MOM: Matrix Odin Mash-up

Potential for Automated Grammar Generation from Linguistically Rich Data

David Wax

1. ODIN
2. Grammar Matrix
3. MOM System
4. Additional Resources

Example IGT:

(20) a. Chan toktang hong duay dokmai
I decorate room with flower
'I decorated the room with flowers.'

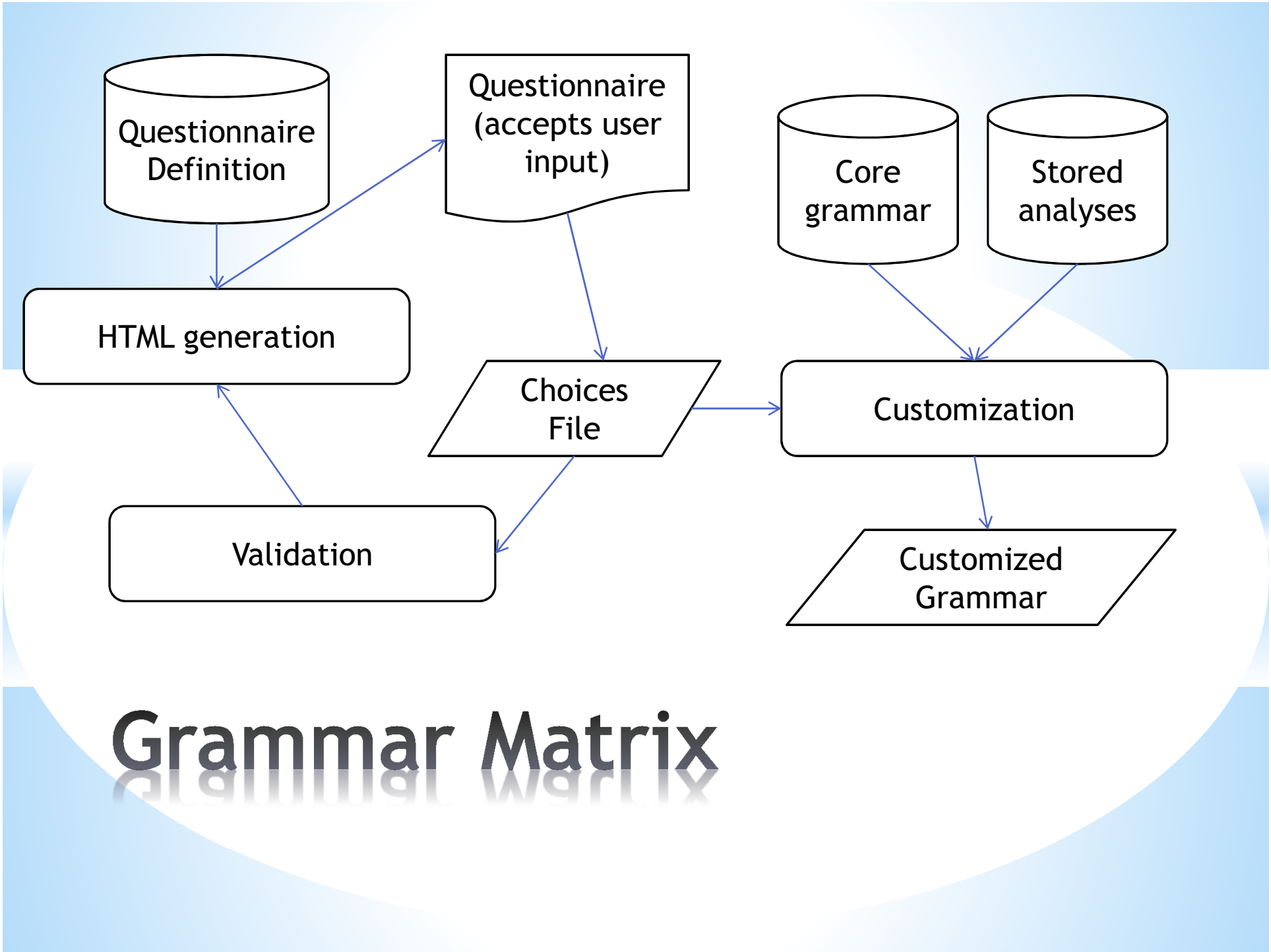
(7) Lu-lutu-in ng lalaki ang adobo.
Asp-cook-Acc CS man ANG adobo
'The man will cook the adobo.'

