

APPENDIX TO

Inclusive CLIL

Pre-vocational pupils' target language oral proficiency, fluency, and Willingness to Communicate

Jenny Denman¹, Erik van Schooten¹ & Rick de Graaff²

¹ Research Centre Urban Talent | ² University of Utrecht

Appendix

Attachment 1. *Sample speaking test*

Speaking Test (3 min. total) (Intro in Dutch: This is not for a school grade, and it's OK to make mistakes. Try to speak in complete sentences and talk as much as you can. All right?)

Part 1. *General conversation (ca. 2 min.)*

(Present) *What's your name? What class are you in? How old are you?*

Tell me a little bit about yourself.

Prompts:

Tell me about where you live.

Tell me about your family.

Tell me about your school.

Tell me about your hobbies.

Tell me about your friends.

What you like to do with your friends?

(Past) *What did you do last weekend?*

What did you do last summer?

What did you do yesterday evening?

(Future) *Tell me something you would like to do or try in the future.*

Part 2. *Picture task (ca. 1 minute)*

Here are pictures of two different places. I'd like you to describe the pictures and tell me which place you would like to visit, and why.

Thank you. That's the end of your speaking test.



(Repeat in Dutch if necessary.)

Attachment 2. WTC questionnaire items (translated from Dutch)

Would you dare to do these things in English? Choose from: Yes, definitely! / Yes, probably / Maybe / Probably not / Definitely not!

Willingness to Communicate at school (inside the classroom/school)

1. Ask to be excused to go to the toilet
2. Ask the teacher to repeat something
3. Read aloud from a book in class
4. Tell what you did last weekend
5. Ask questions about a film or a story
6. Lead a group discussion
7. Give a tour of your school to a new English teacher

Willingness to Communicate in London (outside the classroom/school)

1. Order a meal in a restaurant
2. Buy something in a shop
3. Ask for help when you get lost
4. Talk to strangers in the bus
5. Ask questions during a city tour
6. Talk on the phone to get information about a fun activity
7. Give a presentation at an English school

Attachment 3. Sample speaking anchor texts for holistic proficiency assessment (scores in unanimous agreement by all assessors in the assessment group)

Speaking anchor texts

Score: 0 (cohort 1, non-CLIL, 37 English words)

My name is X, een B.

I am twelve year old and I'm my hobbies *zijn*, football, badminton, tennis, *mijn familie woont in marokko*, My friends are X, X, X, X, X.

We *ging altijd* football, *zwemming*.

School is good, *is ja ik weet niet, is gewoon een goeie school*.

Attachment 3. (continued)

Speaking anchor texts

swimming, *naar familie, naar vrienden, naar de straat, naar children disco.*

Ook wat je wilt worden?

Ik wil techniek en ik wil I have weg, uitbergen.

Good footballer.

Two, avontuur *en cool sport voor men ja*, and I need some hobby *ook*.

Score: 1 (cohort 3, non-CLIL, 78 English words)

X. 3M2. I'm fourteen years old, I live in X, I play soccer, fourteen years. And that's it.

I have a mother, brother, sister, and lots of uncles and aunts and I have no pets.

I stay in Holland and go play with my friends.

As job, as job I want in the, *ja hoe zeg je dat nou weer. Hoe zeg je dat in het Engels. In de haven werken.* And transport, that will I do.

This, because this is a different culture of in Holland and this, it can be in Holland. This will be fun.

Score: 2 (cohort 2, CLIL, 149 English words)

My name is X. I'm in 2B. I have two little sisters, one big sister and one big brother. I don't have a father anymore and I have a mother, that's just all it. I really like to dance, *ja*, that was it. *Ja*, I like to be creative with, *ja*, clay and, *ja*, that stuff. *Ja*, that was it.

Wat deed ik? Toen was ik. I was at home I think. I don't know anymore. I think just sleep everywhere, by my aunt and nieces and I think that, *ja*.

I don't really know. Just, I think, *ja*. No, I don't know.

The city, because I don't really like the nature. It is really beautiful, but I don't really like it.

Just on pictures, that's all. So I would choose the city, because it's, *ja*, I have been there all my life, so I know what's in there, and I know what, how to do and I know what to expect in there.

Score: 3 (cohort 2, non-CLIL, 107 English words)

My name is X. I am in TwoE. I am thirteen years old and I live in X. I have one brother. His name is X and my mother X is forty-four years old. And my dad is X and he is seventy-four years old [*sic*]. My sport is mountain bike and snowboarding and *ja dat is het*.

I like it very much here because it's, *ik weet niet wat gezellig is*. Fun. Just talk and chill and check the email.

Last night? I made homework and I watch tv and I go on the computer. I want to become a copper later. A copper, you know?

This one because I never seen that one in real life and this, yeah, I like this more.

Score: 4 (cohort 3, CLIL, 263 English words)

I'm fourteen years old, I lived on X for three years, on an American school so that's why I chose TTO. I like to sing, my mother is singing teacher so I'm singing like since my six, since I was six.

Attachment 3. *(continued)*

Speaking anchor texts

The girl that was here *net* is my best friend and I have another friend who was in the first class with me, this year but she had to go a level lower so she's not on this school any more. I have friends in Rotterdam and friends all over, actually.

I was outside with a girlfriends, with my sister, the girlfriend of my sister, my friends and another friend of me. We ate with each other and then we went to Cookers, like a restaurant to get some ice. I think I was home like nine o'clock or something.

I like to be, I like to do something with music because I like singing. I like to go to Codarts? A school in Rotterdam. Last year I did audition and I was, I was allowed to go there so I did audition and they accept me. But my level was too low so now this year I'm going to try the same thing. I hope that this time I can go.

Definitely this one. I like forests, I like to walk, I like to do activities. Every summer I go to, I do active things, in the mountains, climbing. Because my mother, she's also doing that. I just really like nature. This would also be fun, I think but I've never done it so I don't know.

Table 2. Cronbach's alphas, Means and Standard deviations for WTC measurement instruments

Variable	Nr of items	Alphas	N	Means (se)	sd
0-M WTC London ¹	7	.88	489	26.52 (6.38)	.94
1-M WTC London ¹	7	.89	490	27.32 (.29)	6.42
2-M WTC London ¹	7	.89	351	28.20 (.33)	6.18
0-M WTC School ¹	7	.87	489	16.45 (3.60)	1.09
1-M WTC School ¹	7	.87	488	16.94 (.16)	3.62
2-M WTC School ¹	7	.88	351	17.145 (.19)	3.51

1. n.b. WTC *London* items are scored 1–5 (max. 35); WTC *School* items are scored 1,2,3 (max. 21)

Table 3. Correlations for paired assessors for speaking proficiency tests

Task and measurement	Number of tasks per pair of raters for Pearson's correlation and/or Spearman's rho	Pearson's correlation (range)	Spearman's rho (range)
Speaking task	min. 55, max. 88	.701 - .876	.707 - .898
0-measurement			
Speaking task	min. 14, max. 24	.796 - .961	.720 - .970
2-measurement			

Table 3. 0-0.20 poor; 0.21-0.40 fair; 0.41-0.60 moderate; 0.61-0.80 substantial; 0.81-1 excellent (Albers, 2017)

Table 4. Inter Rater Reliability: Speaking Proficiency o-measurement: Correlations parametric (r) and non-parametric (r_s); r = Pearson correlation (below the diagonal); r_s = Spearman's rho (above the diagonal)

