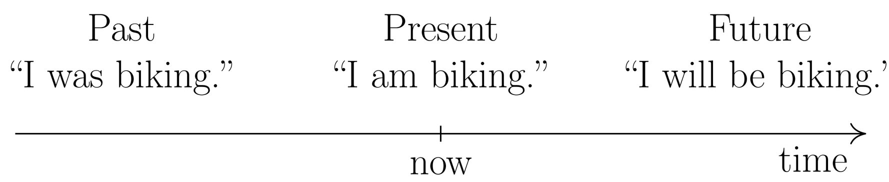
Aspectual classes as lexically-conditioned predictors of aspectual choice

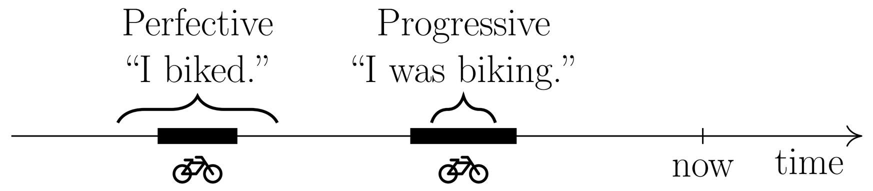
Laurestine Bradford
McGill University
laurestine.bradford@mail.mcgill.ca

Aspectual Choice and Aspectual Class

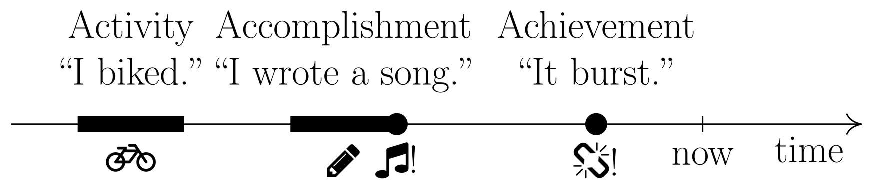
Tense: Location of an event in time.



Grammatical Aspect / Aspectual Choice: Internal or external perspective on the event.



Lexical Aspect / Aspectual Class: Conceptualization of the event's shape in time.



Aspectual class has been used to explain differences in grammaticality and felicity (distinctions based on Smith, 1991):

(1) (a) I'm running. (Event)
(b) # I'm knowing French. (State)

(2) (a) # I swatted a fly for ten minutes. (Punctual)
(b) I danced for ten minutes. (Durative)

(3) (a) I ran to the park (in/#for) two minutes. (Telic)
(b) I read (for/#in) three hours. (Atelic)

Core Intuition: Different verbs describe situations that are different shapes in time. This is reflected in their compatibility, and thus in their distribution, with temporal modifiers and morphology. (c.f. Nerbonne & Van de Cruys, 2009)

Theories of Aspectual Class get Messy

Verbs may fall into multiple classes, or switch class.

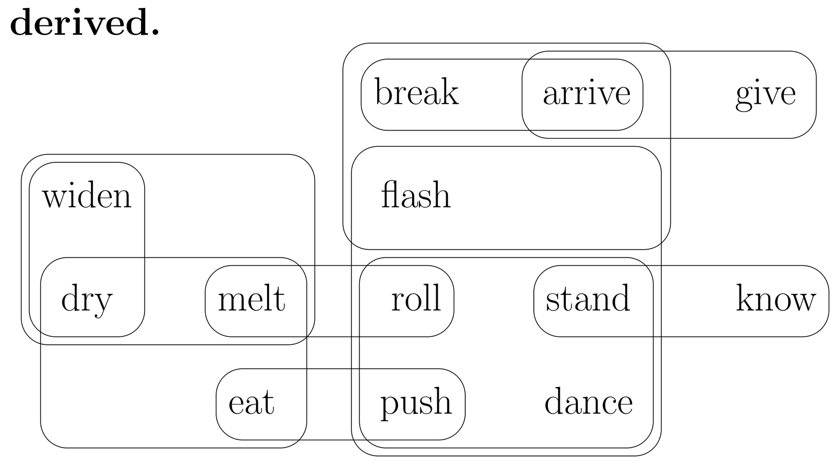
(4) The tank emptied (for/in) two minutes. (Kennedy & Levin, 2008)

(5) I am liking you #(more and more). (Zucchi, 1998)

(6) The light flashed (# once) for two hours.

(7) Susan (was winning for/won in) two minutes.

Disagreement as to which distinctions are essential vs. derived.



