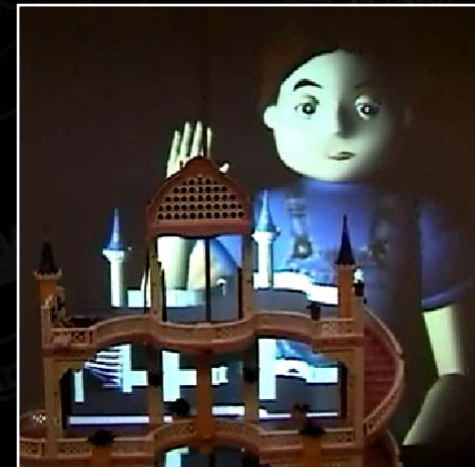
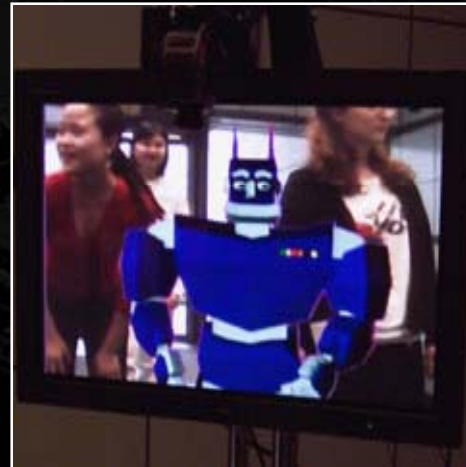


Study Human Conversation



Study Human Conversation

- **Goal**
 - Understand and model “Embodied Conversation”
- **Because we want to (1 of 3)**
 - Create autonomous characters that can have natural, engaging and effective face-to-face conversations with humans (embodied conversational agents)



Study Human Conversation

- Goal
 - Understand and model “Embodied Conversation”
- Because we want to (2 of 3)
 - Create systems that allow remote people to communicate with each other in online environments as if they were meeting face-to-face (avatars)



Building a System: BEAT

- **Approach**

Takes as input the lines to be spoken by characters and a description of a scene, **tries to understand** the conversation functions being carried out and then **produces supporting nonverbal behavior**. The output is an animation script with gesture synchronized with speech.

BEAT Video

(GA) It is apparently (GT) { some [1 "kind]
of] { [2 "virtual [3 "actor] }).

TAGGED

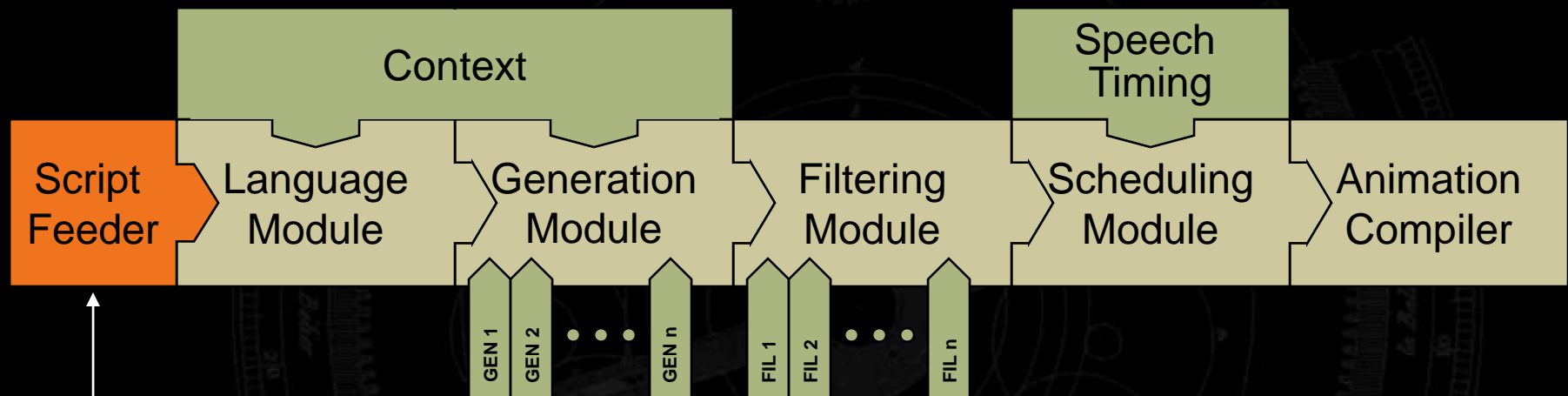
You just need to type in a line like "This is
Bradley Beat live from SIGGRAPH", and the
actor is able to talk and gesture by itself!