		Rater 22	Rater 23	Rater 24	Rater 25	Rater 26	Rater 27
Rater 22	r/r_s	1	.745	.831	.820	.732	.715
	Sig. (2-tailed)		.000	.000	.000	.000	.000
	N	149	58	65	74	69	55
Rater 23	r/r_s	.723	1	.735	.817	.750	.707
	Sig. (2-tailed)	.000		.000	.000	.000	.000
	N	58	152	60	83	58	67
Rater 24	r/r_s	.842	.749	1	.898	.772	.864
	Sig. (2-tailed)	.000	.000		.000	.000	.000
	N	65	60	160	81	68	70
Rater 25	r/r_s	.844	.831	.876	1	.769	.855
	Sig. (2-tailed)	.000	.000	.000		.000	.000
	N	74	83	81	303	88	82
Rater 26	r/r_s	.752	.774	.783	.822	1	.745
	Sig. (2-tailed)	.000	.000	.000	.000		.000
	N	69	58	68	88	165	66
Rater 27	r/r_s	.730	.701	.847	.844	.732	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	
	N	55	67	70	82	66	157

Table 5. Inter Rater Reliability: Speaking Proficiency 2-measurement: Correlations parametric (r) and non-parametric (r_s); $r =$ Pearson correlation (below the diagonal); $r_s =$ Spearman's rho (above the diagonal)

	Rater 28	Rater 29	Rater 30	Rater 31	Rater 32	Rater 33	Rater 34	Rater 35
Rater 28	r/r_s 1	.904	.878	.915	.863	.970	.938	.962
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000
	N	91	23	17	19	18	15	19
Rater 29	r/r_s .892	1	.915	.901	.720	.877	.826	.929
	Sig. (2-tailed)	.000	.000	.000	.002	.000	.000	.000
	N	22	93	20	16	19	19	21
Rater 30	r/r_s .886	.905	1	.848	.934	.931	.899	.930
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000
	N	23	20	98	16	19	22	22
Rater 31	r/r_s .914	.897	.924	1	.966	.861	.950	.913
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000
	N	17	20	20	86	18	19	21
Rater 32	r/r_s .876	.796	.943	.963	1	.876	.899	.943
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000
	N	19	16	14	70	15	15	19
Rater 33	r/r_s .961	.883	.926	.926	.862	1	.901	.872
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000
	N	18	19	18	15	91	24	22
Rater 34	r/r_s .940	.860	.896	.951	.900	.894	1	.887

Table 5. (continued)

	Rater 28	Rater 29	Rater 30	Rater 31	Rater 32	Rater 33	Rater 34	Rater 35
Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000
N	15	19	22	19	15	24	89	19
r/ r_s	.937	.911	.935	.898	.928	.872	.896	1
Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000
N	19	21	22	21	19	22	19	97

Table 6. Results multi-level models with pretest-scores sum Speaking proficiency (models 1-4) and Speaking fluency (number of words in English, models 5-8) as dependent variables (standard errors between brackets)

Model	1	2	3	4	5	6	7	8
Fixed part	Sumspeaking_o	Sumspeaking_o	Sumspeaking_o	Sumspeaking_o	sp.o.words.english	sp.o.words.english	sp.o.words.english	sp.o.words.english
Intercept	-.332 (.059)	-.279 (.160)	-.279 (.160)	.037 (.153)	84.817 (3.670)	86.800 (10.352)	86.800 (10.352)	104.569 (10.525)
Prevoc level (gm)	-.113 ^{**} (.037)	-.169 [#] (.095)	-.169 [#] (.095)	-.246 ^{**} (.079)	-4.898 [*] (2.303)	-8.054 (6.150)	-8.054 (6.150)	-12.400 [*] (5.404)
(1 = high; 5 = low)								
grade 8 (ref.=grade 7)	.437 ^{***} (.100)	.416 (.272)	.416 (.272)	.289 (.218)	26.367 ^{***} (6.237)	26.668 (17.592)	26.668 (17.592)	19.488 (14.993)
grade 9	.734 ^{***} (.111)	.861 ^{**} (.291)	.861 ^{**} (.291)	.796 ^{**} (.231)	49.575 ^{***} (6.921)	62.175 ^{**} (18.798)	62.175 ^{**} (18.798)	58.379 ^{**} (15.884)
non-CLIL (0 = CLIL; 1 = non-CLIL)								-37.581 ^{**} (12.389)
Random part								
School variance			.000 (.000)				.000 (.000)	
Class variance		.253 (.081)	.253 (.081)	.147 (.051)		1064.332 (337.985)	1064.332 (337.985)	718.590 (239.874)
Pupil variance	.785 (.051)	.590 (.040)	.590 (.040)	.591 (.040)	3028.412 (198.611)	2330.531 (157.155)	2330.531 (157.155)	2333.870 (157.393)

Table 6. (continued)

Model	1	2	3	4	5	6	7	8
Fixed part	Sumspeaking_o	Sumspeaking_o	Sumspeaking_o	Sumspeaking_o	sp.o.words.english	sp.o.words.english	sp.o.words.english	sp.o.words.english
Total variance	.785	.843	.843	.738	3028.412	3394.863	3394.863	3052.460
Deviance	1207.116	1128.411	1128.411	1117.311	5046.957	4980.571	4980.571	4972.711
Model of reference and fit improvement	model 1 $\chi^2=78.705$ $df=1$ $p<.001$	model 2 $\chi^2=.000$ $df=1$ $p=n.s.$	model 2 $\chi^2=.000$ $df=1$ $p=n.s.$	model 2 $\chi^2=11.100$ $df=1$ $p<.001$		model 5 $\chi^2=66.386$ $df=1$ $p<.001$	model 6 $\chi^2=.000$ $df=1$ $p=n.s.$	model 6 $\chi^2=7.860$ $df=1$ $p<.01$
prop. expl. var. class level				.419				.325
prop. expl. var. Pupil level				-				-
Total				.125				.101
N	Npupils=465; Nclasses=25; Nschools=7				Npupils=465; Nclasses=25; Nschools=7			

=sig at 10% (=5% one sided) * =sig. at 5% ** sig. at 1% *** =sig. at 0.1%. (n.s.= non significant)

Table 7. Results longitudinal multi-level models for Willingness to Communicate in London, effects of CLIL. Standard errors between brackets

Model	1	2	3	4	5	6	7	8	9	10
Fixed part	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)
Intercept	26.686 (.264)	26.729 (.300)	26.729 (.300)	27.334 (.341)	27.502 (.357)	26.952 (.442)	27.137 (.454)	26.773 (.541)	26.769 (.557)	26.779 (.557)
Time (in weeks)	.020*** (.004)	.020*** (.004)	.020*** (.004)	.020*** (.004)	.014** (.005)	.021*** (.004)	.015*** (.005)	.016*** (.006)	.016* (.007)	.016* (.007)
non-CLIL (0 = cili; 1 = non-clil)				-1.385** (.464)	-1.775** (.526)	-1.472** (.480)	-1.914** (.543)	-1.187 (.741)	-1.190 (.781)	-1.192 (.781)
Grade 8 (grade 7 = refcat.)						.614 (.590)	.605 (.590)	.490 (.943)	.132 (1.037)	.225 (1.047)
Grade 9						1.020 (.623)	1.053 (.623)	3.316* (1.103)	3.225# (1.181)	3.215 (1.181)
Pre-voc level (gm (1 = high; 5 = low)						-2.33 (.216)	-2.57 (.216)	-1.61 (.356)	-0.82 (.383)	-0.099 (.384)
2-way interactions										
non-CLIL*Time					.012# (.007)		.013# (.007)	.010 (.008)	.016 (.022)	.018 (.023)
Time*pre-voc level								-0.002 (.003)	-0.003 (.005)	-0.003 (.005)
pre-voc level*grade 8								-0.668 (.882)	-1.543 (1.154)	-1.340 (1.197)

Table 7. (continued)

