PNAS
www.pnas.org

Supplementary Information for

Nouns slow down speech across structurally and culturally diverse languages

Frank Seifart, Jan Strunk, Swintha Danielsen, Iren Hartmann, Brigitte Pakendorf, Søren Wichmann, Alena Witzlack-Makarevich, Nivja H. de Jong, and Balthasar Bickel

Corresponding author: Frank Seifart
Email: f.c.seifart@uval.nl

This PDF file includes:

www.pnas.org/cgi/doi/10.1073/pnas.1800708115

Supplementary Text

Additional results articulation rate study. Here we provide discussion of additional results from our study of articulation rate, also taking into account interactions of the effect of word class with position and alternative analyses in which auxiliaries are included as verbs. In our main study, word class and its interaction with position are significant predictors (Tables S8 and S9). Articulation rate is predicted to be higher before verbs than before nouns, but the difference is smaller towards the end of an utterance than at the beginning: While contexts before verbs are predicted to be about half a segment per second (0.5113 segments per second) faster than contexts before nouns at the beginning of utterances, contexts before verbs are only 0.1375 segments per second faster than contexts before nouns at the end of utterances. Crucially, the interaction of word class and position only reduces but never reverses the articulation rate advantage for verbal vs. nominal contexts. The strength of the word class effect, however, depends on the particular language, as indicated by the significant interaction of word class and language. Note that the coefficients of the word class:language interaction is positive for most languages (Table S9), which indicates that the word class effect is stronger in these languages than in Bora (which is used as the reference for the language factor here). In fact, in a model without language and the word class:language interaction as fixed effects, contexts before verbs are predicted to be 0.8556 segments per second faster than contexts before nouns at the beginning of an utterance and 0.4730 segments per second faster at the end of an utterance. In addition to the effect of word class, there is a significant effect of word length, in that articulation is slightly faster before longer words. This effect is even more pronounced towards the end of an utterance (significant interaction of word length and position). Languages, of course, also vary greatly with regard to the predicted average articulation rate (highly significant effect of language). However, both standardized coefficients and likelihood ratio tests indicate that the effect of word class on articulation rate is stronger than the effects of the other predictor variables.

Linear mixed-effects analyses of the nine individual languages show that in seven of these languages—Baure, Chintang, Dutch, English, Even, Hoocąk, Texistepec—there is a significant effect of word class on articulation rate (Tables S10-18, cf. also Fig. 3A in the main text), while the effect of word class on articulation rate is not significant in Bora and N‖ng. In Baure, Chintang, Even, and Hoocąk, there is a significant positive effect of word class, indicating higher articulation rates before verbs than before nouns and no significant interaction of word class and position. The interaction of word class and position is, however, required for a good model fit in Dutch, English, and Texistepec. In Texistepec, the interaction indicates a reduction (not a reversion) of the articulation rate advantage of contexts before verbs compared to contexts before nouns towards the end of the utterance (maximal difference of 3.5472 segments per second at the beginning of utterances, minimal difference of 0.3763 segments per second at the end of utterances). In Dutch, by contrast, the interaction reverses the effect. Contexts before verbs are faster than contexts before nouns only for positions toward the beginning of utterances (< 0.89). These represent 74.83% of the Dutch data (maximal difference of 1.2582 segments per second at position 0). In later positions, which make up 25.17% of the Dutch data,

contexts before nouns are faster than contexts before verbs (predicted articulation rate difference of -0.1594 segments per second at position 1), against the cross-linguistic pattern. For English, in contrast, the coefficient of the word class effect is negative, indicating faster contexts before nouns than before verbs at position 0 (difference = -0.3136 segments per second), but this effect is reversed after position 0.33 (that is, for 75.65% of the English data) with a higher predicted articulation rate before verbs than before nouns towards the end of utterances (difference = 0.6897 segments per second). We currently do not have an explanation for these interactions, but we note that our main results still hold for the vast majority of the contexts, even in Dutch and English. Note that in Fig. 3 in the main text we show the overall effect of word class, averaging over positions, in order to simplify the visual representation.

In an additional study we included auxiliaries with verbs. Here, word class is again a highly significant predictor in the final model, as reported in Tables S19 and S20. In addition, there is again a significant interaction of position and word class. While at the beginning of utterances contexts before verbs are about half a segment per second (0.5260 segments per second) faster than contexts before nouns, at the end of utterances contexts before verbs are only 0.1349 segments per second faster than contexts before nouns. The interaction of word class and position thus reduces but never reverses the articulation rate advantage for verbs compared to nouns. However, positive coefficients for most languages in the word class:language interaction again suggests that the effect of word class is even stronger in some individual languages (compared to Bora, which is again used as reference language here). In fact, in a model without language and the word class:language interaction as fixed effects, contexts before verbs are predicted to be 0.8767 segments per second faster than contexts before nouns at the beginning of an utterance and 0.4753 segments per second faster at the end of an utterance. In addition to the effects of word class, there is again a significant effect of word length in that articulation is slightly faster before longer words. This effect is even more pronounced towards the end of an utterance (significant interaction of word length and position). Languages again also vary greatly with regard to the predicted average articulation rate (highly significant effect of language) and also with regard to the strength of the word class effect on articulation rate (significant interaction of word class and language).

Individual linear-mixed effects analyses of the four languages in which auxiliaries are annotated as a separate category—English, Even, Hoocąk, and Texistepec—yield very similar results: The effect of word class is still statistically significant according to likelihood ratio tests when these auxiliaries are included in the category of verbs (see Tables S21-24). In Even and Hoocąk, contexts before verbs always exhibit a higher articulation rate than contexts before nouns, also when including known auxiliaries. The interaction between word class and position is still significant in English and Texistepec when auxiliaries are included. In Texistepec, this interaction does not reverse the effect of word class, so contexts before verbs are always faster than contexts before nouns, regardless of position: the maximal difference obtains at the beginning of utterances (3.6225 segments per second), and the minimal difference at the end of utterances (0.3501 segments per second). In English, in contrast, the word class effect is again reversed at the beginning of utterances (positions < 0.29), in that contexts before nouns are predicted to be 0.2720 segments per seconds faster than contexts before verbs at the beginning of utterances (23.20% of the English data), while contexts before verbs are

predicted to be 0.7165 segments per second faster than contexts before nouns for positions > 0.29 (76.80% of the data).

To summarize the overall results of our different analyses of articulation rate, predictions for the beginning and end of utterances for individual languages (Table S25) show that in almost all cases, articulation rates in contexts before nouns are slower than in contexts before verbs. The exceptions are Bora and N‖ng, where the word class effect is not statistically significant, and partly also Dutch and English, where the word class effect is reversed for some positions, at the end of utterances in Dutch and at the beginning of utterances in English. Including known auxiliaries in the analyses does not change the results substantially. Overall, the predicted articulation rate is about 0.6 segments per second slower before nouns than before verbs. While this may seem like a relatively minor difference, it has to be taken into account that the mean number of segments per pre-word window in the whole corpus is 6.98 (std. dev. 3.36 segments) and the overall mean articulation rate is just 17.16 segments per second (std. dev. 6.57 segments per second). The predictions of our models thus correspond roughly to a 3.56% difference in articulation rates when comparing contexts before nouns and verbs. This can be compared to articulation rate differences of about 10.56% between different speakers (estimated from the standard deviation of the speaker random intercept in the cross-linguistic model, see Table S9).

Additional results pause probability study. Here we provide discussion of additional results from our study of pause probability, also taking into account the difference between silent and filled pauses and alternative analyses in which auxiliaries are included as verbs. Descriptive statistics on the probability of silent and filled pauses separately (Tables S29-31) show that silent pauses follow essentially the same pattern as silent and/or filled pauses taken together in that the probability of silent pauses before nouns is higher than before a verb, both when including or excluding auxiliaries, except in English. Filled pauses are more likely before nouns than verbs in all languages.

In our main study of pause probabilities, word class is a highly significant predictor (Tables S32-33). Cross-linguistically, there is a significantly lower probability of encountering a pause before a verb than before a noun (odds ratio: 0.5800). Moreover, a higher length of the target word and a later position both increase the probability of a pause in the context. The standardized coefficients, however, suggest that these two effects are weaker than the effect of word class. In separate analyses for individual languages (Tables S34-42), the probability of pauses is significantly lower before verbs than before nouns in six out of nine languages: Bora, Chintang, Even, Hoocąk, N‖ng, and Texistepec. In Baure and Dutch, the word class effect goes in the same direction (more pauses before nouns than before verbs) but is not statistically significant according to the likelihood-ratio tests. For English, however, the effect is reversed: The probability of pauses is significantly higher before verbs than before nouns. This needs further research, but for now we conclude that English deviates from the patterns reflected in all other languages.

In our alternative study of pause probabilities (including auxiliaries with verbs), word class is again a highly significant predictor of pause probability, as reported in Tables S42 and S44. Cross-linguistically, the model predicts a significantly lower probability of encountering a pause before a verb than before a noun (odds ratio =

0.5796). Moreover, the significant effects of word length and position again indicates that pauses are more likely to occur before longer words than shorter words and towards the end of utterances. In separate models for the individual languages (Tables S45-48) the probability of pauses remains significantly lower before verbs than before nouns after inclusion of auxiliaries in three out of the four languages that have auxiliaries (Even, Hoocąk, Texistepec). For English, however, the effect is again reversed: The probability of pauses is significantly higher before verbs than before nouns.

To summarize the overall results of our different analyses of pause probabilities, the predicted probabilities of pauses before nouns and verbs and the corresponding probability ratios (relative risks) and odds ratios for the individual languages (Table S49) show the following pattern: All odds ratios, except those for English, are above one, which means that pauses are more probable before nouns than before verbs in all languages except English (the factor word class is, however, not statistically significant in the generalized linear mixed-effects models for Baure and Dutch). Averaging across the nine languages, we obtain a probability ratio of about 1.6 and an odds ratio of about 1.7 for pauses before nouns versus pauses before verbs. There are thus on average about 60% more pauses before nouns than before verbs. In many languages, the probability ratio is almost 2, indicating that pauses before nouns are twice as likely as pauses before verbs (corresponding to a 100% increase in pause probability for nouns compared to verbs).

Supplementary Tables

Table S1: Basic corpus statistics

	Language information				Corpus information			
Language	**639-3 code**	**Glottocode**	**Family**	**Area**	**Speakers**	**Texts**	**Words**	**Reference**
Baure	brg	baur1253	Arawakan	Amazonia	12	37	17,652	Danielsen et al. (2009) (1)
Bora	boa	bora1263	Boran	Amazonia	46	37	29,802	Seifart (2009) (2)
Chintang	ctn	chhi1245	Sino-Tibetan	Himalayas	74	40	37,737	Bickel et al. (2011) (3)
Dutch	nld	dutc1256	Indo-European	Europe	42	17	39,519	CGN-consortium (2003) (4)
English	eng	stan1293	Indo-European	North America	80	47	56,135	Calhoun et al. (2009) (5)
Even	eve	lamu1253	Tungusic	Siberia	32	67	37,430	Pakendorf et al. (2010) (6)
Hoocąk	win	hoch1243	Siouan	North America	28	62	23,191	Hartmann (2013) (7)
N‖ng	ngh	nuuu1241	!Ui-Taa	Southern Africa	7	33	26,061	Güldemann et al. (2011) (8)
Texistepec	poq	texi1237	Mixe-Zoquean	Central America	1	6	21,321	Wichmann (1996) (9)
Sum					**322**	**346**	**288,848**	

Notes: The corpus of Even documents the Lamunkhin variety of Even, and the Texistepec language is also sometimes ambiguously referred to as (Texistepec) Popoluca. Unambiguous language codes are provided with ISO 639-3/Ethnologue (10) and Glottolog (11) standards.

