Grammar and Aggression in Human Cognitive Evolution

Words – so innocent and powerless as they are, as standing in a dictionary, how potent for good and evil they become, in the hands of one who knows how to combine them

N. Hawthorne
Acknowledgments

I am grateful to my co-authors on this topic: John L. Locke, Antonio Benítez-Burraco, Natalia Rakhlin, Noa Ofen, David Gil, Francesco Ferretti, and Ines Adornetti,

as well as to Martha Ratliff, Michel DeGraff, Eugenia Casielles, Dorit Bar-On, Brady Clark, and many, many others, for great discussions and feedback.

These include the audiences at various recent presentations, including at MIT; Harvard; USC, Los Angeles; University of Tokyo and Kyoto University, Japan; Poznań, Poland; Pavia, Italy; Ljubljana, Slovenia; Ghent, Belgium; Leipzig, Germany; Crete, Greece; Novi Sad and Belgrade, Serbia; etc.

Some of the most recent work reported here is based on joint work with Antonio Benítez-Burraco.

This project has been supported by multiple WSU grants, including Humanities Center Marilyn Williamson’s and Charles Gershenson’s Endowed Distinguished Faculty Fellowships.

Thank you to the AoS for the invitation to join and to present.
Highlights of today’s talk

- To investigate the role of (verbal) aggression in human evolution, focusing on highly expressive pejorative compounds (and ideophones)
- To explain why such expressions are good approximations/proxies of the earliest forms of language (reconstructed proto-grammar stage)
  - To consider how these and other proxies are processed by the brain (fMRI experiments)
- To pursue a gradualist scenario for the evolution of syntax/grammar, subject to selection, contra saltationist views
- To consider how and why physical (reactive) aggression has gradually decreased in human evolution, correlated with the emergence of early language (cross-fertilization with Human Self-Domestication Hypothesis)
  - To seek a common evolutionary cause for why altered aggression, altered cross-modality (relevant for metaphoricity), and altered syntax cluster together in cognitive disorders affecting language
- To highlight how and why linguistic detail and specific linguistic proxies are essential for advancing testable and cross-fertilizable proposals regarding linguistic and cognitive evolution
Gradualist vs. saltationist views

- Darwin’s (1874, 595) view was that language evolved gradually through sexual selection, as an instinct to acquire a particular method of **verbal display similar to music**.
  - “The sensations and ideas thus excited in us by **music** or expressed by the **cadences of oratory** appear, from their **vagueness, yet depth**, like mental reversions to the emotions and thought of a long-past age.”

- **Considering expressive language allows us to see linguistic (and cognitive) evolution as competitive, and not solely cooperative.**

- Some commonly quoted examples of **expressive language**:  
  - gradational (reduplicative) compositions (aka **ideophones**):  
    - wishy-washy, willy-nilly, dilly-dally, teeny-weeny, nitty-gritty, zig-zag  
  - emotionally charged words: **pejoratives/swearing**

- In contrast, based on their view of syntax/grammar as an undecomposable/unnegotiable block, Chomsky (2005) and Berwick and Chomsky (2011, 2016) proposed that language/syntax emerged suddenly and recently, in its full complexity: **saltationist view**.
  - For them, syntax/language evolved as a result of one single random mutation; and it evolved for the purposes of thinking, rather than communicating.
Proto-syntax and insult

- I have considered in detail a gradualist evolutionary scenario for grammar, which includes:
  - creating vivid innovative insults with the crudest of proto-grammars, and the most basic of vocabulary (Progovac & Locke 2009; Progovac 2015; 2016),
  - where verbal aggression is identified as one kind of utility of the simplest grammar (there are many other beneficial uses)
- But first, why insult?
- And second, how can we know/hypothesize what simplest proto-grammars were like?
- The answers to these questions turn out to be related,
  - as the best proto-grammar approximations in present-day languages happen to specialize for insult

Precision comes from relying on syntactic theory and on linguistic detail

- Avoiding impressionistic proposals
- According to the syntactic theory of Minimalism and predecessors (e.g. Chomsky 1995; Adger 2003), a modern clause/sentence is characterized by the following *partial hierarchy of sentential layers*:
  
  (i) \( CP > TP > vP > VP/SC \)

  - The inner VP/SC (verb phrase/*small clause*) layer accommodates the verb/predicate and *one argument* (noun)
  - The little verb phrase vP layer supports transitivity by accommodating *an additional argument*, such as agent
  - TP (Tense Phrase, or Sentence) accommodates the expression of e.g. tense and finiteness
  - CP (Complementizer Phrase) accommodates e.g. embedding and question formation.
Hierarchical sentence is built on the foundation of a (flat) Small Clause (SC).
The inner small clause (SC) still “lives” in both transitive and intransitive sentences.

