

# Appendix 8:

## Model selection for RELATED-WORDS (EXP7) in Ch. 10

Gunn Inger Lyse

19th April 2011

**Content:** Results from model selection (cross-validation) in Chapter 10 with knowledge source=RELATED-WORDS. The target words are ordered alphabetically. Evaluated with 5-fold cross validation and Overall Accuracy (measured as total recall). The best accuracy in each group is marked in bold-face (in case of ties, the model with the smallest context window is selected).

*friskAJ* cross-validation results (baseline: 0.683)

	SF-DEN									
	0	4	5	10	15	20	30	40	50	
(50.4)	.439	.451	.463	.390	.439	.439	.476	.476	.476	
(30.20)	<b>.524</b>	.500	.500	.463	.463	<b>.524</b>	.512	.537	<b>.549</b>	
(30.4)	.488	<b>.524</b>	.500	.488	.500	.488	.463	.488	.512	
(20.20)	.549	.561	.561	.537	.573	.549	.512	.512	.537	
(20.4)	.500	.524	.537	.488	.488	.500	.476	.500	.476	
(10.30)	.634	<b>.671</b>	<b>.683</b>	<b>.671</b>	.622	.634	.659	.659	<b>.671</b>	
(4.30)	.585	.671	.659	.695	.622	.585	.634	.610	.610	
(2.30)	<b>.646</b>	<b>.707</b>	.695	<b>.707</b>	<b>.646</b>	<b>.646</b>	<b>.646</b>	.634	.634	
(1.30)	.598	.671	.683	.671	.610	.598	.622	.610	.598	

Table 1: friskAJ

*fullAJ* cross-validation results (baseline: 0.941)

	SF-DEN									
	0	4	5	10	15	20	30	40	50	
(10.1)	.823	.775	.784	.814	.807	.823	.848	<b>.850</b>	.843	
(4.4)	.832	.807	.811	.830	.830	.830	.834	.832	.832	
(4.2)	<b>.852</b>	.836	.836	.841	.848	<b>.850</b>	<b>.852</b>	<b>.859</b>	<b>.852</b>	
(4.1)	<b>.866</b>	.845	.845	.852	<b>.861</b>	<b>.868</b>	.852	<b>.861</b>	.852	
(2.10)	.823	.802	.793	.814	.816	.825	.814	.823	.814	
(2.4)	.839	.830	.830	.825	.827	.836	.848	.841	.848	
(2.2)	.857	.832	.834	.848	.850	.855	<b>.861</b>	<b>.866</b>	<b>.866</b>	
(1.10)	.843	.809	.811	.816	.825	.848	.818	.830	.836	
(1.4)	<b>.864</b>	.830	.830	.843	.850	<b>.861</b>	<b>.864</b>	<b>.861</b>	<b>.866</b>	

Table 2: fullAJ

*fyr*N cross-validation results (baseline: 0.789)

SF-DEN									
(30.10)	<b>.825</b>	<b>.842</b>	<b>.842</b>	<b>.825</b>	<b>.807</b>	<b>.825</b>	<b>.807</b>	.789	.772
(30.1)	.754	.789	<b>.807</b>	<b>.825</b>	.789	.754	.789	.772	<b>.807</b>
(20.50)	.596	.579	.632	.632	.614	.596	.632	.632	.649
(20.10)	.772	.772	.789	<b>.860</b>	.754	.772	<b>.807</b>	.737	.737
(20.2)	.754	<b>.842</b>	<b>.842</b>	.825	.789	.754	.789	.772	.772
(10.4)	.684	.789	.789	.789	.737	.702	.702	.719	.754
(10.2)	.702	.754	.754	.754	.737	.702	.754	.737	.737
(4.4)	.807	.807	.807	<b>.842</b>	.807	.807	<b>.789</b>	.754	.754
(2.4)	<b>.789</b>	<b>.825</b>	<b>.825</b>	<b>.825</b>	<b>.789</b>	<b>.789</b>	<b>.789</b>	.772	<b>.789</b>
	0	4	5	10	15	20	30	40	50

Table 3: *fyr*N

*galAJ* cross-validation results (baseline: 0.776)

SF-DEN									
(20.20)	.629	.716	.716	.672	.638	.629	.629	.638	.638
(10.30)	.655	<b>.698</b>	.690	.655	.655	.664	.690	.672	<b>.698</b>
(10.20)	.672	<b>.724</b>	.716	<b>.698</b>	<b>.698</b>	.672	.638	.647	.638
(10.10)	.629	.655	.664	.621	.621	.629	.638	.638	.612
(10.4)	.647	<b>.681</b>	<b>.681</b>	<b>.681</b>	.638	.647	.655	.621	.621
(10.2)	.664	.647	.655	.655	.647	.664	.672	<b>.707</b>	.672
(10.1)	.655	.664	.664	.664	.647	.655	.681	<b>.759</b>	<b>.707</b>
(4.20)	.664	<b>.707</b>	<b>.707</b>	.681	.664	.672	.672	.681	.672
(2.20)	.664	.655	.655	.664	<b>.690</b>	.664	.647	.672	.664
	0	4	5	10	15	20	30	40	50

