

Introduction to ELAN

Bradley McDonnell

March 10, 2020

Goals

1. Navigate ELAN files
2. Produce/edit annotations in ELAN
3. Create/edit tiers in ELAN
4. Conceptualize & annotate your own recordings

Goals

1. Navigate ELAN files
2. Produce/edit annotations in ELAN
3. Create/edit tiers in ELAN
4. Conceptualize & annotate your own recordings

Goals

1. Navigate ELAN files
2. Produce/edit annotations in ELAN
3. Create/edit tiers in ELAN
4. Conceptualize & annotate your own recordings

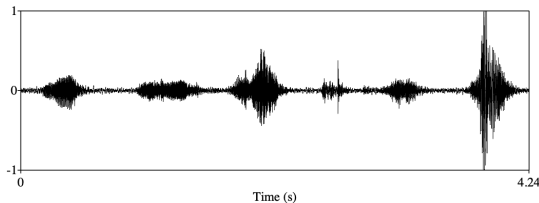
Goals

1. Navigate ELAN files
2. Produce/edit annotations in ELAN
3. Create/edit tiers in ELAN
4. Conceptualize & annotate your own recordings

what is a time-aligned transcript?

1. Time-aligned transcriptions match:

- ▶ sections of text (ANNOTATIONS) to
- ▶ pieces of audio and/or video recordings (MEDIA).

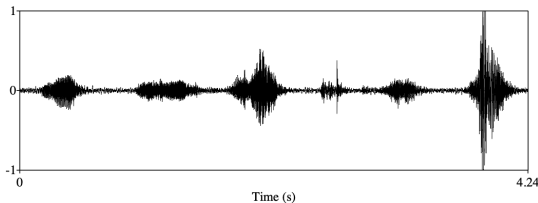


Jack: Hola.
Kristen: Hey!
Jake: What up.
Jack: How's it going?
Jake: Good.

what is a time-aligned transcript?

1. Time-aligned transcriptions match:

- ▶ sections of text (ANNOTATIONS) to
- ▶ pieces of audio and/or video recordings (MEDIA).

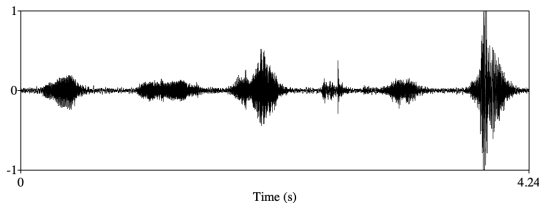


Jack: Hola.
Kristen: Hey!
Jake: What up.
Jack: How's it going?
Jake: Good.

what is a time-aligned transcript?

1. Time-aligned transcriptions match:

- ▶ sections of text (ANNOTATIONS) to
- ▶ pieces of audio and/or video recordings (MEDIA).

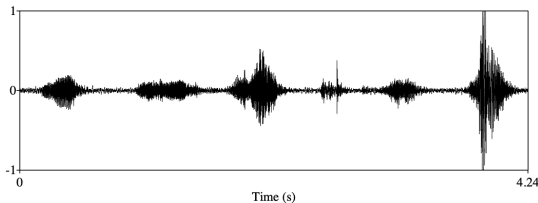


Jack: Hola.
Kristen: Hey!
Jake: What up.
Jack: How's it going?
Jake: Good.

what is a time-aligned transcript?

1. Time-aligned transcriptions match:

- ▶ sections of text (ANNOTATIONS) to
- ▶ pieces of audio and/or video recordings (MEDIA).

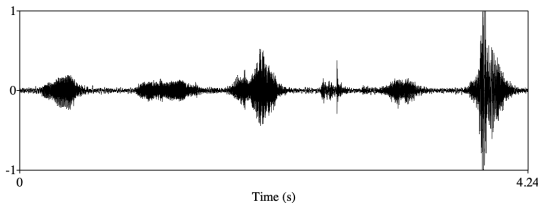


Jack: Hola.
Kristen: Hey!
Jake: What up.
Jack: How's it going?
Jake: Good.

what is a time-aligned transcript?

1. Time-aligned transcriptions match:

- ▶ sections of text (ANNOTATIONS) to
- ▶ pieces of audio and/or video recordings (MEDIA).

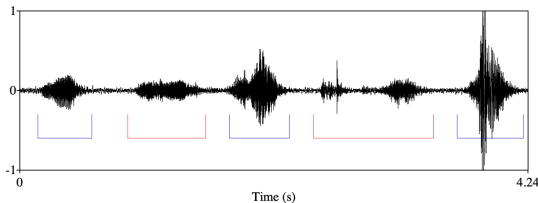


Jack:	Hola.
Kristen:	Hey!
Jake:	What up.
Jack:	How's it going?
Jake:	Good.

what is a time-aligned transcript?

1. Time-aligned transcriptions match:

- ▶ sections of text (ANNOTATIONS) to
- ▶ pieces of audio and/or video recordings (MEDIA).



Jack:	Hola.
Kristen:	Hey!
Jake:	What up.
Jack:	How's it going?
Jake:	Good.

1. What are some common time-aligned transcripts?

- ▶ closed captioning
- ▶ subtitles
- ▶ ...

1. What are some common time-aligned transcripts?

- ▶ closed captioning
- ▶ subtitles
- ▶ ...

1. What are some common time-aligned transcripts?

- ▶ closed captioning
- ▶ subtitles
- ▶ ...

1. What are some common time-aligned transcripts?

- ▶ closed captioning
- ▶ subtitles
- ▶ ...

why bother?

Your language resources will be much richer. Your resources:

1. are searchable
2. are verifiable
3. reinforce your language as a living language
4. match orthography to pronunciation for literacy
5. ...

why bother?

Your language resources will be much richer. Your resources:

1. are searchable
2. are verifiable
3. reinforce your language as a living language
4. match orthography to pronunciation for literacy
5. ...

why bother?

Your language resources will be much richer. Your resources:

1. are searchable
2. are verifiable
3. reinforce your language as a living language
4. match orthography to pronunciation for literacy
5. ...

why bother?

Your language resources will be much richer. Your resources:

1. are searchable
2. are verifiable
3. reinforce your language as a living language
4. match orthography to pronunciation for literacy
5. ...

why bother?

Your language resources will be much richer. Your resources:

1. are searchable
2. are verifiable
3. reinforce your language as a living language
4. match orthography to pronunciation for literacy
5. ...

why bother?

Your language resources will be much richer. Your resources:

1. are searchable
2. are verifiable
3. reinforce your language as a living language
4. match orthography to pronunciation for literacy
5. ...

why bother?

Your language resources will be much richer. Your resources:

1. are searchable
2. are verifiable
3. reinforce your language as a living language
4. match orthography to pronunciation for literacy
5. ...

what is ELAN?

ELAN is a tool for producing time-aligned transcripts of audiovisual recordings.

1. Developed & maintained by the Max Planck Institute for Psycholinguistics in Nijmegen.
2. ELAN stands for **E**UDICO Linguistic **A**nnotator.
3. EUDICO stands for the **E**uropean **D**istributed **C**orpus Project.
4. It is developed alongside a number of other software tools that work in concert with ELAN (i.e., Arbil for metadata & Lexus for lexical databases).

what is ELAN?