(1) Maria will grow corn.  Maria will bathe the baby.
(2) Corn will grow.  The baby will bathe.
(3) Maria will grow.  Maria will bathe.
A precise reconstruction of proto-syntax

- **Internal reconstruction of** the initial syntactic stage(s) (Progovac 2015):
  (i) **Structure X is considered to be (evolutionarily) primary relative to Structure Y if X can be composed independently of Y, but Y can only be built upon the foundation of X.**
    - While SCs/VPs can be composed without a TP and little vP (transitivity) layers, the vP and TP can only be built upon the foundation of a SC/VP.

- **One can thus reconstruct a vP-less and TP-less (intransitive and tenseless) Small Clause stage in the evolution of language.**

- **Best approximations of this grammar**, or “living fossils” in the sense of Jackendoff (1999), are **verb-noun compounds**:
  - rattle-snake; dare-devil; kill-joy; turn-table; turn-coat; tumble-weed; tumble-dung

- **No possibility for subject/object distinction in this stage** (novel insight, as other approaches, even gradualist (Jackendoff’s), assume that there always was a subject-object distinction.

- **Just one verb and one noun, whether subject-like or object-like**

Verb-noun compounds specialize for insult: “unquotable coarseness”

- kill-joy, turn-skin (cf. turn-coat), hunch-back, wag-tail, tattle-tale, scatter-brain, cut-throat, mar-wood (bad carpenter), heck-wood, busy-body, cry-baby, break-back, catch-fly (plant), cut-finger (plant), fill-belly (glutton), lick-spit, pinch-back (miser), shuffle-wing (bird), skin-flint (miser), spit-fire, swish-tail (bird), tangle-foot (whiskey), tumble-dung (insect), crake-bone (crack-bone), shave-tail (shove-tail), wipe-tail, wrynge-tail, fuck-ass, fuck-head, shit-ass, shit-head.

- These compounds tend to have transient lives, with many of them now lost or obsolete, and with different generations being familiar with different ones;
  - there were thousands of them created in medieval times;
  - they do not get preserved in dictionaries or grammar books because they often show “unquotable coarseness” (Weekley 1916)

VN compounds in Serbian – not so lame either

- ispi-čutura (drink.up-flask—drunkard), guli-koža (peel-skin—who rips you off), cepi-dlaka (split-hair—who splits hairs), muti-voda (muddy-water—trouble-maker); vrti-guz (spin-butt—fidget); tuži-baba (whine-old.woman; tattletale); pali-drvice (ignite-stick, matches), jedi-vek [eat-life = one who constantly annoys], kosi-noga [skew-leg = person who limps], mami-para [lure-money = what lures you to spend money]; pali-kuća [burn-house = one who burns houses]; podvi-rep [fold-tail = someone who is crestfallen]; priši-petlja [sow-loop = who clings onto another]; probi-svet [break-world = wanderer]; raspi-kuća (waste-house = who spends away property], vuci-batina [pull-whip = good-for-nothing]; kaži-prst (say-finger = index finger)
- jebi-vetar (fuck-wind—charlatan); češi-guz ‘scratch-butt;’ deri-muda ‘rip-balls’ (place name, a steep hill); gladi-kur ‘stroke-dick’ (womanizer); kapi-kur ‘drip-dick’ (name of a slow water spring); liz-guz ‘lick-butt;’ nabi-guz ‘shove-butt;’ piš-kur ‘piss-dick;’ plači-guz ‘cry-butt;’ poj-kurić ‘sing-dick’ (womanizer); seri-vuk ‘shit-wolf’

- Of note is also the imperative-like form of the verb in Serbian
Transitivity evolves in diverging directions, but with a common SC denominator (minimal tweaks/tinkering with structure)