Table 4: *galAJ*

*lagN* cross-validation results (baseline: 0.703)

SF-DEN									
(50.4)	.784	.757	.757	.757	.757	<b>.811</b>	<b>.811</b>	<b>.811</b>	<b>.811</b>
(50.2)	.784	<b>.811</b>	.784	.757	.784	.784	<b>.811</b>	<b>.811</b>	<b>.811</b>
(50.1)	.784	<b>.811</b>	<b>.811</b>	.757	.784	.784	<b>.811</b>	<b>.811</b>	.784
(30.30)	.811	.838	.838	.784	.811	.811	.811	<b>.892</b>	<b>.892</b>
(30.10)	.838	<b>.919</b>	<b>.919</b>	<b>.946</b>	.838	.838	.865	.838	.811
(30.4)	.865	<b>.919</b>	<b>.919</b>	<b>.892</b>	.811	.865	.865	.865	.865
(10.50)	.784	.811	.784	.784	.784	.784	.892	<b>.919</b>	<b>.919</b>
(2.50)	.838	<b>.946</b>	<b>.946</b>	<b>.946</b>	.838	.865	.865	<b>.919</b>	<b>.919</b>
(1.50)	.838	<b>.946</b>	<b>.946</b>	<b>.919</b>	.838	.865	.865	.892	<b>.919</b>
	0	4	5	10	15	20	30	40	50

Table 5: *lagN*

*livN* cross-validation results (baseline: 0.981)

SF-DEN									
(10.1)	.806	.695	.702	.749	.778	.805	.801	.779	.780
(4.4)	.826	.761	.760	.787	.809	.825	.857	.841	.847
(4.2)	<b>.893</b>	.870	.864	.867	.887	<b>.892</b>	<b>.918</b>	.908	.911
(4.1)	.909	.881	.882	.891	.901	.909	.923	.918	.924
(2.4)	.900	.846	.844	.859	.886	.898	.923	.923	<b>.928</b>
(2.2)	<b>.927</b>	.911	.907	.920	<b>.928</b>	.925	.941	<b>.945</b>	.943
(2.1)	.929	.915	.916	.922	.929	.929	.933	.941	.940
(1.4)	.909	.874	.867	.882	.898	.909	.925	.934	.940
(1.2)	<b>.943</b>	.909	.913	.925	.934	<b>.942</b>	.951	<b>.952</b>	<b>.952</b>
	0	4	5	10	15	20	30	40	50

Table 6: *livN*

planN cross-validation results (baseline: 0.872)

SF-DEN									
(30.2)	.826	.789	.798	.835	.844	.826	.817	.817	.817
(10.30)	.817	<b>.862</b>	<b>.862</b>	.844	.817	.817	.872	<b>.862</b>	.807
(10.20)	.844	<b>.862</b>	.844	<b>.899</b>	.872	.853	.872	<b>.881</b>	.853
(4.30)	.807	<b>.862</b>	<b>.862</b>	<b>.853</b>	.826	.807	.798	.826	.789
(2.30)	.844	<b>.862</b>	<b>.862</b>	.844	.807	.844	.780	.807	.761
(2.20)	.789	<b>.862</b>	<b>.853</b>	.835	.798	.789	<b>.862</b>	<b>.853</b>	.817
(2.10)	.789	<b>.899</b>	.881	.826	.826	.789	.807	.844	.844
(1.30)	.817	.844	.862	<b>.853</b>	.817	.826	.789	.807	.780
(1.10)	.817	.890	.890	.826	.835	.817	.817	<b>.853</b>	<b>.853</b>
	0	4	5	10	15	20	30	40	50

Table 7: planN

rotN cross-validation results (baseline: 0.804)

