

1. Experimentación

1.1. Primer experimento

El primer experimento consiste en medir (generando una matriz de confusión) la información extraída de COBUILD contra la misma información generada a partir de un etiquetador automático (TnT). Es decir, la información extraída de COBUILD, como se mencionó anteriormente, es la unión de definiciones y ejemplos, con la información gramatical correspondiente a la palabra definida. A continuación se presenta un pequeño extracto:

A	are
cat NN	often
is	kept
a	as
small	pets
furry	.
animal	She
with	put
a	out
tail	a
,	hand
whiskers	and
,	stroked
and	the
sharp	cat NN
claws	softly
that	...
kills	...
smaller	domestic
animals	animals
such	such
as	as
mice	dogs
and	and
birds	cats NNS
.	.
Cats NNS	

Esta es la información extraída de COBUILD para la palabra *cat*; la unión de la definición:

A cat is a small furry animal with a tail, whiskers, and sharp claws that kills smaller animals such as mice and birds. Cats are often kept as pets.

y los ejemplos

She put out a hand and stroked the cat softly...
...domestic animals such as dogs and cats.

Se puede notar la información gramatical expresada mediante las etiquetas NN y NNS para las palabras *cat* y *cats* respectivamente. La idea de este experimento será comparar estas etiquetas contra las etiquetas asignadas por el etiquetador automático TnT. Entonces se tomará este corpus plano (sin etiquetas), se lo etiquetará utilizando TnT entrenado con el corpus de entrenamiento Wall Street Journal (de ahora en más WSJ) ¹ y luego se realizará la comparación. La matriz de confusión² generada a partir de dicha comparación es la siguiente:

Cuadro 1: *Matriz de confusión para etiquetas extraídas de COBUILD vs generadas por TnT*

COBUILD \ TnT	NN	VB	JJ	VCN	RB	VBG	NNP	IN	VBZ	NNS
NN	-	556	1953	52	86	276	-	8	-	-
VB	2616	-	614	-	42	-	77	15	-	5
JJ	1577	96	-	1361	634	555	281	30	-	16
VCN	-	-	-	-	-	-	-	-	-	-
RB	219	23	408	10	-	9	34	249	-	11
VBG	-	-	-	-	-	-	-	-	-	-
NNP	-	-	-	-	-	-	-	-	-	-
IN	-	-	-	-	-	-	-	-	-	-
VBZ	-	-	-	-	-	-	-	-	-	-
NNS	83	1	17	-	1	2	104	3	192	-

Porcentaje de aciertos: 99,16 %

Cantidad de errores: 13082

Se puede apreciar un alto porcentaje de aciertos entre las etiquetas extraídas de COBUILD (99,16 %) y las etiquetas asignadas por TnT. Este porcentaje indica que la información de etiquetas extraídas de COBUILD es consistente con las producidas por TnT. La mayoría de los errores se da en etiquetas VB, NN y JJ de COBUILD cuando son etiquetadas como NN, JJ y NN por TnT respectivamente.

¹Wall Street Journal es un corpus anotado, parte del Penn Treebank

²Las matrices de confusión presentadas de aquí en adelante contienen las primeras 10 etiquetas de mayor error

1.2. Segundo experimento: Etiquetar el corpus WSJ

El segundo experimento realizado tiene como objetivo evaluar la nueva fuente de información obtenida (NFI) como corpus de entrenamiento. Para esto se entrenarán 2 etiquetadores gramaticales (Stanford Tagger y TnT) y se etiquetará con ellos el corpus Wall Street Journal (WSJ). Posteriormente se realizarán mediciones de desempeño pertinentes.

1.2.1. Etiquetar el corpus WSJ con TnT

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

Cuadro 2: *WSJ original contra WSJ etiquetado con TnT (entrenado con WSJ)*

WSJ \ TnT(WSJ)	JJ	NNP	VBN	RB	IN	RP	NNPS	VBD	NN	VB
NN	2592	1649	28	79	18	7	-	39	-	316
VBD	70	5	1642	-	-	-	-	-	42	41
IN	108	30	-	1476	-	1376	-	-	4	2
RB	726	45	4	-	1459	857	-	1	200	37
VBN	1262	22	-	-	-	-	-	1146	36	53
NNP	484	-	4	46	23	1	1236	3	378	7
JJ	-	699	916	601	69	23	1	42	903	70
VBG	424	20	-	-	-	-	-	-	884	-
VBP	26	10	33	11	10	1	1	51	349	872
WDT	-	-	-	-	722	-	-	-	-	-

Aciertos: 1.226.484 (97,10 %)

Errores: 36.636 (2,90 %)