- There is significant variation across languages as to how they build their transitive structures, distinguishing subjects from objects;
- With rare exceptions, transitive structures across languages typically add only one extra piece to the foundational intransitive structure, whether it is on top (ergative) or on the bottom (accusative), and serial verb patterns tend to string together a limited number of (small) clauses, often just two.
  - (i) VN (proto-grammar – intransitive absolutive foundation)*
  - (ii) VNN (accusative grammar: adding just one additional argument (e.g. patient/theme) to (i) from below)
  - (iii) NVN (ergative grammar: adding just one additional argument (e.g. agent) to (i) from above)
  - (iv) VN VN (serial verb grammar: duplicating the small clause)

[*Word order VN or NV not of relevance here.*]
Engaging the hominin timeline

- This led me to propose that these widely diverging hierarchical solutions were a later add-on, superimposed upon the common proto-syntactic foundation, and that the hierarchical layers of language did not emerge only once and uniformly (in Africa) in our species,
  - but instead multiple times, and independently, either within Africa, or after the dispersion from Africa (Progovac, 2015, 2016).

- At least under the uniregional view of human origins, this would suggest that hierarchical syntax emerged no earlier than 100–50 kya, with humans.

- On the other hand, the proto-grammar stage could have been present much earlier, with other species as well.
  - It has been proposed that some forms of language with grammar may have been in place as early as 500 kya, based on the skeletal and genetic evidence among Homo heidelbergensis’ descendants, including Neanderthals and Denisovans (e.g., Dediu and Levinson, 2013; see also Johansson, 2005; Zilhão, 2011).
The Growth of Grammar

The reconstructed **proto-grammar provides the foundation** for building hierarchical syntax. It also provides a **common denominator** for cross-linguistic variation.

1. **one-word stage** (no syntax); includes also non-combinable “proto-words”: *Psst, Tsk-tsk, Ouch, Ugh, Grr* (see e.g. Jackendoff 1999, 2002)

2. the **earliest (two-slot, small clause, intransitive) syntactic stage** (reconstructed based on syntactic theory)

3. **hierarchical, transitive syntax** (subjects vs. objects differentiation, tense, embedding, etc.)
Gradual emergence of hierarchical layers

Proposal: complex hierarchical syntax *emerges gradually* (e.g. Progovac 2015), building on the flat SCs foundation, with each layer adding another grammaticalized, explicit expression of some category, including:
Child development: Parallel development of language and cognition

- Parallel to the arguments for the evolution of grammar, the development of sentential structure in children between the ages of 2-5 also seems to proceed in these rough stages:
  - (i) **Intransitive SC** stage
  - (ii) **Transitivity** (vP), featuring both subjects and objects; adding explicitness to the expression of “who does what (to whom);”
  - (ii) **Tense and finiteness** (TP), featuring explicit ways of signaling tense and/or aspect, as well as subjecthood (through e.g. agreement or case);
  - Also: (iii) **Embedding and question formation** (CP), featuring explicit ways of embedding one sentence (point of view) within another, including recursively.

This provides a **representational tool for stage-like cognitive advancements of early childhood** (Rakhlin and Progovac 2020).

- Adding increasingly complex syntactic layers to a child’s grammar, layer by layer, **expands children’s cognitive capacity** by giving them more precise and efficient ways to represent (and discover) complex information than is possible relying on non-verbal means.
Grammar as “cognitive technology”

- Humans, as non-human species, are equipped with innate basic knowledge in a number of domains, which allows infants to make sense of objects and agents, and their relations in space and social sphere, long before they acquire language (Spelke & Kinzler, 2007).
- However, these areas of core knowledge are limited.
- Our proposal is that layers of syntactic structure represent “cognitive technology” that supplements and augments the non-linguistic representations.
- (Adult) humans are also quite capable of reasoning about abstract or psychologically distant (i.e., beyond one’s direct sensory access) concepts and events, such as:
  - hidden causes (transitivity, vP)
  - distant past and future (TP)
  - counterfactual situations (TP and CP)
  - other people’s beliefs and desires (Theory of Mind) (CP)
- While reasoning about such constructs may be somewhat possible without language/grammar, it is at best inefficient (slow, imprecise, and inconsistent).
- Various experiments show correlations between syntactic and cognitive development along these lines; these claims are testable and falsifiable.
The (rough) stages of linguistic/cognitive evolution (Benítez-Burraco and Progovac, 2020)

1. **The first stage**, occurring roughly in the period prior to 200 kya: self-domestication (SD) processes only start to emerge, with reactive physical aggression still high; possible emergence and use of simple, proto-language forms.

