Aspectual classes as lexically-conditioned predictors of aspectual choice

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

1 Introduction

I investigate a distribution-based characterization of lexical aspectual classes.

The *grammatical aspect* of a verb is morphology which reflects either an internal perspective or an external perspective on the time course of an event. For example progressive aspect *was laughing* in *Mary was laughing when I arrived* takes an internal perspective, while perfective aspect *laughed* in *Mary laughed when I arrived* takes an external perspective. Not all verbs are felicitous in all aspects. For example, verbs denoting static situations usually sound worse in the progressive: **I am knowing French*. It has long been theorized that each verb in a language has an *aspectual class* which captures something about the temporal shape of the described event and thereby explains its compatibilities with different grammatical aspects (for an overview, see Filip, 2020). Indeed, past statistical work explicitly comparing aspectual class labels with the distribution of grammatical aspects has found strong statistical effects (Wulff et al., 2009; Hundt et al., 2020; Bardovi-Harlig, 1998; Andersen and Shirai, 1994), and these are thought to facilitate verb morphology acquisition (Wulff et al., 2009; Shirai and Andersen, 1995).

I propose to flip the script, asking to what extent statistical association with grammatical aspect is an adequate *characterization* of aspectual class. I propose that aspectual class is precisely the lexical information which contributes to aspectual choice. Therefore, it should be detectable by statistically computing each lexical item's contribution to aspectual choice.

This builds on proposals by Brent (1991) and Klavans and Chodorow (1992) to treat the stative-nonstative aspectual class distinction as gradient based on association with the progressive. It is similar in spirit, but orthogonal, to the work of Nerbonne and Van de Cruys (2009) who treat aspectual class as characterized by compatibility with temporal adverbials.

2 Method

I fit a Bayesian mixed-effects logistic regression to a corpus of natural spoken and written text in English (Zeldes, 2017). I did not presuppose any lexical aspectual classes for any verbs. Rather, I fit a model predicting aspectual choice (progressive or perfective), and I included lexical item as a predictor. I then used the fit weights for each lexical item to characterize its lexical aspectual class. A regression model allowed me to include other known predictors of aspectual choice in order to balance out their effects - namely, tense, matrix verb aspect, preceding verb aspect, subject type (singular, plural, mass, or none), object type, perfect morphology, voice, adverbs, subordinating conjunctions, verbal particles, genre, "for"/"in" preposition modifiers, and specific document/author. Mixed-effects regression allowed me to take advantage of frequent lexical items without them overpowering the analysis. A Bayesian model allowed me to obtain estimates for the effects of not just lexical item in general, but each individual lexical item. I also allowed effects varying by lexeme of tense, "for"/"in", and subject/object type, as these are known to affect aspectual class behaviour.

I fit the model using the R package BRMS (Bürkner, 2017, 2018, 2021) with four chains of 7,500 sampling steps. Intercept and linear coefficient priors were normal with standard deviation 2.5 and mean either -2.5 (intercept) or 0 (coefficients). Contact the author for data and code.

3 Results

3.1 Non-lexical predictors

I report results for a subset of predictors.

I replicated some results of Hundt et al. (2020): present tense verbs are more often progressive than past ($p_d > 0.999, p_{ROPE} < 0.001, 95\%CI = [1.10, 1.88]$) or future ($p_d > 0.999, p_{ROPE} < 0.001, 95\%CI = [1.51, 3.18]$) tense. On the other hand, I found a more consistent result of voice than Hundt et al.: active voice facilitates progressive aspect more than passive ($p_d > 0.999, 95\%CI = [1.69, 2.89]$).

Matching findings of Rautionaho and Hundt (2022), verbs immediately preceded by a progressive verb showed more progressive aspect than those preceded by a perfective ($p_d > 0.999, p_{ROPE} < 0.001, 95\%CI = [0.53, 1.16]$). A temporal adverbial headed by "in" decreased the probability of progressive compared with no temporal phrase ($p_d = 0.982, p_{ROPE} < 0.001, 95\%CI = [-5.76, -0.43]$), but a "for" adverbial was not clearly distinguishable from none ($p_d = 0.760, p_{ROPE} = 0.003, 95\%CI = [-1.39, 1.83]$).