ELAN is a tool for producing time-aligned transcripts of audiovisual recordings.

1. Developed & maintained by the Max Planck Institute for Psycholinguistics in Nijmegen.
2. ELAN stands for **E**UDICO Linguistic **A**nnotator.
3. EUDICO stands for the **E**uropean **D**istributed **C**orpus Project.
4. It is developed alongside a number of other software tools that work in concert with ELAN (i.e., Arbil for metadata & Lexus for lexical databases).

what is ELAN?

ELAN is a tool for producing time-aligned transcripts of audiovisual recordings.

1. Developed & maintained by the Max Planck Institute for Psycholinguistics in Nijmegen.
2. ELAN stands for **E**UDICO Linguistic **A**nnotator.
3. EUDICO stands for the **E**uropean **D**istributed **C**orpus Project.
4. It is developed alongside a number of other software tools that work in concert with ELAN (i.e., Arbil for metadata & Lexus for lexical databases).

what is ELAN?

ELAN is a tool for producing time-aligned transcripts of audiovisual recordings.

1. Developed & maintained by the Max Planck Institute for Psycholinguistics in Nijmegen.
2. ELAN stands for **E**UDICO **L**inguistic **A**nnotator.
3. EUDICO stands for the **E**uropean **D**istributed **C**orpus Project.
4. It is developed alongside a number of other software tools that work in concert with ELAN (i.e., Arbil for metadata & Lexus for lexical databases).

what is ELAN?

ELAN is a tool for producing time-aligned transcripts of audiovisual recordings.

1. Developed & maintained by the Max Planck Institute for Psycholinguistics in Nijmegen.
2. ELAN stands for **E**UDICO **L**inguistic **A**nnotator.
3. EUDICO stands for the **E**uropean **D**istributed **C**orpus Project.
4. It is developed alongside a number of other software tools that work in concert with ELAN (i.e., Arbil for metadata & Lexus for lexical databases).

why use ELAN?

1. ELAN uses non-proprietary, standards-compliant output.
 - ▶ It uses **unicode** so that your fonts display correctly.
 - ▶ It uses **XML** so that your document is long-lasting and archive ready. It conforms to well documented XML schema and is human eyeball-readable.
2. ELAN documents can be exported to a wide range of useful products for a wide range of audiences.
 - ▶ ELAN can be exported as a subtitled video for communities.
 - ▶ ELAN can also be exported to Praat, FLEx, Toolbox, Subtitled text, etc...for linguists.

why use ELAN?

1. ELAN uses non-proprietary, standards-compliant output.
 - ▶ It uses **unicode** so that your fonts display correctly.
 - ▶ It uses **XML** so that your document is long-lasting and archive ready. It conforms to well documented XML schema and is human eyeball-readable.
2. ELAN documents can be exported to a wide range of useful products for a wide range of audiences.
 - ▶ ELAN can be exported as a subtitled video for communities.
 - ▶ ELAN can also be exported to Praat, FLEx, Toolbox, Subtitled text, etc...for linguists.

why use ELAN?

1. ELAN uses non-proprietary, standards-compliant output.
 - ▶ It uses **unicode** so that your fonts display correctly.
 - ▶ It uses **XML** so that your document is long-lasting and archive ready. It conforms to well documented XML schema and is human eyeball-readable.
2. ELAN documents can be exported to a wide range of useful products for a wide range of audiences.
 - ▶ ELAN can be exported as a subtitled video for communities.
 - ▶ ELAN can also be exported to Praat, FLEx, Toolbox, Subtitled text, etc...for linguists.

why use ELAN?

1. ELAN uses non-proprietary, standards-compliant output.
 - ▶ It uses **unicode** so that your fonts display correctly.
 - ▶ It uses **XML** so that your document is long-lasting and archive ready. It conforms to well documented XML schema and is human eyeball-readable.
2. ELAN documents can be exported to a wide range of useful products for a wide range of audiences.
 - ▶ ELAN can be exported as a subtitled video for communities.
 - ▶ ELAN can also be exported to Praat, FLEx, Toolbox, Subtitled text, etc...for linguists.

why use ELAN?

1. ELAN uses non-proprietary, standards-compliant output.
 - ▶ It uses **unicode** so that your fonts display correctly.
 - ▶ It uses **XML** so that your document is long-lasting and archive ready. It conforms to well documented XML schema and is human eyeball-readable.
2. ELAN documents can be exported to a wide range of useful products for a wide range of audiences.
 - ▶ ELAN can be exported as a subtitled video for communities.
 - ▶ ELAN can also be exported to Praat, FLEx, Toolbox, Subtitled text, etc...for linguists.

why use ELAN?

1. ELAN uses non-proprietary, standards-compliant output.
 - ▶ It uses **unicode** so that your fonts display correctly.
 - ▶ It uses **XML** so that your document is long-lasting and archive ready. It conforms to well documented XML schema and is human eyeball-readable.
2. ELAN documents can be exported to a wide range of useful products for a wide range of audiences.
 - ▶ ELAN can be exported as a subtitled video for communities.
 - ▶ ELAN can also be exported to Praat, FLEx, Toolbox, Subtitled text, etc...for linguists.

why use ELAN?

1. ELAN uses non-proprietary, standards-compliant output.
 - ▶ It uses **unicode** so that your fonts display correctly.
 - ▶ It uses **XML** so that your document is long-lasting and archive ready. It conforms to well documented XML schema and is human eyeball-readable.
2. ELAN documents can be exported to a wide range of useful products for a wide range of audiences.
 - ▶ ELAN can be exported as a subtitled video for communities.
 - ▶ ELAN can also be exported to Praat, FLEx, Toolbox, Subtitled text, etc...for linguists.

how is ELAN better than a printed transcript?

1. There are many uses for printed transcripts.
2. In the past, the printed transcript was considered the primary source:
 - ▶ Whatever the printed transcript could not capture was not included in the discussion.
 - ▶ An accompanying recording was just considered to be an added bonus.
3. In ELAN the text & the recording are unified.
 - ▶ This opens up a number of possibilities for research & learning.
4. The printed transcript is one of the many possible exports in ELAN.

how is ELAN better than a printed transcript?

1. There are many uses for printed transcripts.
2. In the past, the printed transcript was considered the primary source:
 - ▶ Whatever the printed transcript could not capture was not included in the discussion.
 - ▶ An accompanying recording was just considered to be an added bonus.
3. In ELAN the text & the recording are unified.
 - ▶ This opens up a number of possibilities for research & learning.
4. The printed transcript is one of the many possible exports in ELAN.

how is ELAN better than a printed transcript?

1. There are many uses for printed transcripts.
2. In the past, the printed transcript was considered the primary source:
 - ▶ Whatever the printed transcript could not capture was not included in the discussion.
 - ▶ An accompanying recording was just considered to be an added bonus.
3. In ELAN the text & the recording are unified.
 - ▶ This opens up a number of possibilities for research & learning.
4. The printed transcript is one of the many possible exports in ELAN.

how is ELAN better than a printed transcript?

1. There are many uses for printed transcripts.
2. In the past, the printed transcript was considered the primary source:
 - ▶ Whatever the printed transcript could not capture was not included in the discussion.
 - ▶ An accompanying recording was just considered to be an added bonus.
3. In ELAN the text & the recording are unified.
 - ▶ This opens up a number of possibilities for research & learning.
4. The printed transcript is one of the many possible exports in ELAN.

how is ELAN better than a printed transcript?