Cuadro 3: *WSJ original contra WSJ etiquetado con TnT (entrenado con WSJ + NFI)*

WSJ \ TnT(WSJ+NFI)	JJ	NNP	VBN	IN	RP	NNPS	RB	VBD	VB	NN
NN	2794	2081	46	16	7	-	87	43	315	-
VBD	73	10	1716	-	-	-	-	-	37	22
RB	799	62	4	1664	1128	-	-	1	50	204
IN	107	51	-	-	1559	-	1222	-	3	4
VBN	1369	27	-	-	-	-	-	1219	42	33
NNP	358	-	4	19	1	1257	44	3	9	263
JJ	-	1013	893	87	28	1	634	47	74	864
VBP	30	12	33	10	1	1	8	56	885	379
VBG	518	27	-	-	-	-	-	-	-	816
VBZ	-	11	-	-	-	2	-	-	-	1

Aciertos: 1.226.967 (97,14 %)

Errores: 36.153 (2,86 %)

Se puede observar que el rendimiento del etiquetador TnT entrenado con WSJ+NFI es un poco mejor (97,14 %) que el rendimiento de TnT entrenado con WSJ (97,1 %). La mayoría de los errores para TnT entrenado con WSJ se da en etiquetas NN del gold standard cuando son etiquetadas como JJ y NNP por TnT. Para TnT 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 NNP.

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

Cuadro 4: 1 mitad WSJ original contra 1 mitad WSJ etiquetado con TnT (entrenado con 2 mitad de WSJ)

TnT(2WSJ) WSJ	JJ	NNP	VBN	NN	VBD	IN	RB	RP	VB	NNPS
NN	1959	1154	26	-	24	5	60	2	269	2
VBD	76	12	1129	19	-	-	1	-	29	-
JJ	-	545	801	1039	52	19	313	9	62	1
VBN	617	23	-	24	819	-	-	-	36	-
RB	432	25	3	91	2	808	-	318	19	-
IN	71	24	1	3	-	-	634	615	1	-
VBP	26	19	19	285	33	6	4	-	613	1
NNP	419	-	8	534	11	19	43	-	20	600
VBG	276	22	-	577	-	-	-	-	-	-
NNPS	26	549	-	-	-	-	-	-	-	-

Aciertos: 607.876 (96,25 %)

Errores: 23.695 (3,75 %)

Cuadro 5: 1 mitad WSJ original contra 1 mitad WSJ etiquetado con TnT (entrenado con 2 mitad de WSJ + NFI)

TnT(2WSJ+NFI) WSJ	JJ	NNP	VBN	IN	NN	RP	VBD	NNPS	VBG	VB
NN	1759	1287	29	6	-	3	27	1	556	213
VBD	54	17	1039	-	10	-	-	-	-	17
RB	434	30	2	872	81	511	1	-	-	23
JJ	-	689	612	37	838	6	29	-	219	45
IN	65	33	-	-	3	749	-	-	2	1
VBN	654	24	-	-	16	-	708	-	-	21
NNP	334	-	6	15	356	-	4	558	23	18
VBP	14	19	20	6	248	-	28	1	-	534

Cuadro 5: 1 mitad WSJ original contra 1 mitad WSJ etiquetado con TnT (entrenado con 2 mitad de WSJ + NFI)

TnT(2WSJ+NFI) WSJ	JJ	NNP	VBN	IN	NN	RP	VBD	NNPS	VBG	VB
NNPS	20	510	-	-	1	-	-	-	-	-
VBG	322	22	-	-	460	-	-	-	-	-

Aciertos: 609.255 (96,47 %)

Errores: 22.316 (3,53 %)

Cuadro 6: 2 mitad WSJ original contra 2 mitad WSJ etiquetado con TnT (entrenado con 1 mitad de WSJ)

TnT(1WSJ) WSJ	JJ	VBN	NNP	NN	RB	VBD	NNPS	IN	RP	VB
NN	1826	35	1089	-	51	27	-	11	6	256
VBD	73	1097	12	37	-	-	-	-	-	19
JJ	-	609	559	1085	360	69	2	44	18	78
IN	44	-	19	6	881	-	-	-	693	4
VBN	838	-	30	22	-	859	-	-	-	33
NNP	457	17	-	458	40	6	855	20	2	18
RB	384	3	45	163	-	-	-	741	517	19
VBP	35	17	14	187	8	31	-	7	1	560
VBG	294	-	21	552	1	-	-	-	-	1
VB	74	25	49	405	9	23	-	7	-	-

Aciertos: 607.593 (96,21 %)

Errores: 23.956 (3,79 %)

Cuadro 7: 2 mitad WSJ original contra 2 mitad WSJ etiquetado con TnT (entrenado con 1 mitad de WSJ)