SF-DEN									
(50.20)	.554	.482	.446	.429	.464	.554	.786	.804	.804
(50.10)	.714	.607	.661	.625	.607	.714	.821	.821	.821
(30.20)	<b>.857</b>	<b>.857</b>	<b>.857</b>	<b>.875</b>	.839	<b>.857</b>	<b>.857</b>	<b>.857</b>	<b>.857</b>
(30.1)	.821	.804	.821	<b>.875</b>	.804	.821	.821	.821	.786
(20.30)	.839	.839	.839	.839	<b>.857</b>	.839	.821	<b>.857</b>	<b>.893</b>
(20.20)	<b>.857</b>	<b>.857</b>	<b>.857</b>	<b>.857</b>	<b>.875</b>	<b>.857</b>	<b>.875</b>	<b>.857</b>	.839
(10.30)	.839	.804	.804	.839	.839	.839	.893	<b>.875</b>	<b>.875</b>
(2.30)	<b>.857</b>	.839	.839	<b>.875</b>	<b>.857</b>	<b>.857</b>	<b>.875</b>	<b>.929</b>	<b>.875</b>
(1.30)	<b>.857</b>	.821	.821	<b>.857</b>	<b>.857</b>	<b>.857</b>	<b>.929</b>	.911	<b>.875</b>
	0	4	5	10	15	20	30	40	50

Table 8: rotN

slagN cross-validation results (baseline: 0.556)

SF-DEN									
(50.4)	.195	.195	.180	.158	.165	.195	.218	.248	<b>.233</b>
(50.2)	.218	.203	.188	.158	.165	.218	.248	<b>.293</b>	.271
(50.1)	<b>.233</b>	.218	.203	.158	.165	<b>.233</b>	.241	.248	.248
(30.10)	<b>.361</b>	.338	.316	.323	.353	.353	.331	<b>.368</b>	.346
(30.4)	<b>.361</b>	.346	<b>.361</b>	<b>.361</b>	<b>.391</b>	<b>.361</b>	<b>.368</b>	<b>.361</b>	.353
(20.30)	.211	.195	.195	.233	.211	.211	.233	.256	.256
(4.30)	.436	.436	.444	.451	.459	.436	<b>.489</b>	.481	.466
(2.30)	<b>.489</b>	.421	.444	.451	.466	<b>.496</b>	.504	<b>.511</b>	<b>.489</b>
(1.10)	.459	.474	.481	.474	.429	.459	<b>.489</b>	.504	.481
	0	4	5	10	15	20	30	40	50

Table 9: slagN

stemmeN cross-validation results (baseline: 0.922)

SF-DEN									
(30.1)	.749	.689	.677	.692	.728	.749	.838	.883	.904
(20.10)	.793	.716	.716	.701	.749	.796	.859	.904	.922
(20.4)	<b>.961</b>	.943	.943	.940	.949	<b>.961</b>	.964	<b>.970</b>	<b>.970</b>
(20.2)	.973	.970	.970	<b>.976</b>	.970	.970	.967	<b>.979</b>	.973
(20.1)	<b>.976</b>	.973	<b>.976</b>	<b>.976</b>	<b>.976</b>	<b>.976</b>	.964	.973	.967
(10.20)	.901	.841	.832	.859	.889	.901	.928	.940	.940
(10.10)	<b>.973</b>	<b>.973</b>	<b>.973</b>	.979	<b>.982</b>	<b>.973</b>	.964	.964	.964
(4.20)	.958	.970	.955	.967	.970	.958	.967	.970	<b>.973</b>
(2.20)	.961	<b>.976</b>	.961	.967	.967	.961	.964	.967	.967
	0	4	5	10	15	20	30	40	50

Table 10: stemmeN

takN cross-validation results (baseline: 0.475)

SF-DEN									
(20.4)	.518	.545	.521	.510	.502	.506	.486	.482	.479
(4.4)	.619	.615	.619	<b>.638</b>	.615	.619	<b>.599</b>	.595	.584
(4.2)	.591	<b>.638</b>	.626	.626	<b>.599</b>	<b>.599</b>	.580	.584	.591
(4.1)	.588	.584	.584	.576	.568	.588	.545	.541	.541
(2.4)	.646	.623	.619	.634	.630	<b>.642</b>	.611	.595	.588
(2.2)	.634	<b>.650</b>	.638	<b>.642</b>	.626	<b>.642</b>	.607	<b>.615</b>	<b>.615</b>
(1.10)	.533	.545	.545	.549	.537	.533	.486	.486	.494
(1.4)	.611	.611	.603	.619	.588	.607	.580	.580	.564
(1.2)	.615	<b>.646</b>	<b>.642</b>	<b>.642</b>	.615	.615	.580	<b>.584</b>	.568
	0	4	5	10	15	20	30	40	50

Table 11: takN

trykkeV cross-validation results (baseline: 0.804)