1. There are many uses for printed transcripts.
2. In the past, the printed transcript was considered the primary source:
 - ▶ Whatever the printed transcript could not capture was not included in the discussion.
 - ▶ An accompanying recording was just considered to be an added bonus.
3. In ELAN the text & the recording are unified.
 - ▶ This opens up a number of possibilities for research & learning.
4. The printed transcript is one of the many possible exports in ELAN.

how is ELAN better than a printed transcript?

1. There are many uses for printed transcripts.
2. In the past, the printed transcript was considered the primary source:
 - ▶ Whatever the printed transcript could not capture was not included in the discussion.
 - ▶ An accompanying recording was just considered to be an added bonus.
3. In ELAN the text & the recording are unified.
 - ▶ This opens up a number of possibilities for research & learning.
4. The printed transcript is one of the many possible exports in ELAN.

how is ELAN better than a printed transcript?

1. There are many uses for printed transcripts.
2. In the past, the printed transcript was considered the primary source:
 - ▶ Whatever the printed transcript could not capture was not included in the discussion.
 - ▶ An accompanying recording was just considered to be an added bonus.
3. In ELAN the text & the recording are unified.
 - ▶ This opens up a number of possibilities for research & learning.
4. The printed transcript is one of the many possible exports in ELAN.

what does ELAN NOT do?

1. ELAN is NOT a media editor: it cannot change your audio/video files.
 - ▶ This is actually a good thing ...we won't inadvertently change our precious recordings.
 - ▶ There are many other programs that can do this (i.e., Audacity).
2. ELAN is NOT a text editor: its annotations and outputs are in plain text.
 - ▶ In order to format text with underlining, *italics*, SMALL CAPS, or **bolding**, you will need to use a text editor after you export your annotations.

what does ELAN NOT do?

1. ELAN is NOT a media editor: it cannot change your audio/video files.
 - ▶ This is actually a good thing ...we won't inadvertently change our precious recordings.
 - ▶ There are many other programs that can do this (i.e., Audacity).
2. ELAN is NOT a text editor: its annotations and outputs are in plain text.
 - ▶ In order to format text with underlining, *italics*, SMALL CAPS, or **bolding**, you will need to use a text editor after you export your annotations.

what does ELAN NOT do?

1. ELAN is NOT a media editor: it cannot change your audio/video files.
 - ▶ This is actually a good thing ...we won't inadvertently change our precious recordings.
 - ▶ There are many other programs that can do this (i.e., Audacity).
2. ELAN is NOT a text editor: its annotations and outputs are in plain text.
 - ▶ In order to format text with underlining, *italics*, SMALL CAPS, or **bolding**, you will need to use a text editor after you export your annotations.

what does ELAN NOT do?

1. ELAN is NOT a media editor: it cannot change your audio/video files.
 - ▶ This is actually a good thing ...we won't inadvertently change our precious recordings.
 - ▶ There are many other programs that can do this (i.e., Audacity).
2. ELAN is NOT a text editor: its annotations and outputs are in plain text.
 - ▶ In order to format text with underlining, *italics*, SMALL CAPS, or **bolding**, you will need to use a text editor after you export your annotations.

the panes of ELAN

The screenshot displays the ELAN software interface with several panes labeled:

- Menu Bar:** Located at the top, containing File, Edit, Annotation, Tier, Type, Search, View, Options, Window, and Help.
- Grid:** A table showing a list of annotations with columns for Start Time, End Time, and Duration.
- Specific Viewers:** A label pointing to the Grid pane.
- Video Viewer:** A central pane displaying a video of two people in a room.
- Recognizers and Metadata Viewer:** A pane on the right showing settings for the Audio Recognizer (Silence Recognizer MPI-PL) and Video Recognizer.
- Media Player controls:** A set of controls on the right side of the Video Viewer pane.
- Waveform Viewer:** A pane at the bottom showing a waveform of the audio signal.
- Annotation Density Viewer:** A pane at the bottom showing a density plot of the annotations.
- Timeline Viewer:** A pane at the bottom showing a timeline of the annotations with text and phonetic transcriptions.

The interface also includes a central timeline with a selection bar and various playback controls.

some foundational concepts

1. To use ELAN effectively, you must understand:

- ▶ Your own goals for the structure of your transcript.
- ▶ ELAN's built in concepts about different kinds of annotations.
- ▶ How your goals maps onto ELAN's concepts.



2. With this in mind, let's ask ourselves:

- ▶ *What kind of information do I want in my time-aligned transcript?*

3. Let's consider what we see in some printed transcripts.

some foundational concepts

1. To use ELAN effectively, you must understand:
 - ▶ Your own goals for the structure of your transcript.
 - ▶ ELAN's built in concepts about different kinds of annotations.
 - ▶ How your goals maps onto ELAN's concepts.



2. With this in mind, let's ask ourselves:
 - ▶ *What kind of information do I want in my time-aligned transcript?*
3. Let's consider what we see in some printed transcripts.

some foundational concepts

1. To use ELAN effectively, you must understand:
 - ▶ Your own goals for the structure of your transcript.
 - ▶ ELAN's built in concepts about different kinds of annotations.
 - ▶ How your goals maps onto ELAN's concepts.



2. With this in mind, let's ask ourselves:
 - ▶ *What kind of information do I want in my time-aligned transcript?*
3. Let's consider what we see in some printed transcripts.

some foundational concepts

1. To use ELAN effectively, you must understand:
 - ▶ Your own goals for the structure of your transcript.
 - ▶ ELAN's built in concepts about different kinds of annotations.
 - ▶ How your goals maps onto ELAN's concepts.



2. With this in mind, let's ask ourselves:
 - ▶ *What kind of information do I want in my time-aligned transcript?*
3. Let's consider what we see in some printed transcripts.

some foundational concepts

1. To use ELAN effectively, you must understand:
 - ▶ Your own goals for the structure of your transcript.
 - ▶ ELAN's built in concepts about different kinds of annotations.
 - ▶ How your goals maps onto ELAN's concepts.



2. With this in mind, let's ask ourselves:
 - ▶ *What kind of information do I want in my time-aligned transcript?*
3. Let's consider what we see in some printed transcripts.

some foundational concepts

1. To use ELAN effectively, you must understand:
 - ▶ Your own goals for the structure of your transcript.
 - ▶ ELAN's built in concepts about different kinds of annotations.
 - ▶ How your goals maps onto ELAN's concepts.



2. With this in mind, let's ask ourselves:
 - ▶ *What kind of information do I want in my time-aligned transcript?*
3. Let's consider what we see in some printed transcripts.

some foundational concepts

1. To use ELAN effectively, you must understand:
 - ▶ Your own goals for the structure of your transcript.
 - ▶ ELAN's built in concepts about different kinds of annotations.
 - ▶ How your goals maps onto ELAN's concepts.



2. With this in mind, let's ask ourselves:
 - ▶ *What kind of information do I want in my time-aligned transcript?*
3. Let's consider what we see in some printed transcripts.

some foundational concepts

1. To use ELAN effectively, you must understand:
 - ▶ Your own goals for the structure of your transcript.
 - ▶ ELAN's built in concepts about different kinds of annotations.
 - ▶ How your goals maps onto ELAN's concepts.



