

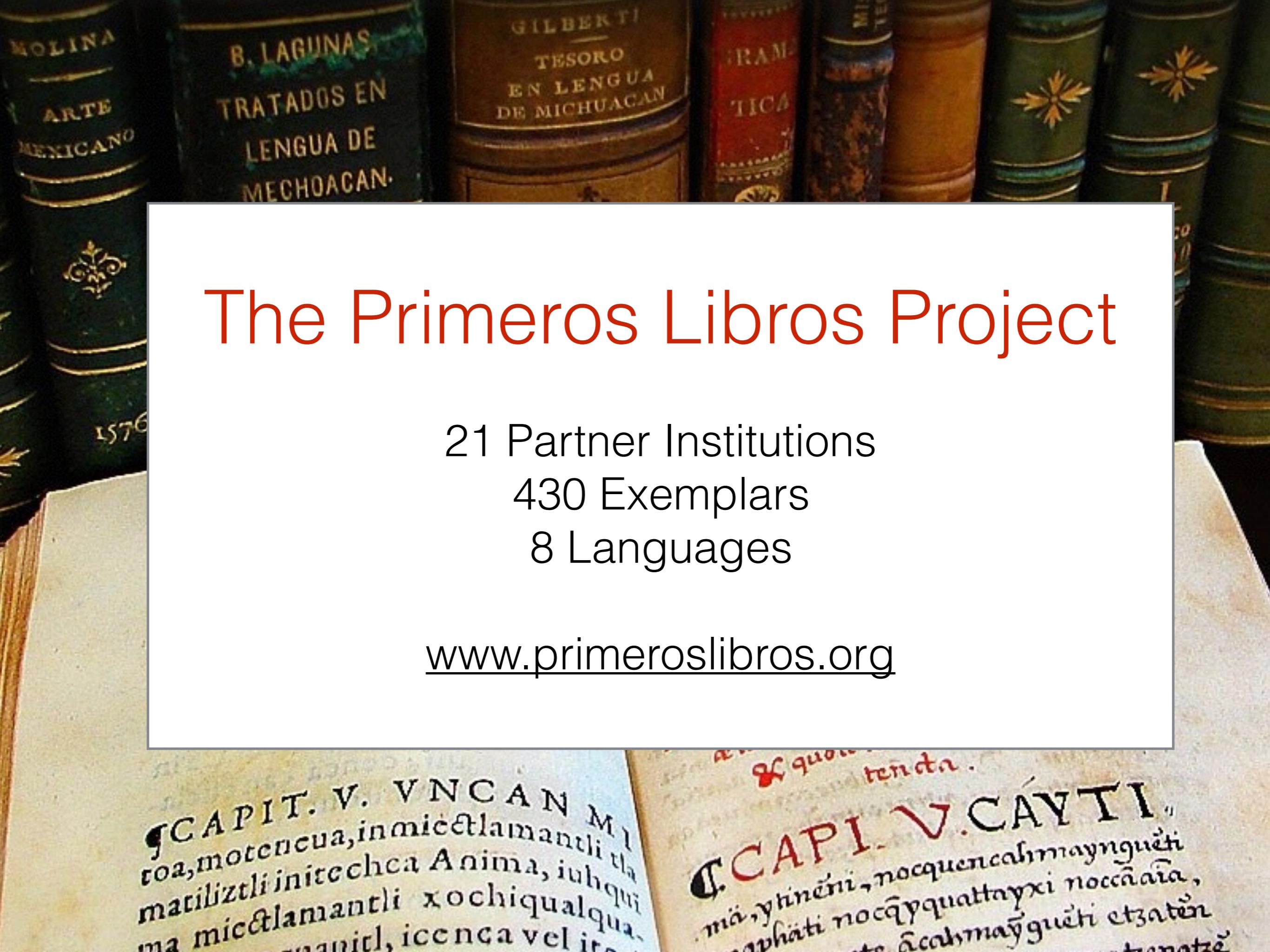
# Reading the First Books

Multilingual, early-modern OCR  
for the Primeros Libros

Hannah Alpert-Abrams

Dic huel nelli teutl Dios in huel imeixtintzitzsancto tre  
q. d. Dios es Padre, Hijo, y Spō |ancto tre  
per|onas, vn |clo Dios verdadero todas tres  
cō la qual reduplicacion |e quita toda dubdoda dubda  
También |e quita con e|tás propo|siciones. ssiones.  
In DTO S.ca Tettatzin, Tepiltzin, Spiritu |  
tto . çan huel iceltzin teutl Dios tlahtohu, Spiritus  
In Dios, cá Tettatzin, Tepiltzin, Spiritu |  
cto, in ieixtin per|onas çan huel iceltzin  
tlahtohuani . Ca in imeixtin per|onas me ca piritus an  
çan huel iceltzin teutl Dios tlahtohuani in ieixtin  
ca piritus an





# The Primeros Libros Project

21 Partner Institutions

430 Exemplars

8 Languages

[www.primeroslibros.org](http://www.primeroslibros.org)

CAPIT. V. VN CAN M  
toa, motenua, in micētlamantli th  
matiliztli in itechca Anima, iuhqui  
ma micētlamantli xochiqualqua  
yanitl, icenca vel it

“<sup>af quo</sup> tenda.  
C. CAPI. V. CÁYTI.  
mä, y hinéni, nocquencahimaynguëti  
nayhiati nocqyquattayxi noccâaia,  
ste ácoh mayguëti etzaten  
tanotzé

¶ Hic. ca y pampa inic yehuantin quicui  
lozque ininemiliztin ynictlaltipac mo  
nemitico in totecuyo Jesu xpo: intley n  
quimochihuilico in quimotemachtilico:  
inictehuantin tictotepotztoquili que y  
nitemachtiltzin. ¶ No. Latlehuatl ino  
quicuiloque in na huintin in Eu angelis-  
tame.

¶ Hic. La yehuatl y nauhtlamantli  
yn Euangilio yn ipan mopia yn ixquich  
totech monequi in tcneltocazque yn tic  
chihuazq. Yhuá intley n tictlalcahuizq  
inticelchihuazque. Yhuán intley n tichi  
huazq yee yxqch totech monequi. Eluh  
yntlacamo tcneltocazque in tlein qui-  
cuilotiaque can nimā ahueltitomaquia-  
tizqui.

¶ No. nicmatiznequi ynici tiazque yn il-  
huicac cui x ga ye ygo tcneltocazque yn  
oquicuilotiaque.

