0.0.1. Etiquetar el corpus BNC con Stanford Tagger

La segunda evaluación de este experimento consiste en entrenar el etiquetador gramatical Stanford Tagger con WSJ como corpus de entrenamiento y con WSJ + NFI. Luego se procede a etiquetar el BNC plano (sin etiquetas gramaticales) con estos dos modelos. Por último se contruye la matriz de confusión:

Cuadro 1: *BNC original contra BNC etiquetado con MaxEnt (entrenado con WSJ)*

BNC \ MaxEnt(BNC)	**NNP**	**JJ**	**NN**	**VBN**	**NNS**	**WRB**	**RB**	**CD**	**VBG**	**NNPS**
NN1	**26141**	**4045**	-	115	533	-	253	16	1143	7
AJ0	**8675**	-	**2860**	**3276**	30	-	1033	3	2054	12
DT0	119	**8132**	192	-	5	-	443	-	-	-
AV0	982	2314	**3753**	159	236	-	-	160	85	4
NN0	567	1115	-	19	**3132**	-	10	2605	2	5
NN2	**2982**	90	750	-	-	-	6	47	-	1959
CJS	60	334	247	57	104	**2901**	522	1	35	2
AVP	43	17	137	-	-	-	2691	-	-	-
UNC	1754	255	540	9	199	-	2	426	1	18
CJT	-	1	-	-	-	-	-	-	-	-

Aciertos: 1.856.979 (92,86 %)
Errores: 142.739 (7,14 %)

Cuadro 2: *BNC original contra BNC etiquetado con MaxEnt (entrenado con WSJ + NFI)*

BNC \ MaxEnt(WSJ+NFI)	**NNP**	**JJ**	**NN**	**CD**	**NNS**	**VBN**	**WRB**	**RB**	**NNPS**	**VBG**
NN1	**26206**	**3864**	-	22	663	108	-	277	1	1166
AJ0	**8263**	-	2099	4	25	**3145**	-	876	12	1707
DT0	109	**7836**	188	-	4	-	-	793	-	-
AV0	840	2210	**3640**	195	279	123	-	-	2	70
NN0	469	863	-	**3222**	**3146**	6	-	-	8	9
CJS	102	374	357	-	147	37	**2901**	776	2	39
NN2	**2643**	68	783	74	-	1	-	8	1844	-
AVP	39	6	140	-	-	-	-	2101	-	-
PRP	497	1680	887	1	572	115	-	711	1	624
VVN	76	461	87	-	1	-	-	1	-	7

Aciertos: 1.859.888 (93,01 %)
Errores: 139.830 (6,99 %)

Se puede observar que el rendimiento del etiquetador entrenado con WSJ es un poco mejor (93,01 %) que cuando es entrenado con WSJ + NFI (92,86 %). La mayoría de los errores para Stanford Tagger entrenado con WSJ se da en

etiquetas NN del gold standard cuando son etiquetadas como JJ y NNP. Para Stanford Tagger entrenado con WSJ + NFI la mayoría de los errores se da en las mismas etiquetas, pero con cantidad de errores mayor, sobre todo para NN etiquetado como JJ.

La segunda evaluación de este experimento consiste en entrenar Stanford Tagger con la mitad de WSJ y con la mitad de WSJ + NFI. Posteriormente con estos dos modelos se etiqueta BNC y se construye la matriz de confusión. Se realiza la misma operación para cada mitad:

Cuadro 3: *BNC original contra BNC etiquetado con MaxEnt (entrenado con 2 mitad de WSJ)*

MaxEnt(2WSJ) / BNC	**NNP**	**JJ**	**NN**	**NNS**	**VBN**	**WRB**	**RB**	**CD**	**VBG**	**NNPS**
NN1	**26061**	**4689**	-	620	97	-	248	6	1296	8
AJ0	**8570**	-	**3574**	49	**3001**	-	1100	6	2190	10
DT0	133	**8068**	224	1	-	-	478	1	-	-
AV0	985	2428	**3611**	475	143	-	-	164	77	9
NN0	554	1404	-	**3128**	7	-	14	2633	7	6
NN2	**2984**	146	855	-	-	-	4	46	-	2054
CJS	98	210	237	146	34	**2901**	568	3	21	2
AVP	47	6	152	-	-	-	2793	-	-	-
UNC	1763	268	454	212	3	-	3	518	1	20
CJT	-	1	-	-	-	-	-	-	-	-