2. With this in mind, let's ask ourselves:
 - ▶ *What kind of information do I want in my time-aligned transcript?*
3. Let's consider what we see in some printed transcripts.

the structure of your transcript

Example 1: Single language with one speaker
(simple transcription system).

((JAY-Z Fresh Air Interview))¹

JAY-Z: “Well I had a - I grew up in the Marcy Projects in Brooklyn, and my mom and pop had an extensive record collection. So Michael Jackson and Stevie Wonder, and all those sounds and souls - and Motown etc., etc. - filled the house. So I was very familiar with the song when Kanye bought me the sample. It was just such an interesting and fresh take on it that I immediately was drawn to it.”

¹<http://www.npr.org/2010/11/16/131334322/the-fresh-air-interview-jay-z-recorded>

the structure of your transcript

Example 1: Single language with one speaker
(simple transcription system).

((JAY-Z Fresh Air Interview))¹

JAY-Z: “Well I had a - I grew up in the Marcy Projects in Brooklyn, and my mom and pop had an extensive record collection. So Michael Jackson and Stevie Wonder, and all those sounds and souls - and Motown etc., etc. - filled the house. So I was very familiar with the song when Kanye bought me the sample. It was just such an interesting and fresh take on it that I immediately was drawn to it.”

¹<http://www.npr.org/2010/11/16/131334322/the-fresh-air-interview-jay-z-decoded> ◀ ≡ ▶ ≡ 🔍 ↺

the structure of your transcript

Example 2: Single language with multiple speakers
(more informative transcription system).

1. Prosodic phrasing
2. Speaker overlap
3. Pause length
4. Laughter & other non-language oral noises
5. Phrase accent/stress

Note: The following example is in the UCSB transcription system.

the structure of your transcript

Example 2: Single language with multiple speakers
(more informative transcription system).

1. Prosodic phrasing
2. Speaker overlap
3. Pause length
4. Laughter & other non-language oral noises
5. Phrase accent/stress

Note: The following example is in the UCSB transcription system.

the structure of your transcript

Example 2: Single language with multiple speakers
(more informative transcription system).

1. Prosodic phrasing
2. Speaker overlap
3. Pause length
4. Laughter & other non-language oral noises
5. Phrase accent/stress

Note: The following example is in the UCSB transcription system.

the structure of your transcript

((*Fish Tacos*, English))

```
50 JACK;      This is ^um;
51 ANNA;      [Is this] the ^sauce [ɹ you ^made]?
52 JACK;      [This is]--
53              [ɹ the ^sau:]ce.
54 JACK;      %^Yeah.
55 ANNA;      .. @[[@]
56 BRAD;      <WHISPER> [^W]o:w </WHISPER>.
57              (0.7)
58 JACK;      Pink ^sau:ce,
59              (0.8)
60              ^u:m,
61              (1.3)
62              we have to (0.7) ^dip the ^fish,
63              (0.3)
64              in: ^e:gg.
65              before ^we:,
66              .. ^dip it in the ^breeding,
67              which,
68              (0.4)
69              is,
70              .. <%> ^still in ^process </%>.
71              (0.3)
72 BRAD;      ^Nice:.
73 JACK;      And ^this,
74              (0.3)
75              is ^lemon juice,
76              ^which,
77              (0.3)
78              the ^fish needs to be ^wa:shed,
79              (0.3)
80              or rin--
81              (0.4)
82              I don't ^know.
83              (0.4)
84 BRAD;      [(H)^]
85 JACK;      [^rinsed],
86              [ɹ ^whatever].
87 ANNA;      [s@@@]
88              (0.5)
89 JACK;      ^Something with ^lemon juice,
```

the structure of your transcript

Example 3: Two languages

1. Subject language
2. Translation language
3. Sentence level translation

((*Polygamous Marriage*, Shona))

1. Ah kwáivápò némùrùmè làivá nèvákàdzì vávìrì.
'There was a man who had two wives.'
2. Pàvákàdzì vávìrì váhòsì Lváínggè vásínggàjgè mbìzì.
'Of the two women, the elder wife was the one who doesn't eat zebra meat.'
3. Nèrímwè zùvá mùrúmé ùyù ákáèndà kùndòvhímá.
'One day this man went to go hunting.'

the structure of your transcript

Example 3: Two languages

1. Subject language
2. Translation language
3. Sentence level translation

((*Polygamous Marriage*, Shona))

1. Ah kwáivápò némùrùmè làivá nèvákàdzì vávìrì.
'There was a man who had two wives.'
2. Pàvákàdzì vávírí váhòsì Lváínggè vásínggàjgè mbìzì.
'Of the two women, the elder wife was the one who doesn't eat zebra meat.'
3. Nèrímwè zùvá mùrúmé ùyù ákáèndà kùndòvhímá.
'One day this man went to go hunting.'

the structure of your transcript

Example 3: Two languages

1. Subject language
2. Translation language
3. Sentence level translation

((*Polygamous Marriage*, Shona))

1. Ah kwáivápò némùrùmè làivá nèvákàdzì vávìrì.
'There was a man who had two wives.'
2. Pàvákàdzì vávìrì váhòsì Lvàínggè vásínggàjgè mbìzì.
'Of the two women, the elder wife was the one who doesn't eat zebra meat.'
3. Nèrímwè zùvá mùrúmé ùyù ákáèndà kùndòvhímá.
'One day this man went to go hunting.'

the structure of your transcript

Example 4: Two languages with full IGT:

1. Subject language
2. Translation language
3. Prosodic phrases
4. Morphemes
5. Morpheme glosses
6. Sentence level translations

the structure of your transcript

Example 4: Two languages with full IGT:

1. Subject language
2. Translation language
3. Prosodic phrases
4. Morphemes
5. Morpheme glosses
6. Sentence level translations

the structure of your transcript

((*Guava*, Besemah))

BESEMAH SENTENCE:

ngibal-ngibal di pinggir ghimbe ni tadi, adinge ini tadi tekinak ngaghi jambu, katah besake.

INTONATION UNIT:	ngibal-ngibal	di	pinggir	ghimbe	ni	tadi,
MORPHEME:	ngibal-ng-ibal	di	pinggir	ghimbe	ni	tadi,
GLOSS:	RED-AV-walk	at	edge	forest	this	earlier

INTONATION UNIT:	adinge	ini	tadi	tekinak	ngaghi	jambu,
MORPHEME:	ading-e	ini	tadi	te-kinak	ngaghi	jambu,
GLOSS:	younger.brother-3	this	earlier	INTR-see	with	guava

INTONATION UNIT:	katah	besake.
MORPHEME:	katah	besak-e.
GLOSS:	INTENS	big-3

FREE TRANSLATION:

‘Playing at the edge of the forest, the younger brother saw a really big guava.’

tiers & types

1. A tier is everything from a transcript that is the same kind of information.
 - ▶ everything that Jay-Z said in example 1.
 - ▶ everything that Jack said in example 2.
 - ▶ everything labeled Intonation Unit (or Gloss, etc.) in example 3.
2. With printed transcripts we are forced to put different parts of a single tier on different lines.