ODIN

Range of IGT Instances	Number of Languages	Number of IGT Instances
10,000+	3	36,691
1,000-9,999	37	97,158
100-999	122	40,260
10-99	326	12,822
1-9	838	2,313

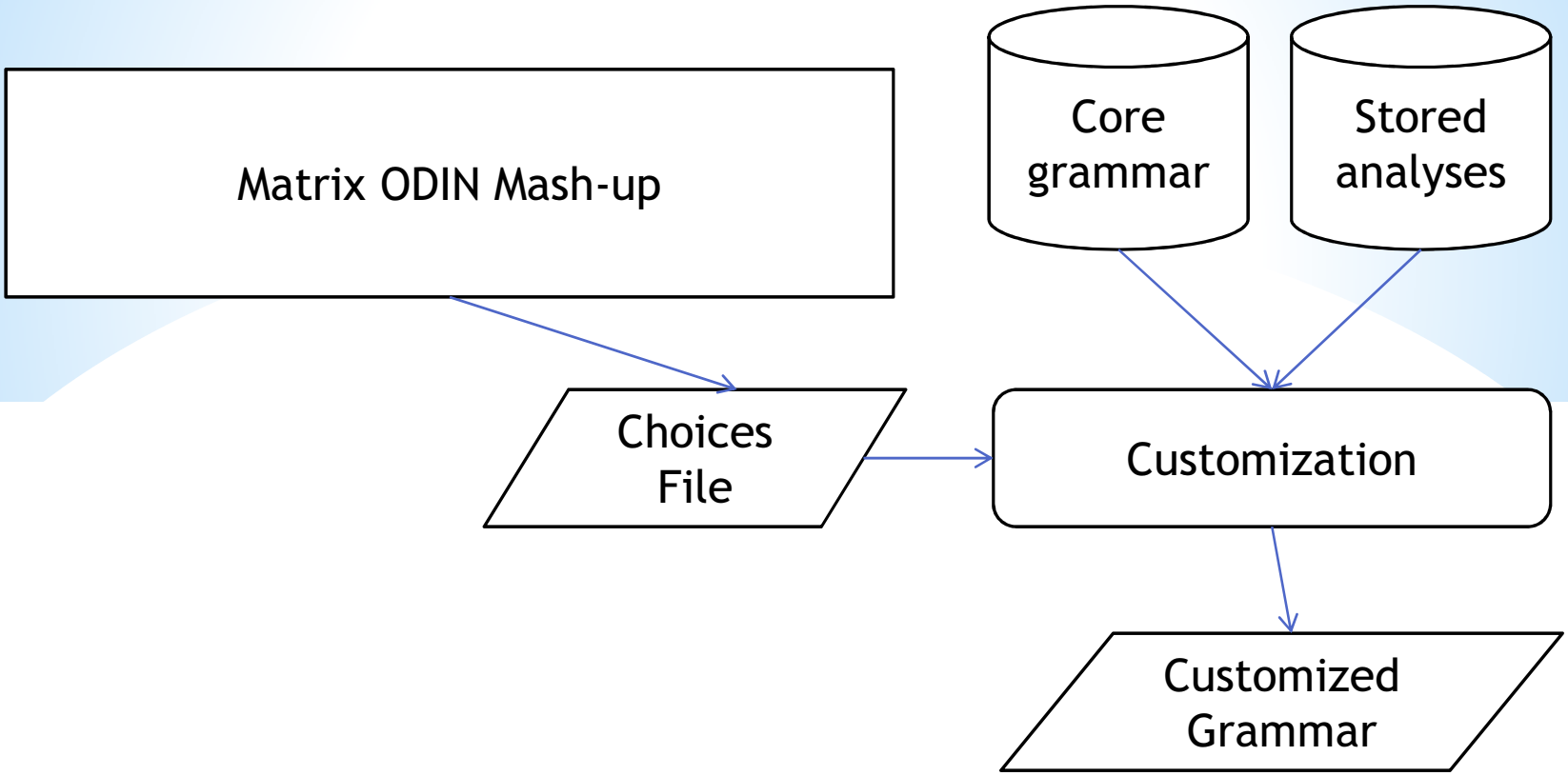
http://www.csufresno.edu/odin/

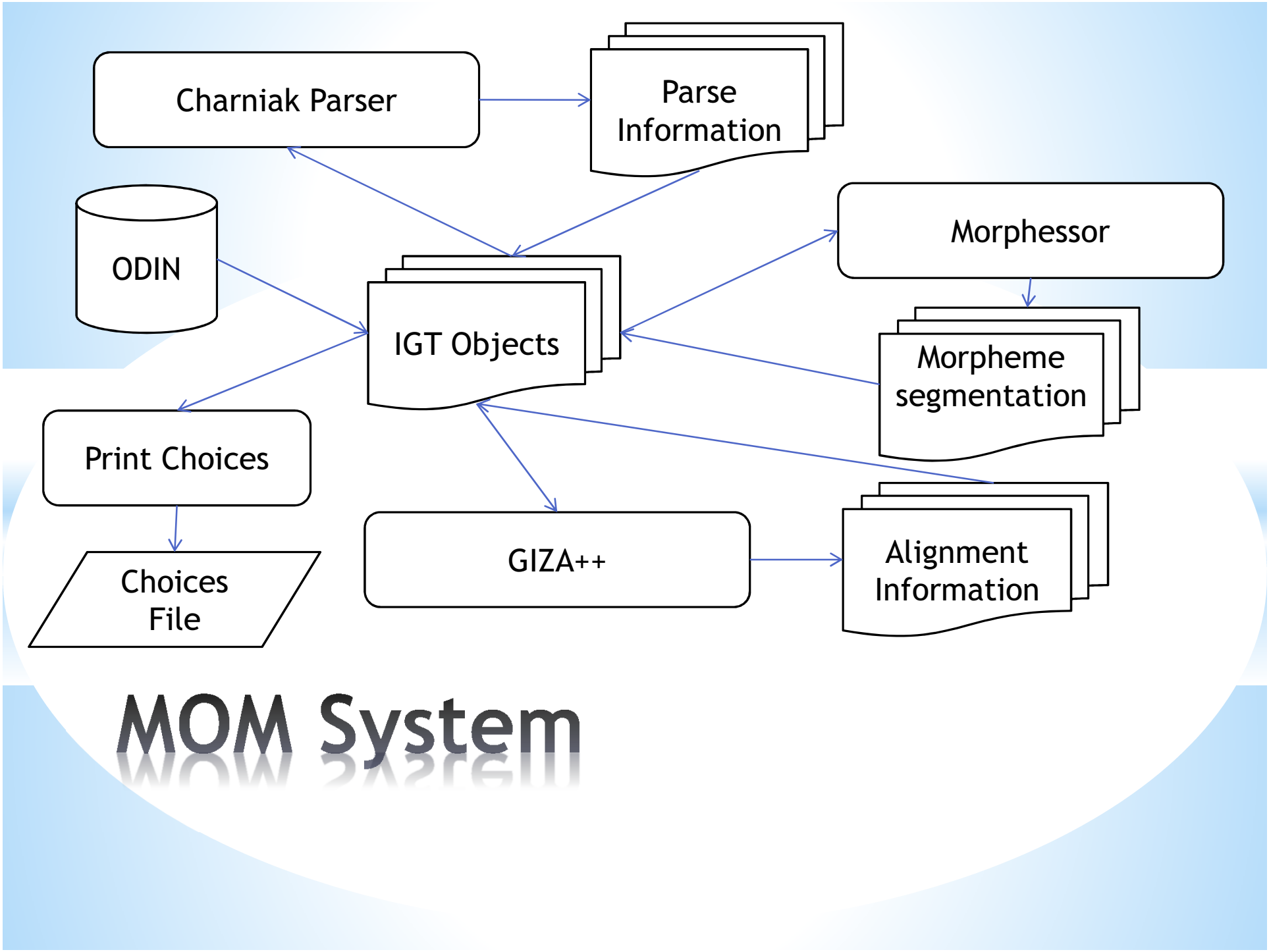
ODIN



Grammar Matrix

Grammar Matrix




Charniak Parser
Parse Information
ODIN
Morphessor
IGT Objects
Morpheme segmentation
Print Choices
GIZA++
Alignment Information
Choices File

MOM System

Additional Resources

French data:

Development set of 15,657 forms of the verb “to do”

Works as an ideal case for IGT:

ʒ(ə)-fɛ-(z)

NOM.1SG-do.PRS-1SG

