

Pruebas realizadas para datos sintéticos:

Prueba #1:

Secuencia original:

Palabras en diccionario = 120

Número de muestras en el tiempo = 6000

Numero de speakers = 5

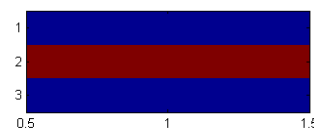
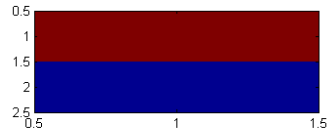
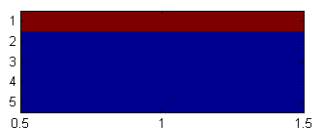
Estimación de parámetros para speakers = {2, 3}

Modelo original

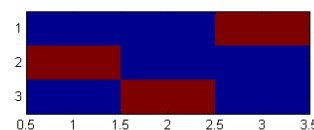
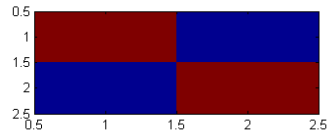
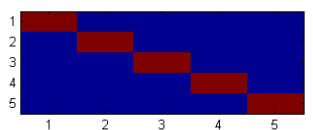
Modelo obtenido para spkr = 2

Modelo obtenido para spkr = 3

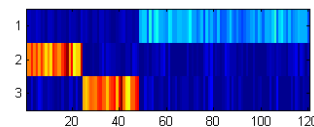
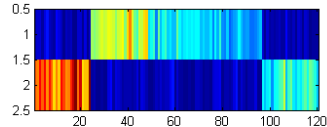
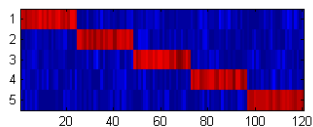
Matriz de speaker inicial



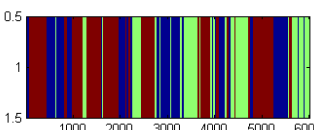
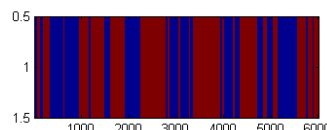
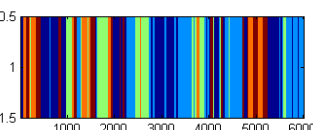
Matriz de transición entre speakers



Matriz de emisión speakers/words

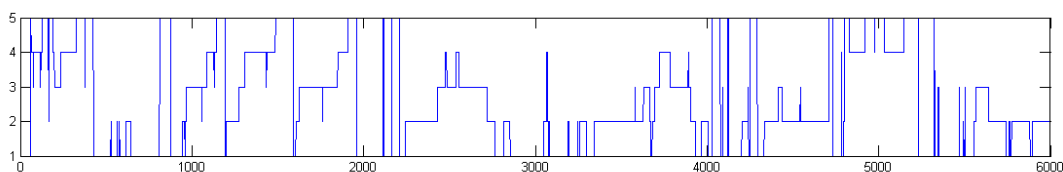


Variable oculta/Segmentación

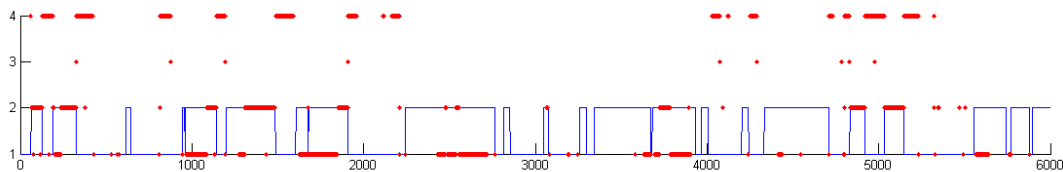


Segmentación/cambio de speakers en el tiempo

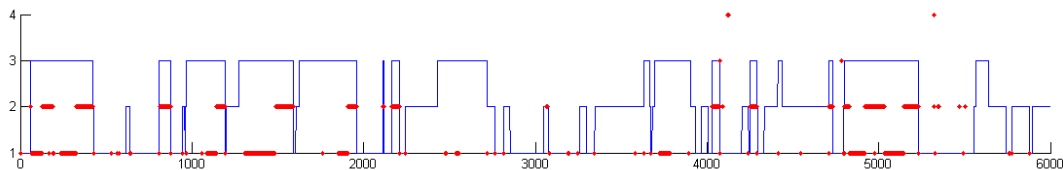
Ground truth



Segmentación para spkr = 2



Segmentación para spkr = 3



(En rojo son los FP y FN)