TnT(1WSJ+NFI) WSJ	JJ	NNP	VBN	RP	IN	NN	NNPS	VBD	RB	VBG
NN	1831	1261	33	4	10	-	-	28	41	518
VBD	39	12	1065	-	-	14	-	-	-	-
IN	43	26	-	870	-	4	-	-	679	1
RB	411	51	2	682	861	137	-	-	-	1
VBN	829	26	-	-	-	22	-	726	-	-
JJ	-	664	495	21	51	819	2	32	372	125
NNP	333	-	17	1	11	321	790	4	28	15
VBP	22	9	15	1	7	194	-	31	4	-
VBG	322	30	-	-	-	462	-	-	1	-
VBZ	1	14	-	-	-	1	7	-	-	-

Aciertos: 608.633 (96,37 %)

Errores: 22.916 (3,63 %)

Se puede apreciar una leve mejoría en el porcentaje de etiquetas acertadas para el modelo que incorpora NFI; 96,25 % contra 96,47 % y 96,21 % contra 96,37 % para cada mitad respectivamente. Los errores más comunes son producidos en etiquetas NN del gold standard cuando son etiquetadas como JJ y NNP por TnT, 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 TnT entrenado con el modelo que incorpora NFI.

A continuación se presentan las matrices de confusión entre las mitades de WSJ etiquetado con TnT entrenado con la mitad restante con y sin NFI.

Cuadro 8: 1 mitad WSJ etiquetado por TnT (entrenado con 2 mitad WSJ) vs 1 mitad WSJ etiquetado con TnT (entrenado con 2 mitad de WSJ + NFI)

TnT(2WSJ+NFI) TnT(2WSJ)	NN	JJ	VBN	NNP	VBD	VBG	VBP	VB	RP	IN
JJ	592	-	149	353	38	64	11	62	1	26
NN	-	580	30	488	16	338	97	315	-	3
VBD	8	37	506	13	-	-	6	7	-	-
VBN	17	377	-	7	483	-	4	8	-	-
VB	203	55	25	12	22	-	321	-	1	-
RB	29	147	1	14	1	-	-	2	292	229
VBP	99	3	-	1	11	-	-	223	-	2
VBG	167	217	-	36	1	-	-	1	-	-
VBZ	-	-	-	6	-	-	-	1	-	-
NNS	56	9	-	115	-	-	-	-	-	-

Aciertos: 621.391 (98,39 %)

Errores: 10.184 (1,61 %)

Cuadro 9: 2 mitad WSJ etiquetado por TnT (entrenado con 1 mitad WSJ) vs 2 mitad WSJ etiquetado con TnT (entrenado con 1 mitad de WSJ + NFI)

TnT(1WSJ+NFI) TnT(1WSJ)	JJ	VBN	NN	VBD	NNP	RP	IN	VB	VBP	VBG
NN	765	28	-	27	436	1	4	292	72	270
VBD	59	525	11	-	2	-	-	8	3	-
JJ	-	193	497	47	335	-	11	48	23	82
VBN	286	-	22	439	4	-	-	11	5	-
RB	160	-	38	-	30	326	301	15	3	-
VB	44	18	173	16	17	1	1	-	275	1
VBZ	-	-	-	-	5	-	-	1	-	-
VBP	14	3	101	7	4	-	1	219	-	-
VBG	190	1	156	2	18	-	-	6	2	-
NNS	9	-	50	-	72	-	-	-	-	-

Aciertos: 621.749 (98,45 %)

Errores: 9.802 (1,55 %)

La tercer evaluación de este experimento consiste en entrenar TnT con un cuarto de WSJ y con un cuarto de WSJ + NFI. Posteriormente con estos dos modelos se etiqueta los 3/4 restantes de WSJ y se construye la matriz de confusión. Se realiza la misma operación para cada uno de los cuartos:

Cuadro 10: *Rendimiento de TnT entrenado con cuartos de WSJ con y sin NFI*

Evaluación	Porcentaje de aciertos
TnT entrenado con el primer 1/4 de WSJ	95.93 %
TnT entrenado con el primer 1/4 de WSJ + NFI	96.26 %
TnT entrenado con el segundo 1/4 de WSJ	95.89 %
TnT entrenado con el segundo 1/4 de WSJ + NFI	96.26 %
TnT entrenado con el tercer 1/4 de WSJ	95.91 %
TnT entrenado con el tercer 1/4 de WSJ + NFI	96.29 %
TnT entrenado con el cuarto 1/4 de WSJ	95.9 %
TnT entrenado con el cuarto 1/4 de WSJ + NFI	96.30 %

En todos los casos se puede apreciar una mejora en el acierto de etiquetas para el modelo que incorpora NFI.