SF-DEN									
(75.1)	.739	.717	.739	.717	<b>.783</b>	.739	.717	.761	<b>.783</b>
(50.30)	.587	.543	.500	.435	.478	.587	.739	.739	.739
(50.20)	<b>.783</b>	.696	.630	.565	.717	<b>.783</b>	<b>.783</b>	<b>.848</b>	.826
(50.4)	<b>.804</b>	<b>.761</b>	.783	<b>.761</b>	.783	.783	.696	.739	.739
(50.2)	<b>.804</b>	<b>.761</b>	.783	<b>.761</b>	<b>.761</b>	<b>.804</b>	.717	.717	<b>.761</b>
(50.1)	<b>.804</b>	<b>.761</b>	<b>.804</b>	.783	.739	<b>.804</b>	.696	.696	.739
(20.50)	.391	.413	.348	.348	.413	.391	.478	<b>.543</b>	.500
(10.50)	<b>.543</b>	.500	.500	.435	<b>.565</b>	<b>.543</b>	.522	.522	<b>.587</b>
(2.75)	.304	.174	.196	.174	.217	.304	.326	.391	.500
	0	4	5	10	15	20	30	40	50

Table 12: trykkeV

utsetteV cross-validation results (baseline: 0.675)

SF-DEN									
(30.10)	.727	.688	.701	.675	.714	.727	<b>.753</b>	<b>.740</b>	<b>.766</b>
(20.20)	.714	.688	.714	<b>.740</b>	.714	.714	.714	.727	.714
(20.10)	.714	.714	<b>.753</b>	<b>.740</b>	.714	.714	.714	<b>.740</b>	.727
(20.4)	<b>.766</b>	<b>.766</b>	.753	<b>.805</b>	.753	<b>.766</b>	<b>.740</b>	<b>.740</b>	.727
(10.20)	.714	.753	<b>.766</b>	.727	.727	.714	.688	.727	<b>.740</b>
(10.10)	.753	.753	<b>.766</b>	.701	.753	.753	<b>.740</b>	.714	<b>.740</b>
(10.2)	<b>.805</b>	<b>.805</b>	<b>.805</b>	.779	.779	<b>.805</b>	.779	.792	<b>.805</b>
(10.1)	.753	.766	.766	.753	.740	.753	.701	.714	.701
(4.2)	.779	.792	.792	<b>.805</b>	.792	.779	.779	.779	.766
	0	4	5	10	15	20	30	40	50

Table 13: utsetteV

utvalgN cross-validation results (baseline: 0.609)

SF-DEN									
(75.4)	<b>.761</b>	<b>.761</b>	<b>.826</b>	.674	.696	<b>.761</b>	.848	<b>.870</b>	<b>.870</b>
(75.2)	<b>.761</b>	<b>.761</b>	<b>.826</b>	.717	.696	<b>.761</b>	.848	.848	<b>.870</b>
(20.50)	.739	<b>.761</b>	.804	.696	.739	.739	.804	<b>.870</b>	<b>.870</b>
(20.30)	.870	.913	.935	.935	.935	.870	.891	.913	.913
(20.20)	<b>.957</b>	<b>.957</b>	<b>.957</b>	<b>.957</b>	.935	<b>.957</b>	<b>.957</b>	.891	.891
(20.10)	.913	<b>.957</b>	<b>.957</b>	.935	.935	.913	.913	.913	.891
(10.75)	.587	.543	.587	.543	.565	.587	.696	.761	.761
(4.75)	<b>.609</b>	.587	<b>.609</b>	.587	<b>.609</b>	<b>.609</b>	.674	.739	.783
(1.75)	<b>.609</b>	<b>.609</b>	<b>.609</b>	.587	<b>.609</b>	<b>.609</b>	.696	.804	<b>.826</b>
	0	4	5	10	15	20	30	40	50

Table 14: utvalgN

valgN cross-validation results (baseline: 0.606)

	SF-DEN									
	0	4	5	10	15	20	30	40	50	
(30.20)	.510	.510	.519	.519	.481	.510	.510	.519	.529	
(30.10)	.692	.635	.654	.635	.663	.692	.673	.702	.731	
(30.4)	.779	<b>.817</b>	.779	.779	<b>.808</b>	.779	<b>.817</b>	.827	<b>.856</b>	
(30.2)	.808	<b>.865</b>	<b>.856</b>	<b>.846</b>	.808	.808	.817	<b>.846</b>	<b>.846</b>	
(30.1)	.817	<b>.865</b>	<b>.846</b>	.837	<b>.846</b>	.817	<b>.846</b>	<b>.846</b>	<b>.856</b>	
(20.20)	.654	.683	.663	.663	.663	.654	.721	.740	.760	
(20.10)	.846	.885	.856	.875	.875	.846	.904	.904	.913	
(10.20)	.875	.827	.846	.875	.885	.885	.904	<b>.923</b>	<b>.923</b>	
(10.10)	.875	<b>.894</b>	<b>.894</b>	<b>.894</b>	<b>.894</b>	.875	.913	.904	<b>.923</b>	

Table 15: valgN