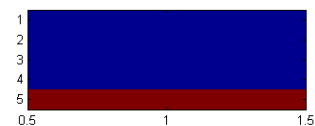
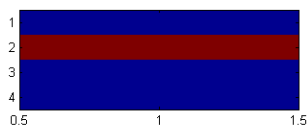
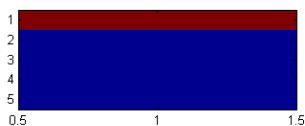
Estimación de parámetros para speakers = {4, 5}

Modelo original

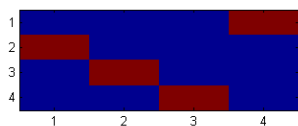
Modelo obtenido para spkr = 4

Modelo obtenido para spkr = 5

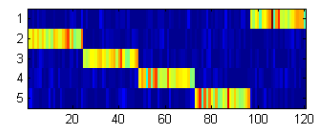
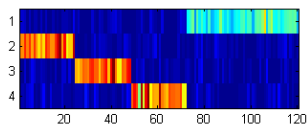
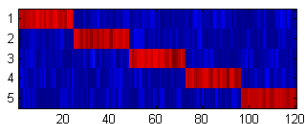
Matriz de speaker inicial



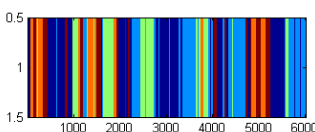
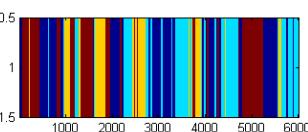
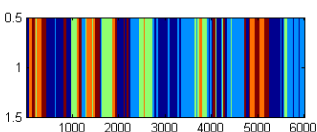
Matriz de transición entre speakers