La cuarta evaluación de este experimento consiste en entrenar TnT con un décimo de WSJ y con un décimo de WSJ + NFI. Posteriormente con estos dos modelos se etiqueta los 9/10 restantes de WSJ y se presentan los resultados:

- 95.32 % de acierto de etiquetas para el etiquetado de 9/10 de WSJ con TnT entrenado con 1/10 WSJ
- 96.1 % de acierto de etiquetas para el etiquetado de 9/10 de WSJ con TnT entrenado con 1/10 WSJ+NFI

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

1.2.2. Etiquetar el corpus WSJ 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 WSJ plano (sin etiquetas gramaticales) con estos dos modelos. Por último se contruye la matriz de confusión:

Cuadro 11: *WSJ original contra WSJ etiquetado con MaxEnt (entrenado con WSJ)*

MaxEnt(WSJ)	JJ	IN	NN	NNP	VBD	RB	VRN	VBP	RP	JJR
WSJ										
NN	1726	15	-	1132	16	61	18	28	1	2
RB	736	1593	189	139	-	-	3	1	293	36
JJ	-	60	1276	632	51	515	762	7	-	5
VRN	894	-	44	25	1052	1	-	5	-	-
NNPS	40	-	-	997	-	-	-	-	-	-
IN	87	-	4	22	-	959	-	2	527	-
VBG	196	-	829	14	-	-	1	1	-	-
VBD	40	-	26	8	-	-	806	14	-	-
RP	4	628	-	1	-	230	-	-	-	-
VB	58	8	365	36	37	12	22	544	-	6

Aciertos: 1.236.647 (97,90 %)

Errores: 26.477 (2,10 %)

Cuadro 12: *WSJ original contra WSJ etiquetado con MaxEnt (entrenado con WSJ + NFI)*

MaxEnt(WSJ+NFI)	JJ	IN	NNP	NN	RB	VBD	RP	VRN	VBP	NNPS
WSJ										
NN	1918	16	1314	-	73	21	3	19	30	-
RB	742	1403	141	177	-	-	527	3	3	-
JJ	-	68	685	1260	583	45	2	851	8	-
IN	107	-	29	3	1062	-	1005	-	2	-
VRN	980	-	29	49	-	1049	-	-	6	-
NNPS	39	-	935	-	-	-	-	-	-	-
VBD	34	-	10	24	-	-	-	923	15	-
VBG	294	-	25	817	-	-	-	1	1	-
NNP	555	29	-	458	21	4	1	9	4	546
VB	62	10	29	361	13	50	-	38	555	-

Aciertos: 1.234.495 (97,73 %)

Errores: 28.629 (2,27 %)

Se puede observar que el rendimiento del etiquetador entrenado con WSJ es un poco mejor (97,9 %) que cuando es entrenado con WSJ + NFI (97,73 %). 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 la mitad restante de WSJ y se construye la matriz de confusión. Se realiza la misma operación para cada mitad:

Cuadro 13: 1 mitad WSJ original contra 1 mitad WSJ etiquetado con MaxEnt (entrenado con 2 mitad de WSJ)

MaxEnt(2WSJ) WSJ	JJ	NN	NNP	IN	VBN	VBD	RB	NNS	VBG	JJR
NN	1558	-	1027	6	22	15	38	133	403	8
JJ	-	1309	606	32	746	39	299	65	263	5
RB	512	104	31	989	2	1	-	4	1	14
NNPS	31	-	943	-	-	-	-	246	-	-
VBN	545	28	36	-	-	722	-	-	-	-
VBG	192	614	22	-	1	-	-	-	-	-
VBD	41	26	9	-	604	-	-	-	-	-
NNP	401	542	-	23	8	10	38	156	37	-
IN	72	5	26	-	1	-	489	-	2	-
RP	2	3	1	449	-	-	179	-	-	-

Aciertos: 610.045 (96,59 %)

Errores: 21.529 (3,41 %)

Cuadro 14: 1 mitad WSJ original contra 1 mitad WSJ etiquetado con MaxEnt (entrenado con 2 mitad de WSJ + NFI)

MaxEnt(2WSJ+NFI) WSJ	JJ	NN	NNP	IN	VBN	VBD	RP	RB	NNS	VB
NN	1434	-	1033	6	13	14	1	41	124	124
JJ	-	1119	571	31	625	31	1	334	68	26
RB	470	81	73	851	2	1	251	-	1	19
NNPS	27	3	834	-	-	-	-	-	215	-
VBD	35	16	14	-	748	-	-	-	-	10
VBN	564	26	34	-	-	642	-	-	-	11
IN	80	7	27	-	-	-	573	531	-	1
VBG	262	531	26	-	1	-	-	-	-	-
NNP	445	497	-	19	3	6	-	24	141	16
VBZ	-	1	17	-	-	-	-	-	412	-

Aciertos: 611.099 (96,76 %)