I found significant variation by document (st. dev. $p_{ROPE} < 0.001, 95\%CI = [0.46, 0.84]$) and genre (st. dev. $p_{ROPE} < 0.001, 95\%CI = [0.58, 1.44]$), indicating that style and genre affect aspectual choice, as explored in theoretical (e.g. Smith, 2003; Egetenmeyer, 2021) and corpus (e.g. Mavridou et al., 2015) literature.

3.2 Lexical aspectual classes

I found statistically significant variation by lexical item in all measured effects (all st. devs. $p_{ROPE} < 0.001$).

I extracted fit estimates of the effect of each lexical item. All plots in this section are computed for verbs with at least 25 occurrences, and due to computational constraints, they use a subset of 1,000 samples from the model's posterior distribution.

Figure 1 shows random intercepts representing lexical effect on log-odds of progressive on one axis, and on the other, lexical effect on the present tense vs. past tense contrast. Remarkably, canonically stative verbs exactly coincide with those that strongly disfavor progressive aspect (from *believe*, downward). Moreover, in this plot and others not pictured, stative verbs cluster together in the lexical effects of any other predictors. Thus the lexical aspectual property of dynamicity emerges readily as a predictor of aspectual choice. The verbs which most favor progressive aspect are all standard examples of activities - dynamic verbs with duration but no endpoint.

Meanwhile, the verbs *try*, *think*, and *keep* on the far left of Figure 1, meanwhile, highlight a limitation of my approach. These verbs appear in the past progressive as readily as (or more readily than) in the present progressive. I expect that these verbs are often used to set up background information in a story, a primary use of the progressive aspect (Hopper, 1979). I was not able to control for communicative intent, and it may have contributed to the behaviour of these lexemes.

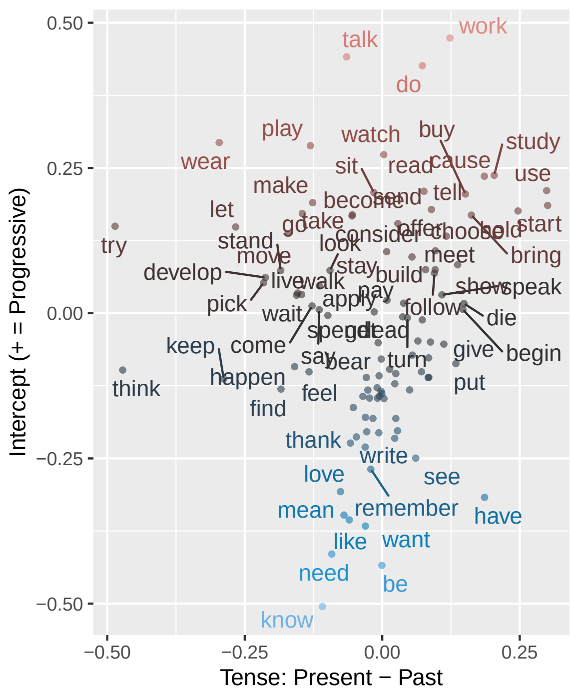

Figure 1: A scatterplot of verbs with at least 25 occurrences. Lexical effect on log-odds of progressive aspect is on the vertical axis and also represented with the color scheme; the horizontal axis shows the lexical effect on the difference in log-odds of progressive between present and past tense contexts.

Not pictured here, lexical items showed a very tight direct relationship between their effect on the present tense vs. past tense contrast and their effect on the future tense vs. past tense contrast. So, the most important axis of lexical variation in tense effect captures how much the past tense specifically favors or disfavors progressive aspect. This is counter to the prediction of a standard theory of aspectual class in which the present perfective (which conceptually forces an event to take place at a single instant) is the most restrictive.

The Bayesian nature of the model allows me to represent uncertainty in its predictions. Figure 2 shows 66% and 95% confidence intervals for lexical effect on log-odds of progressive for a subset of verbs. Due to the small corpus size, the model is not highly confident in any lexical effects.