Matriz de emisión speakers/words

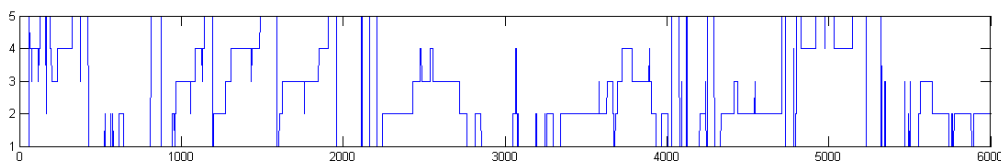


Variable oculta / Segmentación

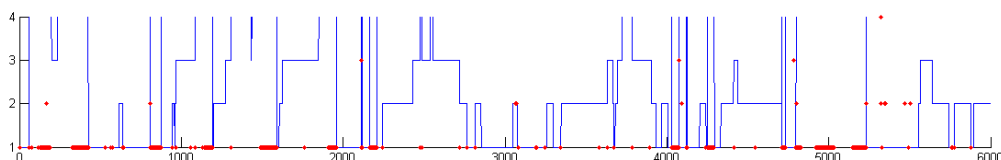


Segmentación/cambio de speakers en el tiempo

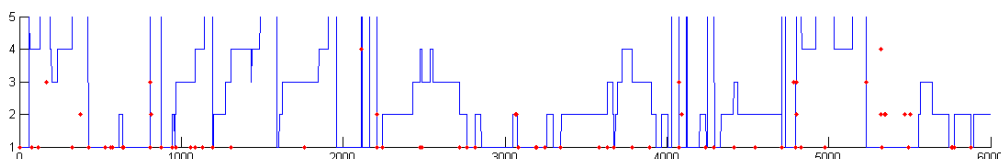
Ground truth



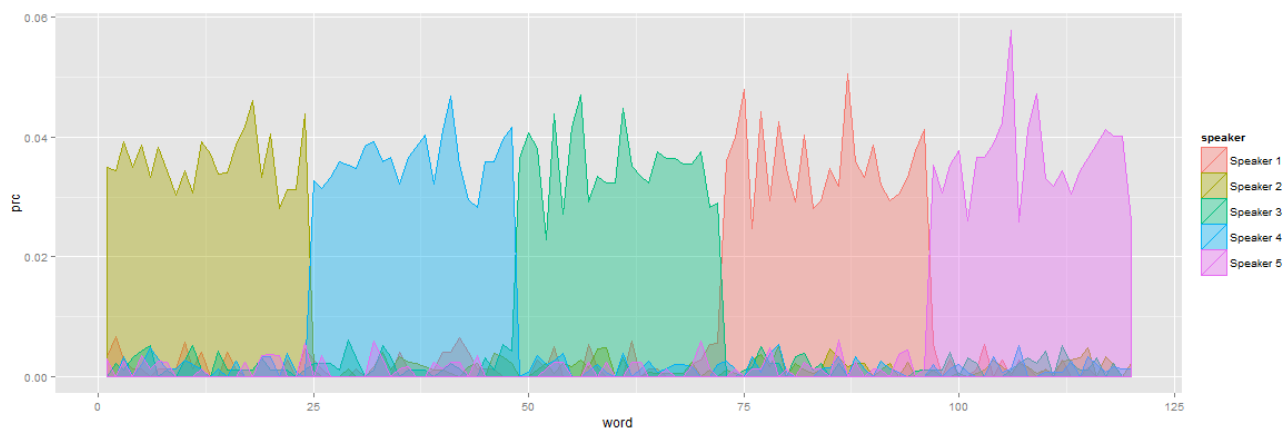
Segmentación para spkr = 4



Segmentación para spkr = 5



Probabilidades de emisión de palabras para cada speaker (speaker = 5)

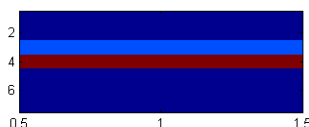
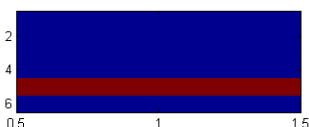
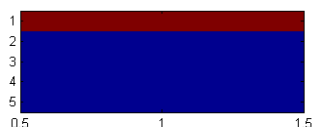


Estimación de parámetros para speakers = {6, 7} **Modelo obtenido para spkr=6**

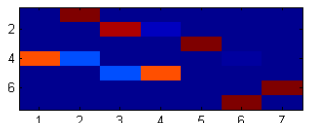
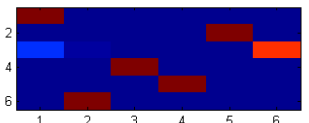
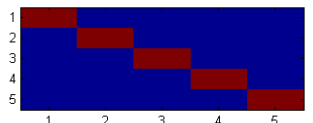
Modelo obtenido para spkr = 7

Modelo original

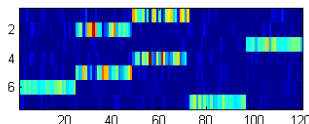
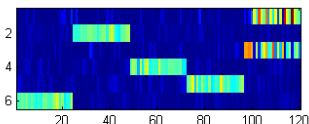
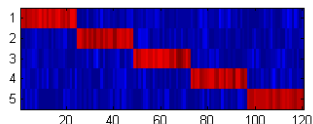
Matriz de speaker inicial



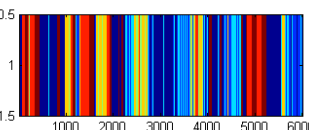
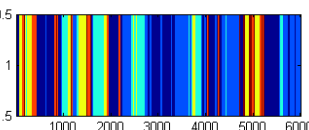
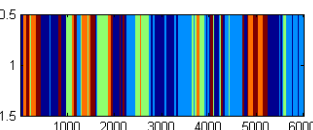
Matriz de transición entre speakers