¶ Hic. ca itlacamo tictopielicā tictone  
miliztican in teotenahuatiltzin, y yehua-  
tin in quicauhtiaque caniman ahuel tiaz  
que yn ilhuicac.

christiana. Fo 10.

¶ Nic. ca ypampa inic yehuantin quicui  
lozque ininemiliztin ynictlaltecpacmo  
nemitico in totecuyo Jesu xpo: intleyn  
quimochihuilico in quimotemachtilico:  
inictehuantin tictotepotztoquilizque y  
nitemachtiltzin. ¶ No. Catlehuatl ino  
quicuiloque in nahuintin in Euangeli-  
tame.

¶ Nic. La yehuatl ynauhtlamantli  
yn Euangilio yn ipan mopia yn ixquich  
totech monequi in ticneltocazque yn tic  
chihuazq. Yhuá intleyn tictlalcahuizq  
intic telchihuazque. Yhuan intley tichi  
huazq yye yx'qch totech monequi. Tluh  
yntlacamo ticneltocazque in tlein qui-  
cuitotiaque çan nimá ahueltitomaquiax-  
tizqui.

¶ No. nicmatiznequi yn ic tiazque yn il-  
huicac cuix ça yeyyo ticneltocazque yn  
oquicuitotiaque.

¶ Nic. ca itlacamo tictopielicā tictone  
milizticā initeotenahuatiltzin, yyehuá  
tin inquicauhtiaque caniman ahuel tiaz  
que yn ilhuicac.

b ij

chrijtiana. Fo 10.

¶ Nic. ca ypampa inic yehuantin quicui  
lozque ininemilztzin yn ictlaltecpacmo  
nemitico in totecuyo Jeju x\~po: intleyn  
quimochihuilico in quimotemachtilico:  
inictehuantin tictotepotztoquilizque y  
nitemachtiltzin. ¶ No. Catlehuatl ino  
quicuiloque in nahuintin in Euangeli-  
tame.

¶ Nic. Ca yehuatl yn auhtlamantli  
yn Euangilio yn ipan mopia yn ixquich  
totech monequi in ticneltocazque yn tic  
chihuaz\~q. Yhu\~a intleyn tictlalcahuiz\~q  
intic telchihuazque. Yhuan intley tichi  
huaz\~q yye yx\~qch totech monequi. Tluh  
yntlacamo ticneltocazque in tlein qui-  
cuitotiaque çan nim\~a ahueltitomaquiax-  
tizqui.

¶ No. nicmatiznequi yn ic tiazque yn il-  
huicac cuix ça yeyyo ticneltocazque yn  
oquicuitotiaque.

¶ Nic. ca \~itlacamo tictopielic\~a tictone  
milizticā initeotenahuatiltzin, yyehu\~a  
tin inquicauhtiaque caniman ahuel tiaz  
que yn ilhuicac.

b ij

# Reading the First Books

NEH Office of Digital Humanities  
Digital Humanities Implementation Grant



NATIONAL ENDOWMENT FOR THE  
**Humanities**



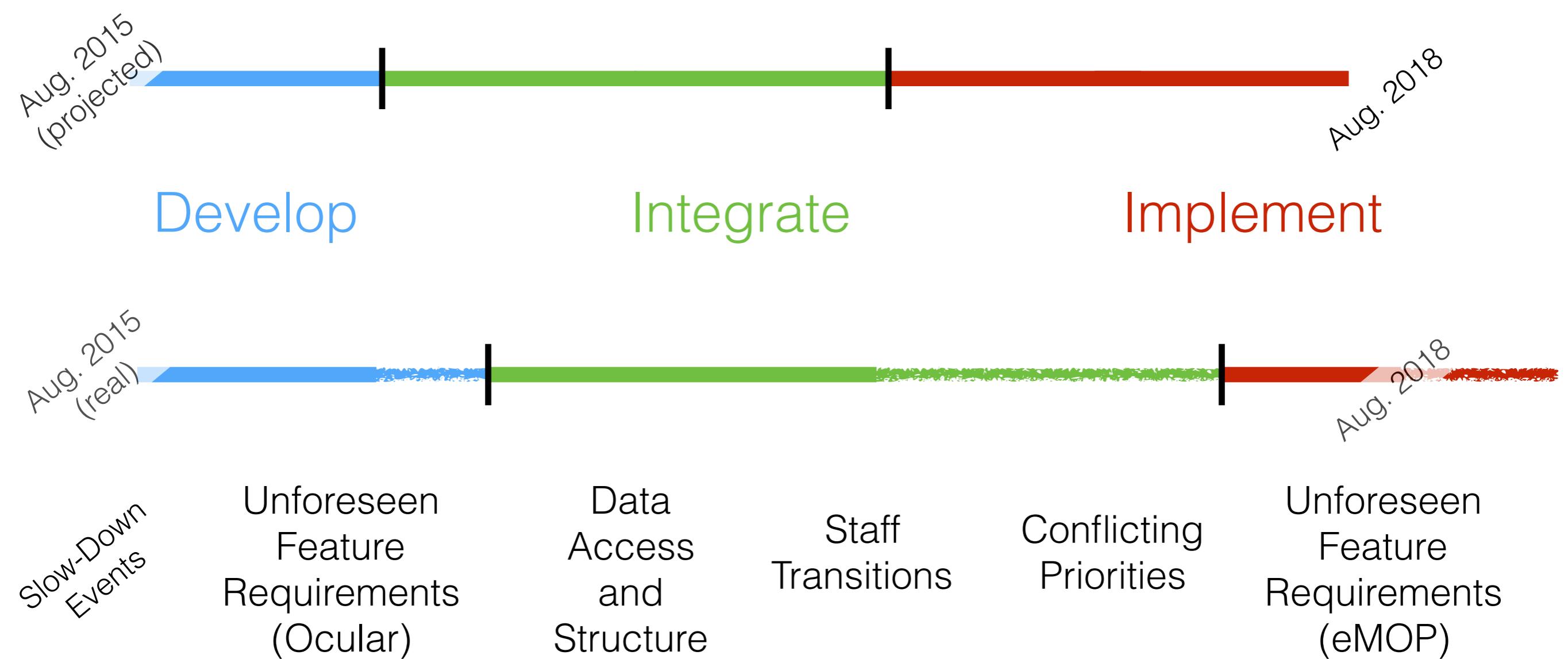
# Reading the First Books

## Project Goals:

1. **Develop:** transcription tools for multilingual early-modern texts
2. **Integrate:** tools with the Early Modern OCR Project interface.
3. **Implement:** transcribe the books in the Primeros Libros project.