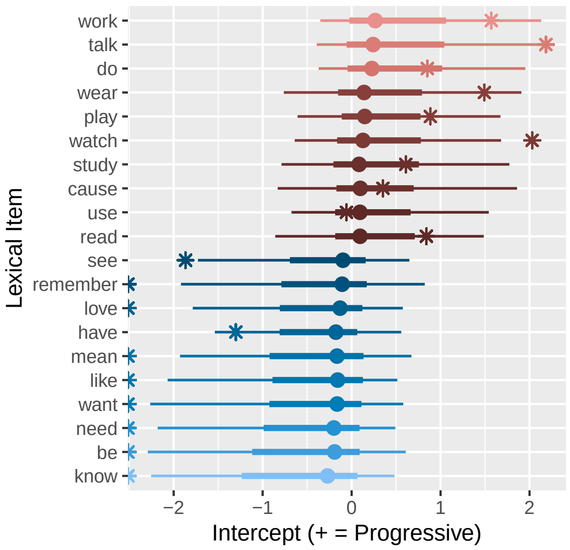

Figure 2: Interval plots showing the ten verbs with highest and lowest fit lexical effect on log-odds of progressive. Intervals show 66% and 95% credible intervals for each lexeme's effect. Star shapes show empirical log-odds-effect of each lexical item computed from raw corpus counts.

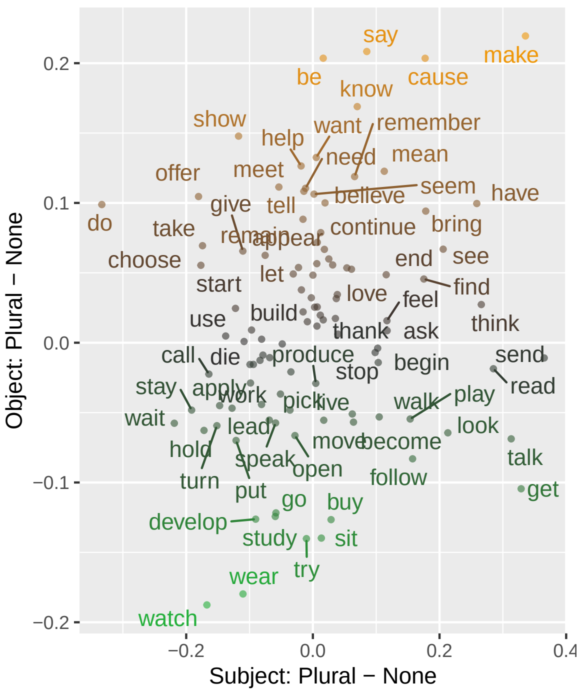

Figure 3: A scatterplot of verbs with more than 25 occurrences. The vertical axis shows lexical effect the difference in log-odds of progressive between contexts with plural objects or with no object, which is also represented with the color scheme; the horizontal axis shows the lexical effect on the difference in log-odds of progressive between plural subjects and no subject.

Figure 2 also shows, with asterisks, the empirical log-odds-ratio of progressive computed for each verb. We see the regularizing effect of modelling. For example, the linear model was able to abstract away from many confounding predictors and identify that *have* lexically behaves like the other stative verbs, despite its differing counts.

Figure 3 shows the lexical effect on the plural object vs. no object contrast on the vertical axis, and the lexical effect on the plural subject vs. no subject contrast on the horizontal. Verbs for which plural objects strongly favor the progressive are mostly of two kinds: stative verbs and verbs of creation and presentation (e.g. *cause* and *bring*). For the former, it is possible that plural objects facilitated an eventive coercion which allowed these verbs to be progressive, possibly by making them gradable. For the latter, this fits with their traditional classification as incremental theme verbs whose aspectual class is linked to their object. Of note, however, is the fact that incremental consumption verbs like *read* do not pattern in the same way.

Not pictured, the words *watch* and "wear" had especially negative plural object vs. singular object contrasts. This suggests that these two words specifically disfavor progressive when they have plural objects. Since each of these usually describes a long sustained interaction with a single object, their use with plural objects may have been restricted to habitual contexts, which disfavor progressive. This, again, is a place where not controlling for communicative intent may have created unexpected results.

For subjects, the lexical effect on plural subject - no subject contrast was closely tied to the lexical effect on the singular subject - no subject contrast (not pictured). This suggests that the largest lexical effect on subject behaviour was in the effect of having no subject. This may have been an oversight on my part: I did not include model the possibility of lexemes varying in the effect of passive voice. Verbs which are on the left in Figure 3 (e.g. *watch*, *do*, *develop*) may just be ones for which the progressive-disfavoring effect of passive voice is less strong.

Finally, the lexical aspectual behaviour of these verbs never appears discretized. We see continuous variation between verbs on all axes. Verbs are known to be able to shift between aspectual classes (Filip, 2020). My data suggest that verbs have different propensities to do this, placing them on a continuum of aspectual behaviour.

