

PATTERN MINING

	Retranscription_1	Retranscription_2	Retranscription_3	...	Retranscription_N
ENFANT_1	Phrase_1_1 Phrase_1_2 . . . Phrase_1_m	Phrase_2_1 Phrase_2_2 . . . Phrase_2_p			Phrase_N_1 Phrase_N_2 . . . Phrase_N_q
ENFANT_2	Phrase_1_1 Phrase_1_2 . . . Phrase_1_m'				
ENFANT_3					
ENFANT_4					
ENFANT_5					
ENFANT_6					



	Sequences
E1_groupe1	< (Phrase_1_1) (Phrase_2_1) ... (Phrase_N_1) >
E1_groupe2	< ((Phrase_1_1)) (Phrase_2_1) ... (Phrase_N_1) >
.	↑ ITEMSET
.	
.	
E1_groupeT	< (Phrase_1_1) (Phrase_2_1) ... (Phrase_N_1) >
E2_groupe1	< (Phrase_1_1) (Phrase_2_1) ... (Phrase_N_1) >
.	
.	
.	
E2_groupeT'	< (Phrase_1_1) (Phrase_2_1) ... (Phrase_N_1) >
E3_groupe1	< (Phrase_1_1) (Phrase_2_1) ... (Phrase_N_1) >
.	
.	
.	
E6_groupeT ^N	< (Phrase_1_1) (Phrase_2_1) ... (Phrase_N_1) >

Chaque Prise_i_j est en réalité écrite sous forme d’une suite de phonèmes
*voir exemple plus bas

T = Nombre maximal de phrases

Choix :
Faire l’étude sur les données avec correspondance « a prononcer » et «
aurait dû prononcer »

Critères :
Phonème + sa localisation dans le mot : Pho1_d (début de mot)
Pho1_m (milieu de mot)
Pho1_f (fin de mot)

Phonème + valeur prononciation -> Pho1_c (prononciation correcte)
Pho1_i (prononciation incorrecte)

Combinaison de ces deux critères : Pho1_d_c
Pho3_f_i
...

Exemple:
Maman vient = mamã vjẽ
(Phrase_2_1) = (m_d_c a_m_c m_m_c ã_f_c v_d_i j_m_c
ẽ_f_i)