Tier 1: ...	Tier 1: ...
Tier 2: ...	Tier 2: ...
Tier 3: ...	Tier 3: ...
Tier 1: ...	Tier 1: ...
Tier 2: ...	Tier 2: ...
Tier 3: ...	Tier 3: ...

tiers & types

1. A tier is everything from a transcript that is the same kind of information.
 - ▶ everything that Jay-Z said in example 1.
 - ▶ everything that Jack said in example 2.
 - ▶ everything labeled Intonation Unit (or Gloss, etc.) in example 3.
2. With printed transcripts we are forced to put different parts of a single tier on different lines.

Tier 1: ...	Tier 1: ...
Tier 2: ...	Tier 2: ...
Tier 3: ...	Tier 3: ...
Tier 1: ...	Tier 1: ...
Tier 2: ...	Tier 2: ...
Tier 3: ...	Tier 3: ...

tiers & types

1. A tier is everything from a transcript that is the same kind of information.
 - ▶ everything that Jay-Z said in example 1.
 - ▶ everything that Jack said in example 2.
 - ▶ everything labeled Intonation Unit (or Gloss, etc.) in example 3.
2. With printed transcripts we are forced to put different parts of a single tier on different lines.

Tier 1: ...	Tier 1: ...
Tier 2: ...	Tier 2: ...
Tier 3: ...	Tier 3: ...
Tier 1: ...	Tier 1: ...
Tier 2: ...	Tier 2: ...
Tier 3: ...	Tier 3: ...

tiers & types

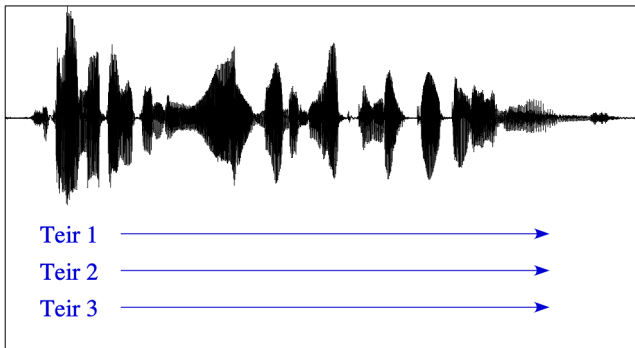
1. A tier is everything from a transcript that is the same kind of information.
 - ▶ everything that Jay-Z said in example 1.
 - ▶ everything that Jack said in example 2.
 - ▶ everything labeled Intonation Unit (or Gloss, etc.) in example 3.
2. With printed transcripts we are forced to put different parts of a single tier on different lines.

Tier 1: ...	Tier 1: ...
Tier 2: ...	Tier 2: ...
Tier 3: ...	Tier 3: ...
Tier 1: ...	Tier 1: ...
Tier 2: ...	Tier 2: ...
Tier 3: ...	Tier 3: ...

tiers & types

With a time-based tool like ELAN, we can conceive of each tier as existing

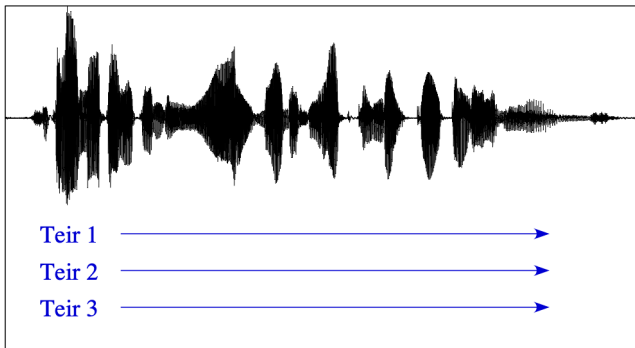
- ▶ as a **continuous** stream on a single line,
- ▶ **contiguous** to the timeline of the media.



tiers & types

With a time-based tool like ELAN, we can conceive of each tier as existing

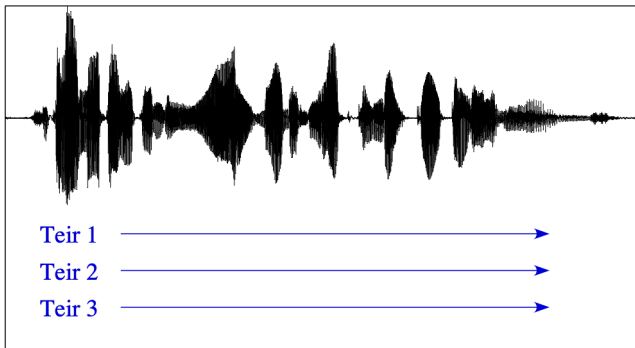
- ▶ as a **continuous** stream on a single line,
- ▶ **contiguous** to the timeline of the media.



tiers & types

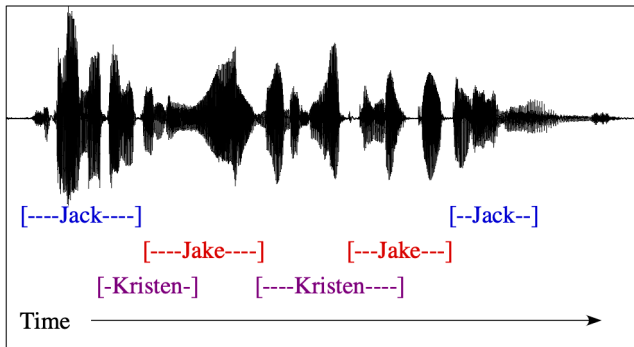
With a time-based tool like ELAN, we can conceive of each tier as existing

- ▶ as a **continuous** stream on a single line,
- ▶ **contiguous** to the timeline of the media.



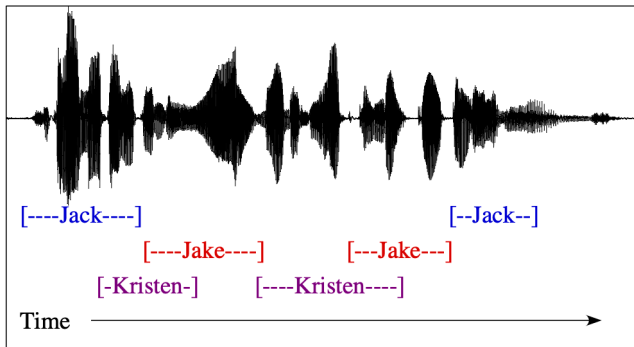
tiers & types

With multiple continuous tiers, we can capture overlap between speakers.



tiers & types

With multiple continuous tiers, we can capture overlap between speakers.



tiers & types

1. By conceiving of tiers as progressing forward through time, we can begin to establish some relationships between different tiers.
2. The parent-child relationship between tiers:
 - Parents: tiers that control the behavior of other (child) tiers in some way.
 - Children: tiers that are associated with and dependent upon other (parent) tiers.
3. This parent-child relationship is recursive, so a child of one tier could be the parent of another tier.
4. Certain changes to parent tiers will affect child tiers.

1. By conceiving of tiers as progressing forward through time, we can begin to establish some relationships between different tiers.
2. The parent-child relationship between tiers:
 - Parents: tiers that control the behavior of other (child) tiers in some way.
 - Children: tiers that are associated with and dependent upon other (parent) tiers.
3. This parent-child relationship is recursive, so a child of one tier could be the parent of another tier.
4. Certain changes to parent tiers will affect child tiers.