# Project Timeline



# Develop

## **Collaborators:**

Taylor Berg-Kirkpatrick, Dan Garrette

## **Goal:**

OCR for multilingual and orthographically variable texts

## **Outcomes:**

Multilingual Tagging

Joint Diplomatic and Normalized Transcription

# Multilingual Tagging

"¶ Quid dicendum de Ocnamacaque. f. de  
las pulqueras, o taberneros, qui vendunt vi-  
num indorum, indis, quod dicitur Octli .&  
etiam de iis qui eis venduntur nō verum His

¶ Quid dicendum de Ocnamacaque. f. de  
las pulqueras, o taberneros, qui vendunt vi-<sup>27</sup>  
num Indorum, indis, quod dicitur Octli :&  
etiam de iis qui eis venduntur viñū verum His

# Diplomatic & Normalized Transcriptions

Normalized text: illustres. E logo os Reitores **que** taõ juntos, derão o mes**mo**  
Historical text: illustres. E logo os Reitores \~qe taõ juntos, derão o mefmo

Normalized text: aviso a seus Collegios, **sendo** em todos, & cada hum delles  
Historical text: avifo a feus Collegios, fendo em todos, & cada hum delles

illustres. E logo os Reitores \~q eraõ juntos, derão o mesmo  
aviso a feus Collegios, sendor em todos, & cada hum delles

# Integrate

## **Collaborators:**

Matt Christy, Trey Dockendorf, Bryan Tarpley,  
Elizabeth Grumbach

## **Goal:**

Integrate Ocular with the Early Modern OCR  
Project

## **Outcomes:**

Ocular interface for eMOP

# Ocular Interface for eMOP

Chrome File Edit View History Bookmarks People Window Help

DH Dashboard dh-db02.tamu.edu/corpus-manager/ Hannah

DH Dashboard Corpus Manager Job Manager Hello, Hannah! Logout

All Documents

Show 25 out of 18 rows

	Document ID	FB ID	Title	Year	Training Set	View
<input type="checkbox"/>	1059	pl_tamu_015	<i>Advertencias</i>	600	45	Pages Details
<input type="checkbox"/>	1060	pl_tamu_017	<i>Advertencias</i>	600	47	Pages Details
<input type="checkbox"/>	930	pl_tamu_016	<i>Advertencias</i>	600	48	Pages Details
<input type="checkbox"/>	931	pl_tamu_018	<i>Advertencias</i>	600	50	Pages Details
<input type="checkbox"/>	1056	pl_tamu_012	<i>Arte mexicano</i>	595	20	Pages Details
<input type="checkbox"/>	974	pl_tamu_009	<i>Arte y diccionario</i>	574	undefined	Pages Details
<input checked="" type="checkbox"/>	1057	pl_tamu_013	<i>Confesionario</i>	599	43	Pages Details
<input type="checkbox"/>	1058	pl_tamu_014	<i>Confesionario en lengua mexicana y castellana</i>	M. Ocharte	1599	44
<input type="checkbox"/>	1175	pl_tamu_010	<i>De constructione octo partium orationis</i>	A. Ricardo	1579	41
<input type="checkbox"/>	994	pl_tamu_006	<i>Doctrina cristiana en lengua mixteca</i>	P. Ocharte	1568	undefined
<input type="checkbox"/>	1150	pl_tamu_011	<i>Estatutos generales de Barcelona, para la familia cismontana, de la orden de n... </i>	P. Ocharte	1585	42
<input type="checkbox"/>	927	pl_tamu_003	<i>Phisica speculatio</i>	J. Pablos	1557	undefined
<input type="checkbox"/>	928	pl_tamu_004	<i>Reverendi patris fratris Bartholomaei à Ledesma ordinis praedicatorum et sacr...</i>	A. Espinosa	1566	undefined
<input type="checkbox"/>	925	pl_tamu_001	<i>Speculum conjugiorum</i>	J. Pablos	1556	undefined
<input type="checkbox"/>	926	pl_tamu_002	<i>Speculum conjugiorum</i>	J. Pablos	1556	undefined
<input type="checkbox"/>	929	pl_tamu_005	<i>Tabula privilegiorum, quae sanctissimus Papa Pius Quintus, concessit fratribus ...</i>	A. Espinosa	1568	undefined
<input type="checkbox"/>	1054	pl_tamu_007	<i>Vocabulario en lengua castellana y mexicana (Volume 1)</i>	A. Espinosa	1571	undefined
<input type="checkbox"/>	1055	pl_tamu_008	<i>Vocabulario en lengua mexicana y castellana (Volume 2)</i>	A. Espinosa	1571	undefined

With Selected: Run a Task

Run a Task

Job Name: pl\_tamu\_013-OCR\_5.29-haa-v03

Job Site: Brazos Supercomputing Cluster

Task: OCR Document with Ocular

Cancel Go

# Ocular Interface for eMOP

A screenshot of a web browser window showing the Ocular Interface for eMOP. The title bar includes the browser name (Chrome), menu items (File, Edit, View, History, Bookmarks, People, Window, Help), system status (43% battery, Mon 4:33 PM, halperta), and a search bar. The address bar shows the URL [dh-db02.tamu.edu/corpus-manager/document-details/?corpus-id=7&document-id=1056](http://dh-db02.tamu.edu/corpus-manager/document-details/?corpus-id=7&document-id=1056). The page header includes links for DH Dashboard, Corpus Manager, Job Manager, Hello, Hannah!, and Logout. The main content area displays the document details for 'Arte mexicana'.