2. **The second stage**, roughly from 200 to 50 kya: an **accelerated feedback loop** between SD and the solidification of the early forms of language/grammar, both promoting a reduction in reactive physical aggression (all relying on the evolution of the same brain circuits).

3. **The third stage**, 50–10 kya (the Upper Paleolithic): SD reaches its peak, with more cooperation and socialization and less reactive aggression; a suitable niche for language and cognition to complexify.

4. **The fourth stage**, especially 10kya onward, after the onset of Neolithic, the rise of **proactive aggression, especially warfare** (Wrangham et al., 2006), correlated with the complexification of language and syntax, including the emergence of higher, more abstract layers of syntactic structure.
   - At this stage, hostile intergroup encounters in the form of raids and ambushes (as also observed in present-day hunter-gatherer societies, Allen and Jones, 2014), were being replaced by **escalated, coordinated battles** (Kissel and Kim, 2019)
Fig. 1. A graphical summary of the four-stage model of language evolution in our species under the effect of self-domestication.
Continuity of (proto-)grammar on both ends: with modern grammars, and with other species

- Looking forward, the structure of these proto-grammars (e.g. compounds) foreshadows/predetermines the very nature of modern human grammars, in three crucial respects:
  - their **binary** combinatorial nature
  - their **small clause intransitive inner layer**
  - their reliance on **nouns and verbs** to express who does what (to whom), i.e. **basic predication**
Recall that hierarchical sentences are built on the foundation of a (flat) Small Clause (SC).

Modern syntax posits binary branching only; only two units can be merged at a time, forcing hierarchy, and utilizing what is already available: the two-slot SC.
Continuity with other species

- On the other end, looking backward, comparatively speaking, other primates seem capable of simple two-slot combinations (with no subject/object differentiation):
  - such as *hide peanut* and *hide Kanzi* (see e.g. Greenfield & Savage-Rumbaugh (1990: 161) regarding bonobo Kanzi);
  - according to Patterson & Gordon (1993), *gorilla Koko* is not only capable of producing novel two-word metaphorical combinations (e.g. *cookie rock,* for a stale bun), but also of insult, playfulness, and humor.
  - Darwin (1872) observed that **strong emotions** expressed in animals are those of lust and hostility, and that they may have been **the first verbal threats and intimidations uttered by humans** (see also Code 2005)

- If so, then **proto-syntactic two-slot compositions, expressing verbal aggression**, would have provided a more graceful transition from animal communication to human language, both in terms of combinatorial abilities, and in terms of emotional load.

So, why insult? Adaptiveness of verbal aggression

First, I was taken to this evolutionary scenario by simply following the syntactic reconstruction I performed, and the data it led me to (i.e. hundreds of verb-noun compounds across languages, which turn out to specialize for insult when referring to humans).

Also, considering insult (i.e. verbal aggression) reveals direct and immediate selection/survival benefits.

- Bergen (2016: 7): the most potent words of all are swear words and expletives, especially those that are highly taboo, which “elicit the strongest measurable psychological reactions: the fastest pulse, the sweatiest palms, the shallowest breathing.”

- Insult (verbal aggression) is also neurobiologically “real”/tangible, leading to testable hypotheses.

- It also leads to cross-fertilization with Human Self-Domestication hypothesis, whose main ingredient is also aggression, i.e. physical reactive aggression.

Even though my focus here is on verbal aggression and insult, the benefits and uses of this kind of proto-grammar would have been much broader, including: naming animals (tumble-dung; swish-tail (bird); stink-bug), plants (tumble-weed; catch-fly), objects, and places; expressing basic commands (e.g. Catch fly! Turn snake!) and basic statements (e.g. Bug stink; Monkey see).
Cognitive, at the expense of physical contest

- The emergence of proto-grammars affords a new, **more adaptive way to compete for status and sex** in ancient times.
- Power to create **many novel expressions**, never heard before, and to capture a trait of a person, and to discredit a person, with only two basic proto-words.
- Words can hit just as hard as stones, but only in the hands of the one who knows how to combine them (cognitive evolution).
- Their successful use would have **enhanced relative status**:
  - first by derogating existing rivals and placing prospective rivals on notice (aggressive rivalry),
  - and second by demonstrating verbal skill and quick-wittedness (mate choice) (Progovac and Locke, 2009).
  - both types are relevant for sexual selection (Darwin 1874).