Model	1	2	3	4	5	6	7	8	9	10
Fixed part	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)
vmbo*grade 9								-1.708 (1.117)	-2.174 (1.251)	-2.157 (1.251)
non-CLIL*pre-voc level								.050 (.450)	.058 (.538)	.096 (.541)
Time*grade 8								.002 (.009)	-.050 (.039)	-.031 (.049)
Time*grade 9								-.006 (.011)	.038 (.055)	-.037 (.055)
non-CLIL*grade 8								-.720 (1.254)	-3.30 (1.989)	-7.35 (2.089)
non-CLIL*grade 9								-4.488 [#] (1.860)	-4.347 (2.084)	-4.319 (2.085)
3-way interaction										
pre-voc level*time*grade 8									.022 (.016)	.014 (.021)
pre-voc level*time*grade 9									.014 (.017)	.014 (.017)
pre-voc level*time*non-CLIL									-.001 (.007)	-.002 (.007)
pre-voc level*grade 8*non-CLIL									-.935 (1.860)	.400 (2.041)

Table 7. (continued)

Model	1	2	3	4	5	6	7	8	9	10
Fixed part										
pre-voc level*grade 9*non-CLIL									.000 (.000)	.000 (.000)
time *grade 8*non-CLIL									-.013 (.020)	-.028 (.068)
time *grade 9*non-CLIL									-.013 (.030)	-.014 (.030)
4-way interaction										
pre-voc level*time*grade 8*non-CLIL										.021 (.033)
pre-voc level*time*grade 9*non-CLIL										.000 (.000)
Random part										
School variance			.000 (.000)							
Class variance		.499 (.522)								
Pupil variance	24.393 (1.904)	23.908 (1.914)	23.908 (1.914)	23.919 (1.876)	23.887 (1.873)	23.550 (1.856)	23.489 (1.851)	23.182 (1.832)	23.241 (1.832)	23.247 (1.833)
Repeated measures variance	15.745 (.815)	15.748 (.815)	15.748 (.815)	15.744 (.815)	15.710 (.813)	15.767 (.816)	15.734 (.814)	15.702 (.813)	15.597 (.807)	15.588 (.807)

Table 7. (continued)

Model	1	2	3	4	5	6	7	8	9	10
Fixed part										
Deviance	8303,948	8302,952	8302,952	8295,122	8292,636	8289,521	8286,517	8278,913	8274,085	8273,680
Model of reference and fit improvement	model 1 $\chi^2 = .996$ $df=1$ $p = n.s.$	model 2 $\chi^2 = .000$ $df=1$ $p = n.s.$	model 1 $\chi^2 = 8.826$ $df=1$ $p < .01$	model 4 $\chi^2 = 2.486$ $df=1$ $p = n.s.$	model 4 $\chi^2 = 5.601$ $df=3$ $p = n.s.$	model 6 $\chi^2 = 3.004$ $df=1$ $p = n.s.$	model 6 $\chi^2 = 10.608$ $df=9$ $p = n.s.$	model 6 $\chi^2 = 4.828$ $df=7$ $p = n.s.$	model 8 $\chi^2 = 4.828$ $df=7$ $p = n.s.$	model 9 $\chi^2 = .405$ $df=2$ $p = n.s.$
prop. expl. var. Pupil level				.019						
prop. expl. var. Rep. meas. level				.000						
Nrepmeas = 1330; Npupils = 590; Nclasses = 25; Nschools = 7										

=sig at 10% (=5% one sided) * =sig. at 5% ** sig. at 1% *** =sig. at 0.1%. (n.s. = non significant)

Table 8. Results longitudinal multi-level models for Willingness to Communicate at School, effects of CLIL. Standard errors between brackets

Model	1	2	3	4	5	6	7	8	9	10
Fixed part										
Intercept	16.531 (.149)	16.604 (.223)	16.604 (.223)	17.085 (.236)	17.158 (.245)	16.838 (.264)	16.923 (.270)	16.637 (.304)	16.540 (.310)	16.540 (.310)
Time (in weeks)	.009*** (.002)	.010*** (.002)	.010*** (.002)	.010*** (.002)	.007* (.003)	.010*** (.002)	.008** (.003)	.010* (.004)	.012** (.004)	.012** (.004)
non-CLIL (0 = cili; 1 = non-cili)				-1.115** (.338)	-1.323** (.369)	-1.294*** (.289)	-1.494*** (.323)	-1.003* (.417)	-.963 (.437)	-.963* (.437)
Grade 8 (grade 7 = refcat.)						.467 (.354)	.462 (.350)	.637 (.527)	.585 (.577)	.589 (.584)
Grade 9						.602# (.375)	.614 (.371)	1.758* (.614)	1.982# (.658)	1.982 (.658)
Pre-voc level (gm) (1 = high; 5 = low)						-.312* (.130)	-.321* (.129)	-.240 (.198)	-.043 (.213)	-.043 (.437)
2-way interactions										
non-CLIL*Time					.005 (.004)		.006 (.004)	.004 (.005)	-.016 (.013)	-.015 (.014)
Time*pre-voc level								-.001 (.002)	-.006** (.003)	-.006** (.003)
pre-voc*grade 8								-.438 (.488)	-.984 (.639)	-.976 (.667)

Table 8. (continued)

Model	1	2	3	4	5	6	7	8	9	10
Fixed part	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)
vmbo*grade 9								-0.972 (.616)	-1.711 [#] (.695)	-1.710 (.696)
non-CLIL*pre-voc								.026 (.250)	-.297. (.301)	-.296 (.303)
Time*grade 8								-.005 (.006)	-.025 (.023)	-.024 (.029)
Time*grade 9								.000 (.007)	-.062 [#] (.033)	-.062 [#] (.033)
non-CLIL*grade 8								-.639 (.696)	-.154 (1.099)	-.170 (1.162)
non-CLIL*grade 9								-2.456 [*] (1.029)	-3.060 [#] (1.160)	-3.059 (1.160)
3-way interaction										
pre-voc level*time*grade 8									.006 (.010)	.006 (.012)
pre-voc level*time*grade 9									.022 [*] (.010)	.022 [*] (.010)
pre-voc level*time*non- CLIL									.007 [#] (.004)	.007 [#] (.004)

Table 8. (continued)

Model	1	2	3	4	5	6	7	8	9	10
Fixed part	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)
pre-voc level*grade 8*non-CLIL									1.128 (1.018)	1.108 (1.135)
pre-voc level*grade 9*non-CLIL									.000 (.000)	.000 (.000)
time *grade 8*non-CLIL									.012 (.012)	.011 (.040)
time *grade 9*non-CLIL									.012 (.018)	.012 (.018)
4-way interaction										
pre-voc level*time*grade 8*non-CLIL										.001 (.020)
pre-voc level*time*grade 9*non-CLIL										.000 (.000)
Random part										
School variance			.000 (.000)							
Class variance		.686 (.311)	.686 (.311)	.292 (.198)	.298 (.200)	.077 (.133)	.069 (.130)	.011 (.112)	.003 (.109)	.003 (.109)
Pupil variance	7.279 (.596)	6.610 (.570)	6.610 (.570)	6.649 (.572)	6.634 (.571)	6.577 (.567)	6.563 (.566)	6.517 (.563)	6.494 (.560)	6.494 (.560)

Table 8. (continued)

Model	1	2	3	4	5	6	7	8	9	10	
Fixed part	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	
Repeated measures variance	5.598 (.290)	5.614 (.291)	5.614 (.291)	5.609 (.290)	5.606 (.290)	5.619 (.291)	5.616 (.291)	5.606 (.290)	5.563 (.288)	5.563 (.288)	
Deviance	6842.089	6828.504	6828.504	6819.188	6817.880	6807.165	6805.375	6797.410	6788.365	6788.364	
Model of reference and fit improvement	model 1 $X^2 = 13.585$ $df = 1$ $p < .001$	model 2 $X^2 = .000$ $df = 1$ $p = n.s.$	model 2 $X^2 = 9.316$ $df = 1$ $p < .01$	model 2 $X^2 = 9.316$ $df = 1$ $p < .01$	model 4 $X^2 = 1.308$ $df = 1$ $p = n.s.$	model 4 $X^2 = 12.023$ $df = 3$ $p < .01$	model 6 $X^2 = 1.790$ $df = 1$ $p = n.s.$	model 6 $X^2 = 9.755$ $df = 9$ $p = n.s.$	model 6 $X^2 = 9.045$ $df = 7$ $p = n.s.$	model 8 $X^2 = 9.045$ $df = 7$ $p = n.s.$	model 9 $X^2 = .001$ $df = 2$ $p = n.s.$
prop. expl. var. class level					.574	.736					
prop. expl. var. Pupil level				-		.011					
prop. expl. var. Rep. meas. level					.001	-					