Bubbles showing selected aspectual classes hypothesized by Bach (1998); Kennedy and Levin (2008); Levin (1999); Maienborn (2005); Ramchand (2008); and Smith (1991).

Proposal: Aspectual class is precisely the lexical information which contributes to aspectual choice. No more, no less.

Dataset

GUM English corpus (Zeldes, 2017) 2024-10-19 release.

1 Path: GUM > GUM_interview_brotherhood (tokens 172 - 182)

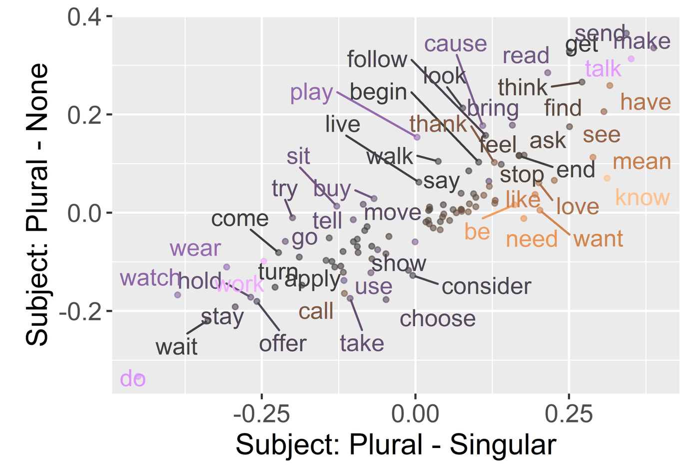
That	started	me	out	on	books	and	I	have	amassed	quite	a	few	since	then	.
DT	VVD	PP	RP	IN	NNS	CC	PP	VHP	VVN	PDT	DT	JJ	IN	RB	SENT
DT	VBD	PRP	RP	IN	NNS	CC	PRP	VBP	VBN	PDT	DT	JJ	IN	RB	.
that	start	I	out	on	book	and	I	have	amass	quite	a	few	since	then	.
PRON	VERB	PRON	ADP	ADP	NOUN	CCONJ	PRON	AUX	VERB	DET	DET	ADJ	ADP	ADV	PUNCT
DT0	VVD	PNP	AVP	PRP	NN2	CJC	PNP	VHB	VVN	AV0	AT0	DT0	CJS	AV0	PUN
_	start-ed	_	_	_	book-s	_	_	_	a-mass-ed	_	_	_	_	_	_
nsubj	root	obj	compound:prt	case	obl	cc	nsubj	aux	conj	det:predet	det	obj	case	obl	punct

Extracted indicative verbs that head clauses: 14,622 tokens.

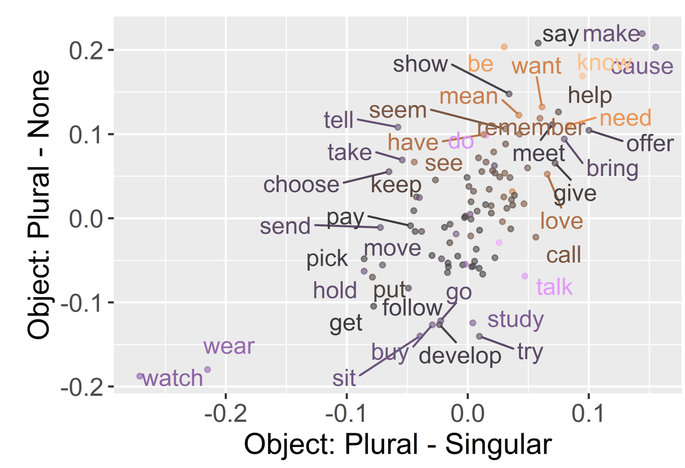
Aspect	Lemma	Perfect	Voice	Tense	Matrix	Preceding	Subject	Object	Time Span	Adverbs	Particle	Conjunction	Document	Genre
Perfective	start	FALSE	Active	Past	None	Perfective	Singular	Singular	None		out	None	GUM_int...	interview
Perfective	amass	TRUE	Active	Present	Perfective	Perfective	Singular	Singular	None		None	None	GUM_int...	interview
Perfective	think	FALSE	Active	Present	None	Perfective	Singular	None	None		None	None	GUM_int...	interview
Perfective	be	FALSE	Active	Present	None	Perfective	Plural	None	None		None	None	GUM_int...	interview
Perfective	be	FALSE	Active	Future	None	Perfective	Singular	None	None	always	None	None	GUM_int...	interview
Perfective	help	TRUE	Active	Present	None	Perfective	Singular	Singular	None		None	None	GUM_int...	interview
Progressive	read	FALSE	Active	Present	Perfective	Perfective	Plural	Plural	None		None	None	GUM_int...	interview
Perfective	have	FALSE	Active	Present	None	Progressive	Plural	None	None		None	None	GUM_int...	interview