- **Dramatic increase** in the variety and expressivity of insults relying on proto-grammar compounds, vs. just isolated words.
Compounding the insult: The power of two-slot combinatorial syntax (over just single words)
The poetic dimension of verb-noun compounds

- Weekley (1916): this is a very expressive way of naming
- Mihajlović (1992), who devoted his career to collecting over 500 Serbian place and people names (VN compounds), calls them condensed compositions which pack in them ... frozen fairy tales, proverbs, and ancient wisdoms and metaphors (1992: 8-9)
- For Darmesteter (1934: 443), the artistic beauty and richness of these compounds (in French) is inexhaustible:
  - “At the time of Renaissance, Ronsard introduced [VN compounds] in a new and original manner as epithets: Jupiter lance-tonnerre, le soleil donne-vie, Hercule porte-massue...
  - It would be well could French poets again make use in lofty poetry of this class of epithets; for they may attain Homeric breadth...”
The audacity and the poetry of insult

• Samarin (1969) collected 3,000 ideophones in Gbeya (Central African Republic), many of which are humorous insults; reduplicative and poetic

Acc. to Samarin: people who are more prone to use insults than others might set a model for the other villagers by their “skill, imagination, and audacity” (sexual selection argument)

• One can say that someone is short or tall, using regular adjectives, without getting very much reaction, but the ideophones invariably arouse laughter.

• the initiation of humor itself is considered to involve strong assertiveness (i.e. audacity, in Samarin’s sense) (see e.g. McGhee 1976)
What are ideophones?

- The structure of ideophones is also typically a two-slot mold.
- Prototypical examples are reduplicative and iconic in imitating the sounds (tick-tock) or the sights (zig-zag), directly relevant for cross-modality considerations (Cuskley & Kirby 2013).
- They are also playful and creative.
  - (1) tick-toc; zig-zag; flip-flop; willy-nilly; wishy-washy; hanky-panky, okey-dokey, mumbo-jumbo, teeny-weeny, nitty-gritty
  - (2) tika-taka; cik-cak; trte-mrte (aha, you are scared!); apa-drapa (dressed in an unruly, disorderly manner); kuku-riku (rooster’s call) (Serbian)
  - (3) mî mê (mosquitoes buzzing); plì -plôn (empty bottle submerged in water filling up) (Hmong, Martha Ratliff, p.c.; 2010)

Some languages are especially rich in such ideophones, numbering in the thousands, including Gbeya, Japanese, Korean, Hmong.

- They are often used, and effectively so, for ridicule and insult.
- Ideophones are processed differently by the brain:
  - Lockwood and Tuomainen’s (2015) EEG experiment found that ideophones in Japanese elicit a specific brain response not found in non-ideophonic adverbs,
  - which is consistent with abundant linguistic literature reporting on the vivid experience’ of ideophones (e.g. Doke, 1935, for Bantu languages).
Verb-noun compounds are also processed differently by the brain

- In an fMRI experiment we contrasted the processing of verb-noun compounds (e.g. kill-joy; pick-pocket; cry-baby) vs. hierarchical compounds (e.g. joy-kill-er; boot-lick-er; whistle-blow-er)
  - and found a robust effect in the fusiform gyrus area (BA 37) (Progovac et al. 2018b)
- BA 37 is the area where visual processing and certain non-compositional semantic processing (e.g. concreteness, metaphor) come together (e.g. Bookheimer, 2002)
- Verb-noun compounds seem to evoke a more vivid, more visceral effect,
  - even though the two compound types were matched in imageability/metaphoricity (see above)

The simpler the grammar, ...

- ... the more expressive the language
- The layer(s) of abstract syntactic structure seem to render -er compounds in the experiment less visceral, less expressive,
  - meaning that the postulated proto-grammars are a better fit for expressive language, including insult
- If so, then it is not a coincidence that derogatory language/insult has been so well preserved in these proto-grammatical compositions, across different languages
- Verb-noun compounds are often:
  - aggressive and humorous (audacity), both of which are relevant for self-domestication hypothesis, and for sexual selection argument,
  - they are also highly metaphorical (relevant for cross-modality, and for the poetical dimension).
  - Aggression and poetry can go together?
What is cross-modality?