Nrepmeas = 1328; Npupils = 590; Nclasses = 25; Nschools = 7

=sig at 10% (=5% one sided) * =sig. at 5% ** sig. at 1% *** =sig. at 0.1%. (n.s. = non significant)

Table 9. Results multi-level models for Speaking proficiency (Ancova), effects of CLIL and other variables. Standard errors between brackets. (posttest=z)

Model	1	2	3	4	5	6	7	8
Fixed part								
Intercept	Ancova (.037)	Ancova (.058)	Ancova (.088)	Ancova (.089)	Ancova (.071)	Ancova (.101)	Ancova (.077)	Ancova (.080)
Speaking pretest (z)	.732*** (.038)	.695*** (.042)	.697*** (.040)	.679*** (.042)	.690*** (.040)	.679*** (.042)	.670*** (.042)	.637*** (.044)
non-CLIL (0 = clil; 1 = non-clil)						-.154 (.106)		-.214* (.094)
Grade 8 (grade 7 = ref.cat.)				.039 (.102)			-.020 (.106)	-.031 (.100)
Grade 9				.218 (.127)			.234# (.123)	.272* (.117)
Pre-voc level (gm) (1 = high; 5 = low)					-.086# (.048)		-.102* (.048)	-.126** (.045)
Random part								
School variance			.041 (.029)	.033 (.024)	.021 (.019)	.046 (.031)	.017 (.016)	.014 (.013)
Class variance		.054 (.024)	.012 (.014)	.009 (.013)	.014 (.015)	.009 (.013)	.009 (.013)	.004 (.011)
Pupil variance	.390 .032)	.343 (.030)	.344 (.030)	.343 (.030)	.343 (.030)	.342 (.030)	.342 (.030)	.341 (.030)
Total	.390	.397	.397	.385	.378	.397	.368	.359

Table 9. (continued)

Model	1	2	3	4	5	6	7	8
Fixed part	Ancova	Ancova	Ancova	Ancova	Ancova	Ancova	Ancova	Ancova
prop. expl School variance								.176
prop. expl Class variance								.556
prop. expl Pupil variance								.003
Prop. Expl total var	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	.024
Deviance	548.306	535.991	529.950	526.989	527.500	527.888	523.171	518.296
Model of reference and fit improvement		model 1 $X^2 = 12.315$ $df = 1$	model 2 $X^2 = 6.041$ $df = 1$	model 3 $X^2 = 2.961$ $df = 2$	model 3 $X^2 = 2.450$ $df = 1$	model 3 $X^2 = 2.062$ $df = 1$	model 3 $X^2 = 6.779$ $df = 3$	model 7 $X^2 = 4.875$ $df = 1$
N		$p < .001$	$p < .05$	$p = n.s.$	$p = n.s.$	$p = n.s.$	$p = n.s.$	$p < .05$

Npupils=289; Nclasses=25; Nschools=7

=sig at 10% (=5% one sided) * =sig. at 5% ** sig. at 1% *** =sig. at 0.1%. (n.s. = non significant)

Table 10. (continued)

Model	1	2	3	4	5	6	7	8
Fixed part		Change	Change	Change	Change	Change	Change	Change
prop. expl. total variance	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	.198	.030
Deviance	594.151	583.422	581.065	583.021	582.847	583.318	525.741	520.266
Model of reference and fit improvement		model 1 $X^2 = 10.729$ $df = 1$	model 2 $X^2 = 2.357$ $df = 1$	model 2 $X^2 = 0.41$ $df = 2$	model 2 $X^2 = 1.74$ $df = 1$	model 2 $X^2 = 1.04$ $df = 1$	model 2 $X^2 = 57.681$ $df = 3$	model 7 $X^2 = 5.475$ $df = 1$
N		$p < .001$	$p = n.s.$	$p = n.s.$	$p = n.s.$	$p = n.s.$	$p < .001$	$p < .05$
		Npupils = 289; Nclasses = 25; Nschools = 7						

=sig at 10% (=5% one sided) * =sig. at 5% ** sig. at 1% *** =sig. at 0.1%. (n.s. = non significant)

Table 13. (continued)

Model	1	2	3	4	5	6	7	8	9	10
Fixed part	Grade-1	Grade-2	Pre-voc levelrec-1	Pre-voc levelrec-2	Gender-1	Gender-2	Cito-1	Cito-2	Home language-1	Home language-2
non-CLIL*cito-gm								.004 (.014)		
Non-CLIL*gender (boy)						-.097 (.141)				
Non-CLIL*pre-voc level				.038 (.074)						
Non-CLIL*grade8		-.184 (.183)								
Non-CLIL*grade9		-.179 (.213)								
Non-CLIL*home language										-.027 (.052)
Random part										
School variance	.014 (.013)	.010 (.011)	.014 (.013)	.011 (.012)	.014 (.013)	.015 (.014)	.012 (.013)	.013 (.014)	.014 (.013)	.013 (.013)
Class variance	.004 (.011)	.004 (.011)	.004 (.011)	.005 (.012)	.003 (.011)	.002 (.011)	.000 (.000)	.000 (.000)	.004 (.012)	.005 (.012)
Pupil variance	.341 (.030)	.340 (.030)	.341 (.030)	.340 (.030)	.340 (.030)	.340 (.029)	.286 (.032)	.285 (.032)	.342 (.030)	.341 (.030)
total variance	.359	.354	.359	.356	.357	.357	.298	.298	.360	.359

Table 13. (continued)

Model	1	2	3	4	5	6	7	8	9	10
Fixed part	Grade-1	Grade-2	Pre-voc levelrec-1	Pre-voc levelrec-2	Gender-1	Gender-2	Cito-1	Cito-2	Home language-1	Home language-2
Deviance	518.296	517.062	518.296	518.068	517.037	516.576	274.267	274.173	517.464	517.191
Model of reference and fit improvement		model 1 $\chi^2 = 1.234$ $df = 2$ $p = n.s.$		model 1 $\chi^2 = .228$ $df = 1$ $p = n.s.$		model 4 $\chi^2 = .461$ $df = 1$ $p = n.s.$		model 6 $\chi^2 = .094$ $df = 1$ $p = n.s.$		model 8 $\chi^2 = .273$ $df = 1$ $p = n.s.$
N							Npupils = 289; Nclasses = 25; Nschools = 7	Npupils = 170; Nclasses = 25; Nschools = 7	Npupils = 288; Nclasses = 25; Nschools = 7	

=sig at 10% (=5% one sided) * =sig. at 5% ** sig. at 1% *** =sig. at 0.1%. (n.s. = non significant)

Table 14. Results multi-level models for Speaking proficiency (Ancova), variables moderating the effect of CLIL. Standard errors between brackets. (PE=Primary school English)