00:00:11:14



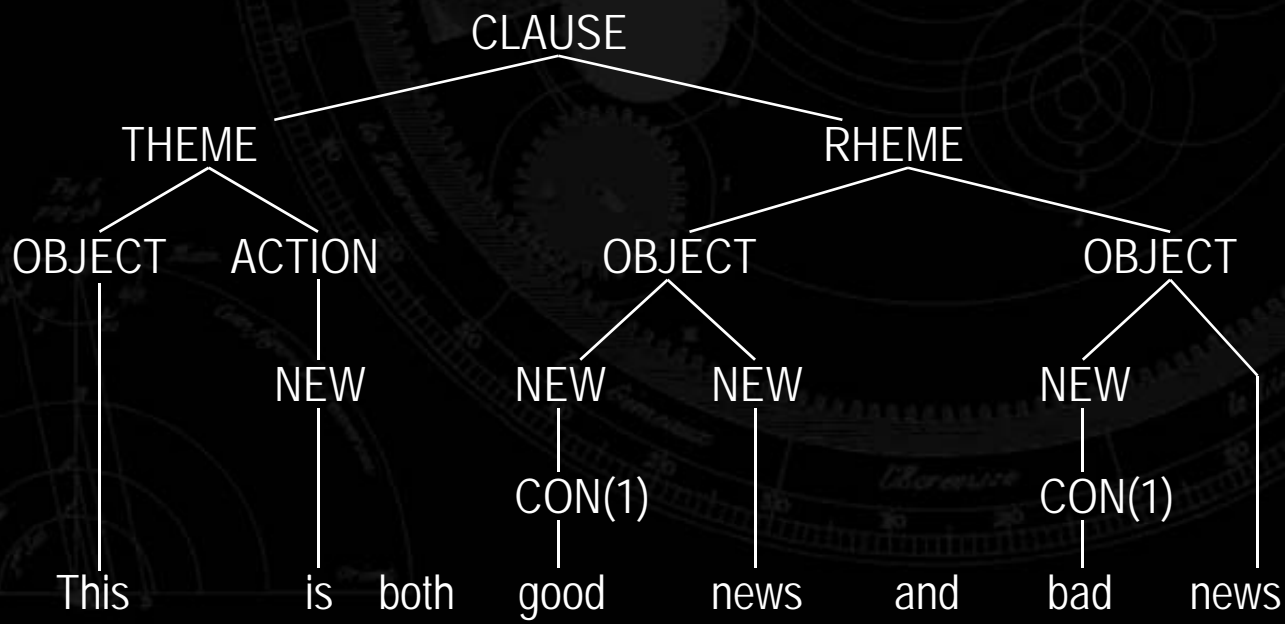
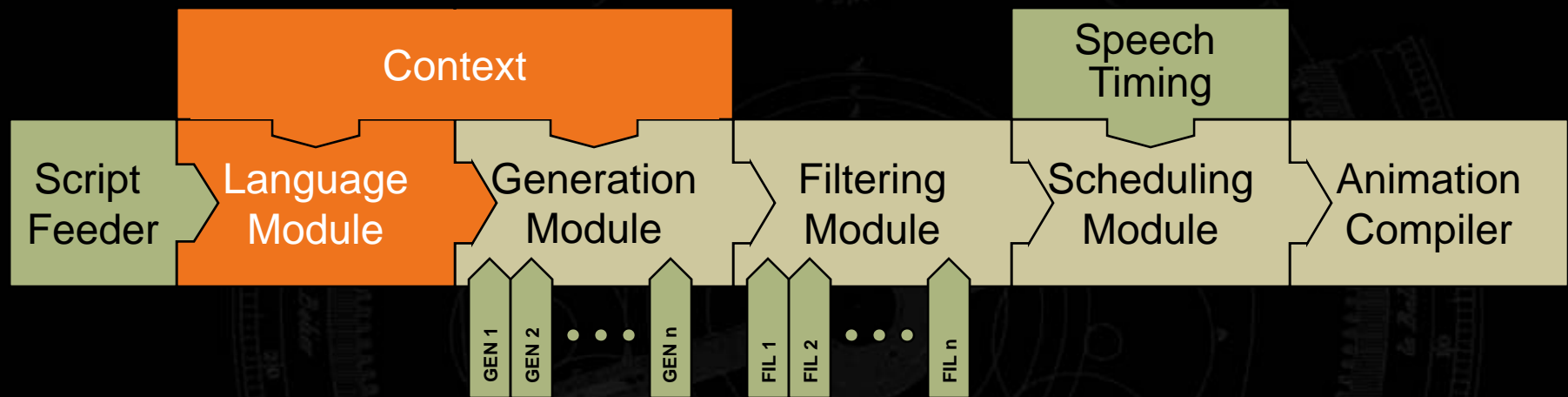
	of	virtual	actor,	
Gaze Towards	Yellow	Light Green	Yellow	
Eyebrows	Dark Green	Dark Green	Dark Green	
Beats			Orange	
Gesture		Brown	Red	

BEAT Architecture: Script Input



"This is both good news and bad news"

