apparatus

In the year 401 of the Armenian era...

In] Com.; I the] B the Armenian, C Armenian, A om.; I year] C era, A om.;

ց՝ Մե գորանելույն գորեն երնան գովեստից մարդկան։ Որոճալով որոճեն զդիւական երա հայաստից մարդկան։ Որոճալով որոճեն զդիւական երա հայաստից մարդկան։

Tara L Andrews | University of Oxford | 18 November 2009

ախա » — թեթաթար» , ար առաարագրագրութեն ըկարգ ժամատեղացն։ լինին և բազումք ուսուն