4 Outlook

I established that different verbs do contribute differently to aspectual choice, and this effect can be seen in a corpus without incorporating prior knowledge of aspectual classes. This lends support to the existence of aspectual classes (or possibly an aspectual continuum) as well as the potential for children to learn them using their associations with different aspects.

This method could be used to discover aspectual class on a new language. Aspectual class is difficult to discover due to sensitivity to context and brittleness under translation. Our statistical technique does not rely on translation, and so could be used to derive language-internally-motivated aspectual classes. I plan to investigate adaptations to smaller corpora to move toward such an application.

My next steps will be creating a more cognitively-grounded model of aspectual choice. This might follow the model which Gantt et al. (2022) use to derive aspectual categories from survey data or the BayesCat model which Fremann and Lapata (2016) use to learn semantic categories of nouns from a corpus.

References

Roger W. Andersen and Yasuhiro Shirai. 1994. Discourse motivations for some cognitive acquisition principles. *Studies in Second Language Acquisition*, 16(2):133–156.

Kathleen Bardovi-Harlig. 1998. Narrative structure and lexical aspect: Conspiring factors in second language acquisition of tense-aspect morphology. *Studies in Second Language Acquisition*, 20(4):471–508.

Michael R. Brent. 1991. Automatic semantic classification of verbs from their syntactic contexts: An implemented classifier for stativity. In *Fifth Conference of the European Chapter of the Association for Computational Linguistics*, Berlin, Germany. Association for Computational Linguistics.

Paul-Christian Bürkner. 2017. brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1):1–28.

Paul-Christian Bürkner. 2018. Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, 10(1):395–411.

Paul-Christian Bürkner. 2021. Bayesian item response modeling in R with brms and Stan. *Journal of Statistical Software*, 100(5):1–54.

Jakob Egetenmeyer. 2021. *Chapter 4 Time Updating Uses of the French Imparfait Extending Across Genres*, pages 56 – 77. Brill, Leiden, The Netherlands.

Hana Filip. 2020. *Lexical Aspect (Aktionsart)*, pages 1–61. John Wiley & Sons, Ltd.

Lea Frermann and Mirella Lapata. 2016. Incremental bayesian category learning from natural language. *Cognitive Science*, 40(6):1333–1381.

William Gantt, Lelia Glass, and Aaron Steven White. 2022. Decomposing and recomposing event structure. *Transactions of the Association for Computational Linguistics*, 10:17–34.

Paul J. Hopper. 1979. *Aspect and Foregrounding in Discourse*, pages 211 – 241. Brill, Leiden, The Netherlands.

Marianne Hundt, Paula Rautionaho, and Carolin Strobl. 2020. Progressive or simple? A corpus-based study of aspect in World Englishes. *Corpora*, 15(1):77–106.

Judith L. Klavans and Martin Chodorow. 1992. Degrees of stativity: The lexical representation of verb aspect. In *COLING 1992 Volume 4: The 14th International Conference on Computational Linguistics*.

Kleio-Isidora Mavridou, Annemarie Friedrich, Melissa Peate Sørensen, Alexis Palmer, and Manfred Pinkal. 2015. Linking discourse modes and situation entity types in a cross-linguistic corpus study. In *Proceedings of the First Workshop on Linking Computational Models of Lexical, Sentential and Discourse-level Semantics*, pages 12–21, Lisbon, Portugal. Association for Computational Linguistics.

J. Nerbonne and T. Van de Cruys. 2009. *Detecting Aspectual Relations Quantitatively*, pages 159 – 182. CSLI Publications. 2009/j.nerbonne/pub003.

Paula Rautionaho and Marianne Hundt. 2022. Primed progressives? predicting aspectual choice in world englishes. *Corpus Linguistics and Linguistic Theory*, 18(3):599–625.

Yasuhiro Shirai and Roger W. Andersen. 1995. The acquisition of tense-aspect morphology: A prototype account. *Language*, 71(4):743–762.

Carlota S. Smith. 2003. *Modes of Discourse: The Local Structure of Texts*. Cambrige University Press.

Stefanie Wulff, Nick C. Ellis, Ute Römer, Kathleen Bardovi-Harlig, and Chelsea J. Leblanc. 2009. The acquisition of tense–aspect: Converging evidence from corpora and telicity ratings. *The Modern Language Journal*, 93(3):354–369.

Amir Zeldes. 2017. The GUM corpus: Creating multilayer resources in the classroom. *Language Resources and Evaluation*, 51(3):581–612.