Aciertos: 1.851.792 (92,60 %)
Errores: 147.926 (7,40 %)

Cuadro 4: *BNC original contra BNC etiquetado con MaxEnt (entrenado con 2 mitad de WSJ+NFI)*

MaxEnt(2WSJ+NFI) / BNC	**NNP**	**JJ**	**NN**	**CD**	**NNS**	**VBN**	**WRB**	**RB**	**NNPS**	**VBG**
NN1	**26036**	**3867**	-	17	657	109	-	285	3	1179
AJ0	**8130**	-	2151	5	24	**3011**	-	897	13	1653
DT0	108	**7742**	215	-	3	-	-	867	-	-
AV0	875	2220	**3538**	203	276	129	-	-	1	68
NN0	452	866	-	**3240**	**3138**	8	-	-	11	8
CJS	123	327	333	2	158	58	**2901**	908	2	17
NN2	**2513**	75	800	81	-	1	-	10	1922	-
AVP	40	5	142	-	-	-	-	2014	-	-
VVD	75	203	89	-	-	1573	-	9	-	13
PRP	495	1528	758	2	689	109	-	723	-	587

Aciertos: 1.859.947 (93,01 %)
Errores: 139.771 (6,99 %)

Cuadro 5: *BNC original contra BNC etiquetado con MaxEnt (entrenado con 1 mitad de WSJ)*

BNC \ MaxEnt(1WSJ)	**NNP**	**JJ**	**NN**	**VBN**	**NNS**	**WRB**	**RB**	**CD**	**VBG**	**VBD**
NN1	**27101**	**4519**	-	146	550	-	241	7	1369	116
AJ0	**9043**	-	**3838**	**3837**	43	-	1025	3	2238	296
DT0	128	**8324**	185	-	1	-	461	-	-	-
AV0	1071	2583	**3445**	141	311	-	-	170	71	78
NN2	**3415**	83	885	-	-	-	7	41	-	1
NN0	573	955	-	24	**3189**	-	5	2732	3	8
CJS	82	412	267	54	106	**2901**	400	1	23	63
AVP	21	26	140	-	-	-	2859	-	-	-
VVN	85	464	126	-	1	-	1	-	1	1986
UNC	1757	181	509	9	242	-	3	442	1	5

Aciertos: 1.848.799 (92,45 %)
Errores: 150.919 (7,55 %)

Cuadro 6: *BNC original contra BNC etiquetado con MaxEnt (entrenado con 1 mitad de WSJ+NFI)*

BNC \ MaxEnt(1WSJ+NFI)	**NNP**	**JJ**	**NN**	**CD**	**VBN**	**NNS**	**WRB**	**RB**	**VBD**	**NNPS**
NN1	**26776**	**3786**	-	20	109	684	-	282	94	1
AJ0	**8368**	-	2113	3	**3249**	32	-	853	215	10
DT0	108	**7768**	192	-	-	1	-	883	-	-
AV0	950	2360	**3475**	194	90	325	-	-	39	2
NN0	473	752	-	**3302**	2	**3193**	-	-	4	4
CJS	151	409	348	-	33	150	**2901**	856	39	1
NN2	**2831**	66	797	74	1	-	-	8	1	1636
AVP	32	6	139	-	-	-	-	2080	-	-
VVN	77	492	94	-	-	1	-	1	1756	-
PRP	605	1613	925	-	129	741	-	802	122	-

Aciertos: 1.857.971 (92,91 %)
Errores: 141.747 (7,09 %)

Se puede apreciar una leve mejoría en el porcentaje de etiquetas acertadas; 92,6 % contra 93,01 % y 92,45 % contra 92,91 % para cada modelo respectivamente. Los errores más comunes son producidos en etiquetas NN del gold standard cuando son etiquetadas como JJ y NNP por , para las dos mitades entrenadas tanto con WSJ como con WSJ + NFI. Se puede notar que el porcentaje de error al etiquetar JJ cuando era NN es menor en la evaluación realizada sobre Stanford Tagger entrenado con WSJ + NFI.

A continuación se presentan las matrices de confusión para BNC etiquetado con Stanford Tagger entrenado con la mitad de WSJ con y sin NFI.