- Cross-modality is crucial for the processing of metaphorical/figurative extensions, upon which language in general heavily relies, including grammar;
  - it often involves interactions among multiple sensory modalities,
    - (cf. e.g. metaphors such as bitter cold; loud shirt; sharp cheese)
  - “Using figurative language allows [one] to be both playful and to communicate information effectively .... It provides tools to paint a picture with words (the words are bringing images to the reader’s mind)” (Merriam-Webster Dictionary)
- According to Ramachandran and Hubbard (2001), the beginning of cross-modality possibly involved a cross-wiring in fusiform gyrus
- The fusiform gyrus area (specifically BA 37) is also implicated in synesthesia, a condition of enhanced cross-modality (due to atypically exuberant brain connectivity),
  - where simulation of one sensory or cognitive pathway leads to involuntary experiences in another (e.g. Cytowic & Eagleman 2009; Cuskley & Kirby 2013)
Root small clauses (mini sentences) are also processed differently by the brain

- In an fMRI study, full sentences with Tense (TPs) (The case is closed; The apology is accepted) were contrasted with TP-less small clauses (Case closed; Apology accepted), another proto-grammatical structure.
- There was reduced activation in (TP-less) Small Clause condition compared to both full TP sentences and the control conditions,
  - both in the Broca’s area (left BA 44) and the right basal ganglia (Progovac et al. 2018b)
  - affirming the relevance of Broca’s–basal ganglia network for the processing of more articulated, more layered syntax)
Full TP sentences > small clauses in English
Broca’s-basal ganglia network for syntax

- **Syntactic processing relies heavily on the cortico-subcortical networks,** which include the Broca’s area, in particular BA 44, and basal ganglia (e.g. the striatum) (Friederici, 2017; Opitz and Friederici, 2007; Teichmann et al., 2009, 2015; Szalisznyó et al., 2017).

- Broca’s area is not the sole center for **syntactic processing**, but rather part of a **larger circuit that involves subcortical structures** (e.g. Gibson, 1996; Lieberman, 2000, 2009; Vargha-Khadem et al., 2005; Ardila et al., 2016a,b; Ullman, 2006).

- There is evidence that **Broca’s-basal ganglia network** has been bolstered via **selection in recent evolution**, in the line of descent of humans (and Neanderthals),
  - resulting in **increased synaptic plasticity and neuronal connectivity**, with **FOXP2** and other genes playing a role (see e.g. Ullman 2006; Dediu 2015; Hillert 2014)

- The emergence and gradual **complexification of syntax** would have played a critical, **active role** in this evolutionary development of the **brain**
  - The emergence of two-slot **proto-grammars** may have been the most **important milestone** in the evolution of human cognition
Language evolution is linked to management of aggression (and vice versa)

- Our proposal is that the emergence of proto-grammar, as well as the enhancement of cross-modality, are **partly a side-effect, and partly a cause**, of human **management of (reactive) aggression**, as it relates to **human self-domestication (HSD)**. (Progovac and Benítez-Burraco, 2019; also: Benítez-Burraco and Progovac, 2020, 2021).

- The **HSD hypothesis** is based on the existence in our species of many of the features found in domesticated animals, including:
  - prolonged juvenile period; reduced sexual dimorphism;
  - **reduced reactive aggression; reduced response (of the HPA axis) to stress** (hypothalamic–pituitary–adrenal axis); **decrease in cortisol levels** (relevance of humor) (Shea, 1989; Somel et al., 2009; Herrmann et al., 2011; Hare et al., 2012; Plavcan, 2012; Stringer, 2016; Hare, 2017; Theofanopoulou et al., 2017; Thomas and Kirby, 2018; Benítez-Burraco et al., 2020).
  - Crucial driver of HSD is considered to be **a gradual reduction in reactive aggression**, 
    - **typically attributed to** (sexual) selection for less aggressive/less reactive partners, and in favor of **pair-bonding** (e.g. Hare et al. 2012; Stanyon and Bigoni 2014; Okanoya 2015; Gleeson 2018) 
  - Later stages of human evolution, featuring highly hierarchical, elaborate syntaxes, correlate with enhanced aggression, but this time, it is **proactive (premeditated) aggression**.
Feedback Loop: Emerging grammars and taming of aggression (HSD)