## Arte mexicana

### Details

ID: 1056  
Corpus: FirstBooks  
Author:  
Path: /data/primeros\_libros/pl\_corpus\_1702/pl\_tamu\_012/1000  
printer: P. Ballí  
ocular\_transcription\_jobs: [{"value": "1094", "name": "pl\_tamu\_012-rincon\_OCR\_5-26-17\_haa"}]  
fb\_work\_id: pl\_tamu\_012  
font\_1: roman  
emop\_work\_id: 165  
ocular\_font\_training\_jobs: [{"name": "pl\_tamu\_012-rincon\_train\_170524", "value": "1088"}, {"value": "1089", "name": "pl\_tamu\_012-rincon\_train\_5-26-17\_haa"}]  
language: nahuatl  
training\_set: 20

### Job Files

pl\_tamu\_012-rincon\_train\_5-26-17\_haa(1089)

Ocular Font: [17-05-24\\_spalatnah\\_500000\\_6-v03.fontser](#)  
Ocular Glyph Substitution Model: [17-05-24\\_spalatnah\\_500000\\_6-v03.gsmser](#)  
Ocular Language Model: [17-05-24\\_spalatnah\\_500000\\_6-v03.lmser](#)

### Pages

Page Num: 1  
ID: 134353 Image: [pl\\_tamu\\_012\\_00001-1000.jpg](#)  
fb\_page\_id: pl\_tamu\_012\_00001

pl\_tamu\_012-rincon\_OCR\_5-26-17\_haa(1094)

Ocular Comparison: [pl\\_tamu\\_012\\_00001-1000\\_comparisons.txt](#)  
Ocular DIPL XML: [pl\\_tamu\\_012\\_00001-1000\\_dipl.alto.xml](#)  
Ocular NORM XML: [pl\\_tamu\\_012\\_00001-1000\\_norm.alto.xml](#)  
Ocular Normalized Transcription: [pl\\_tamu\\_012\\_00001-1000\\_transcription\\_normalized.txt](#)  
Ocular Transcription: [pl\\_tamu\\_012\\_00001-1000\\_transcription.txt](#)

Page Num: 2  
ID: 134354 Image: [pl\\_tamu\\_012\\_00002-1000.jpg](#)  
fb\_page\_id: pl\_tamu\_012\_00002

pl\_tamu\_012-rincon\_OCR\_5-26-17\_haa(1094)

Ocular Comparison: [pl\\_tamu\\_012\\_00002-1000\\_comparisons.txt](#)  
Ocular DIPL XML: [pl\\_tamu\\_012\\_00002-1000\\_dipl.alto.xml](#)  
Ocular NORM XML: [pl\\_tamu\\_012\\_00002-1000\\_norm.alto.xml](#)  
Ocular Normalized Transcription: [pl\\_tamu\\_012\\_00002-1000\\_transcription\\_normalized.txt](#)  
Ocular Transcription: [pl\\_tamu\\_012\\_00002-1000\\_transcription.txt](#)

Page Num: 3  
[pl\\_tamu\\_012\\_00003-1000.jpg](#)

pl\_tamu\_012-rincon\_OCR\_5-26-

Ocular Comparison: [pl\\_tamu\\_012\\_00003-1000\\_comparisons.txt](#)

# Ocular Interface for eMOP

Chrome File Edit View History Bookmarks People Window Help

45% Mon 4:27 PM halperta

Hannah

pl\_tamu\_012\_00022-1000.jpg

dh-db02.tamu.edu/data/prime

Hannah

dh-db02.tamu.edu/data/primeros\_libros/pl\_corpus\_1702/p...  
dh-db02.tamu.edu/data/primeros\_libros/corpus/7/jobs/10...

V. S. en persona los a tomado en si porque a llegado V.S. por la vna parte hasta la mar del norte, y por la otra hasta el mar del Sur, q son los vltimos terminos de su obispado, no perdonado qualquier distacia, o aspereza de caminos, ni a los peligros de los Rios, ni a la diuersidad de tantos templos mal fa- nos y contrarios a la salud de V.S. antes lo da todo por bié empleado, por cultiuar y beneficiar por sus manos tã tas y ta preciosas platas como nuestro Señor lea encomendado. Por lo qual qualquiera ministro se deue cõfundir por vna parte de no imitar a quien tiene obligacion, en padecer algo, y por otra parte se due animar a no huir de este pequeño cuidado y sudor que se le pide en deprender qualquiera legua para abilitarse e hazer su ministerio. Sup.  
2.V  
plico

V. E. en persona los á tomado en si porque a llegado v. S. por la vna parte hasta la mar del norte, y por la otra hasta el mar del Sur, o fon los vitimos terminos de su obispado, no perdonado qualquier distacia, o aspereza de caminos, ni a los peligros de los Rios, ni aja diuersidad de tantos templos mal fa- nos y contrarios a la falud de V. E. antes lo da todo por biëemple ado, por cultiuar y beneficiar por sus manos tã tas y ta preciosas platas como nuestro Ieñor lea encomendado. Por lo qual qualquiera ministro se deue cõfundir por vna parte de no imitar a quien tie ne obligacion, en padecer algo, y por ó tra parte se due animara no huirde fle pequeño cuidado y sudor que fe le pĩ de en de prender qualquieta legua Pa ra abilitarse e hazer firministerio. Su p-  
plicó

# Ocular Interface for eMOP

A screenshot of a Google Chrome browser window. The title bar shows "Chrome" and various menu options. The address bar displays the URL "dh-db02.tamu.edu/data/prime" and "dh-db02.tamu.edu/data/primeros\_libros/corpus/7/jobs/1094/all\_transcriptions/1000/pl\_tamu\_012\_00022-1000\_dipl.alto.xml". The main content area shows an XML document with several nested tags and attributes. The status bar at the bottom right indicates "Mon 4:32 PM halperta".