1. By conceiving of tiers as progressing forward through time, we can begin to establish some relationships between different tiers.
2. The parent-child relationship between tiers:
 - Parents: tiers that control the behavior of other (child) tiers in some way.
 - Children: tiers that are associated with and dependent upon other (parent) tiers.
3. This parent-child relationship is recursive, so a child of one tier could be the parent of another tier.
4. Certain changes to parent tiers will affect child tiers.

1. By conceiving of tiers as progressing forward through time, we can begin to establish some relationships between different tiers.
2. The parent-child relationship between tiers:
 - Parents: tiers that control the behavior of other (child) tiers in some way.
 - Children: tiers that are associated with and dependent upon other (parent) tiers.
3. This parent-child relationship is recursive, so a child of one tier could be the parent of another tier.
4. Certain changes to parent tiers will affect child tiers.

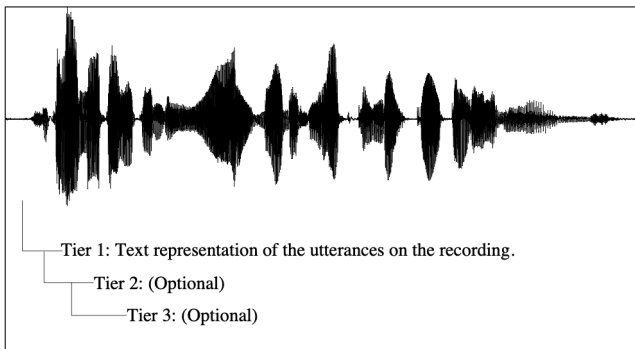
tiers & types

1. By conceiving of tiers as progressing forward through time, we can begin to establish some relationships between different tiers.
2. The parent-child relationship between tiers:
 - Parents: tiers that control the behavior of other (child) tiers in some way.
 - Children: tiers that are associated with and dependent upon other (parent) tiers.
3. This parent-child relationship is recursive, so a child of one tier could be the parent of another tier.
4. Certain changes to parent tiers will affect child tiers.

tiers & types

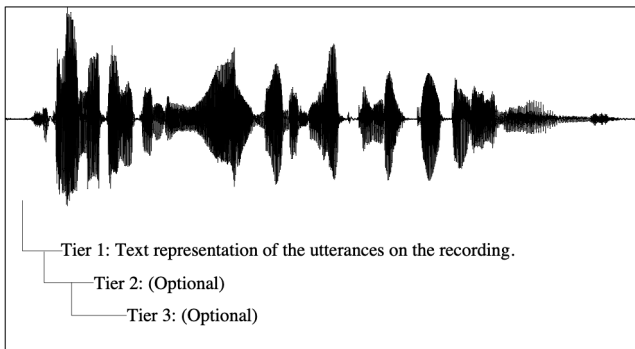
1. Top-level parent tiers are directly associated to the recording.

- ▶ This top-level tier is usually a direct representation of the media.
- ▶ Every ELAN file has at least one of these tiers.



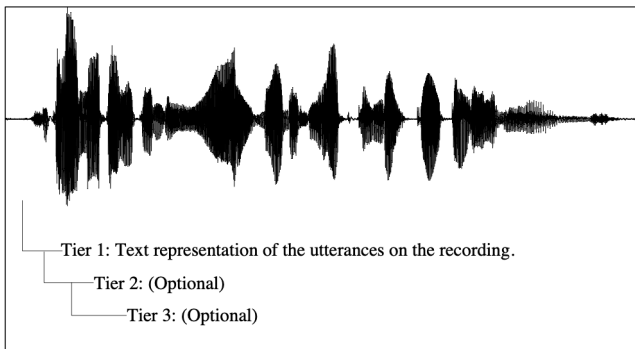
tiers & types

1. Top-level parent tiers are directly associated to the recording.
 - ▶ This top-level tier is usually a direct representation of the media.
 - ▶ Every ELAN file has at least one of these tiers.



tiers & types

1. Top-level parent tiers are directly associated to the recording.
 - ▶ This top-level tier is usually a direct representation of the media.
 - ▶ Every ELAN file has at least one of these tiers.



1. The precise relationship between a parent & a child tier is defined in ELAN by linguistic types.
2. The linguistic type is determined by the kind of data a tier contains:
 - ▶ Utterances
 - ▶ Translations
 - ▶ Linguistic analysis
 - ▶ ...
3. The linguistic type usually determines a one-to-one or one-to-many relationship between the parent & child tier.
4. There are some restrictions on what linguistic types can be parent tiers.

tiers & types

1. The precise relationship between a parent & a child tier is defined in ELAN by linguistic types.
2. The linguistic type is determined by the kind of data a tier contains:
 - ▶ Utterances
 - ▶ Translations
 - ▶ Linguistic analysis
 - ▶ ...
3. The linguistic type usually determines a one-to-one or one-to-many relationship between the parent & child tier.
4. There are some restrictions on what linguistic types can be parent tiers.

tiers & types

1. The precise relationship between a parent & a child tier is defined in ELAN by linguistic types.
2. The linguistic type is determined by the kind of data a tier contains:
 - ▶ Utterances
 - ▶ Translations
 - ▶ Linguistic analysis
 - ▶ ...
3. The linguistic type usually determines a one-to-one or one-to-many relationship between the parent & child tier.
4. There are some restrictions on what linguistic types can be parent tiers.

1. The precise relationship between a parent & a child tier is defined in ELAN by linguistic types.
2. The linguistic type is determined by the kind of data a tier contains:
 - ▶ Utterances
 - ▶ Translations
 - ▶ Linguistic analysis
 - ▶ ...
3. The linguistic type usually determines a one-to-one or one-to-many relationship between the parent & child tier.
4. There are some restrictions on what linguistic types can be parent tiers.

1. The precise relationship between a parent & a child tier is defined in ELAN by linguistic types.
2. The linguistic type is determined by the kind of data a tier contains:
 - ▶ Utterances
 - ▶ Translations
 - ▶ Linguistic analysis
 - ▶ ...
3. The linguistic type usually determines a one-to-one or one-to-many relationship between the parent & child tier.
4. There are some restrictions on what linguistic types can be parent tiers.

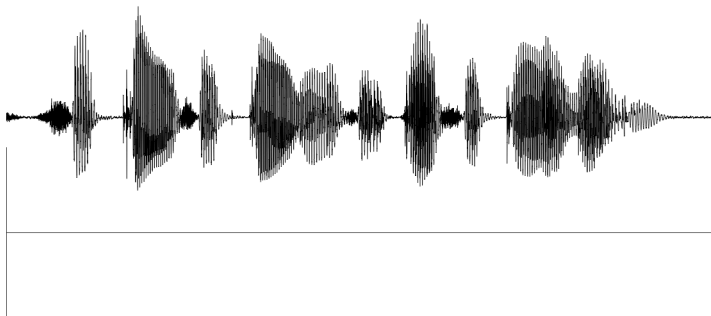
let's get started...