Variables:

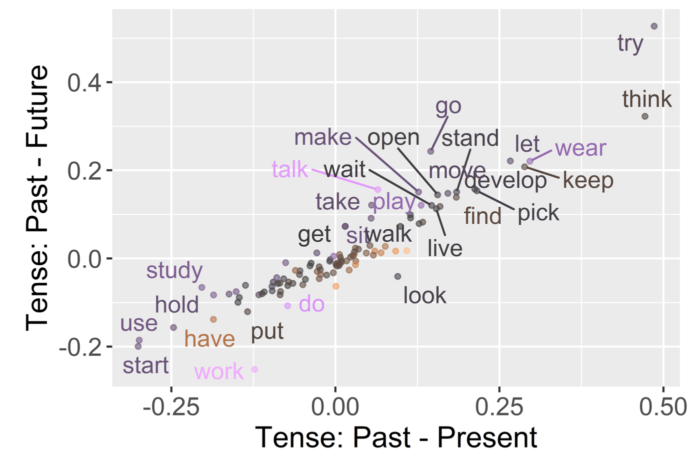
- Perfect aspect (Yes, No)
- Voice (Active, Passive)
- Tense (Past, Present, Future, None)
- Matrix verb aspect (Perfective, Progressive, None)
- Preceding verb aspect (Perfective, Progressive, None)
- Subject type (Singular, Plural, Mass, None)
- Object type (Singular, Plural, Mass, None)
- Time Span adverbial type (For, In, None)
- Adverb, Particle, Conjunction, Document, Genre (random effects)
- **Effect of Subject, Object, Tense, and Time Span vary by lemma**



Lexical effect on relationship between subject type and log-odds of progressive. **Stative verbs still pattern together. Many durative events show especially small effects of plural subjects.**



Lexical effect on relationship between object type and log-odds of progressive. **Verbs where plural objects facilitate progressive are either stative verbs or incremental verbs of creation.**



Lexical effect on relationship between tense and log-odds of progressive. **Compatibility with past progressive is a main axis of lexical variation.**

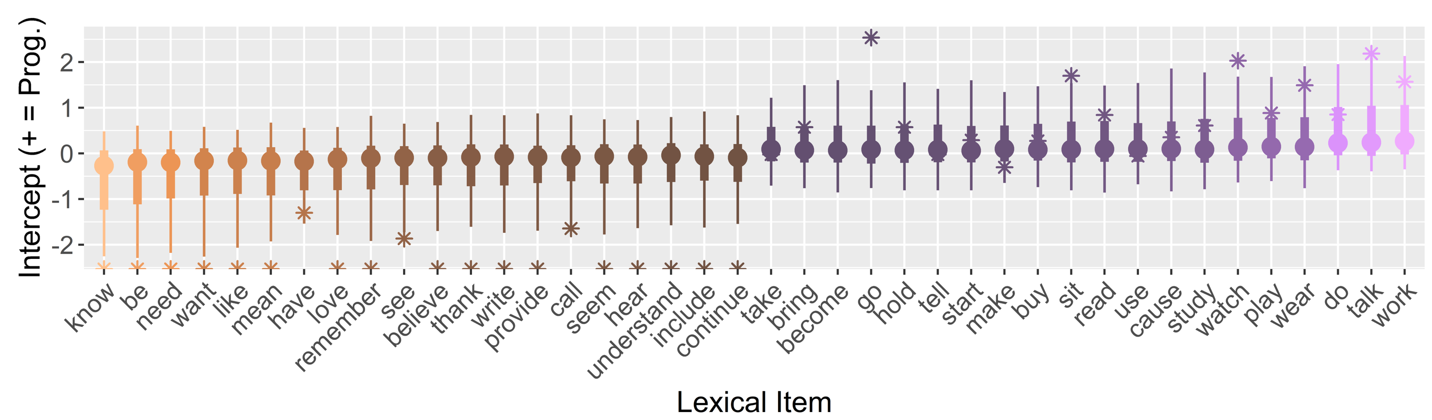
Model

Bayesian: Allows accessing inferred effects of individual lexical items, incorporating uncertainty.
Mixed-effects: Allows including both frequent and infrequent lexical items without frequent ones overpowering analysis.
Logistic Regression: Allows including other predictors known to affect aspectual behaviour to avoid confounding the lexical effects.