This XML file does not appear to have any style information associated with it. The document tree is shown below.

```
<alto xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:xlink="http://www.w3.org/1999/xlink" xmlns="http://www.loc.gov/standards/alto/ns-v3#" xmlns:emop="http://emop.tamu.edu" xsi:schemaLocation="http://www.loc.gov/standards/alto/ns-v3# http://www.loc.gov/standards/alto/v3/alto.xsd">
  <Description>
    <MeasurementUnit>pixel</MeasurementUnit>
    <sourceImageInformation>
      <fileName>pl_tamu_012_00022-1000.jpg</fileName>
    </sourceImageInformation>
    <OCRProcessing ID="Ocular0.0.3">
      <preProcessingStep/>
      <ocrProcessingStep>
        <processingDateTime>2017-05-26T01:08:38</processingDateTime>
        <processingStepSettings>
          -outputPath /fdata/idhmc/firstbooks-input/corpus/7/jobs/1094 -inputDocListPath /fdata/idhmc/firstbooks-input/corpus/7/jobs/1094/t5j1094p22_document_list.txt -inputFontPath
          /fdata/idhmc/firstbooks-input/corpus/7/jobs/1089/17-05-24_spalatnah_500000_6-v03.fontser -inputLmPath /fdata/idhmc/firstbooks-input/corpus/7/jobs/1089/17-05-
          24_spalatnah_500000_6-v03.lmser -inputGsmPath /fdata/idhmc/firstbooks-input/corpus/7/jobs/1089/17-05-24_spalatnah_500000_6-v03.gsmser -allowGlyphSubstitution true -updateLM
          false -updateGsm false -emissionEngine DEFAULT -outputFormats dipl,norm,alto
        </processingStepSettings>
      <processingSoftware>
        <softwareCreator>
          Taylor Berg-Kirkpatrick, Greg Durrett, Dan Klein, Dan Garrette, Hannah Alpert-Abrams
        </softwareCreator>
        <softwareName>Ocular</softwareName>
        <softwareVersion>0.0.3</softwareVersion>
      </processingSoftware>
      </ocrProcessingStep>
    </OCRProcessing>
  </Description>
  <Layout>
    <Page ID="pl_tamu_012_00022" PHYSICAL_IMG_NR="00022">
      <PrintSpace>
        <TextBlock ID="par_1">
          <TextLine ID="line_1">
            <String ID="word_0" WIDTH="21" CONTENT="V" LANG="spa"/>
            <String ID="word_1" WIDTH="5" CONTENT="." LANG="spa"/>
            <SP WIDTH="11"/>
            <String ID="word_2" WIDTH="12" CONTENT="E" LANG="spa"/>
            <String ID="word_3" WIDTH="6" CONTENT="." LANG="spa"/>
            <SP WIDTH="9"/>
            <String ID="word_4" WIDTH="25" CONTENT="en" LANG="spa"/>
            <SP WIDTH="11"/>
            <String ID="word_5" WIDTH="83" CONTENT="perfona" LANG="spa">
              <ALTERNATIVE PURPOSE="Normalization">persona</ALTERNATIVE>
            </String>
            <SP WIDTH="11"/>
            <String ID="word_6" WIDTH="29" CONTENT="los" LANG="spa"/>
            <SP WIDTH="9"/>
            <String ID="word_7" WIDTH="12" CONTENT="á" LANG="spa"/>
            <SP WIDTH="9"/>
            <String ID="word_8" WIDTH="88" CONTENT="tomado" LANG="spa"/>
            <SP WIDTH="9"/>
          </TextLine>
        </TextBlock>
      </PrintSpace>
    </Page>
  </Layout>
</alto>
```

# Implement

## **Collaborators:**

María Victoria Fernández, Melanie Cofield,  
Audrey Templeton, Ting-Benn Chang, Albert  
Palacios

## **Goal:**

Transcribe Primeros Libros collection

## **Outcomes [ongoing]:**

Transcribed PL Corpus  
Searchable PL Website  
Long-term workflow and preservation

# Transcribed PL Corpus



A ante t, & ñ. 29

Atronar se la muger. nino,cuitapan mauhtia.  
Atronada muger. mocuitapan mauhtiqui.  
Atronara otro con ruido. nite,nacaztititza.  
Atronado assi. tlanacaztititztli.  
¶ Audiencia delos juezes. tlacacoya,tlatzonte-  
coya, tecutlatoloya, teccalli.  
Audiencia hazer. nitecutlatoa,nitla,caqui.  
Auelo o aguelo. colli tecol.  
Aullar. nite,coyoca.  
Aullador. tecoyouani.  
Aullido. tecoyoualiztli.  
Aun. noma,oc noma.  
A vna hazerse. ticcemitoa, tictocentequilia.  
A vna parte y a otra. necoc. nenecoc. necoccampa occi-  
pa ixti. yyuccampaixti. yyuntlapalixti.  
Aun aun? cuixçaoc? cuixçac.  
Aun no. ayamo.  
Aun no has buelto. ayate? cenza ayate? ayaiztiuitz? tica-  
tl'i.  
Aunque. ymanel, ymmanel, ymmaçonel, immaçonel,  
iuh,maciui, maçoiui, maçoneliui, aço, yuh.  
Ausentarse, canapa niyauh, nino,yeltia, ni,choloa, nino,  
tlatia,anixpa.  
Ausencia. canapa yaliztli,netlatiliztli, ateixpa.  
Ausente. ayac,canapayaqui, amo ipxa.  
Autor hzedor dios. techuani,tepiquini, teyocoyani.  
Autoridad de persona. teixmauhtiliztli, mauizticayotl.  
Autoridad tener de persona. nite,ixmauhti,ni,mauiztic.  
Autorizada persona. teixmauhti,teixmamauhti.  
Autoridad de scriptura. tlaneltiloni, tlatolneltiloni.  
Autorizada escriptura. tlaneltillilli, tlatollaneltillilli.  
Autorizar escriptura. ni,tlatolneltilla,na,moxtlatolneltilia  
Autoridad tener para hazer algo. ni,nauatile.  
Auaricia tener. ni,teoyeuacati, ni,tlatlameti, ni,tzotzoca  
teuitzti,ni,tzotzocati,atle niccaualiztlamati, & per metà  
phoram, atle niquixcará, aninocotonrlani.

G

# Transcribed PL Corpus

?? ??

Aronarfe la muger, ni no, cuitlapan mauhtia.  
Atró nada muger, mocuitlapan mauhtiqui.  
Atronar á otro con ruido, ni te nacaztitza,  
á tronado allí, tlanacaztitztli,  
á[ Audiencia de los Juezes. tlacaco ya  
coya. tecutlatolya, teccalli,  
Audiencia hazer, ni tecutlatoa. nitla-caqui.  
Aue io o aguelo. colli tecol.  
Aul lar, nite, coycua.  
At illa dor. tecoyouatit-  
Auilido. tecoyoualiztli,  
Aun. Jsoma, oc noma.  
A yn a hazerfe, ticcemitoa, tictocentequilia,  
Ayna parte y a otra, necoc, nenecoc, necocca  
pa ixti yyuccampaixti yyuntlapalixti.  
Aun aunçcuix çä ocçc97xçac.  
Aun no. ayamo.  
Aun no ha abuelto, ayateçcen casayateçayaíz  
tli.  
Aun que, yntla nel, ymmanel, ymmaçonel, ii  
iuh maciui, maçoiui, maçoneltin, aço yuh.  
Aufentat fe. canapa niyauh. nino-yeltia, ni-cl  
tlatia. anñxpa,  
Aufencia. canapa yaliztli, netlatiliztli, atëixpa,  
Aufente. ayac. canapa yaqui amo ixpa.  
Autor hazedor dios. techiuani. tepi quini, te yc  
Autoridad de perfona, teixmauhtiliztli. mau  
Autoridad tener de perfona, ni tesix mauhti, ni  
Autorizada perfona, teixmauhti, teixmamauj  
Autoridad de feriptura. tlaneltililoni, tlatolti  
Autorizada ecriptura, tlaneltililli, tlatollanel  
Autoriza referiptura, ni-tlatol neltilia. na, moxi  
Autoridad tener para hazer algo, ni nauatile,  
Auaticia tener, ni ste oyeuacatl, nist latjanteti,  
teuitzti-ni, tzotzocati. atle niccaualiztlamati.  
photam. atle niquixcatia. anñnocotontlani.