Matriz de emisión speakers/words

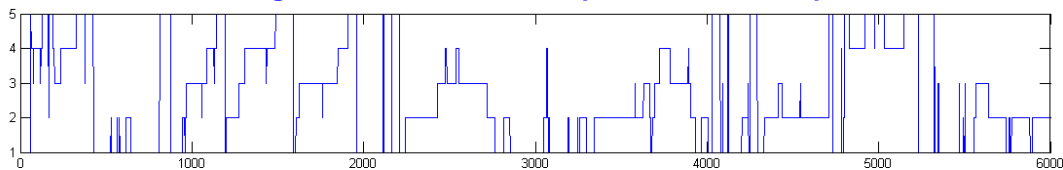


Variable oculta / Segmentación

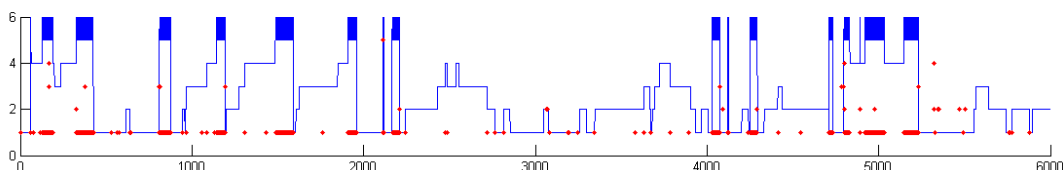


Segmentación/cambio de speakers en el tiempo

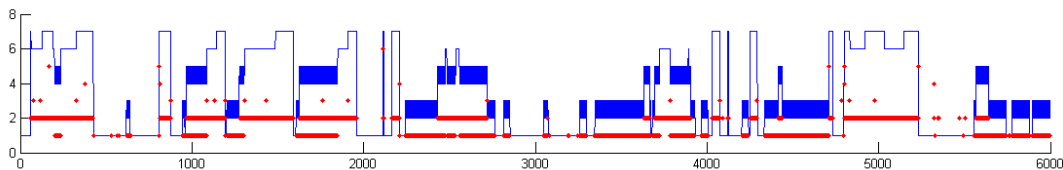
Ground truth



Segmentación para spkr = 6

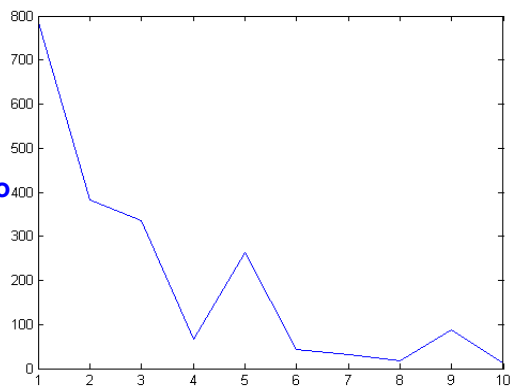


Segmentación para spkr = 7



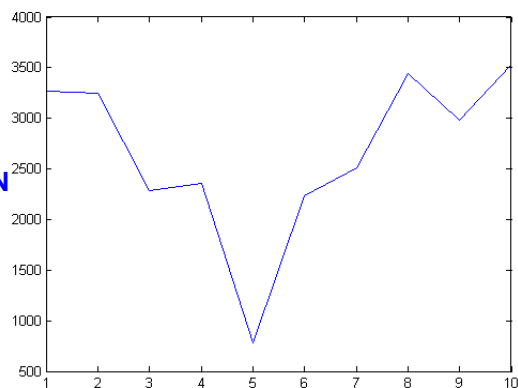
Selección de modelo:

Log-likelihood ratio



No. de speakers

FP + FN



No. de speakers

Prueba #2:

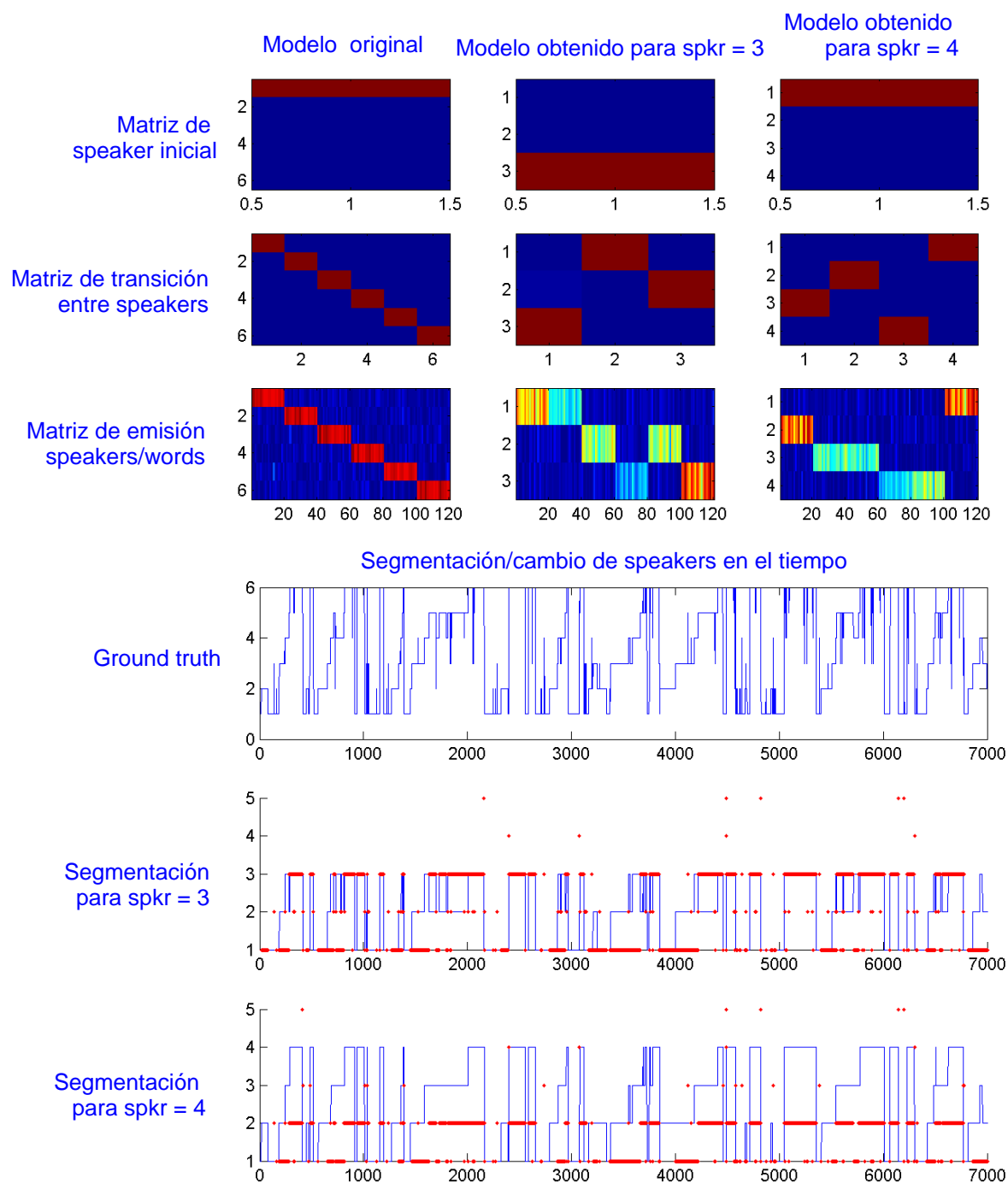
Secuencia original:

Palabras en diccionario = 120

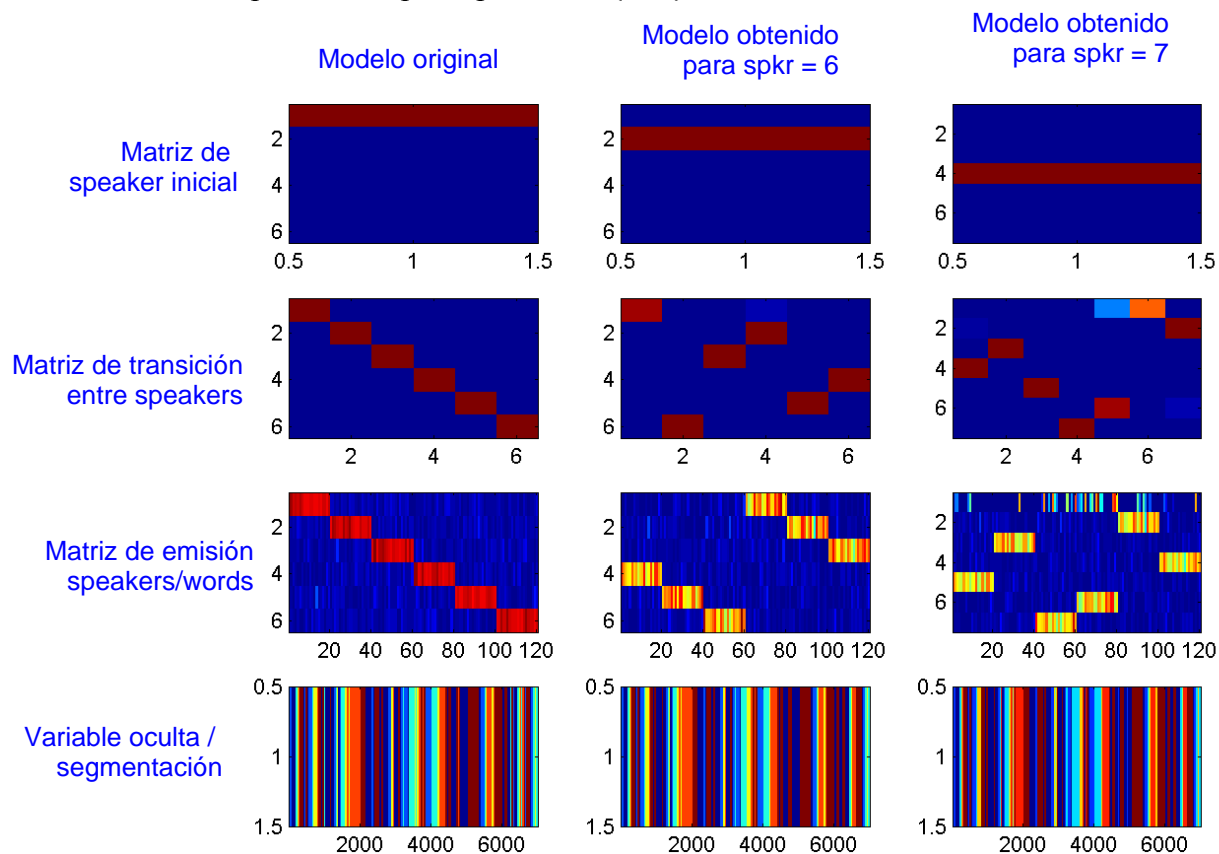
Número de muestras en el tiempo = 7000

Numero de speakers = 6

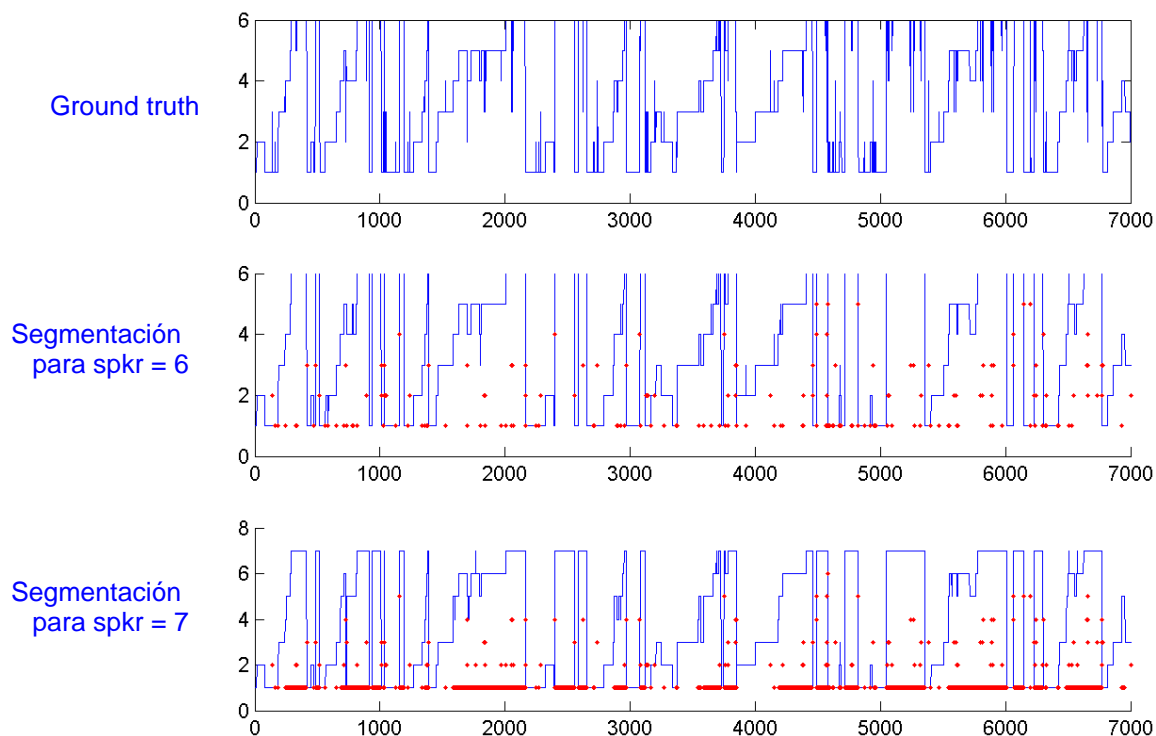
Estimación de parámetros para speakers = {3, 4}



Estimación de parámetros para speakers = {6, 7}



Segmentación/cambio de speakers en el tiempo



Probabilidades de emisión de palabras para cada speaker (speaker = 6)