Errores: 20.475 (3,24 %)

Cuadro 15: 2 mitad WSJ original contra 2 mitad WSJ etiquetado con MaxEnt (entrenado con 1 mitad de WSJ)

MaxEnt(1WSJ) WSJ	JJ	NN	IN	NNP	VBD	RB	VBN	NNPS	NNS	VBG
NN	1604	-	12	916	16	36	21	-	150	381
JJ	-	1197	45	522	37	381	483	-	46	202
RB	466	168	944	62	1	-	1	-	3	1
VBN	863	29	-	32	779	1	-	-	-	-
IN	50	3	-	26	-	698	-	-	-	2
NNPS	16	-	-	651	-	-	-	-	167	-
VBG	198	572	-	19	-	-	-	-	-	-
VBD	66	43	2	16	-	-	570	-	-	-
NNP	462	503	18	-	2	19	19	518	131	16
RP	3	1	426	-	-	129	-	-	-	-

Aciertos: 610.309 (96,64 %)

Errores: 21.241 (3,36 %)

Cuadro 16: 2 mitad WSJ original contra 2 mitad WSJ etiquetado con MaxEnt (entrenado con 1 mitad de WSJ)

MaxEnt(1WSJ+NFI) WSJ	JJ	NNP	NN	IN	VBN	RB	VBD	RP	NNPS	NNS
NN	1482	1011	-	10	18	42	12	2	-	146
JJ	-	522	997	43	444	382	26	8	1	40
RB	438	105	141	819	1	-	-	344	-	1
VBN	810	27	28	-	-	-	706	-	-	-
VBD	40	11	20	-	742	-	-	-	-	-
IN	50	29	3	-	-	727	-	586	-	-
NNPS	13	597	-	-	-	-	-	-	-	171
VBG	256	21	530	-	-	-	-	-	-	-
NNP	475	-	467	16	8	17	3	1	483	140
VBZ	-	9	1	-	-	-	-	-	2	406

Aciertos: 610.874 (96,73 %)

Errores: 20.676 (3,27 %)

Se puede apreciar una leve mejoría en el porcentaje de etiquetas acertadas; 96,23 % contra 96,46 % y 96,20 % contra 96,36 % para cada mitad respectivamente. Los errores más comunes son producidos en etiquetas NN del gold standard cuando son etiquetadas como JJ y NNP por TnT, 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 TnT entrenado con WSJ + NFI.

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

Cuadro 17: 1 mitad WSJ etiquetado por MaxEnt (entrenado con 2 mitad WSJ) vs 1 mitad WSJ etiquetado con MaxEnt (entrenado con 2 mitad de WSJ + NFI)

MaxEnt(2WSJ+NFI) MaxEnt(2WSJ)	JJ	NN	RP	VBN	RB	NNP	VB	IN	VBD	VBG
NN	686	-	2	13	46	291	276	-	21	213
JJ	-	596	1	175	202	183	48	15	19	49
IN	17	6	507	-	318	8	-	-	-	-
VBD	29	9	-	460	-	11	11	-	-	-
VBN	309	18	-	-	-	9	2	-	248	-
NNP	268	242	-	5	8	-	14	4	-	10
WDT	-	-	-	-	-	-	-	252	-	-
VBP	15	123	-	5	2	6	246	-	13	-
RB	128	17	206	1	-	65	14	141	-	-
VBG	199	196	-	-	1	26	-	-	-	-

Aciertos: 622.105 (98,50 %)

Errores: 9.469 (1,50 %)

Cuadro 18: 2 mitad WSJ etiquetado por MaxEnt (entrenado con 1 mitad WSJ) vs 2 mitad WSJ etiquetado con MaxEnt (entrenado con 1 mitad de WSJ + NFI)

MaxEnt(1WSJ+NFI) MaxEnt(1WSJ)	JJ	NN	VBN	RP	NNP	RB	VB	VBD	IN	NNS
NN	611	-	25	1	344	32	244	27	5	45
JJ	-	513	254	-	196	195	48	37	9	23
VBD	28	12	494	-	3	1	2	-	-	-
IN	9	3	-	494	18	326	-	2	-	-
VBP	17	132	6	-	4	-	283	15	2	1
VBN	246	18	-	-	7	-	9	257	-	-
WDT	-	-	-	-	-	-	-	-	255	-
RB	149	13	1	222	70	-	19	-	166	-
NNP	211	215	8	-	-	26	24	4	5	87
VBZ	-	1	-	-	5	1	-	-	-	208

Aciertos: 622.115 (98,51 %)

Errores: 9.435 (1,49 %)

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 los 3/4 restantes de WSJ y se construye la matriz de confusión. Se realiza la misma operación para cada uno de los cuartos:

Cuadro 19: Rendimiento de Stanford Tagger entrenado con cuartos de WSJ con y sin NFI

Evaluación	Porcentaje de aciertos
------------	------------------------

Cuadro 19: 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	96.30 %
Stanford Tagger entrenado con el primer 1/4 de WSJ + NFI	96.57 %
Stanford Tagger entrenado con el segundo 1/4 de WSJ	96.30 %
Stanford Tagger entrenado con el segundo 1/4 de WSJ + NFI	96.52 %
Stanford Tagger entrenado con el tercer 1/4 de WSJ	96.28 %
Stanford Tagger entrenado con el tercer 1/4 de WSJ + NFI	96.57 %
Stanford Tagger entrenado con el cuarto 1/4 de WSJ	96.24 %
Stanford Tagger entrenado con el cuarto 1/4 de WSJ + NFI	96.53 %

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 los 9/10 restantes de WSJ y se presentan los resultados:

- 95.67 % de acierto de etiquetas para el etiquetado de 9/10 de WSJ con Stanford Tagger entrenado con 1/10 WSJ
- 96.27 % de acierto de etiquetas para el etiquetado de 9/10 de WSJ 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.

1.3. Tercer experimento: Etiquetar el corpus BNC

El tercer experimento realizado tiene como objetivo evaluar la nueva fuente de información obtenida (NFI) como corpus de entrenamiento. Para esto se entrenarán 2 etiquetadores gramaticales (Stanford Tagger y TnT) y se etiquetará con ellos el British National Corpus (BNC).

1.3.1. Etiquetar el corpus BNC con TnT

La primer evaluación de este experimento consiste en entrenar el etiquetador gramatical TnT 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 20: *BNC original contra BNC etiquetado con TnT (entrenado con WSJ)*

TnT(BNC) BNC	NNP	JJ	NN	VBN	NNS	WRB	CD	NNPS	VBD	RB
NN1	26585	5739	-	146	732	3	52	13	157	378
AJ0	8608	-	2352	3680	50	1	34	24	327	1092
DT0	83	7771	208	-	1	-	-	-	-	988
AV0	1014	2028	4277	10	242	6	188	-	7	-
NN0	469	342	-	7	3133	-	2601	18	2	8
CJS	217	315	581	20	76	2903	-	-	43	1202
NN2	2472	75	690	2	-	-	77	2578	-	5
VVN	56	361	94	-	-	-	-	-	2365	1
VVD	54	235	87	2159	-	-	-	-	-	6
AVP	25	7	140	-	-	1	-	-	-	2070

Aciertos: 1.849.040 (92,47 %)

Errores: 150.675 (7,53 %)

Cuadro 21: *BNC original contra BNC etiquetado con TnT (entrenado con WSJ + NFI)*

TnT(WSJ+NFI) BNC	NNP	JJ	NN	NNS	VBN	WRB	NNPS	VBD	CD	WDT
NN1	26413	4240	-	663	101	1	8	118	56	1
AJ0	8621	-	1806	33	3361	1	21	291	20	1
DT0	75	7689	206	1	-	-	-	-	-	802
AV0	1068	1756	4347	226	16	2	-	4	125	4
NN0	447	474	-	3366	9	-	14	-	2120	-
CJS	222	252	705	101	18	2903	6	45	-	54
NN2	2328	85	795	-	-	-	2542	-	19	-
VVN	46	458	79	-	-	-	-	2349	-	-
VVD	53	217	76	-	2343	-	-	-	-	-
CJT	-	-	1	-	-	-	-	-	-	1857

Aciertos: 1.854.577 (92,74 %)

Errores: 145.138 (7,26 %)

Se puede observar que el rendimiento del etiquetador TnT entrenado con WSJ+NFI es un poco mejor (97,14 %) que el rendimiento de TnT entrenado con WSJ (97,1 %). La mayoría de los errores para TnT entrenado con WSJ se da en etiquetas NN del gold standard cuando son etiquetadas como JJ y NNP por TnT. Para TnT 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 NNP.

La segunda evaluación de este experimento consiste en entrenar TnT 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 22: BNC original contra BNC etiquetado con TnT (entrenado con 2 mitad de WSJ)

TnT(2WSJ) BNC	NNP	JJ	VBN	NN	NNS	WRB	VBD	RB	VBG	NNPS
NN1	27194	6142	180	-	804	5	191	396	1735	9
AJ0	8813	-	4421	3008	69	1	371	1106	2191	22
DT0	82	7802	-	200	1	-	-	903	-	-
AV0	1008	2145	15	4196	314	5	6	-	41	2
NN0	483	650	8	-	3410	-	1	9	4	19
NN2	2962	76	1	734	-	-	-	19	-	2174
CJS	202	314	27	566	86	2904	45	1152	15	-
VVN	44	432	-	106	-	-	2707	1	7	-
VVD	45	258	2390	122	-	-	-	7	20	-
AVP	18	7	-	144	-	-	-	2335	-	-