A ante f, & ñ. 29

Atronar se la muger. nino,cuitlapan mauhtia.  
Atronada muger, mocuitlapan mauhtiqui.  
Atronara otro con ruido, nite,nacaztitza,  
Atronado assi. tlanacaztitztli.  
¶ Audiencia delos juezes. tlacacoya.tlatzonte-  
coya, tecutlatolya. teccalli.  
Audiencia hazer, nitecutlatoa.nitla,caqui,  
Auelo o aguelo. colli tecol.  
Aullar. nite,coycua.  
Aullador. tecoyouani.  
Aullido. tecoyoualiztli.  
Aun. noma. oc noma,  
A vna hazerse, ticcemitoa, tictocentequilia,  
A vna parte y a otra, necoc. nenecoc necoccampa occ-  
pa ixti.yyuccampaixti.yyuntlapalixti.  
Aun aun? cuixçaoc? cuixçac.  
Aun no. ayamo.  
Aun nohas buelto. ayate? cenza ayate? ayaiztiuitz? tica  
tli.  
Aun que. yntlanel, ymmanel, ymmaçonel, immaçonel-  
iuh.maciui. maçoiui. maçoneliui. aço. yuh.  
Ausentar se, canapa niyauh. nino,yeltia, ni,choloa. nino,  
tlatia.anñxpa.  
Ausencia. canapa yaliztli.netlatiliztli. ateixpa.  
Ausente. ayac.canapayaqui amo ixpa.  
Autor hazedor dios. techiuani.tepiquini. teyocoyani.  
Autoridad de persona. teixmauhtiliztli. mauizticayotl.  
Autoridad tener de persona. nite,ixmauhti,ni,mauiztic.  
Autorizada persona. teixmauhti.teixmamauhti.  
Autoridad de scriptura. tlaneltililoni, tlatolneltihoni.  
Autorizada ecriptura. tlaneltililli, tlatollaneltillli.  
Autorizar ecriptura. ni,tlatolneltilia,na,moxtlatolneltilia  
Autoridad tener para hazer algo, ni.nauatile.  
Auaticia tener, ni,teoyeuacati, ni,tlatlameti. ni,tzotzoca  
teuitzti.ni,tzotzocati,atle niccaualiztlamati. & permeter  
phoram. atle niquixcatia. anñnocotontlani.

G

# Transcribed PL Corpus

*Atronada muger, mocuitapan mauhtiqui.*

Atró nada muger,  
mocuitapan mauhtiqui.

Atrobar a otro con ruido, nite,nacaztitiza,

Atronar á otro con ruido,  
ni te nacaztitza,

Atronado así, tlanacaztitlitztli.

á tronado allí, tlanacaztitlztli,

**Audiencia de los jueces. tlacacoya**

# ñ[ Audiencia de los Juezes. tlacaco ya

# Transcribed PL Corpus

*Atronada muger, mocuitapan mauhtiqui.*

Atró nada muger,  
mocuitapan mauhtiqui.

Atrobar a otro con ruido, nite,nacaztitiza,

Atronar á otro con ruido,  
ni te nacaztitza,

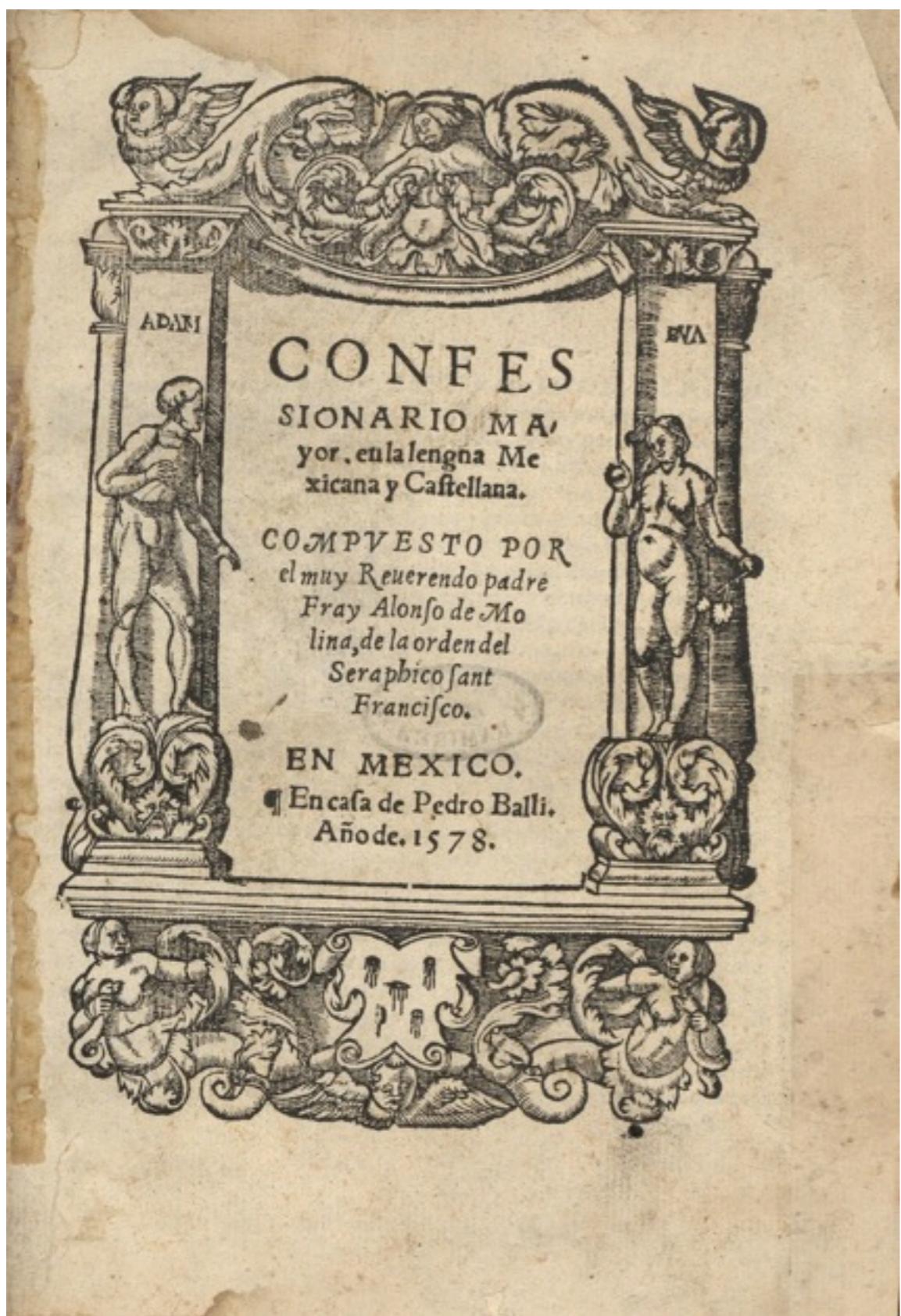
Atronado affi, tlanacaztitlitztli.