BEAT Architecture: Language Tagging



BEAT Language Module

Tokenizer

<W> Anyway </W><W SYN="C">,</W>

POS Tagger

<W POS="PR" LEM="this" SYN="NNHD">this</W>

Chunker

<CLAUSE><OBJECT>it</OBJECT><ACTION>is<ACTION>

Discourse Tagger

Mark New

The tile is <NEW> blue </NEW>

Mark Topic Shift

<TOPICSHIFT TYPE="POP">Anyway</TOPICSHIFT>

Mark Information Structure

<THEME>The tile is</THEME><RHEME> blue</RHEME>

Mark Contrast

<CON ID="1">good</CON><CON ID="1">bad</CON>

Identify Objects

<OBJECT ID="MAN1">the thin guy</OBJECT>

Identify Actions

<ACTION ID="MOVEMENT">walked away</OBJECT>

Mark Interaction Structure

<UTTERANCE HEARER="MAN2">this sucks
Joe</UTTERANCE>

BEAT Context Representation

Domain Knowledge

World and gesture ontology

Scene Description

What and who can I see

Participation Framework

What is everyone's role

Dictionary

How are terms related

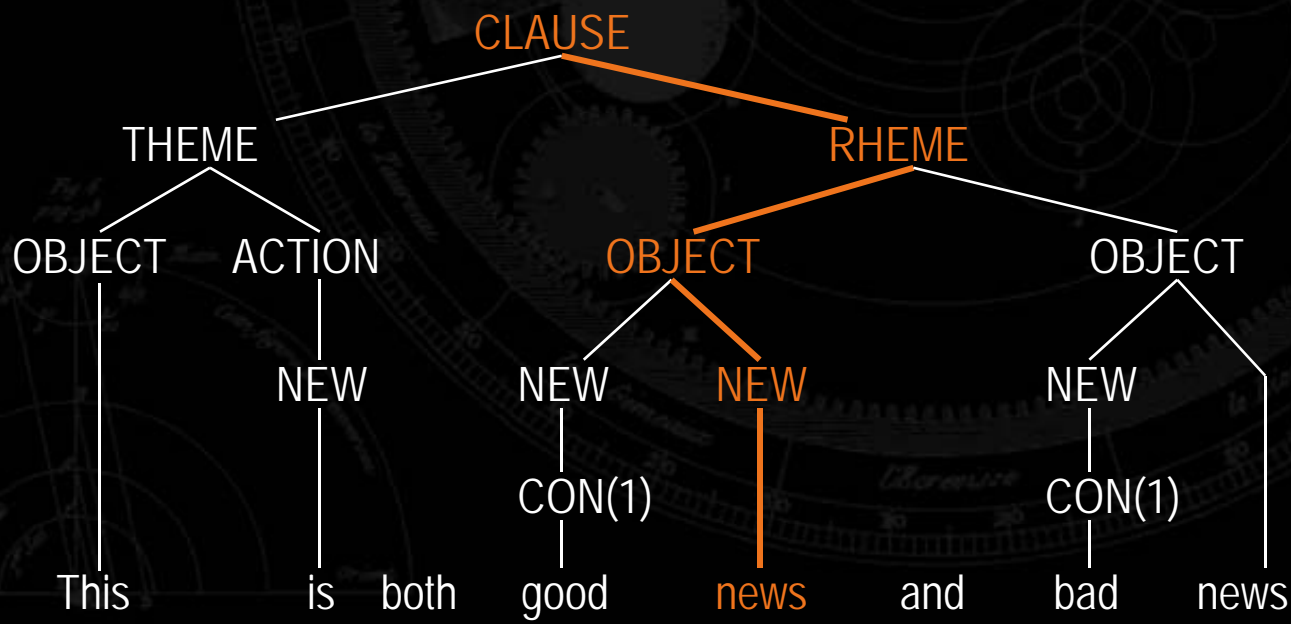
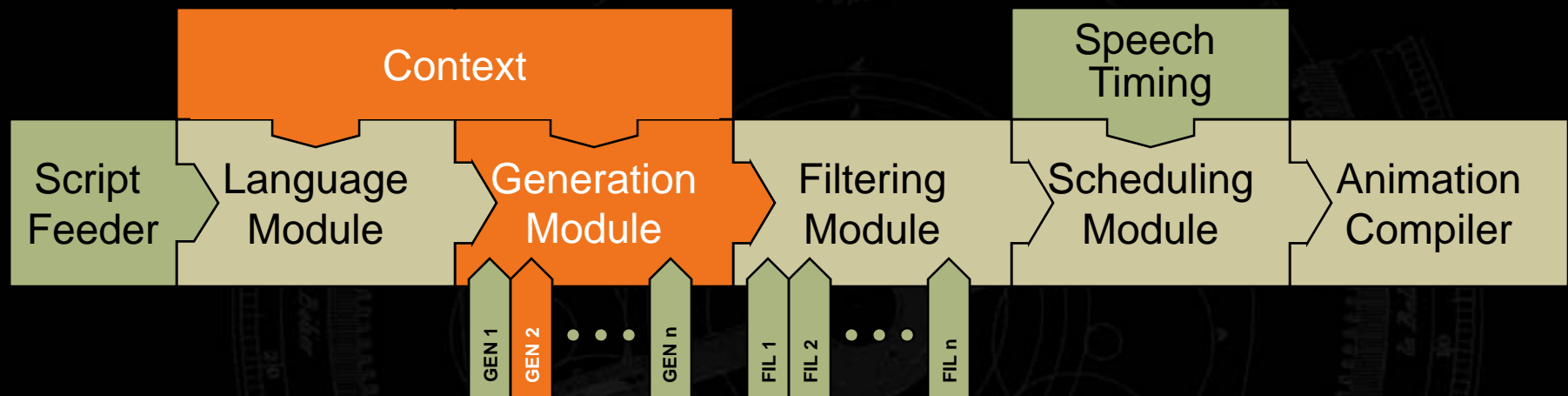
Discourse History