- Our proposal is that these **early stages of grammar** (as approximated by e.g. derogatory verb-noun compounds) engaged in a **mutually reenforcing gene-culture feedback loop with HSD, which was targeting physical aggression** (Progovac & Benítez-Burraco, 2019),
  - because they provided a **more adaptive way** to replace reactive physical aggression with verbal aggression,
  - thus **significantly accelerating** both the evolution of language (including metaphoricity and syntax) and the evolution of HSD, **all of which rely on enhanced connectivity of the cortico-striatal brain networks**
- This **linguistic dimension added to the HSD hypothesis** ensures that human evolution did not just yield a **tame (but mute) phenotype** (perhaps like bonobos),
  - but, in fact, a phenotype excelling at channeling physical aggression/reactivity into verbal behavior.

Altered processing of metaphorical language

- The ability to understand and make use of figurative (metaphorical) language seems to be altered in most, if not all, cognitive conditions, including synesthesia, but also ASD (autism) and SZ (schizophrenia) (Benítez-Burraco 2017)
  - ASD individuals face difficulties establishing connections between two elements of a compound whose combination requires metaphorical “stretching” of meaning (e.g. Riches et al. 2012; Kambanaros et al. 2019)
    - tendency to default to literal interpretations
    - According to e.g. Imke et al. (2008), neurons in fusiform gyrus are fewer and smaller in autistic individuals.
- In contrast, various types of hallucinations experienced by SZ individuals may be attributed to an atypical disinhibition of cross-modality, akin to that found in synesthesia.
  - Synesthetes in fact exhibit an advantage in language abilities, related to a variety of advantages in memory (e.g. Rouw et al., 2011, van Leeuwen 2020),
  - and this also applies to relatives of SZ individuals.
    - Synesthetes also show superior understanding of unfamiliar sound symbolic foreign words, including ideophones (Bankieris and Simner 2015).
Poles on the continuous cline

- **ASD** is characterized by **higher levels of reactive aggression** (and so is Tourette’s Syndrome, **TS**), in comparison to both typical and **SZ** populations (Hill et al. 2014; Fitzpatrick et al. 2016).
  - However, **SZ** exhibits **higher levels of proactive (i.e. premeditated) aggression** (Bo et al. 2013), especially in comparison to ASD.
- **Deficits in language structure and use** are found in both **ASD** and **SZ** (Bailey et al. 1996, Tager-Flusberg et al. 2005, van Os and Kapur 2009, Stephane et al. 2014).
- **ASD** individuals exhibit **difficulties with metaphors and abstract concepts** more generally (Dodd, 2005; Jordan, 2010); tend to be **hyper-systematic and literal** (also the case with **TS**)
  - People with **SZ** can also show difficulties with **novel metaphors** (Rapp et al., 2018), as well as **humor** (Pawełczyk et al., 2018).
  - Nonetheless, SZ individuals exhibiting a mild, nonclinical manifestation of psychotic-affective conditions show **relative strengths**, e.g. when interpreting metaphors, emotions, humor, and irony (Crespi, 2008: 238).
    - In addition, relatives of people with **SZ** were found to **have an advantage** in artistic expression and originality (Fink et al., 2014).

Cognitive conditions as adaptations for language

- **Certain aspects** of the conditions such as ASD (*autism*) and TS (*Tourette Syndrome*) can provide:
  - a glimpse into the earlier stages of linguistic evolution characterized by
    - a higher degree of reactive aggression,
    - emerging grammatical regularity, and
    - emerging cross-modality, i.e. metaphoricity

- In contrast, *synesthesia* (and to some extent SZ (*schizophrenia*)) can provide
  - a window into a later development, when higher levels of (disinhibited) connectivity in the cortico-subcortical and other brain networks would have resulted in:
    - exaggerated, **super cross-modality**, among a significant portion of the population,
      - as an adaptation to acquire language,
      - as well as to suppress reactive aggression
    - with a side-effect of **hallucinations and delusions** in SZ, but also sometimes in synesthesia
      - (based on Benítez-Burraco and Progovac, 2021)
Why do altered aggression, altered cross-modality, and altered language structure cluster together in disorders?