á tronado allí, tlanacaztitlztli,

**Audiencia de los jueces. tlacacoya**

# ñ[ Audiencia de los Juezes. tlacaco ya

# Transcribed PL Corpus



En lengua Mexicana y Castellana. 10

Ynic amo mopinauhtiz, ma  
uilquirtiz, inic amo quich-  
teccatlapiquizque, auhcen  
camomauhtia, quimacaci  
yntlatzacuiltyloca:çanno  
yuh motechmonequi, ynic  
cencatimomociniz, ifquich  
motlapal ticchiuaç inic mu  
chi tictemoz, tiquilnami-  
quiz imotlatlacol, inic vel  
timoyolcuitiz, inic titlapo  
uaç yppantzincototecuiyo  
dios, yuan inixpá sacerdo  
te, inic amo tipinauhtiloz,  
inic amo titlatzacuiltyloz,  
çantinaniçouaz yuan tic  
maquixtizinmanimá. Auh  
intlacamo achtopa, tiquil-  
namiqiz, tictentlaliz imo  
tlatlacol, vel nelli ycti qui  
tlacozi mone yolmelaua  
liç, amoma nelli yeticpale  
uiç imanimá, çan occenca  
ye ictictoliniz, amoma mo  
paleuilocamochiuaz imo  
neyolcuitiliz, çan mopina  
uhtiloca, morelchiualoca  
yez, yuá cemicacmixnaua  
tilocamochiuaz. Auh inic  
amotiquitlacoz, mamocen  
yollocopa achto xictemo,  
yuan tictentlali yvirquich  
uáten q dio ruincueta de si  
y por no ser afretado nica  
er en algúia falta, y por qno  
le tégá por ladron; y tiene  
grá miedo y temor del casti  
go que se le podria dar: así  
tienes necesidad de tener  
grá solicitud y cuidado, y  
dbazertodatu posibildad  
para buscar y pésar todos  
tus pedos para te pfessar d  
uidamente, y para dar buéa  
cuéta dlate nro señor dios  
y dlate el sacerdote, po:q  
no seas auergóçado y casti  
gado, mas átes recibas ho  
ra, y alcáces la saluació de  
tu aia. Y si pmero no truje  
res ala memoria tus pedos  
y los recogieres, ciertame  
te sera tu cōfesió invalida,  
falta y no verdadera, ni me  
nos sera útil y puecposa a  
tu aia, átes la astigiras mu  
chos mas, y nosera para tu  
fauor la tal cōfession mas  
para tu cōfusion, y cōdena  
ció y para q eternamente se  
as desechado. Y por q nosea  
íperfecta la tal cōfesió mu  
ya, busca pmero d todo tu  
corazón y ayúta eno todos  
B 2 tus

# Transcribed PL Corpus

**ynic amo mopīnāuhtij, ma nāten q̄dior uincuētaoe-lī**

ynic amo mopīnāuhtij, ma nāten. q̄Dior uincuētaoe-lī

**uilquirtij, inic amo quicb - y poz no fer atrétado nica**

uilquirtij, inic amo quicb - y poz no fer atrétado nica

**teccatlapiquijque, auh cen eren algūa falta, y poz q̄n.o**

teccatlapiquijque, auh cen eren algua falta, y poz q̄n.o

**ca momaubtia, quimacaci le tēgā poz ladron ; ytiene**

ca momaubtia, quimacaci le tēgā poz ladron ; ytiene

# Transcribed PL Website

The image shows a screenshot of a web browser displaying two versions of a page from a historical document. The document is titled "Arte mexicana" and is attributed to Rincón, Antonio, 1555-1601, published in 1595, and is part of the Benson Latin American Collection. The left side shows the original manuscript page with handwritten Spanish text. The right side shows a transcription of the same text into printed Spanish. Both pages include navigation controls (back, forward, search) and download options (PDF, TEXT). The browser interface at the top includes the title "Detail View - Primeros Libros", the date "Mon 4:48 PM", and the user "halperta".