What has been said

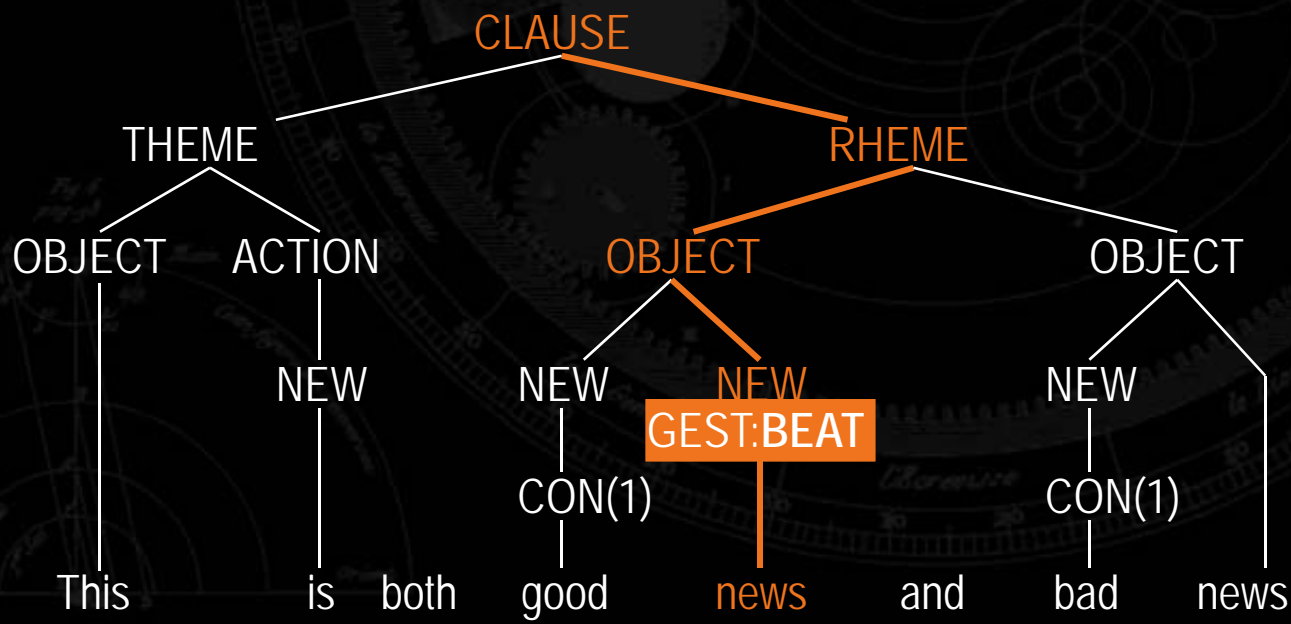
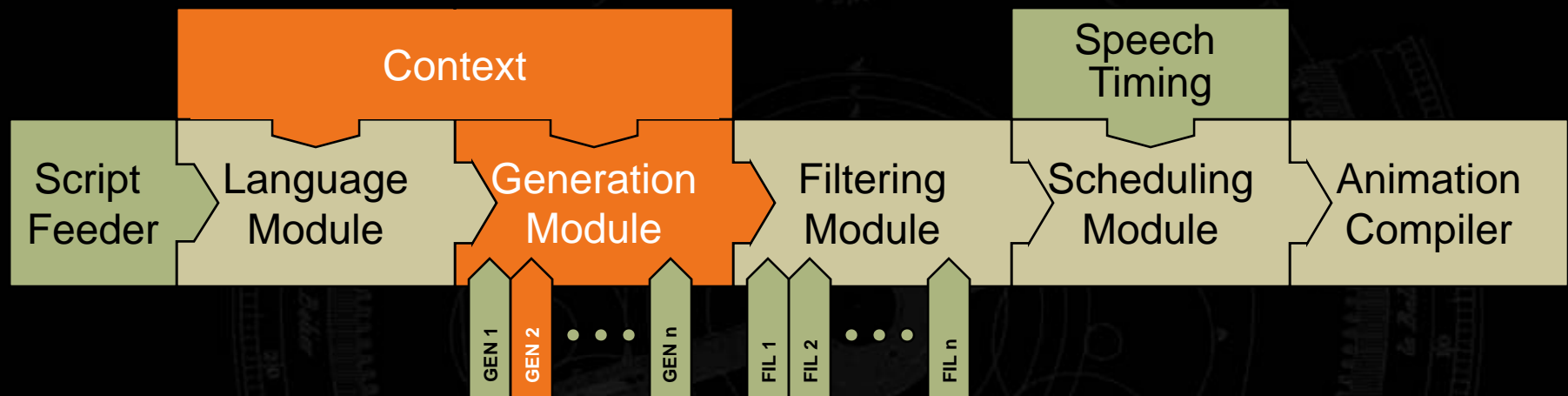
 *static*