- Our finding implicates the cortico-striatal brain networks,
  - whose dense connectivity is instrumental not only for curtailing reactive physical aggression (by enabling cortical structures to exert better control over subcortical structures),
  - but also for metaphoricity, by enhancing cross-modal connections
  - and for processing syntax.

- This suggests that (i) taming of aggression, (ii) cross-modality (metaphoricity), and (iii) language structure/syntax co-evolved, supported by enhanced connectivity in the same brain circuits.

Atypical inhibition in cortico-striatal brain circuits: ASD and TS

On the one hand, **TS and ASD** have been found to exhibit an interneuron dysfunction that gives rise to an **altered degree of inhibition** of specific cortico-striatal circuits, resulting in **reduced control of striatal activity by cortical structures** (e.g. Rapanelli et al. 2017; McBride and Parker 2015)

- contributing to **atypical degrees of reactive aggression** in both ASD and TS, as well as to involuntary verbal aggression (including **coprolalia** in some TS individuals)
  - The damage to these networks also leads to the specific kind of **disinhibition in aphasia**, i.e. the production of **speech automatisms**, including uncontrollable swearing (Code 2005; 2011; Code et al., 2009).

- Worbe et al. (2012) relate TS characteristics to **cortico-basal ganglia network immaturity**, the network also implicated in **physical aggression** and **language processing** Mink 2003; Ganos et al. 2013).

- Specifically, Lischinsky and Lin (2020) found that the suppression of aggression demands an increased control of the striatum, among other subcortical regions, by the prefrontal cortex.
Atypical disinhibition: synesthesia and SZ

- On the other hand, synesthesia and SZ involve **an atypical disinhibition in these networks**
  - According to Grossenbacher and Lovelace (2001), synesthesia is a result of **disinhibited cortical sensory feedback** (see also Cytowic 1993).
    - One possible explanation is to **invoke a failure of adequate pruning** of connections which typically takes place in ontogeny (e.g. Maurer, 1993; Baron-Cohen et al. 1993; Ramachandran and Hubbard, 2001; Mauer and Mondloch, 2006; Ward, 2013)
    - Interestingly, **ASD** has been characterized as involving **over-pruning**, i.e. aggressive synaptic pruning during childhood (Thomas et al., 2016)
  - Silbersweig et al. (1995) ascribe SZ hallucinations in both auditory and visual modalities to **the abnormal disinhibition of cortical–subcortical circuits**.
- In sum, both **enhanced cross-modality and (the suppression of) reactive aggression**, including verbal aggression, rely on **a precise degree of (dis)inhibition of connectivity in the cortico-striatal brain circuits**, the same circuits that are **also essential for the processing of syntax and language more generally**.
Extensive individual variability

- Various cognitive conditions affecting language involve **patterns of inhibition vs. disinhibition** that seem to be **poles on the continuum of cognitive modes**, including also the typically-developing cognition.
- This highlights the **extensive individual variability across all the dimensions relevant for language**, which moreover seems to be **genetically influenced**.
- Such individual variability has been (and will be) providing a **fertile ground for natural/sexual selection** to operate, both in the past, and in the future.
- These considerations suggest that the evolution of human language/syntax cannot be seen as a simple, straightforward, one-mutation step (e.g. Berwick and Chomsky 2016), but rather as a **complex, gradual, ongoing, multi-faceted and multi-gene** phenomenon (e.g. Dediu 2015; Dediu and Ladd 2017), with each new development and innovation potentially subject to selection.
Final thoughts

• Generally speaking, human evolution saw a gradual shift
  • from the more emotional (reactive) to the more rational (premeditated);
  • from primarily physical contest to primarily cognitive/verbal contest and behavior more generally.
• But, human aggression (or competitiveness) have not been eliminated – they only changed shapes and forms...
• Paradoxically perhaps, according to this approach, humans became more tame and more cooperative by practicing verbal aggression,
  • i.e. by replacing physical aggression by verbal aggression;
  • early grammars already afford effective, humorous, playful means for practicing verbal aggression, and deflecting physical aggression.
• Considering expressive language allows us to see linguistic (and cognitive) evolution as gradual, and also as competitive, and not solely cooperative
• In all this, the emergence of language, in particular proto-grammars, would have played a critical, active role,
  • directly influencing the evolution of the brain.

• “A great stride in the development of intellect will have followed, as soon as the half-art and half-instinct of language came into use.”
  Darwin (1874).