Cuadro 7: *BNC etiquetado por MaxEnt (entrenado con 2 mitad WSJ) vs BNC etiquetado con MaxEnt (entrenado con 2 mitad de WSJ + NFI)*

MaxEnt(2WSJ) \ MaxEnt(2WSJ+NFI)	**JJ**	**NN**	**NNP**	**VBN**	**NNS**	**VB**	**RP**	**VBG**	**RB**	**VBD**
NN	**4691**	-	**2413**	152	392	**1859**	11	1579	813	192
JJ	-	**3864**	1168	1524	85	356	2	336	1233	201
NNP	**1688**	**2727**	-	80	938	308	-	117	267	61
VBD	310	107	17	**2345**	-	135	-	6	21	-
VBZ	23	16	35	1	**1922**	6	-	-	24	9
IN	146	247	203	17	32	443	**1854**	75	1408	102
VBP	190	1075	35	28	15	**1693**	-	4	64	123
VBN	1351	99	69	-	13	66	-	12	20	1362
VBG	1220	1166	139	11	-	25	-	-	3	2
VB	300	1125	295	135	11	-	-	17	127	88

Aciertos: 1.933.574 (96,69 %)
Errores: 66.144 (3,31 %)

Cuadro 8: *BNC etiquetado por MaxEnt (entrenado con 1 mitad WSJ) vs BNC etiquetado con MaxEnt (entrenado con 1 mitad de WSJ + NFI)*

MaxEnt(1WSJ) \ MaxEnt(1WSJ+NFI)	**JJ**	**NN**	**NNP**	**VBN**	**NNS**	**VB**	**RP**	**RB**	**VBG**	**VBD**
NN	**5058**	-	**2370**	127	414	**1895**	15	372	1438	115
JJ	-	**3812**	1138	1148	80	360	8	1331	272	130
NNP	**2055**	**2670**	-	44	957	280	-	218	104	24
VBD	384	188	31	**2361**	-	186	-	21	28	-
VBZ	20	63	35	-	**2074**	37	-	43	1	12
VBN	**1973**	166	73	-	5	62	-	30	15	1366
IN	169	302	246	14	49	344	**1807**	1485	63	18
VBP	232	1021	27	27	12	1510	-	77	5	150
RB	1218	453	209	13	34	241	1081	-	21	13
VB	302	1204	269	145	20	-	-	139	23	107

Aciertos: 1.933.672 (96,70 %)
Errores: 66.046 (3,30 %)

La tercer evaluación de este experimento consiste en entrenar Stanford Tagger con un cuarto de WSJ y con un cuarto de WSJ + NFI. Posteriormente con estos dos modelos se etiqueta BNC y se construye la matriz de confusión. Se realiza la misma operación para cada uno de los cuartos:

Cuadro 9: *Rendimiento de Stanford Tagger entrenado con cuartos de WSJ con y sin NFI*

Evaluación	**Porcentaje de aciertos**

Cuadro 9: *Rendimiento de Stanford Tagger entrenado con cuartos de WSJ con y sin NFI*

Evaluación	**Porcentaje de aciertos**
Stanford Tagger entrenado con el primer 1/4 de WSJ	92.09 %
Stanford Tagger entrenado con el primer 1/4 de WSJ + NFI	92.92 %
Stanford Tagger entrenado con el segundo 1/4 de WSJ	92.10 %
Stanford Tagger entrenado con el segundo 1/4 de WSJ + NFI	92.91 %
Stanford Tagger entrenado con el tercer 1/4 de WSJ	92.14 %
Stanford Tagger entrenado con el tercer 1/4 de WSJ + NFI	92.89 %
Stanford Tagger entrenado con el cuarto 1/4 de WSJ	91.98 %
Stanford Tagger entrenado con el cuarto 1/4 de WSJ + NFI	92.83 %

En todos los casos se puede apreciar una mejora en el acierto de etiquetas para el corpus de entrenamiento WSJ + NFI contra WSJ.

La cuarta evaluación de este experimento consiste en entrenar Stanford Tagger con un décimo de WSJ y con un décimo de WSJ + NFI. Posteriormente con estos dos modelos se etiqueta BNC y se presentan los resultados:

- 91.25 % de acierto de etiquetas para el etiquetado de BNC con Stanford Tagger entrenado con 1/10 WSJ
- 92.81 % de acierto de etiquetas para el etiquetado de BNC con Stanford Tagger entrenado con 1/10 WSJ+NFI

Se puede apreciar un aumento del porcentaje de aciertos en el corpus de entrenamiento que incorpora NFI.

1. Conclusiones