 *dynamic*

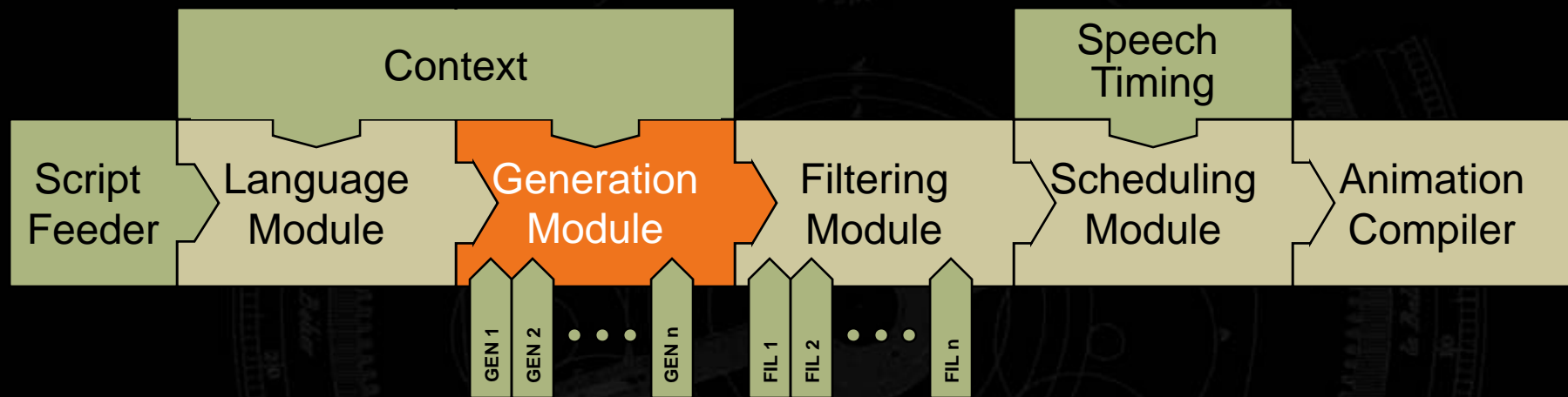
BEAT Architecture: Behavior Generation



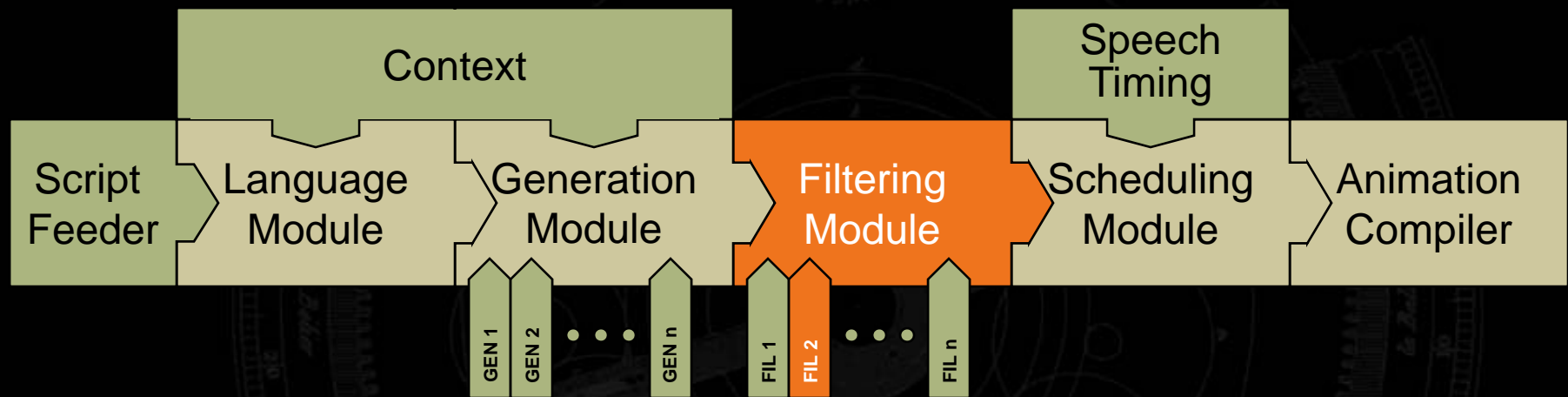