Table S2: Statistics on pre-word context windows per language

Language	AUs	Words per AU	Pre-word windows	Size of pre-word windows (ms)		Excluded first words (%)
				Mean	Std. Dev.	
Baure	4 343	4.06	17 652	477	200	24.60 %
Bora	3 979	7.49	29 802	489	168	13.35 %
Chintang	9 026	4.18	37 737	454	159	23.92 %
Dutch	5 822	6.79	39 519	442	146	14.73 %
English	6 941	8.09	56 135	457	139	12.36 %
Even	4 839	7.74	37 430	485	148	12.93 %
Hoocąk	2 936	7.90	23 191	495	192	12.66 %
N‖ng	7 990	3.26	26 061	357	166	30.66 %
Texistepec	5 461	3.90	21 321	419	165	25.61 %
Sum	**51 337**		**288 848**			
Mean		**5.63**		**456**	**164**	**17.77 %**

Notes: AU = annotation unit; N‖ng has both the shortest annotation units on average and the shortest mean length of pre-word windows. For this reason, N‖ng also has the highest percentage of excluded first words in annotation units compared to the other languages.

Table S3: Differences in mean length of pre-word windows in the nine languages

language	excluding silence (analysis of articulation rate)					including silence (analysis of pause probability)				
	nouns		verbs		difference (ms)	nouns		verbs		difference (ms)
	mean length (ms)	standard deviation (ms)	mean length (ms)	standard deviation (ms)		mean length (ms)	standard deviation (ms)	mean length (ms)	standard deviation (ms)	
Baure	447	193	455	201	-8	467	182	473	189	-6
Bora	453	166	463	186	-10	493	146	500	163	-7
Chintang	427	167	442	154	-15	450	160	460	148	-10
Dutch	440	139	381	173	59	468	125	413	167	55
English	453	138	387	174	66	484	114	438	155	46
Even	450	168	477	155	-27	483	144	498	139	-15
Hoocąk	501	191	510	177	-9	501	155	516	158	-15
N‖ng	393	151	294	151	99	430	140	321	160	109
Texistepec	431	156	360	173	71	460	141	385	172	75

Table S4: Data preparation for main study articulation rate (excl. known auxiliaries).

Language	Word Class	Initial Contexts	Contexts Left After Excluding							Excluded relevant target words (N and V)
			Target words not N or V[a]	Target words with both N and V roots	First words in AU	Disfluencies or filled pauses as target words	Contexts of length zero	Empty contexts (containing no words)	Contexts containing disfluencies or filled pauses	
Baure	All	17 652	7 917	7 843	6 474	6 474	6 364	5 751	5 573	29.61 %
	N	2 778	2 778	2 704	2 454	2 454	2 427	2 167	2 099	24.44 %
	V	5 213	5 213	5 139	4 020	4 020	3 937	3 584	3 474	33.36 %
Bora	All	29 802	14 181	14 180	13 476	13 472	13 316	10 631	10 396	26.69 %
	N	5 913	5 913	5 912	5 566	5 566	5 507	4 105	3 974	32.79 %
	V	8 269	8 269	8 268	7 910	7 906	7 809	6 526	6 422	22.34 %
Chintang	All	37 737	15 263	15 260	12 364	12 347	12 283	11 462	11 091	27.33 %
	N	6 607	6 607	6 604	4 985	4 969	4 942	4 464	4 279	35.24 %
	V	8 659	8 659	8 656	7 379	7 378	7 341	6 998	6 812	21.33 %
Dutch	All	39 519	10 660	10 660	10 164	10 164	10 147	9 965	9 285	12.90 %
	N	3 973	3 973	3 973	3 863	3 863	3 853	3 768	3 466	12.76 %
	V	6 687	6 687	6 687	6 301	6 301	6 294	6 197	5 819	12.98 %
English	All	56 135	18 677	18 674	18 315	18 294	18 282	17 719	17 210	7.85 %
	N	7 777	7 777	7 774	7 542	7 521	7 516	7 246	7 029	9.62 %
	V	10 903	10 903	10 900	10 773	10 773	10 766	10 473	10 181	6.62 %
Even	All	37 430	18 378	18 378	17 215	17 129	17 017	14 223	12 897	29.82 %
	N	8 835	8 835	8 835	7 980	7 948	7 892	6 137	5 342	39.54 %
	V	9 543	9 543	9 543	9 235	9 181	9 125	8 086	7 555	20.83 %
Hoocąk	All	23 191	9 954	9 821	9 149	9 149	8 825	6 412	6 195	37.76 %
	N	3 683	3 670	3 537	3 091	3 091	2 968	1 767	1 646	55.15 %
	V	6 903	6 417	6 284	6 058	6 058	5 857	4 645	4 549	29.11 %
N‖ng	All	26 061	11 881	11 881	9 724	9 630	9 619	9 314	9 029	24.00 %
	N	4 415	4 415	4 415	3 336	3 296	3 290	3 128	2 993	32.21 %
	V	7 466	7 466	7 466	6 388	6 334	6 329	6 186	6 036	19.15 %
Texistepec	All	21 321	8 866	8 723	7 603	7 595	7 586	7 200	7 156	19.29 %
	N	2 738	2 738	2 595	2 393	2 386	2 382	2 154	2 136	21.99 %
	V	6 272	6 271	6 128	5 210	5 209	5 204	5 046	5 020	19.95 %
Overall	**All**	**288 848**	**115 777**	**115 420**	**104 484**	**104 254**	**103 439**	**92 677**	**88 832**	**23.27 %**
	N	**46 719**	**46 706**	**46 349**	**41 210**	**41 094**	**40 777**	**34 936**	**32 964**	**29.42 %**
	V	**69 915**	**69 428**	**69 071**	**63 274**	**63 160**	**62 662**	**57 741**	**55 868**	**19.53 %**

Note: [a] Since words can contain multiple roots of different word class categories, the respective N and V rows in this column (column 4) do not necessarily sum exactly to the value in the respective All row. This is, however, the case for the remainder of the columns after excluding target words containing both N and V roots in column 5. (AU = annotation unit)

Table S5: Data preparation for alternative study articulation rate (incl. auxiliaries).

Language	Word Class	Initial Contexts	Contexts Left After Excluding							Excluded contexts
			Target words not N, V, or AUX[a]	Target words with both N and V roots	First words in AU	Disfluencies or filled pauses as target words	Contexts of length zero	Empty contexts (containing no words)	Contexts containing disfluencies or filled pauses	
Baure	All	17 652	7 917	7 843	6 474	6 474	6 364	5 751	5 573	29.61 %
	N	2 778	2 778	2 704	2 454	2 454	2 427	2 167	2 099	24.44 %
	V	5 213	5 213	5 139	4 020	4 020	3 937	3 584	3 474	33.36 %
Bora	All	29 802	14 181	14 180	13 476	13 472	13 316	10 631	10 396	26.69 %
	N	5 913	5 913	5 912	5 566	5 566	5 507	4 105	3 974	32.79 %
	V	8 269	8 269	8 268	7 910	7 906	7 809	6 526	6 422	22.34 %
Chintang	All	37 737	15 263	15 260	12 364	12 347	12 283	11 462	11 091	27.33 %
	N	6 607	6 607	6 604	4 985	4 969	4 942	4 464	4 279	35.24 %
	V	8 659	8 659	8 656	7 379	7 378	7 341	6 998	6 812	21.33 %
Dutch	All	39 519	10 660	10 660	10 164	10 164	10 147	9 965	9 285	12.90 %
	N	3 973	3 973	3 973	3 863	3 863	3 853	3 768	3 466	12.76 %
	V	6 687	6 687	6 687	6 301	6 301	6 294	6 197	5 819	12.98 %
English	All	56 135	20 551	20 548	20 185	20 161	20 149	19 566	19 027	7.42 %
	N	7 777	7 777	7 774	7 542	7 521	7 516	7 246	7 029	9.62 %
	V	10 903	10 903	10 900	10 773	10 773	10 766	10 473	10 181	6.62 %
	AUX[b]	1 874	1 874	1 874	1 870	1 867	1 867	1 847	1 817	3.04 %
Even	All	37 430	19 003	19 003	17 788	17 700	17 588	14 741	13 379	29.60 %
	N	8 835	8 835	8 835	7 980	7 948	7 892	6 137	5 342	39.54 %
	V	9 543	9 543	9 543	9 235	9 181	9 125	8 086	7 555	20.83 %
	AUX[b]	625	625	625	573	571	571	518	482	22.88 %
Hoocąk	All	23 191	11 504	11 358	10 647	10 647	10 263	7 646	7 413	35.56 %
	N	3 683	3 683	3 537	3 091	3 091	2 968	1 767	1 646	55.31 %
	V	6 903	6 903	6 770	6 531	6 531	6 314	5 034	4 926	28.64 %
	AUX[b]	1 064	1 064	1 051	1 025	1 025	981	845	841	20.96 %
N‖ng	All	26 061	11 881	11 881	9 724	9 630	9 619	9 314	9 029	24.00 %
	N	4 415	4 415	4 415	3 336	3 296	3 290	3 128	2 993	32.21 %
	V	7 466	7 466	7 466	6 388	6 334	6 329	6 186	6 036	19.15 %
Texistepec	All	21 321	10 916	10 773	8 959	8 948	8 939	8 453	8 395	23.09 %
	N	2 738	2 738	2 595	2 393	2 386	2 382	2 154	2 136	21.99 %
	V	6 272	6 272	6 129	5 211	5 210	5 205	5 047	5 021	19.95 %
	AUX[b]	2 049	2 049	2 049	1 355	1 352	1 352	1 252	1 238	39.58 %
Overall	**All**	**288 848**	**121 876**	**121 506**	**109 781**	**109 543**	**108 668**	**97 529**	**93 588**	**23.21 %**
	N	**46 719**	**46 719**	**46 349**	**41 210**	**41 094**	**40 777**	**34 936**	**32 964**	**29.44 %**
	V	**69 915**	**69 915**	**69 558**	**63 748**	**63 634**	**63 120**	**58 131**	**56 246**	**19.55 %**
	AUX[b]	**5 612**	**5 612**	**5 599**	**4 823**	**4 815**	**4 771**	**4 462**	**4 378**	**21.99 %**

Notes: [a] Since words can contain multiple roots of different word class categories, the respective *N* and *V* rows in this column do not necessarily sum exactly to the value in the respective *All* row. This is, however, the case for the remainder of the columns after excluding target words containing both N and V roots in column 5. [b] Words containing both a V root and an AUX root are counted as V in this Table (but are still excluded from the data for the alternative study excluding auxiliaries summarized in Table S4). Words containing only an AUX root but no V root are counted as auxiliaries. Words containing both a V root and an AUX root still only enter as one observation into the statistical analyses. (AU = annotation unit)