Fit in BRMS (Bürkner, 2017, 2018, 2021) with 4 chains of 2,500 warmup and 7,500 sampling iterations each. Priors were Normal(-2.5, 2.5) on intercept, Normal(0, 2.5) on coefficients, Exp(1) on random effect standard deviations, and LKJ(5) on correlations.

Results for Lexical Predictors

Statistically significant variation by lexical item in all predictors ($p_{\text{ROPE}} < 0.001$).
Below plots computed for only verbs with at least 25 occurrences, using a subset of 1,000 posterior samples. Colours show lexical effect on log-odds of progressive (yellow is lower).



Lexical effect on log-odds of progressive for 20 verbs with highest and lowest effects, with 66% and 95% credible intervals. Asterisks show log-odds computed from raw counts. **Stative verbs are consistently on the left. Fit model values regularize away variability in raw counts. Credible intervals are wide.**

Results for Non-Lexical Predictors

Fixed effects (intercept and selected contrasts):

Coefficient	Mean	[2.5%, 97.5%]	p_d	p_{ROPE}	Comment
Intercept	-5.90	[-8.35, -3.86]	>0.999	<0.001	Prog. infrequent
Perfect: No - Yes	0.83	[0.38, 1.30]	>0.999	<0.001	
Voice: Active - Passive	2.28	[1.60, 3.02]	>0.999	<0.001	Mostly replicates Hundt,
Tense: Present - Past	1.47	[1.10, 1.88]	>0.999	<0.001	Rautionaho, and Strobl (2020);
Tense: Past - Future	0.82	[0.06, 1.65]	0.962	0.002	dispreference for stacking auxiliaries.
Tense: Future - None	0.89	[-0.13, 1.95]	0.926	0.003	
Matrix: Prog. - Pfv.	1.39	[0.93, 1.85]	>0.999	<0.001	Replicates Rautionaho and Hundt
Matrix: Pfv. - None	0.27	[0.07, 0.47]	0.986	0.004	(2022) on priming and suggests
Preceding: Prog. - None	0.68	[-0.04, 1.42]	0.940	0.003	additional grammatical constraint on
Preceding: Prog. - Pfv.	0.85	[0.53, 1.16]	>0.999	<0.001	embedding.
Subject: Pl. - Sing.	0.32	[0.04, 0.60]	>0.999	<0.001	Expected direction, but mass
Subject: Pl. - Mass	1.32	[-0.01, 2.92]	0.949	0.002	behaves strangely. Interaction with
Subject: Sing. - None	1.50	[0.93, 2.15]	>0.999	<0.001	Voice?
Object: None - Sing.	0.37	[0.02, 0.73]	0.958	0.006	
Object: None - Pl.	0.48	[0.06, 0.92]	0.969	0.003	Reverse of expected direction.
Object: None - Mass	0.78	[-0.04, 1.76]	0.940	0.003	
Time Span: For - In	3.19	[0.88, 5.96]	0.990	<0.001	Validates methodology. Hints that
Time Span: None - In	2.68	[0.43, 5.76]	0.982	<0.001	atelicity is somehow default.

Random effects with $p_{\text{ROPE}} < 0.001$ were found for Document, Genre (c.f. Mavridou, Friedrich, Peate Sørensen, Palmer, & Pinkal, 2015), Adverb, Particle, and Conjunction.

Abbreviations:

Prog. Progressive
Pfv. Perfective
Sing. Singular
Pl. Plural
p_d Inferred probability that parameter is on the given side of 0.
p_{ROPE} Inferred probability that parameter is small enough not to matter (magnitude ≤ 0.006).
[2.5%, 97.5%] Model is 95% confident the coefficient falls in this interval, with equal chance of being less or more.

Acknowledgments

This project has been helped by comments from Morgan Sonderegger, Massimo Lipari, and Jeanne Brown.
I am supported financially by the Fonds de Recherche du Québec.

Limitations

Omission of discourse-based predictors such as foregrounding vs. backgrounding function (Hopper, 1979).

Dataset size meant that I had to restrict attention to only a subset of verbs, and that conclusions about some predictors were highly uncertain.