Model	1	2	3	4	5	6	7
Fixed part							
Intercept	.055 (.082)	.055 (.083)	.061 (.080)	.053 (.082)	.071 (.081)	.065 (.083)	.055 (.084)
Sumspeaking_0	.634*** (.044)	.633*** (.044)	.641*** (.045)	.638*** (.044)	.639 (.044)	.636*** (.044)	.640*** (.044)
non-CLIL (0 = cili; 1 = non-clil)	-.214* (.096)	-.215* (.097)	-.217* (.094)	-.238* (.095)	-.231* (.093)	-.253* (.094)	-.230* (.096)
Grade 8 (grade 7 = ref.cat.)	-.021 (.102)	-.026 (.103)	-.036 (.099)	-.031 (.097)	.001 (.099)	.002 (.096)	.002 (.096)
Grade 9	.285* (.120)	.288* (.121)	.271* (.117)	.289* (.116)	.321* (.118)	.343** (.117)	.337** (.116)
Pre-voc level (gm) (1 = high; 5 = low)	-.125* (.046)	-.126* (.047)	-.123* (.046)	-.115* (.047)	-.111* (.046)	-.103* (.047)	-.099# (.047)
Years PE (gm)	.025 (.042)	.046 (.051)			.047 (.046)	.064 (.058)	.062 (.058)
Freq. PE (gm)			-.027 (.040)		-.036 (.044)	-.002 (.054)	-.006 (.054)
2-way interactions							
Years PE*Freq. PE					-.067# (.034)	-.074# (.035)	-.052 (.040)
Non-CLIL*Years PE		-.059 (.086)				-.056 (.092)	-.066 (.093)

Table 14. (continued)

Model	1	2	3	4	5	6	7
Fixed part	Years PE	Years PE	Freq. PE	Freq. PE	Years & freq: PE intensity	Years & freq: PE intensity	Years & freq: PE intensity
Non-CLJL*Freq. PE				-1.22 (.084)		-1.14 (.089)	-1.03 (.089)
3-way interaction							
Non-CLJL*Years PE*Freq. PE							-0.83 (.077)
Random part							
School variance	.014 (.014)	.015 (.014)	.014 (.013)	.018 (.015)	.014 (.013)	.019 (.015)	.019 (.015)
Class variance	.005 (.012)	.006 (.012)	.004 (.011)	.002 (.011)	.003 (.011)	.001 (.010)	.001 (.010)
Pupil variance	.339 (.029)	.338 (.029)	.342 (.030)	.339 (.029)	.337 (.029)	.333 (.029)	.332 (.029)
Total variance	.358						
Deviance	517.932	517.465	517.025	515.015	512.369	509.631	508.465
Model of reference and fit improvement		model 1 $X^2 = .467$ $df = 1$ $p = n.s.$		model 3 $X^2 = 2.010$ $df = 1$ $p = n.s.$		model 5 $X^2 = 2.738$ $df = 2$ $p = n.s.$	model 6 $X^2 = 1.166$ $df = 1$ $p = n.s.$
N	Npupils = 289; Nclasses = 25; Nschools = 7		Npupils = 288; Nclasses = 25; Nschools = 7				

=sig at 10% (=5% one sided) * =sig. at 5% ** sig. at 1% *** =sig. at 0.1%. (n.s. = non significant)

Table 15. Results multi-level models for Speaking proficiency (change scores), variables moderating the effect of CLIL. Standard errors between brackets. (Cito = scholastic aptitude test score)

Model	1	2	3	4	5	6	7	8	9
Fixed part	Grade-1/ Pre-voc- levelrec-1	Grade-2	Pre-voc- levelrec-2	Gender-1	Gender-2	Cito-1	Cito-2	Home language-1	Home language-2
Intercept	.098 (.100)	.075 (.111)	.094 (.100)	.079 (.103)	-.056 (.108)	-.049 (.093)	-.056 (.093)	.098 (.100)	.096 (.101)
non-CLIL (0 = clii; 1 = non-clil)	-.001 (.121)	.051 (.164)	-.008 (.121)	-.007 (.120)	.045 (.141)	.104 (.109)	.106 (.108)	-.002 (.121)	.002 (.121)
Grade 8 (grade 7 = ref.cat.)	-.164 (.149)	-.117 (.188)	-.161 (.149)	-.173 (.148)	-.168 (.148)	-.190 (.137)	-.185 (.137)	-.167 (.149)	-.164 (.150)
Grade 9	-.072 (.155)	-.027 (.205)	-.076 (.155)	-.077 (.154)	-.072 (.154)	.041 (.143)	.029 (.143)	-.069 (.155)	-.059 (.157)
Pre-voc level (gm (1 = high; 5 = low)	-.061 (.052)	-.060 (.053)	-.044 (.065)	-.066 (.052)	-.064 (.052)	-.156* (.058)	-.151* (.058)	-.060 (.053)	-.060 (.053)
Gender (boy = 1; girl = 0)			.053 (.079)	.102 (.106)					
Cito-gm									
Home language (gm)						-.011 (.009)	-.015 (.010)	.007 (.029)	.022 (.038)
2-way interactions									
non-CLIL*cito-gm							.015 (.015)		
Non-CLIL*gender (boy)					-.112 (.158)				

Table 15. (continued)

Model	1	2	3	4	5	6	7	8	9
Fixed part	Grade-1/ levelrec-1	Pre-voc- Grade-2	Pre-voc- levelrec-2	Gender-1	Gender-2	Cito-1	Cito-2	Home language-1	Home language-2
Non-CLIL*pre-voc level			-.042 (.095)						
Non-CLIL*grades		-.112 (.269)							
Non-CLIL*grades		-.101 (.307)							
Non-CLIL*home language									-.038 (.059)
Random part									
Class variance	.044 (.023)	.044 (.023)	.043 (.023)	.043 (.023)	.042 (.022)	.016 (.019)	.015 (.018)	.044 (.023)	.045 (.023)
Pupil variance	.409 (.036)	.409 (.036)	.409 (.036)	.409 (.036)	.408 (.036)	.334 (.039)	.332 (.039)	.411 (.036)	.410 (.036)
Deviance	581.631	581.413	581.439	581.191	580.691	302.616	301.599	580.588	580.184
Model of reference and fit improvement	model 1 $X^2 = .218$ $df = 2$	model 1 $X^2 = .192$ $df = 1$	model 1 $X^2 = .192$ $df = 1$	model 4 $X^2 = .500$ $df = 1$	model 4 $X^2 = .500$ $df = 1$	model 6 $X^2 = 1.017$ $df = 1$	model 6 $X^2 = 1.017$ $df = 1$	model 8 $X^2 = .404$ $df = 1$	model 8 $X^2 = .404$ $df = 1$
N	Npupils = 289; Nclasses = 25; Nschools = 7	Npupils = 289; Nclasses = 25; Nschools = 7	Npupils = 289; Nclasses = 25; Nschools = 7	Npupils = 289; Nclasses = 25; Nschools = 7	Npupils = 289; Nclasses = 25; Nschools = 7	Npupils = 170; Nclasses = 25; Nschools = 7	Npupils = 170; Nclasses = 25; Nschools = 7	Npupils = 288; Nclasses = 25; Nschools = 7	Npupils = 288; Nclasses = 25; Nschools = 7

=sig at 10% (=5% one sided) * =sig. at 5% ** sig. at 1% *** =sig. at 0.1%. (n.s. = non significant)

Table 16. Results multi-level models for Speaking proficiency (change scores), variables moderating the effect of CLIL . Standard errors between brackets. (PE=Primary school English)

Model	1	2	3	4	5	6	7
Fixed part							
Intercept	.102 (.100)	.102 (.101)	.104 (.098)	.100 (.099)	.116 (.097)	.116 (.100)	.097 (.099)
non-CLIL (0 = clil; 1 = non-clil)	-.003 (.120)	-.004 (.121)	-.021 (.119)	-.030 (.121)	-.032 (.117)	-.043 (.120)	-.003 (.121)
Grade 8 (grade 7 = ref.cat.)	-.171 (.149)	-.176 (.151)	-.178 (.146)	-.177 (.148)	-.152 (.145)	-.156 (.149)	-.152 (.147)
Grade 9	-.078 (.155)	-.074 (.157)	-.063 (.152)	-.054 (.154)	-.029 (.152)	-.012 (.156)	-.017 (.154)
Pre-voc level (gm) (1 = high; 5 = low)	-.062 (.052)	-.063 (.053)	-.056 (.051)	-.054 (.052)	-.047 (.051)	-.045 (.052)	-.041 (.052)
Years PE (gm)	-.017 (.046)	.002 (.056)			.020 (.050)	.049 (.064)	.045 (.064)
Freq. PE (gm)			-.073 [#] (.044)	-.053 (.052)	-.073 (.048)	-.062 (.060)	-.068 (.060)
2-way interactions							
Years PE* Freq. PE					-.060 (.038)	-.065 (.038)	-.029 (.045)
Non-CLIL*Years PE		-.052 (.095)				-.083 (.103)	-.097 (.103)
Non-CLIL*Freq. PE				-.067 (.094)		-.047 (.101)	-.031 (.101)
3-way interaction							