Table S6: Summary statistics of articulation rate in pre-word windows by language

Language	Variant	Mean	Std. Dev.	Median	Min	Max
Baure		13.91	4.86	13.51	2.20	55.56
Bora		17.21	6.13	16.28	1.72	108.70
Chintang		17.31	6.36	17.07	1.17	104.65
Dutch		20.85	7.25	20.04	1.98	129.03
English	with AUX	19.04	7.38	18.46	1.45	85.11
	without AUX	19.01	7.35	18.42	1.45	85.11
Even	with AUX	14.92	3.83	14.89	2.35	37.63
	without AUX	14.88	3.82	14.83	2.35	37.63
Hoocąk	with AUX	12.90	4.24	12.50	2.48	82.09
	without AUX	12.88	4.33	12.45	2.48	82.09
N‖ng		16.23	6.42	15.38	2.55	80.00
Texistepec	with AUX	19.18	6.68	18.42	3.03	93.75
	without AUX	19.22	6.60	18.43	3.74	93.75
Overall	**with AUX**	**17.17**	**6.59**	**16.33**	**1.17**	**129.03**
	without AUX	**17.17**	**6.57**	**16.33**	**1.17**	**129.03**

Table S7: Articulation rate in pre-word windows depending on word class of target word for individual languages

Language	word class	Articulation rate			Comparison		
		Mean	Std. Dev.	Median	Contrast	Difference	In-/ Decrease
Baure	N	13.60	4.53	13.38	N vs. V	-0.50	-3.56%
	V	14.10	5.04	13.64			
Bora	N	17.02	5.88	16.13	N vs. V	-0.32	-1.83%
	V	17.34	6.28	16.39			
Chintang	N	16.50	6.58	16.13	N vs. V	-1.32	-7.43%
	V	17.82	6.16	17.65			
Dutch	N	20.79	6.37	20.27	N vs. V	-0.10	-0.46%
	V	20.89	7.73	19.84			
English	N	18.77	6.28	18.40	N vs. V	-0.40	-2.09%
	V	19.17	8.00	18.46			
	AUX	19.34	7.62	18.68	N vs. V + AUX	-0.43	-2.24%
	V + AUX	19.20	7.95	18.52			
Even	N	14.51	4.07	14.45	N vs. V	-0.63	-4.15%
	V	15.14	3.61	15.08			
	AUX	16.07	3.96	16.01	N vs. V + AUX	-0.68	-4.51%
	V + AUX	15.19	3.64	15.12			
Hoocąk	N	12.21	4.03	11.98	N vs. V	-0.91	-6.94%
	V	13.12	4.37	12.57			
	AUX	12.96	3.67	12.72	N vs. V + AUX	-0.89	-6.77%
	V + AUX	13.09	4.28	12.59			
N‖ng	N	16.24	5.89	15.63	N vs. V	0.01	0.06%
	V	16.23	6.67	15.38			
Texistepec	N	18.15	5.69	17.74	N vs. V	-1.52	-7.73%
	V	19.67	6.90	18.75			
	AUX	18.98	7.09	18.18	N vs. V + AUX	-1.38	-7.09%
	V + AUX	19.54	6.94	18.64			
Overall	**N**	**16.86**	**6.14**	**16.22**	**N vs. V**	**-0.46**	**-2.66%**
	V	**17.32**	**6.80**	**16.39**			
	AUX	**18.65**	**6.99**	**16.67**	**N vs. V + AUX**	**-0.48**	**-2.77%**
	V + AUX	**17.34**	**6.82**	**16.39**			

Table S8: Comparison of models for predicting articulation rate in all nine languages

Step	Fixed effects structure	Added	AIC	χ^2	df	*p*	sig.
1 (initial)	word class + language + word class:language		567 374				
2a	word class + word length + language + word class:language	word length	567 294	81.64	1	<0.001	***
2b	word class + position + language + word class:language	position	567 366	9.94	1	0.002	**
3a	word class + word length + position + language + word class:language	position	567 282	13.93	1	<0.001	***
3b	word class + word length + word class:word length + language + word class:language	word class:word length	567 295	0.51	1	0.474	n.s.
4a	word class + word length + position + word class:word length + language + word class:language	word class:word length	567 283	0.67	1	0.412	n.s.
4b	word class + word length + position + word class:position + language + word class:language	word class:position	567 277	6.97	1	0.008	**
4c	word class + word length + position + word length:position + language + word class:language	word length:position	567 279	4.65	1	0.031	*
5a	word class + word length + position + word class:word length + word class:position + language + word class:language	word class:word length	567 278	0.91	1	0.340	n.s.
5b (final)	word class + word length + position + word class:position + word length:position + language + word class:language	word length:position	567 278	4.29	1	0.038	*
6	word class + word length + position + word class:word length + word class:position + word length:position + language + word class:language	word class:word length	567 275	0.79	1	0.375	n.s.

Notes: Alternative models are given the same step number followed by different letters. The model structure chosen as the next step is marked in bold.

Table S9: Model predicting articulation rate in all nine languages

Fixed effect	Coefficient	Std. Error	*t* Value	Std. Coefficient	χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(intercept)	16.9708	0.4238	40.040							
word class = verb	0.5147	0.1608	3.200	0.0378	215.89	10	<0.001	***	<0.001	***
word length	0.1327	0.0532	2.500	0.0201	90.96	2	<0.001	***	<0.001	***
position	-0.1011	0.1250	-0.810	-0.0046	25.19	3	<0.001	***	<0.001	***
language = Baure	-3.7686	0.7908	-4.770	-0.1390	295.38	16	<0.001	***	<0.001	***
language = Chintang	-0.7718	0.5233	-1.470	-0.0388						
language = Dutch	3.9246	0.5640	6.960	0.1826						
language = English	1.4745	0.4901	3.010	0.0886						
language = Even	-2.4730	0.5621	-4.400	-0.1325						
language = Hoocąk	-4.4974	0.5997	-7.500	-0.1742						
language = N‖ng	-0.8774	0.8282	-1.060	-0.0403						
language = Texistepec	1.1131	1.8800	0.590	0.0461						
word class = verb : position	-0.3772	0.1468	-2.570	-0.0224	6.60	1	0.010	*	0.014	*
word length : position	0.1444	0.0698	2.070	0.0165	4.29	1	0.039	*	0.050	n.s.
word class = verb : language = Baure	0.2528	0.2260	1.120	0.0075	67.26	8	<0.001	***	<0.001	***
word class = verb : language = Chintang	0.9577	0.1864	5.140	0.0388						
word class = verb : language = Dutch	0.1957	0.2094	0.930	0.0074						
word class = verb : language = English	0.0107	0.1894	0.060	0.0005						
word class = verb : language = Even	0.3130	0.1794	1.740	0.0133						
word class = verb : language = Hoocąk	0.4499	0.2311	1.950	0.0151						
word class = verb : language = N‖ng	-0.2556	0.2180	-1.170	-0.0098						
word class = verb : language = Texistepec	1.1452	0.2256	5.080	0.0402						
Random effect	**Groups**	**Std. Dev.**			χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(1|speaker)	310	1.7863			1085.80	1	<0.001	***	<0.001	***
(1|text)	346	0.9018			484.45	1	<0.001	***	<0.001	***
(1|word type)	29784	1.2211			1838.80	1	<0.001	***	<0.001	***

Table S10: Model predicting articulation rate in Baure

Fixed effect	Coefficient	Std. Error	*t* Value	Std. Coefficient	χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(intercept)	13.4844	0.3523	38.270							
word class = verb	0.5952	0.3897	1.530	0.0593	13.00	2	0.002	**	0.002	**
word length	0.0252	0.1962	0.130	0.0047	5.22	2	0.074	n.s.	0.091	n.s.
position	-0.2806	0.3758	-0.750	-0.0176	1.26	3	0.738	n.s.	0.789	n.s.
word class = verb : position	0.0243	0.4867	0.050	0.0020	0.00	1	0.959	n.s.	0.982	n.s.
word length : position	0.2347	0.2519	0.930	0.0344	0.87	1	0.352	n.s.	0.406	n.s.
Random effect	**Groups**	**Std. Dev.**			χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(1|speaker)	10	0.0000			0.00	1	1.000	n.s.	1.000	n.s.
(1|text)	37	0.8382			56.00	1	<0.001	***	<0.001	***
(1|word type)	2761	1.6110			145.00	1	<0.001	***	<0.001	***

Table S11: Model predicting articulation rate in Bora

Fixed effect	Coefficient	Std. Error	*t* Value	Std. Coefficient	χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(intercept)	17.3305	0.7375	23.498							
word class = verb	0.1450	0.2895	0.501	0.0115	0.90	2	0.639	n.s.	0.697	n.s.
word length	0.4135	0.1513	2.734	0.0632	60.27	2	<0.001	***	<0.001	***
position	-0.4033	0.3329	-1.211	-0.0196	3.92	3	0.270	n.s.	0.320	n.s.
word class = verb : position	-0.0211	0.4111	-0.051	-0.0014	0.00	1	0.960	n.s.	0.982	n.s.
word length : position	0.2160	0.2064	1.046	0.0240	1.10	1	0.295	n.s.	0.345	n.s.
Random effect	**Groups**	**Std. Dev.**			χ^2	**df**	***p***	**sig.**	***p* (BH)**	**sig. (BH)**
(1|speaker)	43	3.8930			89.60	1	<0.001	***	<0.001	***
(1|text)	37	1.0380			87.20	1	<0.001	***	<0.001	***
(1|word type)	5002	1.6850			103.30	1	<0.001	***	<0.001	***

Table S12: Model predicting articulation rate in Chintang

Fixed effect	Coefficient	Std. Error	*t* Value	Std. Coefficient	χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(intercept)	16.1275	0.3750	43.000							
word class = verb	1.3096	0.3188	4.110	0.1002	66.20	2	<0.001	***	<0.001	***
word length	0.2244	0.1713	1.310	0.0347	21.76	2	<0.001	***	<0.001	***
position	-0.0450	0.3538	-0.130	-0.0020	0.62	3	0.892	n.s.	0.936	n.s.
word class = verb : position	-0.2059	0.4564	-0.450	-0.0134	0.20	1	0.652	n.s.	0.704	n.s.
word length : position	0.1268	0.2221	0.570	0.0158	0.33	1	0.568	n.s.	0.623	n.s.
Random effect	**Groups**	**Std. Dev.**			χ^2	**df**	***p***	**sig.**	***p* (BH)**	**sig. (BH)**
(1|speaker)	68	1.6690			169.10	1	<0.001	***	<0.001	***
(1|text)	40	0.9733			31.40	1	<0.001	***	<0.001	***
(1|word type)	4707	1.3520			89.70	1	<0.001	***	<0.001	***