Aciertos: 1.841.617 (92,09 %)

Errores: 158.098 (7,91 %)

Cuadro 23: BNC original contra BNC etiquetado con TnT (entrenado con 2 mitad de WSJ+NFI)

TnT(2WSJ+NFI) BNC	NNP	JJ	NN	VBN	NNS	WRB	NNPS	VBD	CD	WDT
NN1	26728	4264	-	101	663	1	8	119	61	1
AJ0	8795	-	1793	3481	34	-	21	295	22	-
DT0	86	7648	204	-	1	-	-	-	-	869
AV0	1074	1881	4382	15	216	1	-	4	130	4
NN0	445	424	-	9	3380	-	20	-	2177	-
CJS	229	260	738	15	102	2903	1	48	-	17
NN2	2670	76	795	-	-	-	2331	-	22	-
VVD	48	227	87	2458	-	-	-	-	-	-

Cuadro 23: *BNC original contra BNC etiquetado con TnT (entrenado con 2 mitad de WSJ+NFI)*

$\begin{matrix} \text{TnT(2WSJ+NFI)} \\ \text{BNC} \end{matrix}$	NNP	JJ	NN	VBN	NNS	WRB	NNPS	VBD	CD	WDT
VVN	41	481	82	-	-	-	-	2318	-	-
CJT	6	-	1	-	-	-	-	-	-	1857

Aciertos: 1.853.072 (92,67 %)

Errores: 146.643 (7,33 %)

Cuadro 24: *BNC original contra BNC etiquetado con TnT (entrenado con 1 mitad de WSJ)*

$\begin{matrix} \text{TnT(1WSJ)} \\ \text{BNC} \end{matrix}$	NNP	JJ	NN	VBN	NNS	WRB	NNPS	VBD	VBG	RB
NN1	26286	6528	-	187	783	3	36	169	1614	435
AJ0	8503	-	2921	3713	64	1	28	476	2027	1238
DT0	84	7763	232	-	1	-	-	-	-	976
AV0	1157	2031	4455	30	515	1	1	8	42	-
NN0	478	796	-	6	3225	-	21	2	14	4
CJS	193	235	626	24	77	2903	6	51	14	1165
NN2	2504	87	863	2	-	-	2714	-	-	1
VVN	57	615	95	-	-	-	-	2581	13	5
PRP	733	1415	2260	57	499	3	2	-	467	614
VVD	67	317	99	2126	1	-	-	-	27	5

Aciertos: 1.842.527 (92,14 %)

Errores: 157.188 (7,86 %)

Cuadro 25: *BNC original contra BNC etiquetado con TnT (entrenado con 1 mitad de WSJ+NFI)*

$\begin{matrix} \text{TnT(1WSJ+NFI)} \\ \text{BNC} \end{matrix}$	NNP	JJ	NN	NNS	VBN	WRB	NNPS	VBD	CD	WDT
NN1	26360	4345	-	677	107	2	14	127	63	3
AJ0	8729	-	1808	35	3290	1	20	303	24	1
DT0	91	7655	210	1	-	-	-	-	-	870
AV0	1225	1630	4409	215	18	-	1	4	143	4
NN0	442	501	-	3321	8	-	21	-	2174	-
CJS	221	256	696	100	18	2903	6	47	-	63
NN2	2327	85	815	-	-	-	2568	-	17	-
VVD	60	218	76	-	2429	-	-	-	-	-
VVN	62	481	84	-	-	-	-	2381	-	-
CJT	-	-	1	-	-	-	-	-	-	1867

Aciertos: 1.853.701 (92,70 %)

Errores: 146.014 (7,30 %)

Se puede apreciar una leve mejoría en el porcentaje de etiquetas acertadas para el modelo que incorpora NFI; 92,09 % contra 92,67 % y 92,14 % contra 92,7 % para cada mitad respectivamente. Los errores más comunes son producidos en etiquetas NN del gold standard cuando son etiquetadas como JJ y NNP por TnT, 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 TnT entrenado con el modelo que incorpora NFI.

A continuación se presentan las matrices de confusión entre las mitades de WSJ etiquetado con TnT entrenado con la mitad restante con y sin NFI.