BEAT Architecture: Behavior Generation



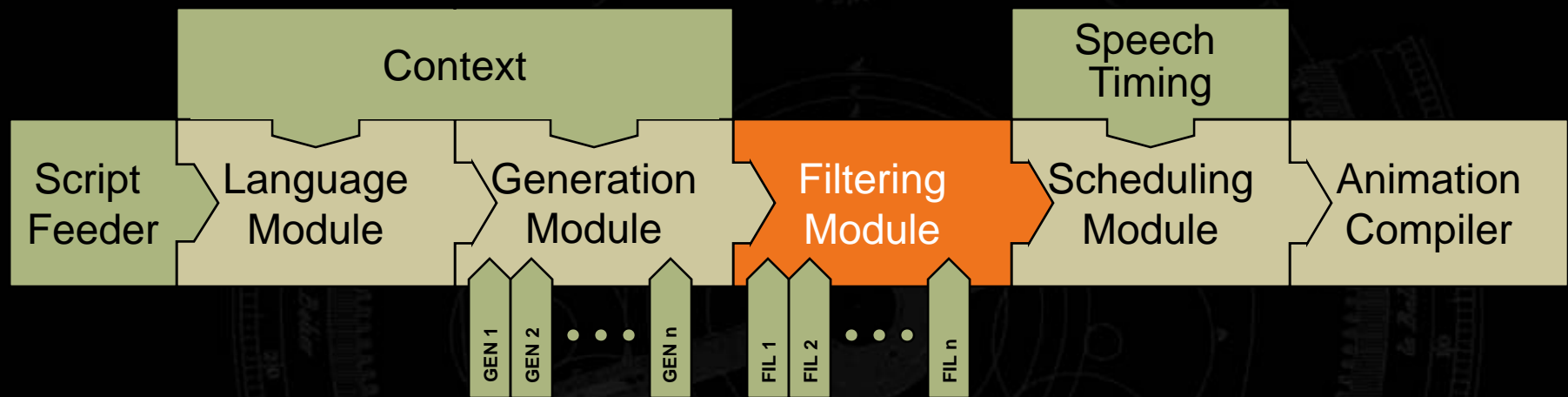
BEAT Architecture: Behavior Generation



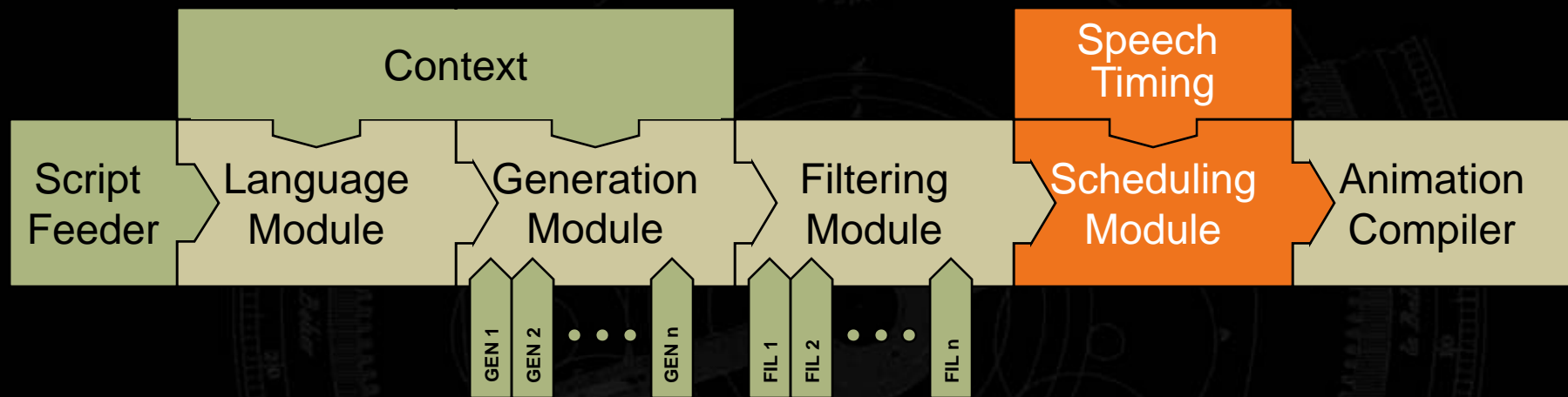