Table S13: Model predicting articulation rate in Dutch

Fixed effect	Coefficient	Std. Error	*t* Value	Std. Coefficient	χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(intercept)	21.0532	0.5675	37.100							
word class = verb	1.2707	0.4111	3.090	0.0848	9.65	2	0.008	**	0.012	*
word length	-0.1204	0.1804	-0.670	-0.0190	7.59	2	0.022	*	0.030	*
position	-0.3159	0.5108	-0.620	-0.0135	30.90	3	<0.001	***	<0.001	***
word class = verb : position	-1.4302	0.5553	-2.580	-0.0718	6.62	1	0.010	*	0.014	*
word length : position	0.4236	0.2365	1.790	0.0504	3.20	1	0.074	n.s.	0.091	n.s.
Random effect	**Groups**	**Std. Dev.**			χ^2	**df**	***p***	**sig.**	***p* (BH)**	**sig. (BH)**
(1|speaker)	42	1.5840			139.42	1	<0.001	***	<0.001	***
(1|text)	17	1.3790			3.23	1	0.070	n.s.	0.090	n.s.
(1|word type)	2382	1.2750			118.82	1	<0.001	***	<0.001	***

Table S14: Model predicting articulation rate in English

Fixed effect	Coefficient	Std. Error	*t* Value	Std. Coefficient	χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(intercept)	18.9596	0.3558	53.290							
word class = verb	-0.3328	0.3085	-1.080	-0.0223	9.54	2	0.008	**	0.012	*
word length	-0.2771	0.1462	-1.900	-0.0399	6.61	2	0.037	*	0.048	*
position	-0.5736	0.3502	-1.640	-0.0230	11.09	3	0.011	*	0.016	*
word class = verb : position	1.0226	0.4158	2.460	0.0463	6.05	1	0.014	*	0.018	*
word length : position	0.4879	0.1936	2.520	0.0493	6.35	1	0.012	*	0.016	*
Random effect	**Groups**	**Std. Dev.**			χ^2	**df**	***p***	**sig.**	***p* (BH)**	**sig. (BH)**
(1|speaker)	80	1.7826			339.10	1	<0.001	***	<0.001	***
(1|text)	47	0.7756			12.30	1	<0.001	***	<0.001	***
(1|word type)	2783	1.5932			846.20	1	<0.001	***	<0.001	***

Table S15: Model predicting articulation rate in Even

Fixed effect	Coefficient	Std. Error	*t* Value	Std. Coefficient	χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(intercept)	14.1926	0.2707	52.440							
word class = verb	0.5619	0.1591	3.530	0.0725	46.66	2	<0.001	***	<0.001	***
word length	0.3132	0.0878	3.570	0.0735	36.71	2	<0.001	***	<0.001	***
position	0.3605	0.1792	2.010	0.0282	7.59	3	0.055	n.s.	0.071	n.s.
word class = verb : position	-0.0339	0.2278	-0.150	-0.0154	0.02	1	0.882	n.s.	0.930	n.s.
word length : position	-0.0919	0.1235	-0.740	-0.0036	0.56	1	0.456	n.s.	0.511	n.s.
Random effect	**Groups**	**Std. Dev.**			χ^2	**df**	***p***	**sig.**	***p* (BH)**	**sig. (BH)**
(1|speaker)	31	1.1312			59.90	1	<0.001	***	<0.001	***
(1|text)	67	0.7574			161.90	1	<0.001	***	<0.001	***
(1|word type)	5720	1.0066			63.40	1	<0.001	***	<0.001	***

Table S16: Model predicting articulation rate in Hoocąk

Fixed effect	Coefficient	Std. Error	*t* Value	Std. Coefficient	χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(intercept)	12.1449	0.4654	26.094							
word class = verb	0.4952	0.2545	1.946	0.0505	18.77	2	<0.001	***	<0.001	***
word length	0.0056	0.1221	0.046	0.0013	1.83	2	0.401	n.s.	0.458	n.s.
position	0.6635	0.3684	1.801	0.0459	23.14	3	<0.001	***	<0.001	***
word class = verb : position	0.1628	0.4220	0.386	0.0144	0.15	1	0.700	n.s.	0.753	n.s.
word length : position	0.1001	0.1711	0.585	0.0170	0.34	1	0.558	n.s.	0.615	n.s.
Random effect	**Groups**	**Std. Dev.**			χ^2	**df**	***p***	**sig.**	***p* (BH)**	**sig. (BH)**
(1|speaker)	28	1.8185			43.90	1	<0.001	***	<0.001	***
(1|text)	62	1.2311			166.50	1	<0.001	***	<0.001	***
(1|word type)	3209	0.9049			33.10	1	<0.001	***	<0.001	***

Table S17: Model predicting articulation rate in N‖ng

Fixed effect	Coefficient	Std. Error	*t* Value	Std. Coefficient	χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(intercept)	16.9507	0.8685	19.517							
word class = verb	0.1444	0.4833	0.299	0.0106	1.27	2	0.529	n.s.	0.587	n.s.
word length	0.3221	0.2710	1.189	0.0437	7.48	2	0.024	*	0.032	*
position	-1.1214	0.5352	-2.095	-0.0486	33.63	3	<0.001	***	<0.001	***
word class = verb : position	-0.3811	0.5577	-0.683	-0.0244	0.47	1	0.493	n.s.	0.550	n.s.
word length : position	-0.0472	0.3045	-0.155	-0.0057	0.02	1	0.876	n.s.	0.928	n.s.
Random effect	**Groups**	**Std. Dev.**			χ^2	**df**	***p***	**sig.**	***p* (BH)**	**sig. (BH)**
(1|speaker)	7	1.8140			41.50	1	<0.001	***	<0.001	***
(1|text)	33	1.2300			222.80	1	<0.001	***	<0.001	***
(1|word type)	1097	1.2610			101.20	1	<0.001	***	<0.001	***

Table S18: Model predicting articulation rate in Texistepec

Fixed effect	Coefficient	Std. Error	*t* Value	Std. Coefficient	χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(intercept)	17.4141	0.4647	37.470							
word class = verb	3.7737	0.5112	7.380	0.2616	62.76	2	<0.001	***	<0.001	***
word length	0.5646	0.2131	2.650	0.0904	10.69	2	0.005	**	0.007	**
position	0.8543	0.5302	1.610	0.0355	68.91	3	<0.001	***	<0.001	***
word class = verb : position	-3.3974	0.6300	-5.390	-0.1968	28.92	1	<0.001	***	<0.001	***
word length : position	-0.4325	0.2733	-1.580	-0.0534	2.52	1	0.113	n.s.	0.136	n.s.
Random effect	**Groups**	**Std. Dev.**			χ^2	**df**	***p***	**sig.**	***p* (BH)**	**sig. (BH)**
(1|text)	6	0.3474			7.69	1	0.006	**	0.008	**
(1|word type)	2123	1.8003			113.56	1	<0.001	***	<0.001	***

Table S19: Comparison of models for predicting articulation rate (incl. auxiliaries) in all nine languages

Step	Fixed effects structure	Added	AIC	χ^2	df	*p*	sig.
1 (initial)	word class + language + word class:language		598 231				
2a	word class + word length + language + word class:language	word length	598 155	78.22	1	<0.001	***
2b	word class + position + language + word class:language	position	598 221	12.40	1	<0.001	***
3a	word class + word length + position + language + word class:language	position	598 140	17.08	1	<0.001	***
3b	word class + word length + word class:word length + language + word class:language	word class:word length	598 157	0.17	1	0.678	n.s.
4a	word class + word length + position + word class:word length + language + word class:language	word class:word length	598 142	0.30	1	0.586	n.s.
4b	word class + word length + position + word class:position + language + word class:language	word class:position	598 134	8.14	1	0.004	**
4c	word class + word length + position + word length:position + language + word class:language	word length:position	598 135	7.40	1	0.007	**
5a	word class + word length + position + word class:word length + word class:position + language + word class:language	word class:word length	598 135	0.48	1	0.488	n.s.
5b (final)	word class + word length + position + word class:position + word length:position + language + word class:language	word length:position	598 129	6.60	1	0.010	*
6	word class + word length + position + word class:word length + word class:position + word length:position + language + word class:language	word class:word length	598 131	0.38	1	0.538	n.s.

Notes: Alternative models are given the same step number followed by different letters. The model structure chosen as the next step is marked in bold.

Table S20: Model predicting articulation rate (incl. auxiliaries) in all nine languages

Fixed effect	Coefficient	Std. Error	*t* Value	Std. Coefficient	χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(intercept)	16.9820	0.4247	39.990							
word class = verb	0.5295	0.1601	3.310	0.0384	224.61	10	<0.001	***	<0.001	***
word length	0.1052	0.0515	2.040	0.0162	90.73	2	<0.001	***	<0.001	***
position	-0.1175	0.1248	-0.940	-0.0053	31.82	3	<0.001	***	<0.001	***
language = Baure	-3.7640	0.7928	-4.750	-0.1351	294.48	16	<0.001	***		
language = Chintang	-0.7646	0.5243	-1.460	-0.0375						
language = Dutch	3.9426	0.5650	6.980	0.1788						
language = English	1.4978	0.4910	3.050	0.0914						
language = Even	-2.4481	0.5632	-4.350	-0.1300						
language = Hoocąk	-4.5612	0.5997	-7.610	-0.1869						
language = N‖ng	-0.8674	0.8306	-1.040	-0.0388						
language = Texistepec	1.0737	1.8862	0.570	0.0465						
word class = verb : position	-0.3946	0.1456	-2.710	-0.0231	7.34	1	0.007	**	0.010	*
word length : position	0.1733	0.0675	2.570	0.0199	6.60	1	0.010	*	0.014	*
word class = verb : language = Baure	0.2472	0.2256	1.100	0.0071	65.00	8	<0.001	***	<0.001	***
word class = verb : language = Chintang	0.9538	0.1860	5.130	0.0376						
word class = verb : language = Dutch	0.1808	0.2086	0.870	0.0066						
word class = verb : language = English	0.0506	0.1870	0.270	0.0026						
word class = verb : language = Even	0.3537	0.1783	1.980	0.0150						
word class = verb : language = Hoocąk	0.4839	0.2258	2.140	0.0177						
word class = verb : language = N‖ng	-0.2585	0.2170	-1.190	-0.0096						
word class = verb : language = Texistepec	1.0986	0.2237	4.910	0.0416						
Random effect	**Groups**	**Std. Dev.**			**χ^2**	**df**	***p***	**sig.**	***p* (BH)**	**sig. (BH)**
(1|speaker)	310	1.7932			1 164.20	1	<0.001	***	<0.001	***
(1|text)	346	0.8994			497.73	1	<0.001	***	<0.001	***
(1|word type)	30526	1.1798			1 808.80	1	<0.001	***	<0.001	***

Table S21: Model predicting articulation rate (incl. auxiliaries) in English

Fixed effect	Coefficient	Std. Error	*t* Value	Std. Coefficient	χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(intercept)	19.0318	0.3562	53.430							
word class = verb	-0.2910	0.3064	-0.950	-0.0190	10.35	2	0.006	**	0.008	**
word length	-0.3498	0.1389	-2.520	-0.0524	10.16	2	0.006	**	0.009	**
position	-0.6450	0.3463	-1.860	-0.0257	13.00	3	0.005	**	0.007	**
word class = verb : position	1.0075	0.4133	2.440	0.0449	5.94	1	0.015	*	0.020	*
word length : position	0.5796	0.1827	3.170	0.0585	10.07	1	0.002	**	0.002	**
Random effect	**Groups**	**Std. Dev.**			**χ^2**	**df**	***p***	**sig.**	***p* (BH)**	**sig. (BH)**
(1|speaker)	80	1.8269			388.30	1	<0.001	***	<0.001	***
(1|text)	47	0.7707			13.10	1	<0.001	***	<0.001	***
(1|word type)	2794	1.5689			843.40	1	<0.001	***	<0.001	***