1

types (stereotypes) vs. tiers

- ▶ Every tier is assigned to a linguistic type
- ▶ The linguistic type specifies the stereotypical constraints of that tier

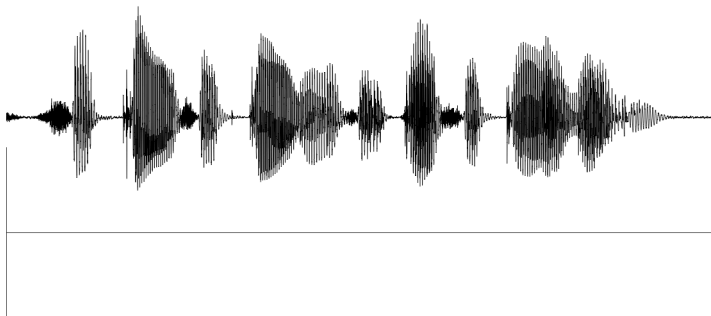
the default stereotype



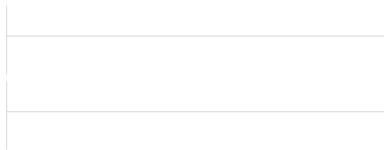
types (stereotypes) vs. tiers

- ▶ Every tier is assigned to a linguistic type
- ▶ The linguistic type specifies the stereotypical constraints of that tier

the default stereotype



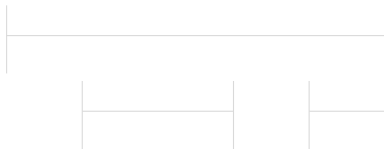
types (stereotypes) vs. tiers



Symbolic Association



Time Subdivision

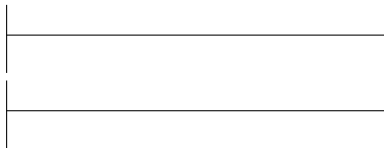


Included In



Symbolic Subdivision

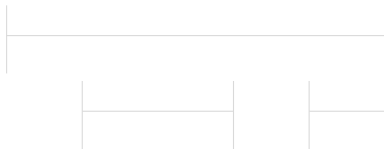
types (stereotypes) vs. tiers



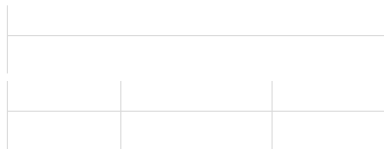
Symbolic Association



Time Subdivision

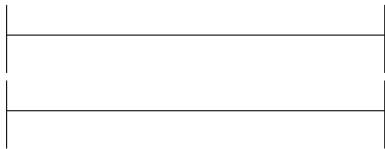


Included In



Symbolic Subdivision

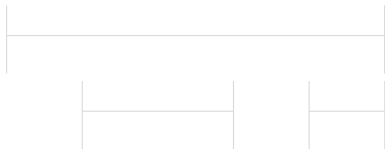
types (stereotypes) vs. tiers



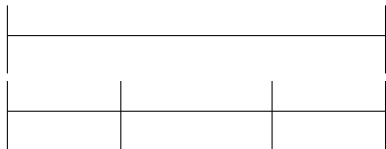
Symbolic Association



Time Subdivision

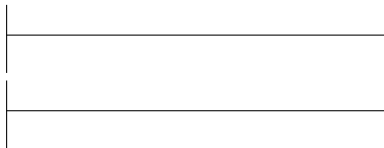


Included In



Symbolic Subdivision

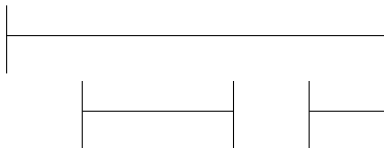
types (stereotypes) vs. tiers



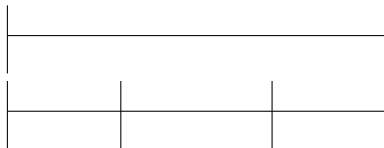
Symbolic Association



Time Subdivision

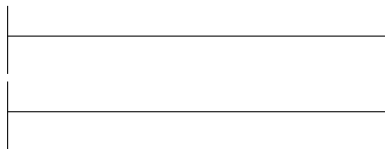


Included In

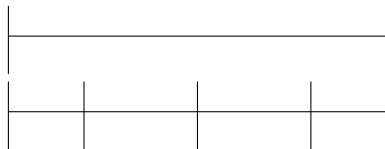


Symbolic Subdivision

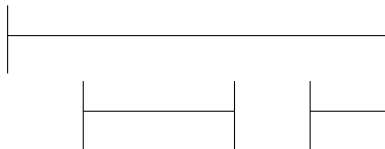
types (stereotypes) vs. tiers



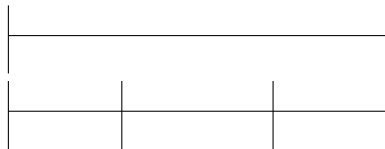
Symbolic Association



Time Subdivision



Included In



Symbolic Subdivision

exercise #1

► In this exercise we will learn:

1. How to create a new ELAN file from a recording
2. How to enter time-aligned text on the default tier
3. How to edit time-aligned text
4. How to save an ELAN file

exercise #1

- ▶ In this exercise we will learn:
 1. How to create a new ELAN file from a recording
 2. How to enter time-aligned text on the default tier
 3. How to edit time-aligned text
 4. How to save an ELAN file

exercise #1

- ▶ In this exercise we will learn:
 1. How to create a new ELAN file from a recording
 2. How to enter time-aligned text on the default tier
 3. How to edit time-aligned text
 4. How to save an ELAN file

exercise #1

- ▶ In this exercise we will learn:
 1. How to create a new ELAN file from a recording
 2. How to enter time-aligned text on the default tier
 3. How to edit time-aligned text
 4. How to save an ELAN file

exercise #1

- ▶ In this exercise we will learn:
 1. How to create a new ELAN file from a recording
 2. How to enter time-aligned text on the default tier
 3. How to edit time-aligned text
 4. How to save an ELAN file

- ▶ In this exercise we will learn:
 1. How to create additional tiers
 2. How to work with overlapping speech
 3. A little about linguistic types

- ▶ In this exercise we will learn:
 1. How to create additional tiers
 2. How to work with overlapping speech
 3. A little about linguistic types

- ▶ In this exercise we will learn:
 1. How to create additional tiers
 2. How to work with overlapping speech
 3. A little about linguistic types

- ▶ In this exercise we will learn:
 1. How to create additional tiers
 2. How to work with overlapping speech
 3. A little about linguistic types

- ▶ In this exercise we will learn:
 1. How to conceive of and create additional linguistic types & tiers
 2. How to work with two languages:
 - ▶ A subject language
 - ▶ A translation language

- ▶ In this exercise we will learn:
 1. How to conceive of and create additional linguistic types & tiers
 2. How to work with two languages:
 - ▶ A subject language
 - ▶ A translation language

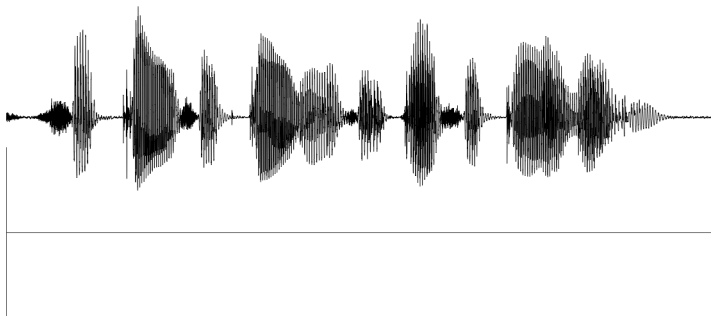
- ▶ In this exercise we will learn:
 1. How to conceive of and create additional linguistic types & tiers
 2. How to work with two languages:
 - ▶ A subject language
 - ▶ A translation language