Cuadro 26: BNC etiquetado por TnT (entrenado con 1 mitad WSJ) vs BNC etiquetado con TnT (entrenado con 1 mitad de WSJ + NFI)

TnT(1WSJ+NFI) TnT(1WSJ)	NN	JJ	VBN	VB	NNP	NNS	VBD	VBP	VBG	RB
JJ	5144	-	1092	693	1701	128	283	82	474	913
NN	-	3630	65	2399	2215	390	58	417	1168	661
VBD	115	366	2487	93	47	1	-	21	-	7
VB	1960	291	131	-	230	1	97	1290	5	103
VBZ	12	65	-	5	38	1931	4	3	-	16
NNP	1870	1178	42	270	-	618	21	32	87	152
VBN	155	1512	-	70	78	1	1865	16	5	12
VBP	641	125	24	1343	41	2	38	-	-	27
RB	491	1215	4	238	279	5	3	9	-	-
VBG	1028	1198	5	21	188	5	23	7	-	5

Aciertos: 1.938.000 (96,91 %)

Errores: 61.726 (3,09 %)

Cuadro 27: BNC etiquetado por TnT (entrenado con 2 mitad WSJ) vs BNC etiquetado con TnT (entrenado con 2 mitad de WSJ + NFI)

TnT(2WSJ+NFI) TnT(2WSJ)	NN	JJ	VBN	VB	NNP	NNS	VBD	VBP	RP	VBG
JJ	5020	-	846	433	1470	73	260	77	2	324
NN	-	4138	67	2355	2027	180	81	455	3	932
VBD	133	304	2442	46	38	2	-	25	-	-
NNP	2220	1337	23	242	-	708	10	22	-	85
VB	2193	359	116	-	140	3	189	1602	2	4
VBN	185	2091	-	53	28	-	1882	18	-	2
VBZ	16	31	-	3	51	1998	-	10	-	-
VBP	599	50	10	1271	37	1	69	-	-	-
RB	539	1146	-	467	237	26	1	24	1177	-
VBG	1166	1172	-	5	200	3	10	-	-	-

Aciertos: 1.938.152 (96,92 %)

Errores: 61.574 (3,08 %)

La tercer evaluación de este experimento consiste en entrenar TnT con un cuarto de WSJ y con un cuarto de WSJ + NFI. Posteriormente con estos dos modelos se etiqueta los 3/4 restantes de WSJ y se construye la matriz de confusión. Se realiza la misma operación para cada uno de los cuartos:

Cuadro 28: Rendimiento de TnT entrenado con cuartos de WSJ con y sin NFI

Evaluación	Porcentaje de aciertos
TnT entrenado con el primer 1/4 de WSJ	91.75 %
TnT entrenado con el primer 1/4 de WSJ + NFI	92.62 %
TnT entrenado con el segundo 1/4 de WSJ	91.74 %
TnT entrenado con el segundo 1/4 de WSJ + NFI	92.63 %
TnT entrenado con el tercer 1/4 de WSJ	91.64 %
TnT entrenado con el tercer 1/4 de WSJ + NFI	92.62 %
TnT entrenado con el cuarto 1/4 de WSJ	91.64 %
TnT entrenado con el cuarto 1/4 de WSJ + NFI	92.58 %

En todos los casos se puede apreciar una mejora en el acierto de etiquetas para el modelo que incorpora NFI.

La cuarta evaluación de este experimento consiste en entrenar TnT 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:

- 90.9 % de acierto de etiquetas para el etiquetado de BNC con TnT entrenado con 1/10 WSJ
- 92.55 % de acierto de etiquetas para el etiquetado de BNC con TnT entrenado con 1/10 WSJ+NFI

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

1.3.2. 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 29: *BNC original contra BNC etiquetado con MaxEnt (entrenado con WSJ)*

MaxEnt(BNC) 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 30: *BNC original contra BNC etiquetado con MaxEnt (entrenado con WSJ + NFI)*

MaxEnt(WSJ+NFI) BNC	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 31: *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 32: *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 33: BNC original contra BNC etiquetado con MaxEnt (entrenado con 1 mitad de WSJ)

MaxEnt(1WSJ) BNC	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 34: BNC original contra BNC etiquetado con MaxEnt (entrenado con 1 mitad de WSJ+NFI)

MaxEnt(1WSJ+NFI) BNC	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 35: BNC etiquetado por MaxEnt (entrenado con 2 mitad WSJ) vs BNC etiquetado con MaxEnt (entrenado con 2 mitad de WSJ + NFI)

MaxEnt(2WSJ+NFI) MaxEnt(2WSJ)	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 36: BNC etiquetado por MaxEnt (entrenado con 1 mitad WSJ) vs BNC etiquetado con MaxEnt (entrenado con 1 mitad de WSJ + NFI)

MaxEnt(1WSJ+NFI) MaxEnt(1WSJ)	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 37: Rendimiento de Stanford Tagger entrenado con cuartos de WSJ con y sin NFI

Evaluación	Porcentaje de aciertos
------------	------------------------

Cuadro 37: 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.

2. Conclusiones