Table S22: Model predicting articulation rate (incl. auxiliaries) in Even

Fixed effect	Coefficient	Std. Error	*t* Value	Std. Coefficient	χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(intercept)	14.2215	0.2754	51.650							
word class = verb	0.6000	0.1569	3.830	0.0767	55.37	2	<0.001	***	<0.001	***
word length	0.2954	0.0859	3.440	0.0696	30.53	2	<0.001	***	<0.001	***
position	0.3710	0.1793	2.070	0.0288	8.55	3	0.036	*	0.047	*
word class = verb : position	-0.0222	0.2251	-0.100	-0.0023	0.01	1	0.922	n.s.	0.957	n.s.
word length : position	-0.1093	0.1208	-0.910	-0.0184	0.82	1	0.365	n.s.	0.420	n.s.
Random effect	**Groups**	**Std. Dev.**			χ^2	**df**	***p***	**sig.**	***p* (BH)**	**sig. (BH)**
(1|speaker)	31	1.1630			67.40	1	<0.001	***	<0.001	***
(1|text)	67	0.7590			163.30	1	<0.001	***	<0.001	***
(1|word type)	5787	1.0230			82.10	1	<0.001	***	<0.001	***

Table S23: **Model predicting articulation rate (incl. auxiliaries) in Hoocąk**

Fixed effect	Coefficient	Std. Error	*t* Value	Std. Coefficient	χ^2	df	*p*	sig.	p (BH)	sig. (BH)
(intercept)	12.0592	0.4422	27.273							
word class = verb	0.6778	0.2414	2.808	0.0664	26.31	2	<0.001	***	<0.001	***
word length	0.0740	0.1112	0.666	0.0176	1.74	2	0.418	n.s.	0.475	n.s.
position	0.6756	0.3604	1.875	0.0479	15.65	3	0.001	**	0.002	**
word class = verb : position	-0.0516	0.4036	-0.128	-0.0046	0.02	1	0.898	n.s.	0.937	n.s.
word length : position	-0.0115	0.1526	-0.076	-0.0021	0.01	1	0.941	n.s.	0.972	n.s.
Random effect	**Groups**	**Std. Dev.**			χ^2	**df**	***p***	**sig.**	**p (BH)**	**sig. (BH)**
(1|speaker)	28	1.7133			59.50	1	<0.001	***	<0.001	***
(1|text)	62	1.1718			212.60	1	<0.001	***	<0.001	***
(1|word type)	3843	0.8674			36.20	1	<0.001	***	<0.001	***

Table S24: Model predicting articulation rate (incl. auxiliaries) in Texistepec

Fixed effect	Coefficient	Std. Error	*t* Value	Std. Coefficient	χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(intercept)	17.4430	0.4674	37.320							
word class = verb	3.7716	0.5036	7.490	0.2461	62.99	2	<0.001	***	<0.001	***
word length	0.4991	0.1990	2.510	0.0812	12.94	2	0.002	**	0.002	**
position	0.7799	0.5374	1.450	0.0321	77.30	3	<0.001	***	<0.001	***
word class = verb : position	-3.5040	0.6231	-5.620	-0.1896	31.47	1	<0.001	***	<0.001	***
word length : position	-0.2839	0.2560	-1.110	-0.0342	1.24	1	0.265	n.s.	0.315	n.s.
Random effect	**Groups**	**Std. Dev.**			χ^2	**df**	***p***	**sig.**	***p* (BH)**	**sig. (BH)**
(1|text)	6	0.3438			8.67	1	0.003	**	0.005	**
(1|word type)	2153	1.6288			100.99	1	<0.001	***	<0.001	***

Table S25: Size of word class effect on articulation rate

language	predicted articulation rate difference (nouns vs. verbs)			
	auxiliaries excluded		auxiliaries included	
	at the beginning of utterances	at the end of utterances	at the beginning of utterances	at the end of utterances
Baure	-0.5967	-0.6195		
Bora	-0.1445	-0.1239		
Chintang	-1.2982	-1.1037		
Dutch	-1.2582	0.1595		
English	0.3136	-0.6897	0.2720	-0.7165
Even	-0.5613	-0.5280	-0.5996	-0.5779
Hoocąk	-0.4970	-0.6580	-0.6772	-0.6262
N‖ng	-0.1254	0.2367		
Texistepec	-3.5472	-0.3763	-3.5872	-0.2676
grand mean[a]	**-0.8572**	**-0.4114**	**-0.8906**	**-0.4043**

Note: [a] We averaged across all nine languages, both when excluding and when including auxiliaries. We did, however, not repeat unchanged results in the *auxiliaries included* columns.

Table S26: Data preparation for main study pause probability (excl. known auxiliaries)

Language	Word Class	Initial Contexts	Contexts Left After Excluding					Excluded relevant target words (N and V)
			Target words of other classes (than N and V)[a]	Target words with both N and V roots	First words in AU	Disfluencies or filled pauses as target words	Contexts of length zero	
Baure	All	17 652	7 917	7 843	6 474	6 474	6 364	19.62 %
	N	2 778	2 778	2 704	2 454	2 454	2 427	12.63 %
	V	5 213	5 213	5 139	4 020	4 020	3 937	24.48 %
Bora	All	29 802	14 181	14 180	13 476	13 472	13 316	6.10 %
	N	5 913	5 913	5 912	5 566	5 566	5 507	6.87 %
	V	8 269	8 269	8 268	7 910	7 906	7 809	5.56 %
Chintang	All	37 737	15 263	15 260	12 364	12 347	12 283	19.52 %
	N	6 607	6 607	6 604	4 985	4 969	4 942	25.20 %
	V	8 659	8 659	8 656	7 379	7 378	7 341	15.22 %
Dutch	All	39 519	10 660	10 660	10 164	10 164	10 147	4.81 %
	N	3 973	3 973	3 973	3 863	3 863	3 853	3.02 %
	V	6 687	6 687	6 687	6 301	6 301	6 294	5.88 %
English	All	56 135	18 677	18 674	18 315	18 294	18 282	2.11 %
	N	7 777	7 777	7 774	7 542	7 521	7 516	3.36 %
	V	10 903	10 903	10 900	10 773	10 773	10 766	1.26 %
Even	All	37 430	18 378	18 378	17 215	17 129	17 017	7.41 %
	N	8 835	8 835	8 835	7 980	7 948	7 892	10.67 %
	V	9 543	9 543	9 543	9 235	9 181	9 125	4.38 %
Hoocąk	All	23 191	9 954	9 821	9 149	9 149	8 825	11.34 %
	N	3 683	3 670	3 537	3 091	3 091	2 968	19.13 %
	V	6 903	6 417	6 284	6 058	6 058	5 857	8.73 %
N‖ng	All	26 061	11 881	11 881	9 724	9 630	9 619	19.04 %
	N	4 415	4 415	4 415	3 336	3 296	3 290	25.48 %
	V	7 466	7 466	7 466	6 388	6 334	6 329	15.23 %
Texistepec	All	21 321	8 866	8 723	7 603	7 595	7 586	14.44 %
	N	2 738	2 738	2 595	2 393	2 386	2 382	13.00 %
	V	6 272	6 271	6 128	5 210	5 209	5 204	17.01 %
Overall	**All**	**288 848**	**115 777**	**115 420**	**104 484**	**104 254**	**103 439**	**10.66 %**
	N	**46 719**	**46 706**	**46 349**	**41 210**	**41 094**	**40 777**	**12.69 %**
	V	**69 915**	**69 428**	**69 071**	**63 274**	**63 160**	**62 662**	**9.75 %**

Note: [a] Since words can contain multiple roots of different word class categories, the respective *N* and *V* rows in this column (column 4) do not necessarily sum exactly to the value in the respective *All* row. This is, however, the case for the remainder of the columns after excluding target words containing both N and V roots in column 5. (AU = annotation unit)

Table S27: Data preparation for alternative study pause probability (incl. auxiliaries)

Language	Word Class	Initial Contexts	Contexts Left After Excluding					Excluded relevant target words (N, V, and AUX)
			Target words of other classes (than N, V, and AUX)[a]	Target words with both N and V roots	First words in AU	Disfluencies or filled pauses as target words	Contexts of length zero	
Baure	All	17 652	7 917	7 843	6 474	6 474	6 364	19.62 %
	N	2 778	2 778	2 704	2 454	2 454	2 427	12.63 %
	V	5 213	5 213	5 139	4 020	4 020	3 937	24.48 %
Bora	All	29 802	14 181	14 180	13 476	13 472	13 316	6.10 %
	N	5 913	5 913	5 912	5 566	5 566	5 507	6.87 %
	V	8 269	8 269	8 268	7 910	7 906	7 809	5.56 %
Chintang	All	37 737	15 263	15 260	12 364	12 347	12 283	19.52 %
	N	6 607	6 607	6 604	4 985	4 969	4 942	25.20 %
	V	8 659	8 659	8 656	7 379	7 378	7 341	15.22 %
Dutch	All	39 519	10 660	10 660	10 164	10 164	10 147	4.81 %
	N	3 973	3 973	3 973	3 863	3 863	3 853	3.02 %
	V	6 687	6 687	6 687	6 301	6 301	6 294	5.88 %
English	All	56 135	20 551	20 548	20 185	20 161	20 149	1.96 %
	N	7 777	7 777	7 774	7 542	7 521	7 516	3.36 %
	V	10 903	10 903	10 900	10 773	10 773	10 766	1.26 %
	AUX[b]	1 874	1 874	1 874	1 870	1 867	1 867	0.37 %
Even	All	37 430	19 003	19 003	17 788	17 700	17 588	7.45 %
	N	8 835	8 835	8 835	7 980	7 948	7 892	10.67 %
	V	9 543	9 543	9 543	9 235	9 181	9 125	4.38 %
	AUX[b]	625	625	625	573	571	571	8.64 %
Hoocąk	All	23 191	11 504	11 358	10 647	10 647	10 263	10.79 %
	N	3 683	3 683	3 537	3 091	3 091	2 968	19.41 %
	V	6 903	6 903	6 770	6 531	6 531	6 314	8.53 %
	AUX[b]	1 064	1 064	1 051	1 025	1 025	981	7.80 %
N‖ng	All	26 061	11 881	11 881	9 724	9 630	9 619	19.04 %
	N	4 415	4 415	4 415	3 336	3 296	3 290	25.48 %
	V	7 466	7 466	7 466	6 388	6 334	6 329	15.23 %
Texistepec	All	21 321	10 916	10 773	8 959	8 948	8 939	18.11 %
	N	2 738	2 738	2 595	2 393	2 386	2 382	13.00 %
	V	6 272	6 272	6 129	5 211	5 210	5 205	17.01 %
	AUX[b]	2 049	2 049	2 049	1 355	1 352	1 352	34.02 %
Overall	**All**	**288 848**	**121 876**	**121 506**	**109 781**	**109 543**	**108 668**	**10.84 %**
	N	**46 719**	**46 719**	**46 349**	**41 210**	**41 094**	**40 777**	**12.72 %**
	V	**69 915**	**69 915**	**69 558**	**63 748**	**63 634**	**63 120**	**9.72 %**
	AUX[b]	**5 612**	**5 612**	**5 599**	**4 823**	**4 815**	**4 771**	**14.99 %**