Table 16. (continued)

Model	1	2	3	4	5	6	7
Fixed part	Years PE	Years PE	Frequency PE	Frequency PE	Years & freq: PE intensity	Years & freq: PE intensity	Years & freq: PE intensity
Years PE* Freq. PE*non-CLIL							-1.137 (.086)
Random part							
Class variance	.043 (.023)	.045 (.023)	.041 (.022)	.043 (.023)	.038 (.021)	.042 (.022)	.041 (.022)
Pupil variance	.409 (.036)	.408 (.035)	.408 (.036)	.406 (.035)	.405 (.035)	.401 (.035)	.398 (.035)
Total variance	.452	.453	.449	.449	.443	.443	.439
Deviance	581.501	581.207	577.772	577.276	575.198	573.946	571.388
Model of reference and fit improvement		model 1 $\chi^2 = .294$ $df = 1$ $p = n.s.$		model 3 $\chi^2 = .496$ $df = 1$ $p = n.s.$		model 5 $\chi^2 = 1.252$ $df = 2$ $p = n.s.$	model 6 $\chi^2 = 2.558$ $df = 1$ $p = n.s.$
N	Npupils = 289; Nclasses = 25; Nschools = 7		Npupils = 288; Nclasses = 25; Nschools = 7				

=sig at 10% (=5% one sided) * =sig. at 5% ** sig. at 1% *** =sig. at 0.1%. (n.s. = non significant)

Table 17. Results multi-level models for Speaking fluency (Ancova), main effects of CLIL and other variables. Standard errors between brackets

Model	1	2	3	4	5	6	7	8
Fixed part	Ancova	Ancova	Ancova	Ancova	Ancova	Ancova	Ancova	Ancova
Intercept	142.284 (2.885)	142.055 (6.082)	137.331 (9.136)	141.164 (9.982)	138.720 (8.471)	144.095 (9.627)	144.490 (8.987)	152.903 (8.756)
Words English pretest-gm	.933 ^{***} (.046)	.910 ^{***} (.050)	.890 ^{***} (.046)	.898 ^{***} (.047)	.890 ^{***} (.046)	.867 ^{***} (.047)	.897 ^{***} (.047)	.864 ^{***} (.049)
non-CLIL (0 = clil; 1 = non-clil)						-16.989 [#] (9.627)		-20.628 ^{***} (8.550)
Grade 8 (grade 7 = ref.cat.)				-13.141 (9.049)			-17.594 (9.585)	-18.828 [*] (8.595)
Grade 9				-4.349 (11.359)			-3.160 (11.309)	.088 (10.262)
Pre-voc level (gm) (1 = high; 5 = low)					-4.750 (5.014)		-8.087 (5.104)	-9.674 [*] (4.610)
Random part								
School variance			470.957 (313.277)	513.610 (328.322)	376.358 (263.675)	454.273 (294.699)	351.361 (241.657)	279.436 (192.745)
Class variance		759.809 (260.881)	203.532 (121.498)	157.429 (105.871)	211.729 (123.851)	147.052 (102.354)	160.591 (106.621)	97.369 (85.283)
Pupil variance	2404.910 (200.062)	1742.464 (151.658)	1741.487 (151.496)	1742.272 (151.545)	1741.682 (151.550)	1744.046 (151.694)	1741.212 (151.476)	1743.424 (151.583)
Total variance	2404.910	2502.273	2415.976	2413.311	2329.769	2345.371	2253.164	2120.229

Table 17. (continued)

Model	1	2	3	4	5	6	7	8	
Fixed part	Ancova	Ancova	Ancova	Ancova	Ancova	Ancova	Ancova	Ancova	
Deviance	3 070.089	3 021.063	3 009.567	3 007.588	3 008.796	3 006.516	3 005.459	3 000.264	
Model of reference and fit improvement	model 1 $\chi^2 = 49.026$ $df = 1$ $p < .001$	model 2 $\chi^2 = 11.496$ $df = 1$ $p < .001$	model 3 $\chi^2 = 1.979$ $df = 2$ $p = n.s.$	model 3 $\chi^2 = 1.979$ $df = 2$ $p = n.s.$	model 3 $\chi^2 = .771$ $df = 1$ $p = n.s.$	model 3 $\chi^2 = 3.051$ $df = 1$ $p < .10$	model 3 $\chi^2 = 3.051$ $df = 1$ $p < .10$	model 7 $\chi^2 = 5.195$ $df = 1$ $p < .05$	model 7 $\chi^2 = 5.195$ $df = 1$ $p < .05$
prop. expl. var. school level						.035		.204	
prop. expl. var. class level						.277		.394	
prop. expl. var. Pupil level									
prop. expl. var. total									
N									

Npupils = 289 ; Nclasses = 25; Nschools = 7

= sig at 10% (=5% one sided) * = sig. at 5% ** sig. at 1% *** = sig. at 0.1%. (n.s. = non significant)

Table 18. Results multi-level models for Speaking fluency (change scores), main effects of CLIL and other variables. Standard errors between brackets

Model	1	2	3	4	5	6	7	8
Fixed part	Change	Change	Change	Change	Change	Change	Change	Change
Intercept	39.083 (2.895)	38.395 (6.023)	34.546 (8.802)	40.233 (9.708)	35.759 (8.298)	38.733 (9.440)	43.530 (8.819)	49.982 (9.071)
non-CLIL (0 = cil; 1 = non-cil)						-10.245 (9.865)		-13.969 (8.879)
Grade 8 (grade 7 = refcat.)				-16.290 (9.094)			-20.857* (9.673)	-22.684* (9.408)
Grade 9				-11.114 (11.099)			-10.200 (11.022)	-9.618 (10.603)
Pre-voc level (gm) (1 = high; 5 = low)					-4.467 (5.076)		-7.897 (5.107)	-9.240# (4.845)
Random part								
School variance			417.087 (291.347)	475.851 (310.479)	345.574 (253.849)	391.820 (274.745)	329.171 (232.004)	263.569 (193.449)
Class variance		740.834 (256.270)	241.459 (134.989)	169.007 (110.559)	247.430 (136.367)	224.022 (129.052)	171.961 (111.072)	150.341 (103.688)
Pupil variance	2422.810 (201.551)	1767.300 (153.811)	1768.819 (153.893)	1769.857 (153.956)	1768.863 (153.945)	1769.778 (153.977)	1768.761 (153.898)	1769.842 (153.975)
Total variance	2422.810	2508.134	2427.365	2414.715	2361.867	2385.620	2269.893	2183.752
Deviance	3072.232	3024.345	3015.056	3012.129	3014.363	3014.005	3010.083	3007.769

Table 18. (continued)

Model	1	2	3	4	5	6	7	8
Fixed part	Change	Change	Change	Change	Change	Change	Change	Change
Model of reference and fit improvement	model 1 $X^2 = 47.887$ $df = 1$ $p < .001$	model 2 $X^2 = 9.289$ $df = 1$ $p < .01$	model 3 $X^2 = 2.927$ $df = 2$ $p = n.s.$	model 3 $X^2 = .693$ $df = 1$ $p = n.s.$	model 3 $X^2 = 1.051$ $df = 1$ $p = n.s.$	model 7 $X^2 = 2.314$ $df = 1$ $p = n.s.$		
N	Npupils = 289; Nclasses = 25; Nschools = 7							