BEAT Architecture: Behavior Filtering



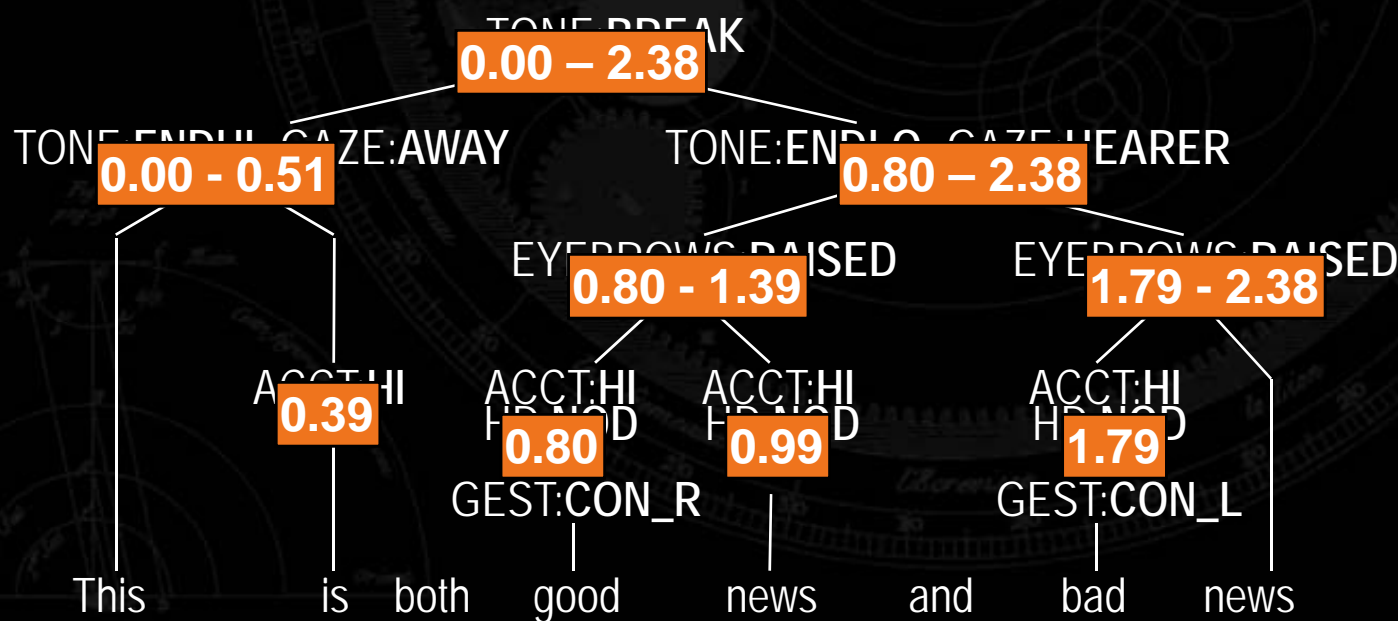
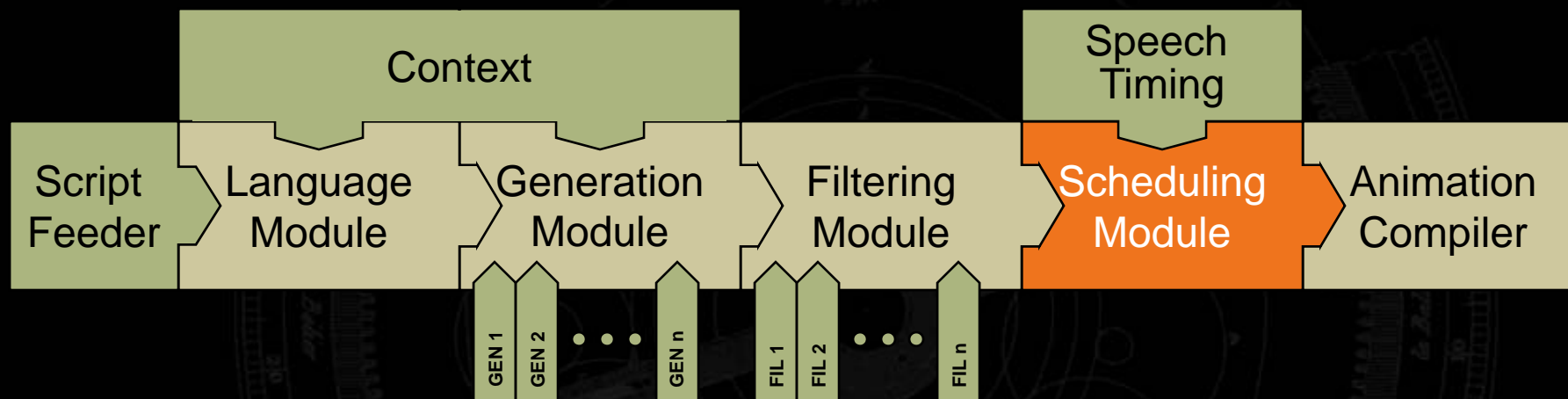
BEAT Architecture: Behavior Filtering