Notes: [a] Since words can contain multiple roots of different word class categories, the respective *N* and *V* rows in this column (column 4) do not necessarily sum exactly to the value in the respective *All* row. This is, however, the case for the remainder of the columns after excluding target words containing both N and V roots in column 5. [b] Words containing both a V root and an AUX root will be counted as V in this table. Words containing only an AUX root but no V root are counted as auxiliaries. Words containing both a V root and an AUX root will still only enter as one observation into the statistical analyses. (AU = annotation unit)

Table S28: Summary statistics of pause probability by language

Language	Variant	All pauses				Silent pauses				Filled pauses			
		Count	Est. prob.	Conf. int. lower	Conf. int. higher	Count	Est. prob.	Conf. int. lower	Conf. int. higher	Count	Est. prob.	Conf. int. lower	Conf. int. higher
Baure		1 022	0.1606	0.1517	0.1698	934	0.1468	0.1382	0.1556	115	0.0181	0.0150	0.0216
Bora		4 093	0.3074	0.2996	0.3153	4 069	0.3056	0.2978	0.3.35	57	0.0043	0.0033	0.0055
Chintang		1 701	0.1385	0.1325	0.1447	1 531	0.1246	0.1189	0.1306	197	0.0160	0.0139	0.0184
Dutch		1 301	0.1282	0.1218	0.1348	890	0.0877	0.0823	0.0933	459	0.0452	0.0413	0.0494
English	w/out AUX	3 165	0.1731	0.1677	0.1786	2 868	0.1569	0.1516	0.1622	381	0.0208	0.0188	0.0230
	with AUX	3 433	0.1704	0.1653	0.1756	3 114	0.1545	0.1496	0.1596	405	0.0201	0.0182	0.0221
Even	w/out AUX	5 139	0.3020	0.2951	0.3089	4 572	0.2687	0.2621	0.2754	797	0.0468	0.0437	0.0501
	with AUX	5 255	0.2988	0.2920	0.3056	4 673	0.2657	0.2592	0.2723	816	0.0464	0.0433	0.0496
Hoocąk	w/out AUX	3 095	0.3507	0.3408	0.3607	2 952	0.3345	0.3247	0.3444	162	0.0184	0.0157	0.0213
	with AUX	3 357	0.3271	0.3181	0.3362	3 206	0.3124	0.3035	0.3214	170	0.0166	0.0142	0.0191
N‖ng		867	0.0901	0.0846	0.0960	863	0.0897	0.0841	0.0956	5	0.0005	0.0001	0.0012
Texistepec	w/out AUX	896	0.1181	0.1110	0.1255	887	0.1169	0.1099	0.1243	13	0.0017	0.0009	0.0029
	with AUX	1 071	0.1198	0.1132	0.1267	1 057	0.1182	0.1117	0.1250	18	0.0020	0.0013	0.0031
Overall	**w/out AUX**	**21 279**	**0.2057**	**0.2032**	**0.2082**	**19 566**	**0.1892**	**0.1868**	**0.1916**	**2 186**	**0.0211**	**0.0202**	**0.0221**
	with AUX	**22 100**	**0.2034**	**0.2010**	**0.2058**	**20 337**	**0.1871**	**0.1848**	**0.1895**	**2 242**	**0.0206**	**0.0198**	**0.0215**

Table S29: Probability of (silent and/or filled) pauses depending on word class of target word for individual languages

Language	Word Class	Probability of (silent and/or filled) pause				Comparison		
		Count	Est. prob.	Conf. int. lower	Conf. int. upper	Contrast	Difference	In-/Decrease
Baure	N	427	0.1759	0.1612	0.1915	N vs. V	0.0248	16.41%
	V	595	0.1511	0.1402	0.1625			
Bora	N	2 039	0.3703	0.3576	0.3831	N vs. V	0.1073	40.80%
	V	2 054	0.2630	0.2534	0.2729			
Chintang	N	889	0.1799	0.1694	0.1908	N vs. V	0.0693	62.66%
	V	812	0.1106	0.1036	0.1179			
Dutch	N	544	0.1412	0.1305	0.1524	N vs. V	0.0209	17.37%
	V	757	0.1203	0.1124	0.1285			
English	N	1 128	0.1501	0.1421	0.1583	N vs. V	-0.0391	-20.67%
	V	2 037	0.1892	0.1819	0.1967			
	AUX	268	0.1435	0.1282	0.1600	N vs. V + AUX	-0.0324	-17.75%
	V + AUX	2 305	0.1825	0.1758	0.1893			
Even	N	3 125	0.3960	0.3852	0.4068	N vs. V	0.1753	79.43%
	V	2 014	0.2207	0.2123	0.2293			
	AUX	116	0.2032	0.1716	0.2376	N vs. V + AUX	0.1763	80.25%
	V + AUX	2 130	0.2197	0.2115	0.2280			
Hoocąk	N	1 479	0.4983	0.4803	0.5163	N vs. V	0.2287	84.83%
	V	1 702	0.2696	0.2587	0.2806			
	AUX	176	0.1794	0.1563	0.2044	N vs. V + AUX	0.2408	93.51%
	V + AUX	1 878	0.2575	0.2475	0.2676			
N‖ng	N	446	0.1356	0.1242	0.1476	N vs. V	0.0691	103.91%
	V	421	0.0665	0.0605	0.0728			
Texistepec	N	415	0.1742	0.1594	0.1898	N vs. V	0.0818	88.53%
	V	481	0.0924	0.0847	0.1005			
	AUX	175	0.1294	0.1123	0.1481	N vs. V + AUX	0.0742	74.20%
	V + AUX	656	0.1000	0.0929	0.1075			
Overall	**N**	**10 492**	**0.2573**	**0.2531**	**0.2615**	**N vs. V**	**0.0850**	**49.33%**
	V	**10 873**	**0.1723**	**0.1694**	**0.1752**			
	AUX	**735**	**0.1541**	**0.1440**	**0.1645**	**N vs. V + AUX**	**0.0863**	**50.47%**
	V + AUX	**11 608**	**0.1710**	**0.1681**	**0.1738**			

Table S30: Probability of silent pauses depending on word class of target word for individual languages

Language	word class	Probability of silent pause				Comparison		
		Count	Est. prob.	Conf. int. lower	Conf. int. upper	Contrast	Difference	In-/Decrease
Baure	N	383	0.1578	0.1437	0.1727	N vs. V	0.0178	12.71%
	V	551	0.1400	0.1294	0.1510			
Bora	N	2 028	0.3683	0.3556	0.3811	N vs. V	0.1069	40.90%
	V	2 041	0.2614	0.2517	0.2712			
Chintang	N	806	0.1631	0.1530	0.1736	N vs. V	0.0643	65.08%
	V	725	0.0988	0.0921	0.1057			
Dutch	N	351	0.0911	0.0823	0.1005	N vs. V	0.0055	6.43%
	V	539	0.0856	0.0789	0.0927			
English	N	991	0.1319	0.1244	0.1397	N vs. V	-0.0424	-24.33%
	V	1 877	0.1743	0.1673	0.1816			
	AUX	246	0.1318	0.1170	0.1477	N vs. V + AUX	-0.0362	-21.53%
	V + AUX	2 123	0.1681	0.1616	0.1746			
Even	N	2 756	0.3492	0.3387	0.3598	N vs. V	0.1502	75.48%
	V	1 816	0.1990	0.1909	0.2073			
	AUX	101	0.1769	0.1472	0.2098	N vs. V + AUX	0.1515	76.63%
	V + AUX	1 917	0.1977	0.1899	0.2057			
Hoocąk	N	1 400	0.4717	0.4538	0.4897	N vs. V	0.2131	82.41%
	V	1 633	0.2586	0.2479	0.2695			
	AUX	173	0.1764	0.1534	0.2011	N vs. V + AUX	0.2241	90.51%
	V + AUX	1 806	0.2476	0.2378	0.2576			
N‖ng	N	444	0.1350	0.1236	0.1470	N vs. V	0.0688	103.93%
	V	419	0.0662	0.0603	0.0725			
Texistepec	N	412	0.1730	0.1582	0.1885	N vs. V	0.0817	89.49%
	V	475	0.0913	0.0837	0.0993			
	AUX	170	0.1257	0.1088	0.1442	N vs. V + AUX	0.0746	75.81%
	V + AUX	645	0.0984	0.0913	0.1057			
Overall	**N**	**9 571**	**0.2347**	**0.2307**	**0.2388**	**N vs. V**	**0.0751**	**47.06%**
	V	**10 076**	**0.1596**	**0.1568**	**0.1625**			
	AUX	**690**	**0.1446**	**0.1349**	**0.1548**	**N vs. V + AUX**	**0.0761**	**47.98%**
	V + AUX	**10 766**	**0.1586**	**0.1558**	**0.1613**			

Table S31: Probability of filled pauses depending on word class of target word for individual languages

Language	Word Class	Probability of filled pause				Comparison		
		Count	Est. prob.	Conf. int. lower	Conf. int. upper	Contrast	Difference	In-/Decrease
Baure	N	57	0.0235	0.0180	0.0301	N vs. V	0.0088	59.86%
	V	58	0.0147	0.0113	0.0189			
Bora	N	32	0.0058	0.0041	0.0081	N vs. V	0.0026	81.25%
	V	25	0.0032	0.0021	0.0046			
Chintang	N	101	0.0204	0.0168	0.0247	N vs. V	0.0073	55.73%
	V	96	0.0131	0.0107	0.0159			
Dutch	N	217	0.0563	0.0494	0.0639	N vs. V	0.0179	46.61%
	V	242	0.0384	0.0339	0.0434			
English	N	176	0.0234	0.0202	0.0270	N vs. V	0.0044	23.16%
	V	205	0.0190	0.0166	0.0217			
	AUX	24	0.0129	0.0084	0.0188	N vs. V + AUX	0.0053	29.28%
	V + AUX	229	0.0181	0.0159	0.0206			
Even	N	513	0.0650	0.0597	0.0706	N vs. V	0.0339	109.00%
	V	284	0.0311	0.0277	0.0349			
	AUX	19	0.0333	0.0208	0.0505	N vs. V + AUX	0.0337	107.67%
	V + AUX	303	0.0313	0.0279	0.0349			
Hoocąk	N	91	0.0307	0.0249	0.0374	N vs. V	0.0187	155.83%
	V	76	0.0120	0.0096	0.0150			
	AUX	3	0.0031	0.0008	0.0083	N vs. V + AUX	0.0199	184.26%
	V + AUX	79	0.0108	0.0086	0.0134			
N‖ng	N	3	0.0009	0.0003	0.0025	N vs. V	0.0006	200.00%
	V	2	0.0003	0.0001	0.0010			
Texistepec	N	5	0.0021	0.0008	0.0046	N vs. V	0.0006	40.00%
	V	8	0.0015	0.0007	0.0029			
	AUX	5	0.0037	0.0014	0.0082	N vs. V + AUX	0.0001	5.00%
	V + AUX	13	0.0020	0.0011	0.0033			
Overall	**N**	**1 195**	**0.0293**	**0.0277**	**0.0309**	**N vs. V**	**0.0135**	**85.44%**
	V	**996**	**0.0158**	**0.0148**	**0.0168**			
	AUX	**51**	**0.0107**	**0.0080**	**0.0139**	**N vs. V + AUX**	**0.0139**	**90.26%**
	V + AUX	**1 047**	**0.0154**	**0.0145**	**0.0163**			

Table S32: Comparison of models for predicting pause probability

Step	Fixed effects structure	Added	AIC	χ^2	df	*p*	Sig.
1 (initial)	word class + language + word class:language		94 736				
2a	word class + word length + language + word class:language	word length	94 614	124.20	1	<0.001	***
2b	word class +position + language + word class:language	position	94 724	14.31	1	<0.001	***
3a (final)	word class + word length + position + language + word class:language	position	94 607	9.57	1	0.002	**
3b	word class + word length + word class:word length + language + word class:language	word class:word length	94 616	0.64	1	0.424	n.s.
4a	word class + word length + position + word class:word length + language + word class:language	word class:word length	94 608	0.54	1	0.464	n.s.
4b	word class + word length + position + word class:position + language + word class:language	word class:position	94 607	1.43	1	0.232	n.s.
4c	word class + word length + position + word length:position + language + word class:language	word length:position	94 609	0.00	1	0.993	n.s.