Variability in existing definitions of aspectual class makes it difficult to objectively assess results.

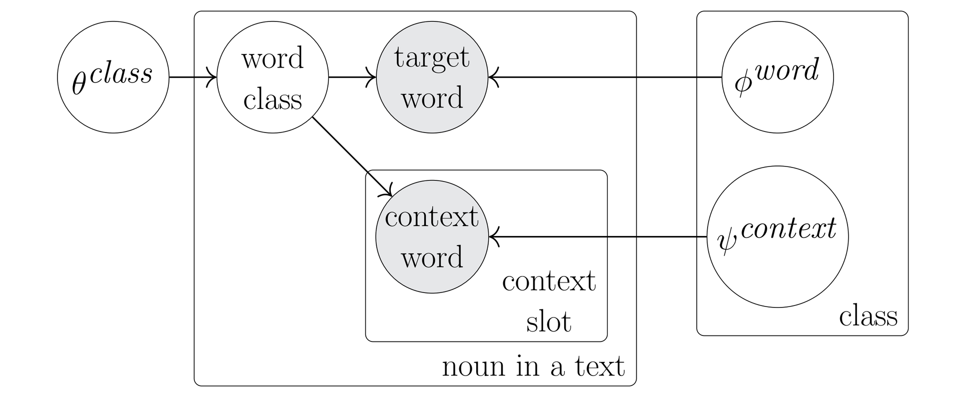
Need for manual inspection of results meant I restricted my attention to English data.

Simplicity of model likely misses important relationships between predictors.

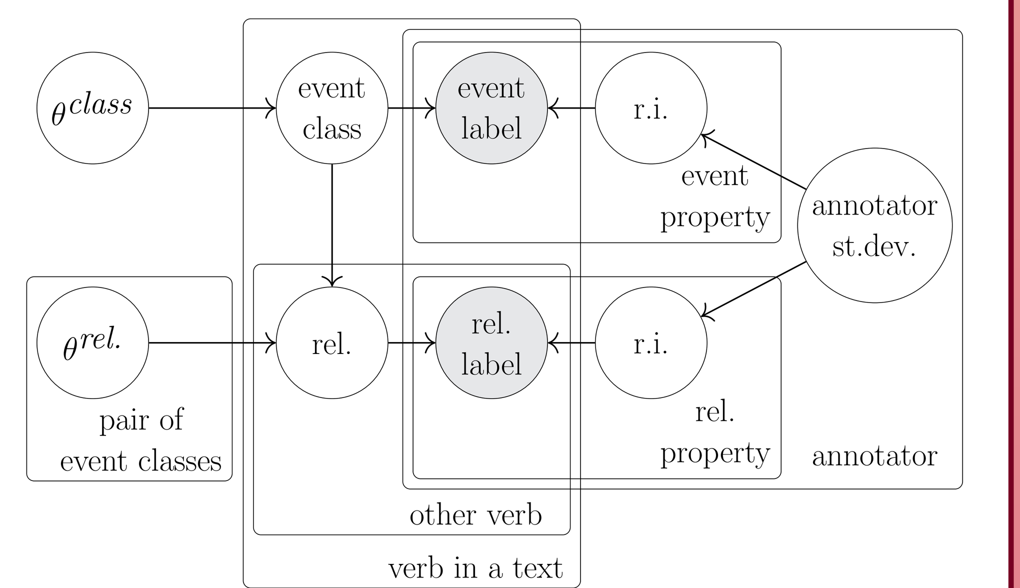
Probabilistic model is a two edged sword, as it can be confused about categorical effects.

Next Steps

Better generative model of aspectual choice process, following Frermann and Lapata (2016) or Gantt, Glass, and White (2022). In below probabilistic graphical models, circles are random variables, and rectangles show number of copies of each variable. Observed (not hidden) variables are shaded grey.



Adapted from Frermann and Lapata (2016). They fit a probabilistic model to a corpus, using covert semantic categories of noun to predict distributions of both target and context words.



Adapted from Gantt et al. (2022). They present verbs in context to annotators and ask questions about an event's temporal properties and relations to other events. Their probabilistic model predicts annotations from covert classes. Here, r.i. = random intercept, st.dev. = standard deviation, and rel. = event relation.

Larger and smaller corpora. Larger corpora to get sufficient data about a wider variety of verbs. Smaller corpora to explore applications to understudied languages.