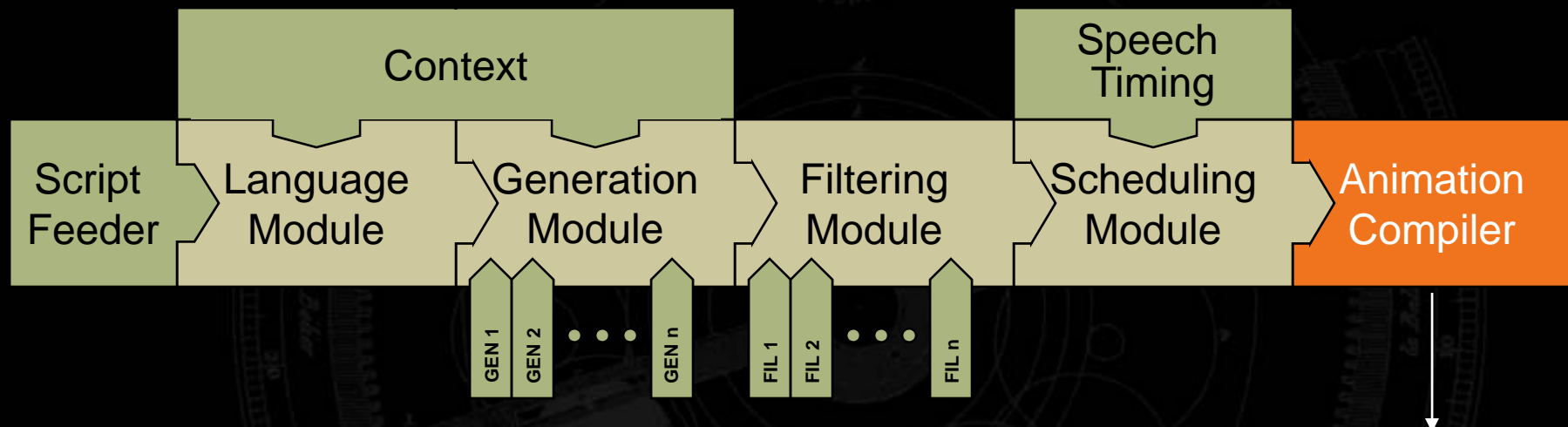
BEAT Architecture: Behavior Scheduling



BEAT Architecture: Behavior Scheduling

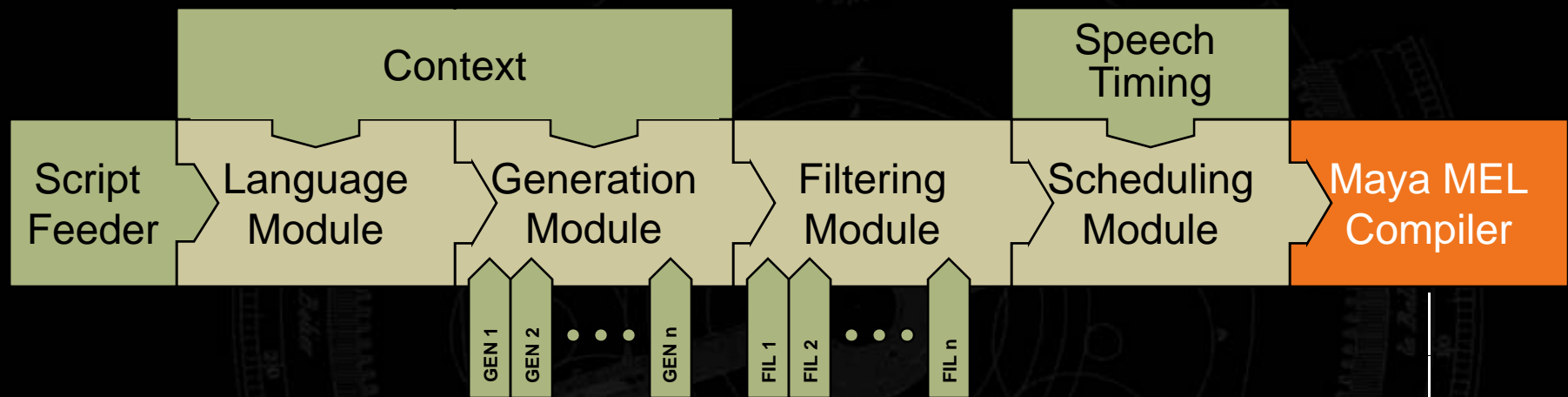


BEAT Architecture: Animation Compilation

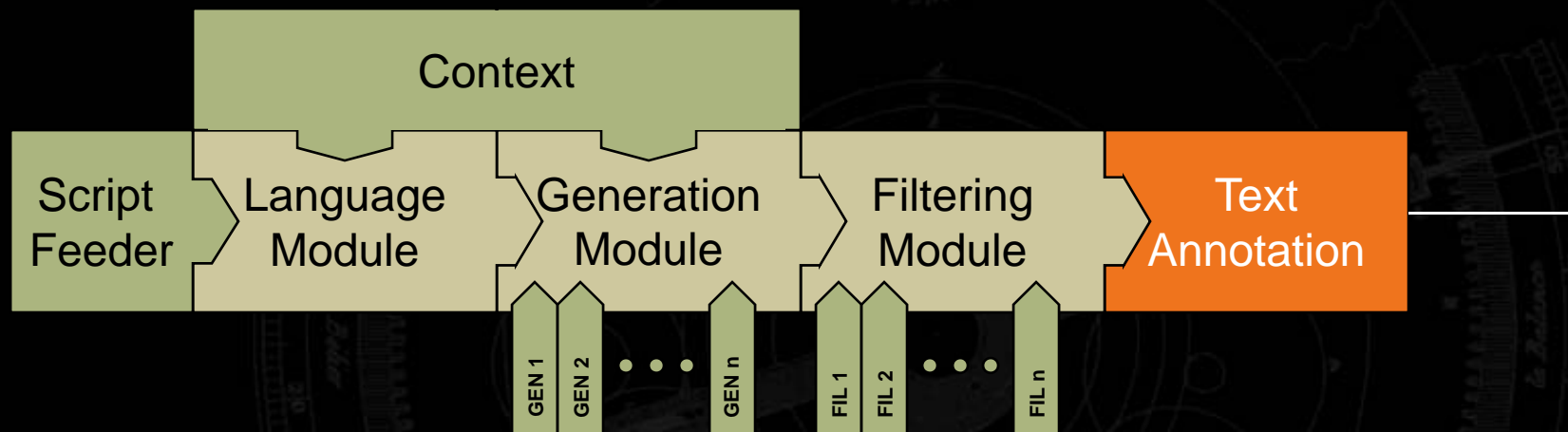


```
<AnimationScript SPEAKER="AGENT" HEARER="USER">  
<START SPEECH="This is both good news and bad news">  
<START ACTION="GAZE" DIRECTION="AWAY_FROM_HEARER" SRT="0.0">  
<START ACTION="VISEME" TYPE="B" SRT="0.0">  
<STOP ACTION="GAZE" DIRECTION="AWAY_FROM_HEARER" SRT="0.511">  
<START ACTION="VISEME" TYPE="A" SRT="0.511">  
<START ACTION="GAZE" DIRECTION="TOWARDS_HEARER" SRT="0.801">  
<START ACTION="EYEBROWS" SRT="0.801">  
<START ACTION="GESTURE_RIGHT" TYPE="CONTRAST_1"  
  RIGHT_TRAJECTORY=CONTRAST_TRAJECTORY"  
  RIGHT_HANDSHAPE="CONTRAST" SRT="0.801">
```

BEAT Example: Maya Compilation



BEAT Example: Animator Instructions



“You just need [1 to *type in] { a [2 *line] like }”

[1] ICONIC - Typing action

[2] BEAT - Emphasis

* Pitch accent

{ } Raised eyebrows

.. Gaze away

— Gaze towards