= sig at 10% (=5% one sided) * = sig. at 5% ** sig. at 1% *** =sig. at 0.1%. (n.s. = non significant)

Table 19. (continued)

Model	1	2	3	4	5	6	7	8	9	10
Fixed part	Grade-1	Grade-2	Pre-voc levelrec-1	Pre-voc levelrec-2	Gender-1	Gender-2	Cito-1	Cito-2	Home language-1	Home language-2
non-CLIL*cito-gm								2.155 [#] (1.188)		
Non-CLIL*gender (boy)					2.288 (10.315)					
Non-CLIL*pre-voc level				-485 (7.601)						
Non-CLIL*grade8		5.443 (15.579)								
Non-CLIL*grade9		24.441 (18.400)								
Non-CLIL*home language										-779 (3.876)
Random part										
School variance	279.436 (192.745)	303.357 (201.160)	279.436 (192.745)	280.270 (193.142)	275.823 (189.773)	274.725 (189.248)	205.741 (166.636)	230.182 (179.550)	283.728 (195.553)	286.419 (196.963)
Class variance	97.369 (85.283)	73.740 (76.995)	97.369 (85.283)	97.078 (85.182)	91.424 (83.267)	91.793 (83.383)	49.969 (111.457)	55.284 (111.303)	99.988 (86.408)	99.871 (86.355)
Pupil variance	1743.424 (151.583)	1741.803 (151.433)	1743.424 (151.583)	1743.436 (151.584)	1744.754 (151.693)	1744.401 (151.662)	1839.522 (214.340)	1793.814 (209.091)	1744.748 (151.992)	1744.222 (151.946)
Total variance	2120.229	2118.900	2120.229	2120.784	2111.648	2110.919	2095.232	2079.280	2128.464	2130.512

Table 19. (continued)

Model	1	2	3	4	5	6	7	8	9	10
Fixed part	Grade-1	Grade-2	Pre-voc levelrec-1	Pre-voc levelrec-2	Gender-1	Gender-2	Cito-1	Cito-2	Home language-1	Home language-2
prop. expl. var. school level							-			
prop. expl. var. class level							-			
prop. expl. var. Pupil level						.025				
prop. expl. var. total							.008			
Deviance	3000.264	2998.581	3000.264	3000.259	2999.945	2999.896	1772.543	1769.318	2990.416	2990.376
Model of reference and fit improvement	model 1 $\chi^2 = 1.683$ $df = 2$ $p = n.s.$	model 1 $\chi^2 = .005$ $df = 1$ $p = n.s.$	model 1 $\chi^2 = .005$ $df = 1$ $p = n.s.$	model 1 $\chi^2 = .005$ $df = 1$ $p = n.s.$	model 1 $\chi^2 = .005$ $df = 1$ $p = n.s.$	model 4 $\chi^2 = .049$ $df = 1$ $p = n.s.$	model 4 $\chi^2 = .049$ $df = 1$ $p = n.s.$	model 6 $\chi^2 = 3.225$ $df = 1$ $p < .10$	model 6 $\chi^2 = 3.225$ $df = 1$ $p < .10$	model 8 $\chi^2 = .040$ $df = 1$ $p = n.s.$
N	Npupils = 289; Nschoools = 7	Npupils = 289; Nschoools = 7	Npupils = 289; Nschoools = 7	Npupils = 289; Nschoools = 7	Npupils = 289; Nschoools = 7	Npupils = 289; Nschoools = 7	Npupils = 170; Nclasses = 25; Nschoools = 7	Npupils = 170; Nclasses = 25; Nschoools = 7	Npupils = 288; Nclasses = 25; Nschoools = 7	Npupils = 288; Nclasses = 25; Nschoools = 7

=sig at 10% (=5% one sided) * =sig. at 5% ** sig. at 1% *** =sig. at 0.1%. (n.s. = non significant)

Table 20. Results multi-level models for Speakingfluency (number of words in English): Ancova, variables moderating the effect of CLLIL. Standard errors between brackets. (PE=Primary school English)

Model	1	2	3	4	5	6	7	8
Fixed part	Years PE	Years PE	Freq. PE	Freq. PE	Years & Freq. PE intensity	Years & Freq. PE intensity	Years & Freq. PE intensity	Years & Freq. PE intensity
Intercept	152.922 (8.784)	152.985 (8.759)	152.796 (8.617)	153.387 (8.476)	153.358 (8.613)	155.495 (8.361)	155.873 (8.258)	155.482 (8.289)
Sumwords English_0	.864*** (.049)	.865*** (.049)	.857*** (.049)	.863*** (.049)	.860*** (.049)	.859*** (.048)	.864*** (.048)	.866*** (.048)
non-CLLIL (0 = clli; 1 = non-clil)	-20.630* (8.550)	-20.617* (8.556)	-19.818* (8.618)	-18.387* (8.686)	-19.682* (8.606)	-21.150* (8.406)	-19.950* (8.490)	-19.040* (8.635)
Grade 8 (grade 7 = refcat.)	-18.863* (8.685)	-18.706* (8.712)	-18.165# (8.727)	-18.666* (8.773)	-19.108* (8.796)	-17.074# (8.625)	-17.462# (8.681)	-17.435# (8.721)
Grade 9	.047 (10.366)	-.162 (10.389)	-.105 (10.357)	-1.723 (10.448)	-1.444 (10.479)	1.583 (10.295)	.302 (10.395)	.095 (10.434)
Pre-voc level (gm (1 = high; 5 = low)	-9.679# (4.614)	-9.707# (4.608)	-10.417* (4.593)	-10.876* (4.548)	-10.723* (4.594)	-10.342* (4.460)	-10.688* (4.429)	-10.580* (4.441)
Years PE (gm)	-.085 (3.002)	-.980 (3.698)			-2.631 (3.290)	-2.419 (3.254)	-2.094 (4.169)	-2.220 (4.170)
Freq. PE (gm)			4.765# (2.895)	2.335 (3.458)	5.822# (3.179)	6.816* (3.162)	4.660 (3.918)	4.513 (3.922)
2-way interactions								
Years PE*freq PE						-6.679* (2.464)	-6.510* (2.482)	-5.548# (2.889)

Table 20. (continued)

Model	1	2	3	4	5	6	7	8
Fixed part	Years PE	Years PE	Freq. PE	Freq. PE	Years & Freq. PE intensity	Years & Freq. PE intensity	Years & Freq. PE intensity	Years & Freq. PE intensity
Non-CLIL*Years PE	2,584 (6.229)						-1.90 (6.646)	-5.66 (6.666)
Non-CLIL*Freq. PE			7.963 (6.176)				6.600 (6.511)	7.036 (6.540)
3-way interactions								
Non-CLIL*Years PE*freq PE								-3.592 (5.548)
Random part								
School variance	279.502 (192.813)	275.677 (190.895)	256.508 (181.712)	237.339 (171.729)	253.443 (179.790)	232.404 (166.488)	218.325 (159.616)	216.608 (159.098)
Class variance	97.336 (85.298)	98.375 (85.592)	105.197 (87.764)	108.657 (88.506)	105.023 (87.418)	96.653 (83.373)	99.871 (84.456)	102.492 (85.241)
Pupil variance	1743.427 (151.579)	1742.318 (151.488)	1731.149 (150.810)	1721.753 (150.014)	1727.614 (150.526)	1689.445 (147.197)	1682.980 (146.609)	1679.427 (146.303)
Total variance	2120.265	2116.370	2092.854	2067.749	2086.080	2018.502	2001.176	1998.527
Prop expl school var						.083		
Prop expl class var						.080		
Prop expl pupil var						.022		
Prop expl total var						.032		
Deviance	3000.263	3000.091	2988.111	2986.465	2987.473	2980.230	2979.102	2978.684

Table 20. (continued)