- ▶ In this exercise we will learn:
 1. How to conceive of and create additional linguistic types & tiers
 2. How to work with two languages:
 - ▶ A subject language
 - ▶ A translation language

types (stereotypes) vs. tiers

- ▶ Every tier is assigned to a linguistic type
- ▶ The linguistic type specifies the stereotypical constraints of that tier

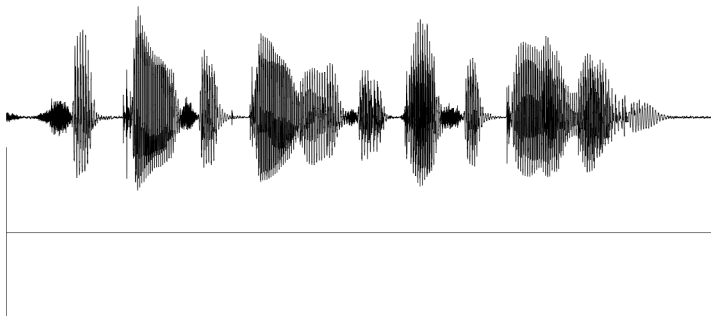
The Default Stereotype



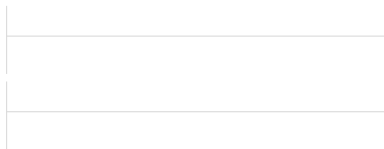
types (stereotypes) vs. tiers

- ▶ Every tier is assigned to a linguistic type
- ▶ The linguistic type specifies the stereotypical constraints of that tier

The Default Stereotype



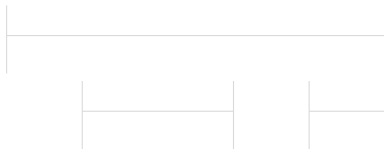
types (stereotypes) vs. tiers



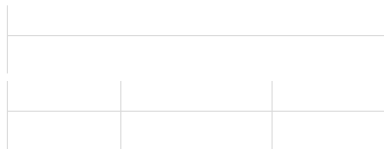
Symbolic Association



Time Subdivision

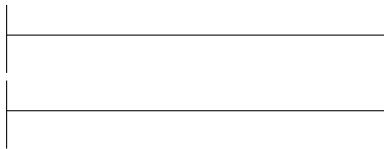


Included In



Symbolic Subdivision

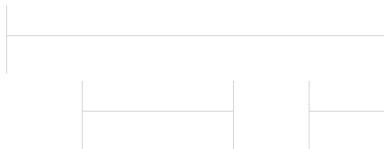
types (stereotypes) vs. tiers



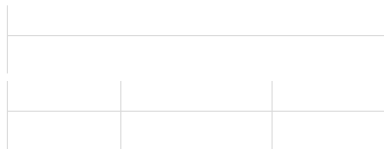
Symbolic Association



Time Subdivision

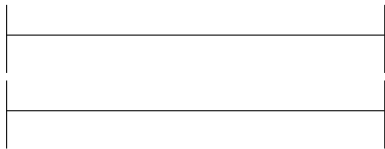


Included In

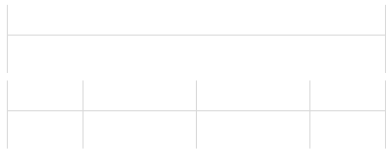


Symbolic Subdivision

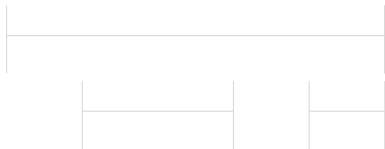
types (stereotypes) vs. tiers



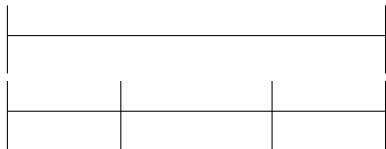
Symbolic Association



Time Subdivision

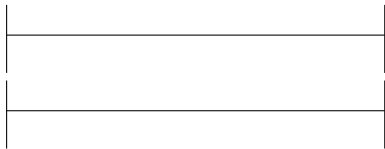


Included In



Symbolic Subdivision

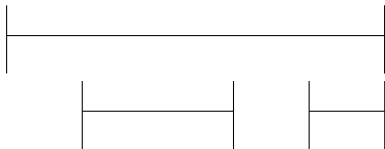
types (stereotypes) vs. tiers



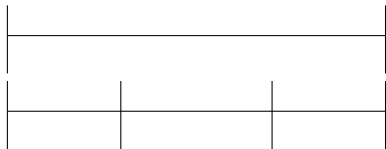
Symbolic Association



Time Subdivision

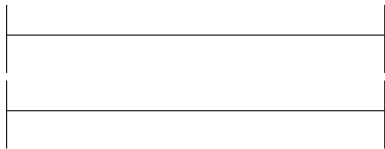


Included In

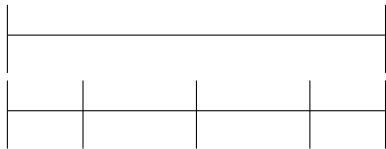


Symbolic Subdivision

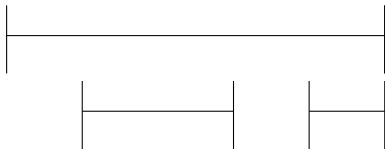
types (stereotypes) vs. tiers



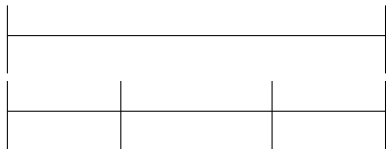
Symbolic Association



Time Subdivision



Included In



Symbolic Subdivision

1. What type would you create for a *Free Translation* tier?
2. What type would you create for a *Word* tier?
3. How could you use the *Included In* type?
4. Could you use the *Time Subdivision* type for gestures?

1. What type would you create for a *Free Translation* tier?
2. What type would you create for a *Word* tier?
3. How could you use the *Included In* type?
4. Could you use the *Time Subdivision* type for gestures?

1. What type would you create for a *Free Translation* tier?
2. What type would you create for a *Word* tier?
3. How could you use the *Included In* type?
4. Could you use the *Time Subdivision* type for gestures?

1. What type would you create for a *Free Translation* tier?
2. What type would you create for a *Word* tier?
3. How could you use the *Included In* type?
4. Could you use the *Time Subdivision* type for gestures?

links to further resources

- ▶ ELAN website with manual:
<http://www.lat-mpi.eu/tools/elan/>
- ▶ InField course on ELAN:
<http://logos.uoregon.edu/infield2010/workshops/aligning-text-elan1/index.php>
- ▶ ELAN tool that incorporates Toolbox functionality:
<http://corpafroas.tge-adonis.fr/tools.html>
- ▶ Review of ELAN by Andrea Berez: <http://scholarspace.manoa.hawaii.edu/handle/10125/1718/>