“ do ”

Additionally we generated 12,212 incorrect French verb forms

Questions?

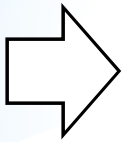
(7) Lu-lutu-in ng lalaki ang adobo.

Asp-cook-Acc CS man ANG adobo

'The man will cook the adobo.'

(S1 (S (NP (DT The) (NN man))
(VP (MD will) (VP (VB cook) (NP (DT the) (NN adobo))))
(. .)

System: Parser

Lulutuin		Lu/PRE + lutu/in + in/SUF
Ng		Ng/STM
Lalaki		Lalaki/STM
Ang		Ang/STM
Adobo		Adobo/STM

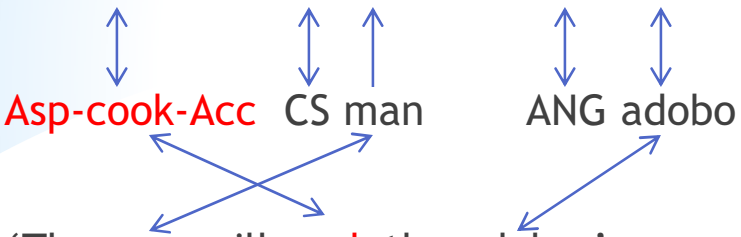
Alignment Part 2:

Lu lutu in

Asp cook Acc

System: Morphemes

(7) Lu-lutu-in ng lalaki ang adobo.



'The man will cook the adobo.'

System: Alignment