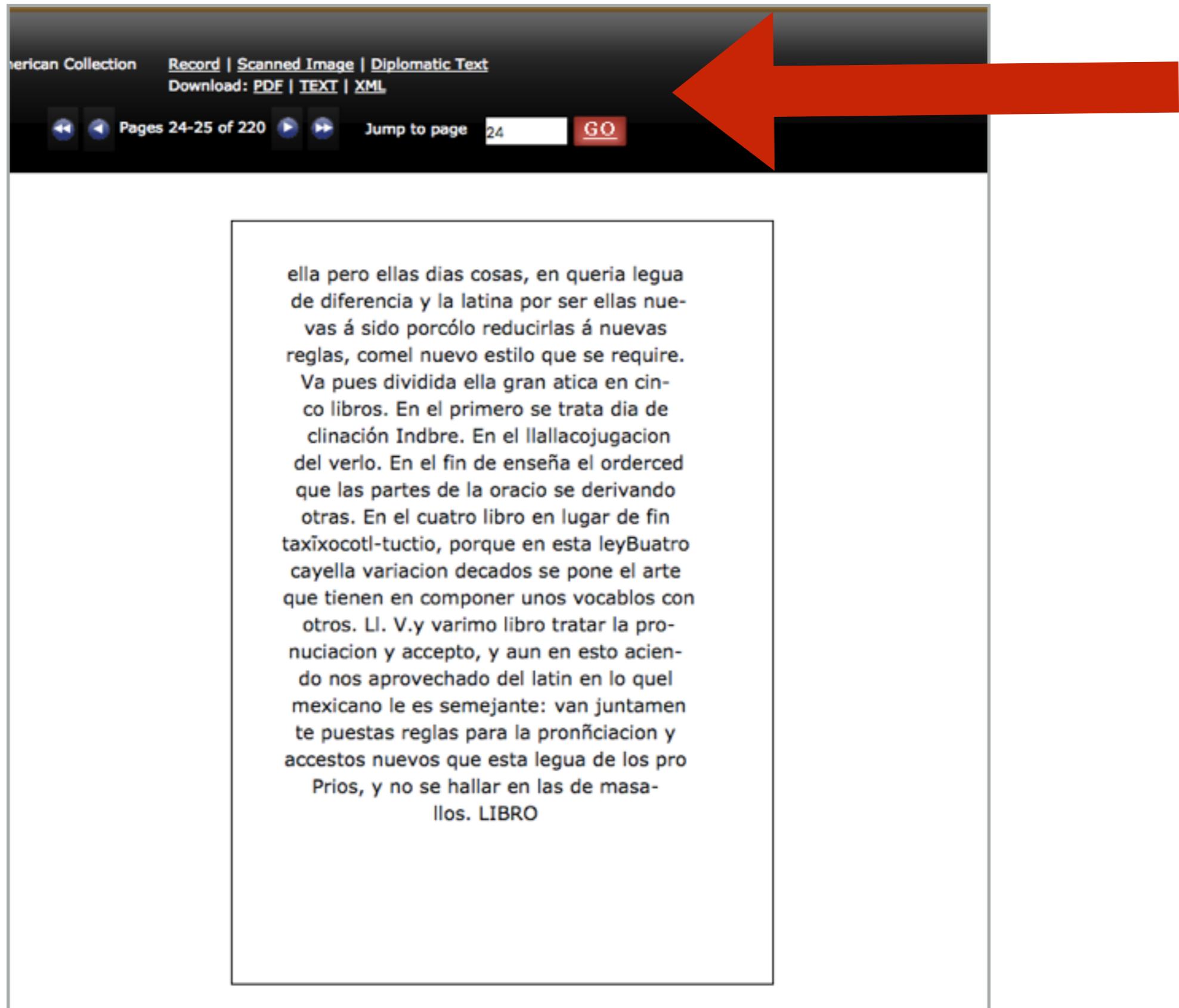
**Arte mexicana**  
Rincón, Antonio, 1555-1601 - Published: 1595 Benson Latin American Collection Record | Normalized Text  
Download: PDF | TEXT

**Arte mexicana**  
Rincón, Antonio, 1555-1601 - Published: 1595 Benson Latin American Collection Record | Normalized Text  
Download: PDF | TEXT

das las partes de la oracion poniendo  
les nóbres conforme a sus officios y ca  
lidades como tambien por hauer mo  
stra do la variacion de las partes, que  
entre ellas son variables, enseñando el  
artificio de jútar, o cōstruir lasvnas cō  
las otras, y finalmēte dādo reglas dela  
propria y buena pronunciació. Por lo  
qual hauiédo yo de escreuir Arte para  
depréder y enseñar la léguia mexicana  
no me parecio apartarme del ordinario  
camino por dōde procede la léguia  
latina, q̄ es mas sabida entre nosotros,

das las partes de la oracion poniendo  
les nobres conforme a sus oficios yel  
lidades como tambien por hauermo  
fixa iola variacton de las partes, que  
entre ellas son variables, enseñando el  
artificio de jurar, o cōstruir las vn ascā  
las otras, y finalmēte dado reglas de la  
propria y buena pronunciacio. Por lo  
qual hauiédo yo de escreuir Atte para  
de prēder y enseñar la legua mexican,  
no me parecio apartar me del ordinario  
camino por dōde procede la legua  
latina, q̄ es mas sabida entre nosotros,  
ni tāpoco me he q̄itido obligata legua  
dltodo sus reglas, porq̄ieria llenar muj  
fuera de propoſitos y como dizē, de los  
cabellos muchas coſas q̄jacapidē mul  
diferētes preceptos. Demanera q̄jear  
ilo q̄mee podido apronechará la gr.  
matica latma fiēpte meyte attim ados  
T "?"

# Transcribed PL Website



A screenshot of a transcribed PL website page. At the top, there is a navigation bar with links for 'American Collection', 'Record | Scanned Image | Diplomatic Text', and download options for 'PDF | TEXT | XML'. Below the navigation bar, there are page navigation buttons ('Pages 24-25 of 220', arrows for previous/next pages, and a 'Jump to page' input field set to '24') and a red 'GO' button. A large red arrow points from the right side of the header towards the main content area. The main content area contains two blocks of text in Spanish. The first block discusses the reduction of Nahuatl words to new Spanish rules. The second block describes the division of the Nahuatl grammar into five books, detailing the treatment of declension, conjugation, and other grammatical features.

ella pero ellas dias cosas, en queria legua  
de diferencia y la latina por ser ellas nue-  
vas á sido porcólogo reducirlas á nuevas  
reglas, comel nuevo estilo que se require.  
Va pues dividida ella gran atica en cin-  
co libros. En el primero se trata dia de  
clinación Indbre. En el llallacojugacion  
del verlo. En el fin de enseña el orderced  
que las partes de la oracio se derivando  
otras. En el cuatro libro en lugar de fin  
taxīxocotl-tuctio, porque en esta leyBuatro  
cayella variacion decados se pone el arte  
que tienen en componer unos vocablos con  
otros. LI. V.y varimo libro tratar la pro-  
nunciacion y accepto, y aun en esto acien-  
do nos aprovechado del latin en lo quel  
mexicano le es semejante: van juntamen-  
te puestas reglas para la pronñciacion y  
accesto nuevos que esta legua de los pro-  
Prios, y no se hallar en las de masa-  
llos. LIBRO

# Reading the First Books

## Project Goals:

1. **Develop:** transcription tools for multilingual early-modern texts
2. **Integrate:** tools with the Early Modern OCR Project interface.
3. **Implement:** transcribe the books in the Primeros Libros project.



# Reading the First Books

## Future:

Complete transcriptions  
Searchable interface  
Downloadable corpus  
Diplomatic digital editions  
Corpus analytics

