

# Towards Incremental Speech Generation in Dialogue Systems

A dark, low-key photograph of a person walking on a sidewalk. The person is in the center, walking away from the camera, with a long, dark shadow cast behind them. To the left is a parking meter. To the right is a bicycle. The background is a plain wall.

Skantze & Hjalmarsson







**Hola. Quiero  
reservar un  
pasaje a Madrid.**





Silence threshold

...





¿?





***¿Para qué día?***









Hola. Quiero reserv...





Hola. Quiero reserv...





Hola. Quiero reservar un  
pasaje a Madrid.





Hola. Quiero reservar un  
pasaje a Madrid. ...





Hola. Quiero reservar un  
pasaje a Madrid. ...



¿Para qué día?





¿Cómo hacemos nosotros?



# Cuando empezamos a hablar...

No sabemos  
qué vamos  
a decir

... ni cómo



# Y lo que hacemos es...

Planear y  
hablar en  
simultáneo

Monitorear  
lo que  
planeamos y  
decimos



¿Y si encontramos un error?





¿Y si encontramos un error?



Lo arreglamos.



Dos casos:



Ya lo dijimos

*(overt repair)*

Aún no lo  
dijimos

*(covert repair)*



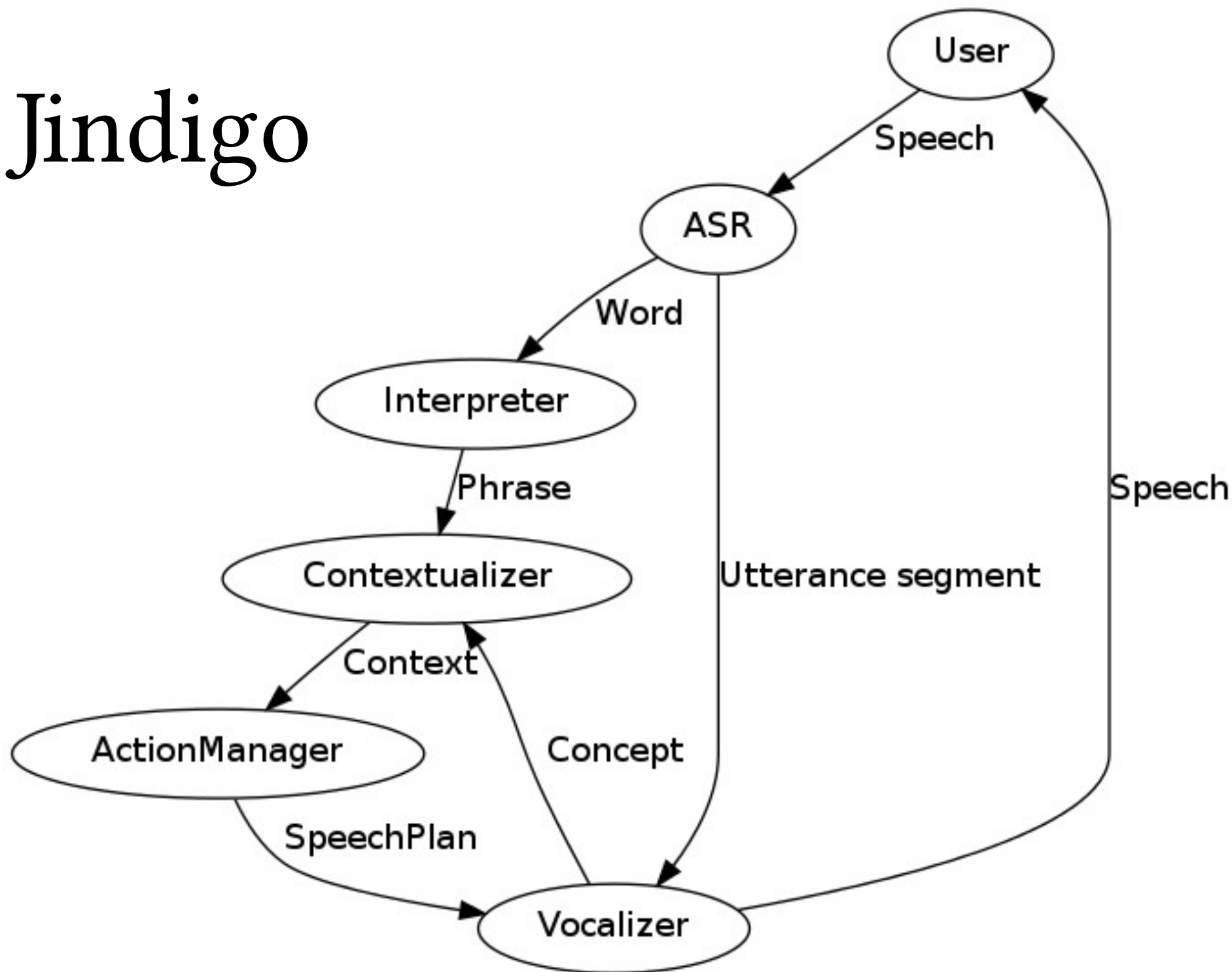
Idea:



Que el sistema se parezca  
más a nosotros



# Jindigo













Experimento:



Incremental

vs.

No Incremental



# Resultados: El incremental...



Comenzaba  
a hablar más  
rápido y lo  
hacía por  
más tiempo

Fue más  
*“polite”*,  
eficiente e  
indicaba mejor  
cuándo hablar

# Créditos:

- Imagen del hombre con el celular, por Ces't June

<http://www.flickr.com/photos/j79/849195237/>

- Imagen de la computadora, por Bruno Cordioli

<http://www.flickr.com/photos/br1dotcom/4736458283/>

- Imagen de las chicas hablando por celular,  
por D. Sharon Pruitt

<http://www.flickr.com/photos/pinksherbet/2796862756/>

- Slides por Pablo Antonio

Licencia: Creative Commons

Attribution-ShareAlike 3.0 Unported