Notes: Alternative models are given the same step number followed by different letters. The model structure chosen as the next step is marked in bold.

Table S33: Model predicting pause probability in all nine languages

Fixed effect	Coefficient	Std. Error	Std. Coefficient	χ^2	df	*p*	sig.	p (BH)	sig. (BH)
(intercept)	-1.1683	0.1366							
word class = verb	-0.5448	0.0460	-0.6586	1 213.90	9	<0.001	***	<0.001	***
word length	0.1126	0.0101	0.2791	119.46	1	<0.001	***	<0.001	***
position	0.0948	0.0302	0.0696	9.57	1	0.002	**	0.003	**
language = Baure	-0.3445	0.2405	-0.2048	548.92	16	<0.001	***	<0.001	***
language = Chintang	-0.5691	0.1701	-0.4554						
language = Dutch	-1.0376	0.1865	-0.7635						
language = English	-0.8220	0.1599	-0.7757						
language = Even	0.7380	0.1727	0.6768						
language = Hoocąk	1.1814	0.1820	0.8164						
language = N‖ng	-0.7615	0.2423	-0.5471						
language = Texistepec	-0.5607	0.5042	-0.3616						
word class = verb : language = Baure	0.3349	0.0921	0.1585	401.26	8	<0.001	***	<0.001	***
word class = verb : language = Chintang	-0.1086	0.0763	-0.0690						
word class = verb : language = Dutch	0.4177	0.0896	0.2470						
word class = verb : language = English	0.8092	0.0762	0.6113						
word class = verb : language = Even	-0.3475	0.0620	-0.2438						
word class = verb : language = Hoocąk	-0.6569	0.0757	-0.3756						
word class = verb : language = N‖ng	-0.2241	0.1020	-0.1329						
word class = verb : language = Texistepec	-0.2003	0.0966	-0.1083						
Random effect	**Groups**	**Std. Dev.**		**χ^2**	**df**	***p***	**sig.**	**p (BH)**	**sig. (BH)**
(1|speaker)	311	0.4560		371.75	1	<0.001	***	<0.001	***
(1|text)	346	0.3642		378.99	1	<0.001	***	<0.001	***
(1|word type)	34 972	0.5430		933.71	1	<0.001	***	<0.001	***

Table S34: Model predicting pause probability in Baure

Fixed effect	Coefficient	Std. Error	Std. Coefficient	χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(intercept)	-1.9587	0.1600							
word class = verb	-0.0970	0.0848	-0.1284	1.27	1	0.260	n.s.	0.310	n.s.
word length	-0.0021	0.0458	-0.0052	0.00	1	0.963	n.s.	0.982	n.s.
position	0.6018	0.1254	0.4941	22.64	1	<0.001	***	<0.001	***
Random effect	**Groups**	**Std. Dev.**		**χ^2**	**df**	***p***	**sig.**	***p* (BH)**	**sig. (BH)**
(1|speaker)	10	0.0000		0.00	1	0.999	n.s.	1.000	n.s.
(1|text)	37	0.6093		188.67	1	<0.001	***	<0.001	***
(1|word type)	3 084	0.4960		23.11	1	<0.001	***	<0.001	***

Table S35: Model predicting pause probability in Bora

Fixed effect	Coefficient	Std. Error	Std. Coefficient	χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(intercept)	-0.9549	0.1786							
word class = verb	-0.5703	0.0489	-0.6087	134.53	1	<0.001	***	<0.001	***
word length	0.1880	0.0249	0.3868	54.81	1	<0.001	***	<0.001	***
position	-0.0947	0.0724	-0.0606	1.68	1	0.194	n.s.	0.233	n.s.
Random effect	**Groups**	**Std. Dev.**		**χ^2**	**df**	***p***	**sig.**	***p* (BH)**	**sig. (BH)**
(1|speaker)	43	0.2218		0.65	1	0.420	n.s.	0.475	n.s.
(1|text)	37	0.6304		123.41	1	<0.001	***	<0.001	***
(1|word type)	6 273	0.6113		170.63	1	<0.001	***	<0.001	***

Table S36: Model predicting pause probability in Chintang

Fixed effect	Coefficient	Std. Error	Std. Coefficient	χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(intercept)	-1.7738	0.1611							
word class = verb	-0.6136	0.0643	-0.8711	88.55	1	<0.001	***	<0.001	***
word length	0.0684	0.0314	0.1946	4.48	1	0.034	*	0.046	*
position	0.0814	0.1057	0.0666	0.57	1	0.450	n.s.	0.507	n.s.
Random effect	**Groups**	**Std. Dev.**		**χ^2**	**df**	***p***	**sig.**	***p* (BH)**	**sig. (BH)**
(1|speaker)	68	0.8344		55.53	1	<0.001	***	<0.001	***
(1|text)	40	0.4132		11.07	1	<0.001	***	0.001	**
(1|word type)	5 120	0.4476		29.52	1	<0.001	***	<0.001	***

Table S37: Model predicting pause probability in Dutch

Fixed effect	Coefficient	Std. Error	Std. Coefficient	χ^2	df	*p*	sig.	p (BH)	sig. (BH)
(intercept)	-2.0995	0.1263							
word class = verb	-0.1229	0.0710	-0.1784	2.96	1	0.085	n.s.	0.104	n.s.
word length	0.0801	0.0282	0.2824	7.81	1	0.005	**	0.008	**
position	0.0978	0.1036	0.0899	0.88	1	0.349	n.s.	"0.406	n.s.
Random effect	**Groups**	**Std. Dev.**		**χ^2**	**df**	***p***	**sig.**	**p (BH)**	**sig. (BH)**
(1|speaker)	42	0.4019		58.20	1	<0.001	***	<0.001	***
(1|text)	17	0.2222		1.15	1	0.284	n.s.	0.334	n.s.
(1|word type)	2 589	0.2579		3.25	1	0.071	n.s.	0.089	n.s.

Table S38: Model predicting pause probability in English

Fixed effect	Coefficient	Std. Error	Std. Coefficient	χ^2	df	*p*	sig.	p (BH)	sig. (BH)
(intercept)	-1.7336	0.0821							
word class = verb	0.2247	0.0564	0.2922	15.42	1	<0.001	***	<0.001	***
word length	0.0865	0.0254	0.2445	11.09	1	<0.001	***	0.001	**
position	-0.1809	0.0726	-0.1401	6.11	1	0.013	*	0.018	*
Random effect	**Groups**	**Std. Dev.**		**χ^2**	**df**	***p***	**sig.**	**p (BH)**	**sig. (BH)**
(1|speaker)	80	0.3646		93.01	1	<0.001	***	<0.001	***
(1|text)	47	0.1500		3.54	1	0.060	n.s.	0.076	n.s.
(1|word type)	2 941	0.3805		231.44	1	<0.001	***	<0.001	***

Table S39: Model predicting pause probability in Even

Fixed effect	Coefficient	Std. Error	Std. Coefficient	χ^2	df	*p*	sig.	p (BH)	sig. (BH)
(intercept)	-0.5166	0.0910							
word class = verb	-0.9113	0.0440	-0.9898	424.28	1	<0.001	***	<0.001	***
word length	0.1061	0.0233	0.2107	20.08	1	<0.001	***	<0.001	***
position	0.2395	0.0654	0.1529	13.29	1	<0.001	***	<0.001	***
Random effect	**Groups**	**Std. Dev.**		**χ^2**	**df**	***p***	**sig.**	**p (BH)**	**sig. (BH)**
(1|speaker)	32	0.3475		30.05	1	<0.001	***	<0.001	***
(1|text)	67	0.2283		39.66	1	<0.001	***	<0.001	***
(1|word type)	7 357	0.6046		296.69	1	<0.001	***	<0.001	***

Table S40: Model predicting pause probability in Hoocąk

Fixed effect	Coefficient	Std. Error	Std. Coefficient	χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(intercept)	0.0722	0.1195							
word class = verb	-1.2180	0.0678	-1.2059	341.25	1	<0.001	***	<0.001	***
word length	0.1493	0.0289	0.3177	26.21	1	<0.001	***	<0.001	***
position	-0.0260	0.0922	-0.0160	0.08	1	0.780	n.s.	0.828	n.s.
Random effect	**Groups**	**Std. Dev.**		**χ^2**	**df**	***p***	**sig.**	***p* (BH)**	**sig. (BH)**
(1|speaker)	28	0.3821		16.09	1	<0.001	***	<0.001	***
(1|text)	62	0.3314		40.29	1	<0.001	***	<0.001	***
(1|word type)	4 231	0.6286		108.22	1	<0.001	***	<0.001	***

Table S41: Model predicting pause probability in N‖ng

Fixed effect	Coefficient	Std. Error	Std. Coefficient	χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(intercept)	-2.0622	0.3137							
word class = verb	-0.7622	0.0937	-1.2626	61.37	1	<0.001	***	<0.001	***
word length	0.0944	0.0498	0.2901	3.33	1	0.068	n.s.	0.086	n.s.
position	0.2893	0.1463	0.2801	3.77	1	0.052	n.s.	0.067	n.s.
Random effect	**Groups**	**Std. Dev.**		**χ^2**	**df**	***p***	**sig.**	***p* (BH)**	**sig. (BH)**
(1|speaker)	7	0.7005		117.33	1	<0.001	***	<0.001	***
(1|text)	33	0.3531		16.84	1	<0.001	***	<0.001	***
(1|word type)	1 162	0.5338		34.40	1	<0.001	***	<0.001	***

Table S42: Model predicting pause probability in Texistepec

Fixed effect	Coefficient	Std. Error	Std. Coefficient	χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(intercept)	-2.2171	0.1388							
word class = verb	-0.7131	0.0941	-1.0256	54.61	1	<0.001	***	<0.001	***
word length	0.1153	0.0416	0.3786	7.37	1	0.007	**	0.010	*
position	0.6141	0.1444	0.5219	18.47	1	<0.001	***	<0.001	***
Random effect	**Groups**	**Std. Dev.**		**χ^2**	**df**	***p***	**sig.**	***p* (BH)**	**sig. (BH)**
(1|text)	6	0.000		0.00	1	1.000	n.s.	1.000	n.s.
(1|word type)	2 215	0.6838		44.12	1	<0.001	***	<0.001	***

Table S43: Comparison of models for predicting pause probability (incl. auxiliaries) in all nine languages

Step	Fixed effects structure	Added	AIC	χ^2	df	*p*	sig.
1 (initial)	word class + language + word class:language		99 209				
2a	word class + word length + language + word class:language	word length	99 084	126.92	1	<0.001	***
2b	word class +position + language + word class:language	position	99 194	16.88	1	<0.001	***
3a (final)	word class + word length + position + language + word class:language	position	99 075	11.38	1	<0.001	***
3b	word class + word length + word class:word length + language + word class:language	word class:word length	99 086	0.44	1	0.506	n.s.
4a	word class + word length + position + word class:word length + language + word class:language	word class:word length	99 077	0.34	1	0.560	n.s.
4b	word class + word length + position + word class:position + language + word class:language	word class:position	99 076	1.04	1	0.308	n.s.
4c	word class + word length + position + word length:position + language + word class:language	word length:position	99 076	0.50	1	0.479	n.s.