Model	1	2	3	4	5	6	7	8
Fixed part	Years PE	Years PE	Freq. PE	Freq. PE	Years & Freq: PE intensity	Years & Freq: PE intensity	Years & Freq: PE intensity	Years & Freq: PE intensity
Model of reference and fit improvement	model 1 $\chi^2 = .172$ $df = 1$ $p = n.s.$	model 2 $\chi^2 = 1.646$ $df = 1$ $p = n.s.$	model 3 $\chi^2 = 1.646$ $df = 1$ $p = n.s.$	model 4 $\chi^2 = 7.243$ $df = 1$ $p < .01$	model 5 $\chi^2 = 1.128$ $df = 2$ $p = n.s.$	model 6 $\chi^2 = .418$ $df = 1$ $p = n.s.$	model 7 $\chi^2 = .418$ $df = 1$ $p = n.s.$	model 8 $\chi^2 = .418$ $df = 1$ $p = n.s.$
N	Npupils = 289; Nclasses = 25; Nschools = 7							

=sig at 10% (=5% one sided) * =sig. at 5% ** sig. at 1% *** =sig. at 0.1%. (n.s. = non significant)

Table 21. Results multi-level models for Speaking fluency (number of words in English): Change scores, variables moderating the effect of CLIL. Standard errors between brackets. (Cito = scholastic aptitude test score)

Model	1	2	3	4	5	6	7	8	9
Fixed part	Grade-1/ Pre-voc- levelrec-1	Grade-2	Pre-voc- levelrec-2	Gender-1	Gender-2	Cito-1	Cito-2	Home language-1	Home language-2
Intercept	49.982 (9.071)	52.857 (9.716)	50.258 (9.090)	49.685 (9.219)	49.985 (9.456)	45.180 (9.320)	45.266 (9.337)	49.878 (9.096)	49.879 (9.104)
non-CLIL (0 = cllil; 1 = non-clil)	-13.969 (8.879)	-21.757 [#] (11.293)	-15.226 (10.115)	-14.035 (8.866)	-14.690 (10.049)	-5.185 (9.908)	-7.893 (9.976)	-14.078 (8.933)	-14.078 (8.984)
Grade 8 (grade 7 = refcat.)	-22.684* (9.408)	-24.943* (11.316)	-22.613* (9.416)	-22.833* (9.433)	-22.908* (9.452)	-32.252** (11.146)	-31.708* (11.047)	-22.336* (9.499)	-22.337* (9.506)
Grade 9	-9.318 (10.603)	-26.320 [#] (13.305)	-9.566 (10.600)	-9.741 (10.603)	-9.810 (10.620)	-1.856 (12.419)	-2.750 (12.345)	-9.452 (10.669)	-9.452 (10.708)
Pre-voc level (gm) (1 = high; 5 = low)	-9.240 [#] (4.845)	-8.408 [#] (4.737)	-8.630 (5.458)	-9.326 [#] (4.868)	-9.338 [#] (4.868)	-13.984* (5.682)	-14.178* (5.679)	-9.250 [#] (4.869)	-9.2516 [#] (4.871)
Gender (boy = 1; girl = 0)				.848 (5.234)	.219 (6.959)				
Cito (gm)						-0.065 (.678)	-0.880 (.795)		
Home language (gm)								-0.433 (1.951)	-0.433 (2.512)
2-way interactions									
Non-CLIL*grade8		7.018 (16.288)							

Table 21. (continued)

Model	1	2	3	4	5	6	7	8	9
Fixed part	Grade-1/ Pre-voc-levelrec-1	Grade-2	Pre-voc-levelrec-2	Gender-1	Gender-2	Cito-1	Cito-2	Home language-1	Home language-2
prop. expl. var. class level			.011						
prop. expl. var. Pupil level			.023						
prop. expl. var. total			.016						
Deviance	3 007.769	3 004.694	3 007.705	3 007.743	3 007.724	1 777.753	1 774.138	2 998.116	2 998.116
Model of reference and fit improvement		model 1 $X^2 = 3.075$ $df = 2$ $p = n.s.$	model 1 $X^2 = .064$ $df = 1$ $p = n.s.$		model 4 $X^2 = .019$ $df = 1$ $p = n.s.$		model 6 $X^2 = 3.615$ $df = 1$ $p < .10$	2 998.116	model 8 $X^2 = .000$ $df = 1$ $p = n.s.$
	Npupils = 289; Nclasses = 25; Nschools = 7			Npupils = 170; Nclasses = 25; Nschools = 7				Npupils = 288; Nclasses = 25; Nschools = 7	

=sig at 10% (=5% one sided) * =sig. at 5% ** sig. at 1% *** =sig. at 0.1%. (n.s. = non significant)

Table 22. (continued)

Model	1	2	3	4	5	6	7	8
Fixed part	Years PE	Years PE	Freq. PE	Freq. PE	Years & freq: PE intensity	Years & freq: PE intensity	Years & freq: PE intensity	Years & freq: PE intensity
Non-CLIL*Freq. PE				9.523 (6.236)			7.917 (6.586)	8.420 (6.610)
3-way interaction								
Non-CLIL*Years PE*freq. PE								-4.456 (5.605)
Random part								
School variance	264.351 (193.767)	259.282 (191.143)	243.477 (184.374)	225.058 (173.571)	239.523 (182.281)	219.854 (169.497)	205.814 (161.272)	202.561 (160.061)
Class variance	149.803 (103.491)	150.360 (103.649)	159.621 (106.841)	155.031 (104.982)	160.233 (106.864)	150.677 (102.387)	147.046 (100.977)	150.605 (102.015)
Pupil variance	1769.251 (153.923)	1767.565 (153.773)	1761.844 (153.562)	1751.335 (152.639)	1755.380 (152.997)	1717.774 (149.725)	1710.889 (149.116)	1705.964 (148.690)
Total variance	2183.405	2177.207	2164.942	2131.424	2155.136	2088.305	2063.749	2059.130
prop. expl. var. school level						.082		
prop. expl. var. class level						.060		
prop. expl. var. Pupil level						.021		
prop. expl. var. total						.031		
Deviance	3007.658	3007.333	2996.355	2994.043	2995.308	2988.361	2986.704	2986.074

Table 22. (continued)

Model	1	2	3	4	5	6	7	8
Fixed part	Years PE	Years PE	Freq. PE	Freq. PE	Years & freq: PE intensity	Years & freq: PE intensity	Years & freq: PE intensity	Years & freq: PE intensity
Model of reference and fit improvement	model 1 $X^2 = .325$ $df = 1$ $p = n.s.$	model 2 $X^2 = 2.312$ $df = 1$ $p = n.s.$	model 3 $X^2 = 2.312$ $df = 1$ $p = n.s.$	model 4 $X^2 = 2.312$ $df = 1$ $p = n.s.$	model 5 $X^2 = 6.947$ $df = 1$ $p < .01$	model 6 $X^2 = 1.657$ $df = 1$ $p = n.s.$	model 7 $X^2 = .630$ $df = 1$ $p = n.s.$	model 8 $X^2 = .630$ $df = 1$ $p = n.s.$
N	Npupils = 289; Nclasses = 25; Nschools = 7							

=sig at 10% (=5% one sided) * =sig. at 5% ** sig. at 1% *** =sig. at 0.1%. (n.s. = non significant)

Address for correspondence

Jenny Denman
Research Centre Urban Talent/Kenniscentrum Talentontwikkeling
Rotterdam University of Applied Sciences/Hogeschool Rotterdam
Postbus 25035
2001 HA Rotterdam
The Netherlands
denmanjl@gmail.com
j.l.Denman@hr.nl
 <https://orcid.org/0000-0001-8624-0125>

Co-author information

Erik van Schooten
Research Centre Urban Talent/
Kenniscentrum Talentontwikkeling
Rotterdam University of Applied Sciences/
Hogeschool Rotterdam
The Netherlands
E.J.van.Schooten@hr.nl

Rick de Graaff
Department of Languages, Literature and
Communication
University of Utrecht
The Netherlands
R.degraaff@uu.nl