Evaluation. Use existing grammaticality-based tests and apply them systematically to the verbs in question; but as noted, these are thorny. Or, find a trustworthy existing dataset.

References

Bach, E. (1998). On time, tense, and aspect: An essay in English metaphysics. In P. Cole (Ed.), *Radical pragmatics* (pp. 63–81). New Yorkk: Academic Press.
Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28. doi: 10.18637/jss.v080.i01
Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, *10*(1), 395–411. doi: 10.32614/RJ-2018-017
Bürkner, P.-C. (2021). Bayesian item response modeling in R with brms and Stan. *Journal of Statistical Software*, *100*(5), 1–54. doi: 10.18637/jss.v100.i05
Frermann, L., & Lapata, M. (2016). Incremental bayesian category learning from natural language. *Cognitive Science*, *40*(6), 1333-1381. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/cogs.12304 doi: https://doi.org/10.1111/cogs.12304
Gantt, W., Glass, L., & White, A. S. (2022, 01). Decomposing and recomposing event structure. *Transactions of the Association for Computational Linguistics*, *10*, 17-34. Retrieved from https://doi.org/10.1162/tacl_a_00445 doi: 10.1162/tacl_a_00445
Hopper, P. J. (1979). Aspect and foregrounding in discourse. In T. Givon (Ed.), *Discourse and syntax* (p. 211 - 241). Leiden, The Netherlands: Brill. doi: 10.1163/9789004368897
Hundt, M., Rautionaho, P., & Strobl, C. (2020). Progressive or simple? A corpus-based study of aspect in World Englishes. *Corpora*, *15*(1), 77-106. Retrieved from https://doi.org/10.3366/cor.2020.0186 doi: 10.3366/cor.2020.0186
Kennedy, C., & Levin, B. (2008, 03). Measure of change: The adjectival core of degree achievements. In L. McNally & C. Kennedy (Eds.), *Adjectives and adverbs: Syntax, semantics, and discourse.* Oxford University Press. Retrieved from https://doi.org/10.1093/oso/9780199211616.003.0007 doi: 10.1093/oso/9780199211616.003.0007
Levin, B. (1999). Objecthood: An event structure perspective. In *Proceedings of CLS 35.* Retrieved from https://web.stanford.edu/ bclevin/cls99.pdf
Maienborn, C. (2005). On the limits of the Davidsonian approach: The case of copula sentences. *Theoretical Linguistics*, *31*(3), 275–316. Retrieved 2025-07-16, from https://doi.org/10.1515/thli.2005.31.3.275 doi: doi:10.1515/thli.2005.31.3.275
Mavridou, K.-I., Friedrich, A., Peate Sørensen, M., Palmer, A., & Pinkal, M. (2015, September). Linking discourse modes and situation entity types in a cross-linguistic corpus study. In M. Roth, A. Louis, B. Webber, & T. Baldwin (Eds.), *Proceedings of the first workshop on linking computational models of lexical, sentential and discourse-level semantics* (pp. 12–21). Lisbon, Portugal: Association for Computational Linguistics. Retrieved from https://aclanthology.org/W15-2702/ doi: 10.18653/v1/W15-2702
Nerbonne, J., & Van de Cruys, T. (2009). Detecting aspectual relations quantitatively. In E. Hinrichs & J. Nerbonne (Eds.), *Theory and evidence in semantics* (pp. 159 – 182). CSLI Publications. (2009/j.nerbonne/pub003)
Ramchand, G. (2008). *Verb meaning and the lexicon.* Cambridge University Press.
Rautionaho, P., & Hundt, M. (2022). Primed progressives? predicting aspectual choice in world englishes. *Corpus Linguistics and Linguistic Theory*, *18*(3), 599–625. Retrieved 2024-12-14, from https://doi.org/10.1515/cllt-2021-0012 doi: doi:10.1515/cllt-2021-0012
Smith, C. S. (1991). *The parameter of aspect.* Springer.
Zeldes, A. (2017). The GUM corpus: Creating multilayer resources in the classroom. *Language Resources and Evaluation*, *51*(3), 581–612. doi: http://dx.doi.org/10.1007/s10579-016-9343-x
Zucchi, S. (1998). Aspect shift. In S. Rothstein (Ed.), *Events and grammar* (pp. 349–370). Dordrecht: Springer Netherlands. Retrieved from https://doi.org/10.1007/978-94-011-3969-4_15 doi: 10.1007/978-94-011-3969-4_15

July 20, 2025