Notes: Alternative models are given the same step number followed by different letters. The model structure chosen as the next step is marked in bold.

Table S44: Model predicting pause probability (incl. auxiliaries) in all nine languages

Fixed effect	Coefficient	Std. Error	Std. Coefficient	χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(intercept)	-1.1696	0.1359							
word class = verb	-0.5454	0.0461	-0.6560	1312.70	9	<0.001	***	<0.001	***
word length	0.1117	0.0099	0.2830	121.42	1	<0.001	***	<0.001	***
position	0.1015	0.0297	0.0747	11.38	1	<0.001	***	0.001	**
language = Baure	-0.3480	0.2392	-0.2030	592.65	16	<0.001	***	<0.001	***
language = Chintang	-0.5712	0.1695	-0.4493						
language = Dutch	-1.0399	0.1859	-0.7517						
language = English	-0.8268	0.1592	-0.7983						
language = Even	0.7402	0.1718	0.6773						
language = Hoocąk	1.1792	0.1809	0.8568						
language = N‖ng	-0.7660	0.2398	-0.5406						
language = Texistepec	-0.5464	0.4987	-0.3730						
word class = verb : language = Baure	0.3362	0.0922	0.1561	451.80	8	<0.001	***	<0.001	***
word class = verb : language = Chintang	-0.1094	0.0764	-0.0682						
word class = verb : language = Dutch	0.4177	0.0897	0.2424						
word class = verb : language = English	0.8076	0.0756	0.6431						
word class = verb : language = Even	-0.3463	0.0618	-0.2453						
word class = verb : language = Hoocąk	-0.7665	0.0743	-0.4766						
word class = verb : language = N‖ng	-0.2235	0.1021	-0.1300						
word class = verb : language = Texistepec	-0.1352	0.0954	-0.0800						
Random effect	**Groups**	**Std. Dev.**		**χ^2**	**df**	***p***	**sig.**	***p* (BH)**	**sig. (BH)**
(1|speaker)	311	0.4511		369.63	1	<0.001	***	<0.001	***
(1|text)	346	0.3669		395.13	1	<0.001	***	<0.001	***
(1|word type)	35 821	0.5474		979.87	1	<0.001	***	<0.001	***

Table S45: Model predicting pause probability (incl. auxiliaries) in English

Fixed effect	Coefficient	Std. Error	Std. Coefficient	χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(intercept)	-1.7760	0.0802							
word class = verb	0.2260	0.0552	0.2906	16.34	1	<0.001	***	<0.001	***
word length	0.0896	0.0246	0.2652	12.62	1	<0.001	***	<0.001	***
position	-0.1215	0.0699	-0.0945	2.98	1	0.084	n.s.	0.103	n.s.
Random effect	**Groups**	**Std. Dev.**		**χ^2**	**df**	***p***	**sig.**	***p* (BH)**	**sig. (BH)**
(1|speaker)	80	0.3598		89.90	1	<0.001	***	<0.001	***
(1|text)	47	0.1486		3.25	1	0.071	n.s.	0.089	n.s.
(1|word type)	2 952	0.3639		240.86	1	<0.001	***	<0.001	***

Table S46: Model predicting pause probability (incl. auxiliaries) in Even

Fixed effect	Coefficient	Std. Error	Std. Coefficient	χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(intercept)	-0.5060	0.0915							
word class = verb	-0.9086	0.0433	-0.9873	434.16	1	<0.001	***	<0.001	***
word length	0.1069	0.0228	0.2144	21.24	1	<0.001	***	<0.001	***
position	0.2296	0.0646	0.1465	12.52	1	<0.001	***	<0.001	***
Random effect	**Groups**	**Std. Dev.**		**χ^2**	**df**	***p***	**sig.**	***p* (BH)**	**sig. (BH)**
(1|speaker)	32	0.3518		30.28	1	<0.001	***	<0.001	***
(1|text)	67	0.2334		43.52	1	<0.001	***	<0.001	***
(1|word type)	7 428	0.5982		297.96	1	<0.001	***	<0.001	***

Table S47: Model predicting pause probability (incl. auxiliaries) in Hoocąk

Fixed effect	Coefficient	Std. Error	Std. Coefficient	χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(intercept)	0.0804	0.1180							
word class = verb	-1.3343	0.0667	-1.2894	423.52	1	<0.001	***	<0.001	***
word length	0.1552	0.0275	0.3373	31.42	1	<0.001	***	<0.001	***
position	-0.0399	0.0868	-0.0252	0.21	1	0.650	n.s.	0.703	n.s.
Random effect	**Groups**	**Std. Dev.**		χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(1|speaker)	28	0.3735		19.26	1	<0.001	***	<0.001	***
(1|text)	62	0.3419		52.77	1	<0.001	***	<0.001	***
(1|word type)	4 968	0.6527		125.98	1	<0.001	***	<0.001	***

Table S48: Model predicting pause probability (incl. auxiliaries) in Texistepec

Fixed effect	Coefficient	Std. Error	Std. Coefficient	χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(intercept)	-2.2601	0.1364							
word class = verb	-0.6477	0.0956	-0.8818	44.67	1	<0.001	***	<0.001	***
word length	0.0825	0.0414	0.2767	3.87	1	0.049	*	0.064	n.s.
position	0.6617	0.1374	0.5599	23.74	1	<0.001	***	<0.001	***
Random effect	**Groups**	**Std. Dev.**		χ^2	df	*p*	sig.	*p* (BH)	sig. (BH)
(1|text)	6	0.0000		0.00	1	0.999	n.s.	1.000	n.s.
(1|word type)	2 245	0.7583		81.82	1	<0.001	***	<0.001	***

Table S49: Size of word class effect on pause probability

language	auxiliaries excluded				auxiliaries included			
	pause probability		probability ratio	odds ratio[b]	pause probability		probability ratio	odds ratio[b]
	before nouns	before verbs			before nouns	before verbs		
Baure	0.1760	0.1624	1.0837	1.1018				
Bora	0.2841	0.1832	1.5508	1.7687				
Chintang	0.1573	0.0918	1.7135	1.8471				
Dutch	0.1207	0.1082	1.1155	1.1307				
English	0.1428	0.1726	0.8273	0.7988	0.1409	0.1705	0.8264	0.7977
Even	0.4232	0.2278	1.8578	2.4876	0.4236	0.2285	1.8538	2.4808
Hoocąk	0.5325	0.2521	2.1123	3.3807	0.5346	0.2322	2.3023	3.7979
N‖ng	0.1445	0.0731	1.9767	2.1432				
Texistepec	0.1508	0.0800	1.8850	2.0404	0.1427	0.0801	1.7815	1.9113
grand mean[a]			**1.5692**	**1.7015[c]**			**1.5782**	**1.7104[c]**

Notes: [a] We averaged across all nine languages, both when excluding and when including auxiliaries. We did, however, not repeat unchanged results in the *auxiliaries included* columns. [b] Odds ratios were calculated by inverting the exponentiated coefficients of the word class factor from the logistic regression models for individual languages. [c] The mean of the odds ratios was calculated on the log scale and subsequently exponentiated.

References

1. Danielsen S, Riedel F, Admiraal F, Terhart L (2009) *Baure Documentation* (The Language Archive, Nijmegen) Available at: https://hdl.handle.net/1839/00-0000-0000-000D-8382-B.
2. Seifart F (2009) Bora documentation. *A Multimedia Documentation of the Languages of the People of the Center. Online Publication of Transcribed and Translated Bora, Ocaina, Nonuya, Resígaro, and Witoto Audio and Video Recordings with Linguistic and Ethnographic Annotations and Descriptions*, eds Seifart F, Fagua D, Gasché J, Echeverri JA (The Language Archive, Nijmegen). Available at: https://hdl.handle.net/1839/00-0000-0000-0008-38E5-2.
3. Bickel B, et al. (2011) *Audiovisual corpus of the Chintang language, including a longitudinal corpus of language acquisition by six children: ca. 650,000 words transcribed and translated, of which ca. 450,000 glossed, plus paradigm sets and grammar sketches, ethnographic descriptions, photographs* (The Language Archive, Nijmegen) Available at: https://hdl.handle.net/1839/00-0000-0000-0005-6F41-C.
4. CGN-consortium, Language and Speech Nijmegen, ELIS Gent (2003) *Corpus Gesproken Nederlands* (Nederlandse Taalunie, Nijmegen).
5. Calhoun S, et al. (2009) *NXT Switchboard Annotations LDC2009T26* (Linguistic Data Consortium, Philadelphia).
6. Pakendorf B, Matić D, Aralova N, Lavrillier A (2010) *Documentation of the dialectal and cultural diversity among Évens in Siberia* (DOBES, MPIP, MPI-EVA, Nijmegen, Leipzig). Available at: https://hdl.handle.net/1839/00-0000-0000-000F-2DFE-0.
7. Hartmann I (2013) *Hoocąk Corpus* (MPI-EVA, Leipzig).
8. Güldemann T, Ernszt M, Siegmund S, Witzlack-Makarevich A (2011) *A Text documentation of N|uu* (ELAR, London). Available at: https://elar.soas.ac.uk/Collection/MPI194591.
9. Wichmann S (1996) *Cuentos y colorados en popoluca de Texistepec* (C.A. Reitzel, Copenhagen). Available at: https://www.ailla.utexas.org/islandora/object/ailla%3A124443.
10. Simons GF, Fennig CD (2017) *Ethnologue: Languages of the World, Twentieth edition* (SIL International, Dallas, Texas) Available at: http://www.ethnologue.com.
11. Hammarström H, Forkel R, Haspelmath M (2017) *Glottolog 3.0* (Max Planck Institute for the Science of Human History, Jena) Available at: http://glottolog.org/.